Oestrogen-dependent regulation of gene expression by cohesin in breast cancer

Tanushree Dasgupta

A thesis submitted for the degree of Doctor of Philosophy at the University of Otago, Dunedin, New Zealand.

May 2015
Abstract

Two-thirds of human breast carcinomas test positive for oestrogen receptor α (ERα), which predominantly mediates oestrogenic actions in the normal breast as well in breast tumour cells. Treatment of hormone-sensitive tumours has generally relied on anti-oestrogen therapy. However, ~30% of such tumours do not respond to hormone therapy, and a considerable proportion of tumours develop treatment resistance, ultimately leading to cancer recurrence. Therefore, understanding the regulatory mechanisms controlling tumour cell proliferation downstream of oestrogen is crucial for development of new treatment strategies.

Oestrogen Receptors (ERs) function as part of multi-protein complexes, and the overall cellular response to oestrogen is dependent on cell type-specific functional partners. A previous study identified that ERα shares coincident chromatin binding with the cohesin protein complex in oestrogen-stimulated breast cancer cells. The overlap was striking at oestrogen-regulated genes. The genome-wide concordance in ERα and cohesin binding led to the hypothesis that cohesin facilitates oestrogen-dependent transcription. Following on from this, my research focused on investigating the transcriptional role of cohesin in ER-positive MCF7 breast cancer cells. Cohesin’s role in ER-transcription was tested by cDNA microarray analysis of cohesin (RAD21)-depleted MCF7 cells. Microarray analysis revealed that cohesin (RAD21) depletion affected transcription of a subset of oestrogen-responsive genes, either activating or impeding transcription depending on the gene target. Oestrogen-responsive genes most significantly influenced by cohesin were over-represented in ErbB1 and PI3K/mTOR, pathways associated with breast cancer progression and survival.

I then went on to investigate the mechanism of cohesin’s divergent transcriptional control for two of the microarray-identified candidate genes. Oestrogen activation of SOX4 was robustly enhanced whereas expression of IL20 was significantly reduced in cohesin (RAD21)-depleted cells. Chromatin immunoprecipitation followed by quantitative PCR analyses revealed that RAD21 depletion enriched ERα binding to putative enhancers of SOX4, whereas ERα and RNA Polymerase II binding were weakened at the promoter of IL20, corresponding with the direction of expression of these genes.
I next investigated if a cohesin inhibitor PCI-34051, which is predicted to target cohesin by blocking SMC3 deacetylation, was able to recapitulate transcriptional dysregulations observed upon RAD21-mediated depletion of cohesin. Despite accumulation of acetylated-SMC3, oestrogen responsiveness of selected cohesin-dependent genes remained unaffected. This highlighted that chemical inhibition of SMC3 deacetylation was not comparable with RNAi-mediated depletion of RAD21. Rather, PCI-34051 treatment led to concentration-dependent cell cycle delay, suppression of cell proliferation and induction of cell death in MCF7 cells.

Taken together, these studies have revealed that cohesin can both positively and negatively influence ER-dependent transcription in a gene-dependent manner. Since cohesin-dependent genes participate in breast cancer signalling networks, results from this study suggest that cohesin function might be important for ER-positive breast cancer. PCI-34051 was unable to target the oestrogen-specific transcriptional role of cohesin in MCF7 cells; however, PCI-34051 was significantly oncostatic. Thus, thorough functional characterisation in additional ER-positive model systems would be necessary before considering PCI-34051 as a potential anti-cancer agent for ER-positive breast cancers.
Acknowledgements

There are a number of people I would like to acknowledge, and thank, as their contribution has meant so much during the pursuit of my Doctoral degree and composition of this thesis.

First and foremost, I would like to express my sincere thanks and appreciation to my PhD supervisor, Associate Professor Julia Horsfield. She has been a tremendous mentor. I want to convey my deepest gratitude for all her time, guidance, encouragement and constructive criticisms. These have helped in letting me grow as a research scientist and have made my PhD experience stimulating and productive. Her passion and enthusiasm for research has been inspirational and motivating.

I would also like to thank my co-supervisor Professor Mike Eccles, and PhD committee member Professor Antony Braithwaite and his group, for their scientific advice and insightful discussions about my research.

Sincere thanks goes out to all the past and present members of the Horsfield lab, for being a great support system to help me survive in Grad School. Thank you all for being an excellent team. Special thanks to post-doctoral fellow Dr. Jisha Antony for her advice and assistance in various aspects of the project. I am also thankful to people of the Department of Pathology in general, who were instrumental in making my time at the department enjoyable as a whole. My stay in Dunedin was also enriched by support from a great group of friends, I made here. I am grateful to each one of them for their friendships.

I will forever be thankful for the constant love and encouragement from my former supervisor Dr. Arvind Rai. It was his affectionate and commanding persona, and excellent science communication skills that fortified my interest in biological science. He was and remains my role model for a scientist and mentor. I would also like to thank Dr. Anjana Sharma for her love and support all along the way that made a difference to my life.
I gratefully acknowledge the funding sources Genesis Oncology Trust, Cancer Society of New Zealand, Health Research Council of New Zealand and Breast Cancer Research Trust that made my PhD work possible. I am also thankful to University of Otago for awarding me with a Prestigious Doctoral Scholarship for successful completion of my PhD.

Finally, yet importantly, I would like to acknowledge my family for their immense love and unconditional care. I would fall short of words to express how fortunate I feel to have been raised by such wonderful parents, who have inspired me in all walks of life and supported me in all my endeavours. Mom Dad, the values you instilled in me, gave me strength and helped me get through the toughest of time, even though in reality we were miles apart. Thank you for always believing in me. I am also thankful to my elder brother Suvojit for his love and support over the years, and also for the challenges, it made me stronger and I love him more.

This thesis wouldn't have been a reality without the unwavering support of my husband Rahul. I am immensely indebted to him for the sacrifices he has made vis a vis his medical career in the past 3.5 years or so, in order for me to complete my PhD studies. When I count my blessings, I count him twice! Thank you for appreciating my quirkiness, for being non-judgmental and for always boosting my morale. Thanks for being my biggest critic and my greatest support, for all your patience and positivity, the love and care in good times and bad, and most of all, I want to thank you for always being there. You are my best friend, mentor, soul mate and husband all wrapped in one! I love you beyond words can convey.
I dedicate this thesis to my loving parents, Mandira (Mom) and Mrinal (Dad), and my beloved husband, Rahul, for their unconditional love and support is my greatest blessing in life!
Publication arising from this study

Results published:

This paper can be found in Appendix N
Table of Contents

Abstract ........................................................................................................................................... iii
Acknowledgements ........................................................................................................................... v
Publication from this study ............................................................................................................... viii
List of Tables .................................................................................................................................... xv
List of Figures .................................................................................................................................... xv
List of Abbreviations ....................................................................................................................... xix

Chapter 1: Introduction

1.1 Cancer results from a dysregulated genome/epigenome ............................................................. 1
1.2 Reprogramming of the transcriptome as a consequence of disorganised cancer genome/epigenome .................................................................................................................. 2
1.3 Breast cancer ............................................................................................................................... 3
1.3.1 Classification of breast cancer ............................................................................................... 4
1.3.1.1 Molecular taxonomy of breast cancer .............................................................................. 5
1.3.2 ER-positive breast carcinomas ............................................................................................ 6
1.3.3 Role of oestrogen in breast carcinogenesis ........................................................................... 7
1.3.4 Oestrogen Receptor ............................................................................................................... 9
1.3.5 Transcriptional activity of ER ............................................................................................... 10
1.3.5.1 Genomic activity of ER ................................................................................................. 11
1.3.5.1.1 Direct, ERE-dependent association of ER with target DNA .................................... 11
1.3.5.1.2 Transcriptional, ERE-independent association of ER with target DNA ............ 11
1.3.5.1.3 Pioneer factors ........................................................................................................... 12
1.3.5.1.4 Transcriptional co-regulators ...................................................................................... 13
1.3.5.2 Non-genomic activity of ER ............................................................................................ 15
1.3.6 ERα target gene networks .................................................................................................... 18
1.3.7 Genome-scale analysis of ERα binding sites ....................................................................... 19
1.3.8 Genome-wide analyses of RNA Pol II binding regions ...................................................... 22
1.3.9 Chromosomal looping and regulation by ERα .................................................................... 22
1.4 Cohesin binds at ERα-anchored chromosome loops ............................................................... 23
1.5 Cohesin is a potential mediator of transcriptional response to oestrogen in breast cancer cells .............................................................................................................................. 24
1.5.1 The cohesin complex .......................................................................................................... 24
1.5.1.1 Cohesion-based functions of cohesin ............................................................................ 25
1.5.1.2 Cohesion-independent functions of cohesin.........................................................27
1.5.2 Cohesin’s implications in breast cancer.................................................................29
1.5.3 Cohesin’s role in transcriptional regulation............................................................35
1.5.3.1 Cohesin is a regulator of gene transcription.........................................................35
1.5.3.2 Mechanisms underpinning cohesin-mediated gene regulation............................37
1.5.3.2.1 Cohesin mediates long-distance chromosomal interactions in a ‘CTCF-dependent’ and also ‘CTCF-independent manner’..........................................................37
1.5.3.2.2 CTCF-dependent interactions...........................................................................37
1.5.3.2.3 CTCF-independent interactions.......................................................................39
1.5.3.2.4 Cohesin directly influences RNA Pol II occupancy and transition from pausing to elongation.........................................................................................................40
1.5.3.2.5 Cohesin interaction with epigenetic silencing proteins.....................................41
1.6 Overall aim of this thesis..........................................................................................44

Chapter 2: Materials and Methods

2.1 Materials...................................................................................................................47
2.1.1 Chemicals and Reagents..........................................................................................47
2.1.2 Consumables............................................................................................................49
2.1.3 Solutions and buffers..............................................................................................50
2.1.4 Antibodies...............................................................................................................54
2.1.4.1 Antibodies used for western blotting.................................................................54
2.1.4.2 Antibodies used for ChIP....................................................................................54
2.1.4.3 Antibodies used for immunofluorescence.........................................................54
2.1.5 Enzymes used in molecular techniques.................................................................55
2.1.6 Kits used in molecular techniques...........................................................................55
2.1.7 Primers used............................................................................................................56
2.1.8 Model system used..................................................................................................57
2.2 Cell culture maintenance and manipulation..............................................................57
2.2.1 Reviving cells from liquid nitrogen stock/thawing frozen cells...........................57
2.2.2 Maintenance of cell line.........................................................................................57
2.2.2.1 Sub-culturing/Splitting/Passaging.................................................................57
2.2.2.2 Trypan blue cell counts using a haemocytometer.............................................58
2.2.3 Freezing cells for liquid nitrogen storage/Cryopreservation...............................59
2.2.4 Making of culture medium.....................................................................................59
2.2.5 Charcoal-dextran stripping of FBS for oestrogen-free culture medium.................59
2.2.6 siRNA Transfection ........................................................................................................60
2.3 Gene expression studies ......................................................................................................61
2.3.1 Total RNA isolation ........................................................................................................61
2.3.1.1 Determination of RNA concentration and purity ......................................................61
2.3.1.2 RNA integrity analysis by Bioanalyzer profiling ......................................................62
2.3.2 cDNA synthesis ..............................................................................................................62
2.3.3 Quantitative reverse transcriptase PCR (qRT-PCR) ....................................................62
2.3.3.1 Primer design ............................................................................................................63
2.3.3.2 Primer validation .......................................................................................................64
2.3.3.2.1 Agarose gel electrophoresis ...............................................................................65
2.3.4 Microarray expression analysis ......................................................................................65
2.3.4.1 Sample preparation ..................................................................................................65
2.3.4.2 Data processing and analysis ..................................................................................65
2.4 Protein expression studies .................................................................................................66
2.4.1 Total cell lysate extraction ............................................................................................66
2.4.2 Protein quantitation, SDS-PAGE and Western blotting ...............................................66
2.4.2.1 Protein quantitation- BCA assay .........................................................................66
2.4.2.2 SDS-PAGE and Western blotting .........................................................................67
2.5 Chromatin immunoprecipitation (ChIP) analysis ............................................................68
2.5.1 Principle of ChIP ..........................................................................................................68
2.5.2 ChIP Protocol ...............................................................................................................69
2.5.2.1 Sonication characterisation ..................................................................................70
2.5.3 Detection and quantification of ChIP-ed DNA ............................................................72
2.5.3.1 ChIP primer design and validation .....................................................................72
2.5.3.2 Standard curve based detection and quantification of protein binding at sites of interest ........................................................................................................72
2.6 Immunofluorescence analysis .........................................................................................73
2.7 Cell cycle analysis ............................................................................................................74
2.7.1 Asynchronous cell cycle analysis ...............................................................................75
2.7.2 Synchronised cell cycle analysis ...............................................................................76
2.8 Functional assays .............................................................................................................76
2.8.1 Assessment of cell proliferation ..................................................................................76
2.8.1.1 MTT assay ..............................................................................................................77
2.8.2.2 IncuCyte assessment of cell proliferation .............................................................77
2.8.2 Assessment of cell death........................................................................................................78
2.8.2.1 YOYO-1 staining of DNA.................................................................................................78
2.8.2.2 Trypan blue exclusion assay............................................................................................79
2.8.2.3 Assessment of DNA fragmentation....................................................................................79
2.8.2.4 Hoechst staining of DNA................................................................................................80
2.9 Statistical analysis....................................................................................................................81

Chapter 3: Identifying cohesin-dependent oestrogen-responsive genes in oestrogen receptor-positive MCF7 breast cancer cells

3.1 Introduction..............................................................................................................................83
3.2 Hypothesis................................................................................................................................84
3.3 Aim..........................................................................................................................................84
3.4 Experiment Design...................................................................................................................84
3.5 Generation of independent biological replicates......................................................................86
3.5.1 Optimisation of RAD21 knockdown by RNAi-mediated gene silencing..............................86
3.5.2 Profiling alterations in transcriptional responses to oestrogen with loss of cohesin required a robust knockdown of cohesin................................................................................89
3.5.3 Verification of oestrogen induction by assessing expression of well-characterised oestrogen-regulated genes..........................................................................................................91
3.5.4 Assessment of RNA quality..................................................................................................92
3.5.5 Depletion of cohesin subunit RAD21 slowed cell cycle progression of MCF7 cells, without markedly affecting cellular viability..................................................................................96
3.5.6 RAD21 depletion did not affect cellular growth and morphology........................................98
3.5.7 RAD21 depletion did not affect ERα transcript and protein levels........................................100
3.6 Identification of oestrogen-responsive genes with altered expression upon cohesin depletion.................................................................................................................................103
3.6.1 Microarray data analysis procedures.....................................................................................103
3.6.2 RAD21 depletion altered oestrogen responsiveness of a discrete subset of genes in MCF7 cells.................................................................................................................................104
3.6.3 Functional annotation of the differentially regulated genes.................................................111
3.6.4 Functional network and pathway analysis of differentially regulated genes.........................113
3.7 Discussion................................................................................................................................116
Chapter 4: How does cohesin affect ER target genes?

4.1 Introduction................................................................................................................123
4.2 Aim.................................................................................................................................124
4.3 Results.............................................................................................................................124
4.3.1 Validation of microarray data by qRT-PCR.................................................................124
4.3.2 Experimentally validated candidate genes co-bind ERα and RAD21.........................130
4.3.3 SOX4 and IL20 as candidate cohesin-regulated genes.................................................133
4.3.4 Recruitment of ERα and RAD21 to candidate genes....................................................134
4.3.4.1 Primers designed to determine RAD21 and ERα binding at SOX4 and IL20...........134
4.3.4.2 Characterisation of primer specificity and efficiency...............................................135
4.3.5 Cohesin (RAD21) binding restricts oestrogen activation of SOX4...............................135
4.3.6 Cohesin (RAD21) binding at the promoter facilitates oestrogen activation of IL20........139
4.4 Discussion.......................................................................................................................142

Chapter 5: Targeting cohesin with an inhibitor that disrupts the acetylation cycle of cohesin subunit SMC3

5.1 Introduction......................................................................................................................145
5.2 Aim....................................................................................................................................148
5.3 Results...............................................................................................................................148
5.3.1 HDAC8 inhibitor PCI-34051 blocks deacetylation of SMC3 subunit of cohesin in MCF7 cells ....................................................................................................................................148
5.3.2 PCI-34051 treatment did not affect expression of oestrogen-responsive genes................151
5.3.3 PCI-34051 delayed cell cycle progression of MCF7 cells without altering ERα levels....................................................................................................................................155
5.3.4 PCI-34051 delayed cell cycle progression of MCF7 cells in a concentration-dependent manner ....................................................................................................................................158
5.3.5 PCI-34051 suppressed proliferation of MCF7 cells.......................................................161
5.3.6 PCI-34051 induced cell death in MCF7 cells.................................................................164
5.3.7 Investigating the mechanism of cell death induced by PCI-34051 in MCF7 cells............167
5.4 Discussion.........................................................................................................................171
Chapter 6: Discussion

6.1 Cohesin depletion leads to bidirectional transcriptional response to oestrogen in MCF7 cells

6.2 Cohesin (RAD21) depletion alters transcription of oestrogen-sensitive genes that actively contribute to breast cancer

6.3 Cohesin modulates chromatin binding of proteins important for oestrogen-responsive gene transcription

6.4 Other potential mechanisms for cohesin-dependent influences on ER-associated transcription

6.5 Chemical disruption of SMC3 deacetylation did not mimic oestrogenic changes caused by RAD21 depletion of cohesin

6.6 PCI-34051 delayed progression of cell cycle, suppressed proliferation, and induced cell death in MCF7 cells

6.7 Future Directions

References

Appendix A
Appendix B
Appendix C
Appendix D
Appendix E
Appendix F
Appendix G
Appendix H
Appendix I
Appendix J
Appendix K
Appendix L
Appendix M
Appendix N
List of Tables

Table 1.1 Summary of genome-scale analyses of ERα binding in breast cancer cells........................................21
Table 1.2 Evidence of cohesin’s involvement in breast cancer.................................................................33
Table 2.1 List of common chemicals and reagents..................................................................................47
Table 2.2 List of consumables...............................................................................................................49
Table 2.3 List of general solutions and buffers.....................................................................................50
Table 2.4 List of solutions/buffers used for DNA techniques.................................................................50
Table 2.5 List of solutions/buffers used for western blotting.................................................................51
Table 2.6 List of solutions used for flow cytometry.................................................................................51
Table 2.7 List of solutions used for ChIP..............................................................................................52
Table 2.8 List of buffers used for ChIP................................................................................................53
Table 2.9 List of solutions used for immunofluorescence.....................................................................53
Table 2.10 List of enzymes................................................................................................................55
Table 2.11 List of kits..........................................................................................................................55
Table 2.12 List of primers..................................................................................................................56
Table 2.13 siRNAs used in this study................................................................................................61
Table 3.1 Summary of RNA integrity assessment of samples from the 3 and 6 hour time points.........................................................................................................................................................95
Table 3.2 Summary of significantly differentially expressed transcripts from the 3 and 6 hours array..............................................................................................................................................................109
Table 3.3 Functional annotation of dysregulated genes in oestrogen-stimulated, RAD21-depleted MCF7 cells at the 3 hour time point.................................................................112
Table 3.4 Molecular pathway enrichment analysis of dysregulated genes in oestrogen-stimulated RAD21-depleted MCF7 cells............................................................................................................116
Table 4.1 Dysregulated genes identified by affymetrix cDNA microarray and experimentally-validated by qRT-PCR, in cohesin (RAD21)-depleted MCF7 cells exposed to 6 hours of oestradiol.........................................................................128
Table 4.2 Proximity of RAD21 and ERα binding sites to genes that are dysregulated in oestrogen-stimulated RAD21-depleted MCF7 cells.................................................................132
Table 6.1 Comparison of the consequences of siRNA depletion of RAD21 or blockade of SMC3 deacetylation by PCI-3-4051 in MCF7 cells.................................................................189
List of Figures

Figure 1.1 Domain organisation and homology of oestrogen receptors α and β ..........10
Figure 1.2 Gene regulatory actions mediated by ER.................................................17
Figure 1.3 Cohesin complex and its role in SCC..........................................................28
Figure 1.4 Gene alteration summary for individual cohesin subunits generated using breast cancer datasets from cBioPortal (Cerami et al, 2012; Gao et al, 2013)..............31
Figure 1.5 Transcription oriented functions of cohesin..................................................43
Figure 2.1 Sonication characterisation in MCF7 cells......................................................71
Figure 3.1 Schematic illustration of the experiment design used for sample preparation for cDNA microarray analysis ..................................................................................86
Figure 3.2 Optimisation of cohesin knockdown in MCF7 cells........................................88
Figure 3.3 Verification of RAD21 knockdown in MCF7 cells...........................................90
Figure 3.4 Verification of oestrogen induction by assessing levels of MYC and GREB1, two well-characterised oestrogen-regulated genes.........................................................92
Figure 3.5 Characterisation of RNA intactness by Bioanalyzer profiling..............................94
Figure 3.6 Depletion of cohesin subunit RAD21 delayed cell cycle progression of MCF7 cells.............................................................................................................................97
Figure 3.7 Growth profiles of MCF7 cells depleted of the RAD21 subunit of cohesin...................99
Figure 3.8 RAD21 depletion did not affect oestradiol-induced transcript and protein levels of ERα..................................................................................................................102
Figure 3.9 Transcriptional changes associated with cohesin knockdown in oestradiol-stimulated MCF7 cells...............................................................................................107
Figure 3.10 Scatterplots showing pattern of expression level changes for significantly dysregulated genes of the 3 and 6 hour time points.........................................................110
Figure 3.11 Correlations of the significantly perturbed cancer-associated genes of the 3 hour time point to the underlying biological functions.................................................115
Figure 3.12 Model of ER and growth factor signalling pathway crosstalk in human breast cancer......................................................................................................................121
Figure 4.1 Depletion of cohesin (RAD21) promotes as well as inhibits expression of oestradiol-responsive genes...............................................................................................126
Figure 4.2 Expression changes of seven additional cohesin-regulated candidate genes identified from the microarray analysis.................................................................127
Figure 4.3 Schematic representation of genomic binding sites for RAD21, oestradiol-stimulated RAD21 and ERα, previously identified by ChIP-seq (Schmidt et al. 2010) for SOX4, IL20, BMPR2 and DKK1 .......................................................................................................................... 131
Figure 4.4 Oncoprints of SOX4 and IL20 .......................................................................................................................... 134
Figure 4.5 RAD21 depletion enhanced oestradiol responsiveness of SOX4 ................................................................. 138
Figure 4.6 RAD21 depletion repressed oestradiol responsiveness of IL20 ................................................................. 141
Figure 4.7 Hypothetical model of how cohesin (RAD21) might influence oestrogenic response of SOX4 and IL20 .......................................................................................................................... 144
Figure 5.1 Functions mediated by cohesin in mitotic and non-mitotic cells ................................................................. 147
Figure 5.2 Treatment of MCF7 cells with HDAC8i PCI-34051 results in the accumulation of ac-SMC3 .......................................................................................................................... 150
Figure 5.3 Accumulation of ac-SMC3 in serum-starved MCF7 cells in response to treatment with PCI-34051 .......................................................................................................................... 153
Figure 5.4 PCI-34051 treatment did not affect expression of oestrogen-responsive genes despite accumulation of ac-SMC3 .......................................................................................................................... 154
Figure 5.5 Assessment of cell cycle progression and level of ERα, in oestrogen-starved MCF7 cells treated with 25 μM of PCI-34051 .......................................................................................................................... 157
Figure 5.6 Characterisation of cell cycle kinetics of PCI-34051-treated synchronised MCF7 cell .......................................................................................................................... 160
Figure 5.7 Characterisation of growth suppression induced by treatment with PCI-34051 .......................................................................................................................... 163
Figure 5.8 PCI-34051 induced cell death in MCF7 cells .......................................................................................................................... 166
Figure 5.9 Evaluating the mechanism of cell death induced by PCI-34051 in MCF7 cells .......................................................................................................................... 170
Figure 6.1 Model of how oestrogenic transcriptional responses of cohesin-dependent genes in MCF7 cells might be maintained despite accumulating ac-SMC3 upon PCI-34051 treatment .......................................................................................................................... 188
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>°C</td>
<td>Degree celcius</td>
</tr>
<tr>
<td>3C</td>
<td>Chromosome conformation capture</td>
</tr>
<tr>
<td>Ac</td>
<td>Acetylated</td>
</tr>
<tr>
<td>AF1/2</td>
<td>Activation function 1/2</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>APC/C</td>
<td>Anaphase promoting complex/cyclosome</td>
</tr>
<tr>
<td>APS</td>
<td>Ammonium persulfate</td>
</tr>
<tr>
<td>Aurora B</td>
<td>Aurora B kinase</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CDK1</td>
<td>Cyclin-dependent kinase 1</td>
</tr>
<tr>
<td>CdLS</td>
<td>Cornelia De Lange Syndrome</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary deoxyribonucleic acid</td>
</tr>
<tr>
<td>ChIA-PET</td>
<td>Chromatin Immunoprecipitation coupled with high throughput Paired-End Tag sequencing</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin Immunoprecipitation</td>
</tr>
<tr>
<td>ChIP-qPCR</td>
<td>Chromatin Immunoprecipitation coupled with quantitative PCR</td>
</tr>
<tr>
<td>ChIP-ed DNA</td>
<td>Chromatin immunoprecipitated DNA</td>
</tr>
<tr>
<td>ChIP-on-chip</td>
<td>Chromatin Immunoprecipitation coupled with DNA microarray</td>
</tr>
<tr>
<td>ChIP-PET</td>
<td>Chromatin Immunoprecipitation coupled with high throughput cloning and sequencing using Paired-End diTags</td>
</tr>
<tr>
<td>ChIP-seq</td>
<td>Chromatin Immunoprecipitation coupled with high throughput sequencing</td>
</tr>
<tr>
<td>Ct</td>
<td>Cycle threshold</td>
</tr>
<tr>
<td>CTCF</td>
<td>CCCTC-binding factor</td>
</tr>
<tr>
<td>CTD</td>
<td>Carboxy terminal domain</td>
</tr>
<tr>
<td>DBD</td>
<td>DNA binding domain</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Double distilled water</td>
</tr>
<tr>
<td>DHS</td>
<td>DNase I hypersensitive sites</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s modified eagle medium</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Dinucleotide triphosphate</td>
</tr>
<tr>
<td>DSBs</td>
<td>Double strand breaks</td>
</tr>
<tr>
<td>DSIF</td>
<td>DRB (5,6-Dicholoro-1-β-D-Ribofuranosylbenzimidazole) sensitivity inducing factor</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EcR</td>
<td>Ecdysone receptor</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylene diamine tetraacetic acid</td>
</tr>
<tr>
<td>EGFR</td>
<td>Epidermal growth factor receptor</td>
</tr>
<tr>
<td>ER</td>
<td>Oestrogen receptor (in general)</td>
</tr>
<tr>
<td>ErbB1/2</td>
<td>Erythroblastic leukemia viral oncogene homolog 1 or 2</td>
</tr>
<tr>
<td>ERE</td>
<td>Oestrogen response element</td>
</tr>
<tr>
<td>ERα/β</td>
<td>Oestrogen receptor alpha/beta</td>
</tr>
<tr>
<td>ESCO1/2</td>
<td>Establishment of sister chromatid cohesion</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal bovine serum</td>
</tr>
<tr>
<td>FC</td>
<td>Fold change</td>
</tr>
</tbody>
</table>
FDR  False discovery rate
FoxA1  Forkhead box A1
G  Gravitational force
gDNA  Genomic DNA
GEO  Gene Expression Omnibus
GO  Gene Ontology
GSS  Gene start site
H3K27me3  Histone 3 Lysine 27 trimethylation
H3K27ac  Histone 3 Lysine 27 acetylation
H3K36me3  Histone 3 Lysine 36 trimethylation
H3K4me1  Histone 3 Lysine 4 monomethylation
H3K4me2  Histone 3 Lysine 4 dimethylation
H3K4me3  Histone 3 Lysine 4 trimethylation
H3K9me3  Histone 3 Lysine 9 trimethylation
HCl  Hydrochloric acid
HDAC  Histone deacetylase
HDACi  Histone deacetylase inhibitor
IKB  Ingenuity knowledge base
IPA  Ingenuity pathway analysis
kDa  Kilodalton
LBD  Ligand binding domain
M  Molar
MAPK  Mitogen activated protein kinase
ml  milliliter
mRNA  Messenger RNA
mTOR  Mammalian target of rapamycin
MTT  3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NaCl  Sodium chloride
NaOH  Sodium hydroxide
NCBI  National Centre for Biotechnology Information
NCI  National Cancer Institute
NELF  Negative elongation factor
NIH-DAVID  National Institute of Health – The Database for Annotation, Visualisation and Integrated Discovery
NIPBL  Nipped-B like
nm  nanometer
nM  nanomolar
NTD  N-terminal domain
NRs  Nuclear receptors
NZGL  New Zealand Genomics Limited
P.I  Propidium iodide
PBS  Phosphate buffer saline
PBS-T  Phosphate buffer saline-Tween
PcG  Polycomb group
PCR  Polymerase chain reaction
PDS5  Precocious dissociation of sisters 5
PFA  Paraformaldehyde
PI3K  Phosphatidylinositol 3-kinase
PLK1  Polo-like kinase 1
PP2A  Protein phosphatase 2A
PPEP  Promoter proximal enrichment of RNA Polymerase II
PR    Progesterone receptor
PRC1  Polycomb repressor complex 1
PRE   Polycomb response element
PRO-seq  Precision global run-on sequencing
qPCR  Quantitative polymerase chain reaction
qRT-PCR  Quantitative reverse transcriptase polymerase chain reaction
RBS   Roberts Syndrome
RIN   RNA Integrity Number
RIPA  Radio Immunoprecipitation assay buffer
RNA   Ribonucleic acid
RNA Pol II  RNA Polymerase II
RNAPIIser-5  RNA Polymerase II phosphorylated at serine 5
Rpm   Revolutions per minute
RT    Room temperature
SCC   Sister chromatid cohesion
SDS   Sodium dodecylsulfate
SDS-PAGE  Sodium dodecylsulfate-Polyacrylamide gel electrophoresis
SEM   Standard error of the mean
SGO1  Shugoshin 1
siRNA  Small interfering RNA
SMC1A/3  Structural maintenance of chromosomes subunit 1A/3
SNP   Single nucleotide polymorphism
SRC   Steroid receptor co-activator
TADs  Topologically associating domains
TCGA  The Cancer Genome Atlas
TF    Transcription factor
TSS   Transcription start site
UCSC  University of California, Santa Cruz
V     Volts
v/v   Volume/volume
w/v   Weight/volume
WAPL  Wings apart like
WebGestalt  Web based Gene set analysis tool kit
WHO   World Health Organisation
µl    Microlitre
µM   Micromolar
CHAPTER 1

Introduction

1.1 Cancer results from a dysregulated genome/epigenome

The aetiology of cancer is multifactorial, however, it has now become apparent that cancer is fundamentally genetic. The idea that cancer could be the outcome of a disordered genome was first coined almost a century ago in 1914 by Theodor Boveri, who speculated that chromosomal abnormalities could be the driving force behind a cell’s ability to divide uncontrollably (quoted in Boveri (2008)). The evidence for cancer being a genetic disease, however, came much later from the laboratory of Bruce Ames in the mid-1970s. Experiments using various chemical compounds and evaluation of their ability to induce tumour development in laboratory animals were correlated. Thereafter, it was inferred that the tumour-causing potency of the mutagenic agents was derived from their ability to damage DNA of cells, which in turn served as a carcinogenic stimulus for tumour initiation and progression in those animals (McCann et al, 1975). This discovery brought about a revolutionary paradigm in cancer biology that was later confirmed by subsequent studies (Stehelin et al, 1976; Tabin et al, 1982), each suggesting that cancer is primarily caused by genetic abnormalities.

Simultaneous recognition that the src gene present in the genome of the Rous Sarcoma virus that transmitted tumourigenic properties to avian cells was actually part of the normal avian genome argued against the ‘viral model of cancer origin’. Furthermore, it was suggested that cancer might emerge as a result of alterations in normal cellular genes (Tabin et al, 1982). This line of thinking matured when several viral oncogenes were found to be variants of normal cellular genes or proto-oncogenes. Such an observation was suggestive that genome perturbation could be the initial trigger that unleashes the tumour-promoting functions of proto-oncogenes, which over time metamorphose to cancer-causing oncogenes in transformed cells. The classic example was the src proto-oncogene itself, which led to the momentous discovery that cellular genes, when removed from their normal chromosomal context, could gain the ability to turn oncogenic. This was backed by further evidence when the HRAS gene was found to be oncogenic in cells derived from bladder cancer; however, the same gene was non-oncogenic in normal bladder cells (Parada et al, 1982). Together with evidence from other concurrent studies (Tabin et al, 1982), the idea that cancer is a genetic malignancy was firmly established.
Work spanning the next decade (by mid-1980s) identified the fundamental mechanisms pertaining to cardinal cancer genomic aberrations. These included copy number alterations (Alitalo et al, 1983; Little et al, 1983), point mutations and chromosomal translocations/rearrangements (de Klein et al, 1982; Taub et al, 1982). However, with modern technology leading to the availability of complete sequences of cancer genomes (owing to the launch of The Cancer Genome Atlas (TCGA)), it is now apparent that oncogenic phenotypes do not exclusively arise by alteration of the DNA sequence of cells but also by modulation of the epigenetic landscape (Mueller et al, 2007; Toyooka et al, 2001; Wu et al, 2005). Epigenetics refers to the process(es) that influence cell behaviour resulting from changes in chromosome(s), without altering the primary sequence of DNA (Berger et al, 2009). Thus, despite having the same genetic information, cells are able to develop distinct identities/phenotypes. Epigenetic modifications of DNA are mediated by processes such as DNA methylation, posttranslational modifications of histones, expression of non-coding RNAs and positioning of nucleosomes. Current understanding of cancer is suggestive of a collaborative synchrony between the genome and epigenome of an organism, which when disrupted gives rise to the distinctive cancer phenotypes. Considering the variability in cancer origin, and genomic/epigenomic profiles across and also within different tumour types, the challenge today lies in making sense of this daunting complexity.

1.2 Reprogramming of the transcriptome as a consequence of disorganised cancer genome/epigenome

Genetic alterations at the DNA sequence level or epigenomic level is collectively manifested via reprogramming the associated transcriptome. One could consider the DNA sequence to be the hard drive of life that stores all information required for determining cellular phenotypes. As such, epigenetic modulations could be regarded as the operating system that guides the DNA sequence for temporal, selective functioning to subsequently facilitate processes such as cellular differentiation and embryonic development. The sum of these events direct genome-wide gene expression through the production of transcripts and effector proteins, which subsequently carry out cellular functions, characterize cellular identity and ultimately elicit a biological response. Thus, cancer can be collectively regarded as a heterogeneous malignancy caused due to heterogeneous alterations resulting in deregulation of discrete gene signatures across the genome.
Classical tumour biology suggests that generalised perturbation of gene expression signatures underlie key properties of abnormal cellular behaviour, invasion and metastasis, directly contributing to neoplastic development and progression. Cancer transcriptomes not only serve as a bridge to correlate causative aberrations and cancerous phenotypes allowing one to understand the explicit molecular basis of gene regulation, but also aid in identifying the most significant of all changes and discover tissue-specific novel candidates. Examples of tissue-specific novel candidates include fusion genes in lung carcinoma (Ju et al, 2012; Takeuchi et al, 2012) and colorectal carcinoma (Lipson et al, 2012). Most cancers harbour multiple aberrations but typically only a few are driver events that influence tumourigenesis and maintenance/viability of a given tumour. It is therefore important to identify the ones that contribute to the phenotype. Analysing the whole transcriptome helps identify those candidates. Investigation of cancer transcriptomes has additional advantages as outlined:

(i) Transcriptomic landscapes help in assigning clinical and pathological subtypes to cancerous tissues as these often demonstrate distinct gene signatures. Heterogeneous cancers can be further subdivided into molecular subclasses (for example, different subtypes of breast carcinomas) (Curtis et al, 2012; Perou et al, 2000).

(ii) Transcriptome analysis can also aid in predicting chances of disease recurrence, metastatic potential, treatment response and overall survival (Chia et al, 2012; Kim & Paik, 2010; Knauer et al, 2010; Prat et al, 2011; Reis-Filho & Pusztai, 2011).

(iii) Furthermore, cancer transcriptomes can also be used as part of integrative analysis (pan cancer meta-analysis) across different cancer types, with the idea to simplify complex differential gene signatures by extracting an overriding regulatory or functional module/gene signature common to all (Brooks et al, 2014; Cabanski et al, 2015).

It has thus emerged that alongside aberrations in the genome and epigenome, dysregulation of cellular transcription is one of the fundamental hallmarks of cancer. Upstream cues (from the genome/epigenome) reprogram the downstream transcriptional machinery, which in turn drives gene expression programs that affect tumour initiation, metastasis and drug response.

1.3 Breast cancer

of all new cases of cancer in the United States in 2014 was represented alone by breast cancer, indicating the prevalence of this class of neoplasia. As per the contemporary population-based estimation study of SEER (which is an estimate based on a series of 135,157 breast cancer cases reported to NCI between 1992 to 2001 (Li et al, 2005)) 232,670 new cases were diagnosed and 40,000 women were estimated to die out of breast cancer in 2014. Breakthroughs in research and treatment advancements in the last two or so decades have contributed considerably to improving early detection, prognosis and survival rates among patients. However, because incidence and mortality are still very high, there is an acute need for developing better screening and treatment strategies to further curb this increasingly alarming number.

1.3.1 Classification of breast cancer

Breast cancer is a clinically diverse, highly heterogeneous disease resulting in the malignant transformation of cells of the normal epithelia (most breast malignancies arise from the epithelial tissue). The high degree of heterogeneity is attributed to the existence of a myriad histopathological feature, different molecular makeup, clinical outcomes and variation in treatment responses. In its current form, breast cancer is categorised mainly based on either parameters such as histopathological grades, tumour microenvironment/size, lymph node status or presence of biomarkers. The biomarkers include hormone receptors, \( HER2 \) oncogene, and markers of proliferation, invasion and metastasis.

The histopathologic classification of breast cancer is based on the morphological appearance of tumours and the broad grading classifies breast carcinoma into carcinoma in-situ (limited to breast ducts and lobules) or invasive carcinoma (also called infiltrating carcinomas for its potential to invade lymph nodes and distant sites). The sub-classification includes 20 major tumour types and a further 18 minor subtypes, according to the World Health Organisation (WHO). A major drawback to this type of tumour grading is that it fails to group or dissect out tumours that may have different molecular or clinical profiles despite having the same morphological phenotype. Hence, this sort of classification is unable to reflect the wider heterogeneity of breast cancer. This limits the prognostic and predictive value of histologic classification from the viewpoint of clinical utility.

Molecular classification is thus better equipped to fulfil the requisites of clinically meaningful breast tumour classification, to better tackle complexities of this heterogeneous
Molecularly defined classes of breast tumours were identified through hierarchical clustering of gene expression profiles (Perou et al., 2000) and were found to exhibit distinct biological, epidemiological and clinical features (Sorlie et al., 2001; Sorlie et al., 2003). The advantage to this type of classification resides in its prognostic value and predictive ability. Molecular classification enables prediction of clinical response, such as response to hormonal therapy, chemotherapy alone or in combination, and calculation of risk of relapse. Molecular classification is carried out using a number of techniques, such as:

(i) Quantitative Polymerase Chain Reaction (qPCR) for multigene predictors of clinical outcome analysis, such as the 21 gene recurrence score (RS, Oncotype Dx assay) and 70 gene prognostic profile (MammaPrint) which accurately identifies and quantifies the major molecular markers.

(ii) Fluorescent in-situ hybridisation (FISH) for determining HER2 amplification particularly for primary invasive carcinomas.

(iii) Standard immunohistochemical approaches, especially for hormone receptors (oestrogen/progesterone receptors (ER/PR)) for hormone-responsive carcinomas, detection of p53 mutation by analysis of mutant p53 protein, keeping in mind guidelines laid out by the American Society of Clinical Oncology (ASCO).

However, for routine treatment decision-making, unlike all molecular classes mentioned here, only hormone receptor (ER/PR) and HER2 statuses are currently uniformly accepted prognostic molecular markers.

1.3.1.1 Molecular taxonomy of breast cancer

Based on gene expression profiling initiatives of Perou et al., two distinct molecular classes had emerged in the year 2000 (Perou et al., 2000): oestrogen receptor (ER)-positive and ER-negative. The ER-positive group of tumours were represented by the relatively high expression of genes expressed by breast luminal cells. Hence, these were categorised as luminal breast cancers. Based on variation in gene expression profiles and prognostic outcomes, the luminal group was further subdivided into luminal A and luminal B subtypes (Hu et al., 2006; Sotiriou et al., 2003). The ER-negative group of tumours were also divided into subclasses. The subclasses include basal epithelial-like represented by characteristics of breast basal epithelial cells, ERBB2 (erb-b2 receptor tyrosine kinase 2)/HER2 (Human epidermal growth factor 2) overexpressing tumours characterised by high levels of the oncogene HER2 and normal breast-like, poorly characterised in comparison to the other subtypes (Sorlie et al., 2001).
Using patterns of gene expression, molecular taxonomy recognised five distinct subtypes of breast cancer, which were confirmed by subsequent follow-on studies using larger cohorts of breast tumours taking into consideration additional factors such as ethnic populations as well (Birnbaum et al, 2004; Naderi et al, 2007). The discovery of intrinsic breast cancer subtypes following gene expression profiling was ground-breaking, considering it brought about a complete new perspective to how breast cancer was perceived and classified by researchers, diagnosed and treated by clinicians. In addition to being originally defined by gene expression profiles, the new molecular taxonomic groupings were reproduced using additional platforms such as DNA methylation, exome sequencing, microRNA sequencing, genomic DNA copy number arrays, mRNA arrays and reverse-phase protein assays, in the recent times as well (TCGA, 2012). Taken together, genomics studies of the last decade and a half revealed that breast tumours have distinct molecular portraits. Molecular subtyping is valuable from both biological and clinical perspectives, as it can greatly help in identifying new predictive, prognostic and therapeutic targets for the effective treatment of breast cancer (Carey, 2010).

1.3.2 ER-positive breast carcinomas

ER-positive tumours belong to the most common subtype of breast cancer, the luminal subtype representing approximately 75% of all breast cancers. Luminal tumours express hormone receptors for oestrogen (ER) and/or progesterone (PR), and these hormones drive proliferation of tumour cells. Such hormone-sensitive tumours can either belong to luminal subtype A (representative of 50-55% of breast cancers) or luminal subtype B (representative of 20-25% of all breast carcinomas). Luminal subtype A is characterised by increased levels of ER and ER-responsive genes such as LIV1 (zinc transporter gene), GATA3 (GATA binding protein 3), FOXA1 (Forkhead box protein A1), XBP1 (X-box binding protein 1), BCL2 (B-cell lymphoma 2), ErbB3, ErbB4. Subtype A tumours are also characterised by lower levels of proliferation-related genes such as v-MYB (avian myeloblastosis viral oncogene homolog), GGH (Gama glutamyl hydrolase), CCNE1 (Cyclin E1), LAPTMB4 (Lysosome-associated transmembrane protein-4 beta). Luminal subtype B is characterised by a high proliferative index represented by the Ki-67 expression status (de Azambuja et al, 2007; Urruticoechea et al, 2005) in addition to expressing lower levels of ER/PR (negative or low).
Luminal subtype A is associated with low histologic grades, although more prevalent, it is less aggressive than luminal subtype B and are mostly treated by hormonal-targeted therapy (Guarneri & Conte, 2009; Kennecke et al, 2010). Luminal A cancers have relatively better prognosis, and disease relapse rate is comparatively lower than that of luminal B cancers. In contrast, luminal type B tumours have an aggressive phenotype, have higher histologic grades, increased rate of proliferation and a higher degree of treatment relapse (Creighton, 2012). Recent studies have suggested that luminal B tumours also express higher levels of genes that are participants of the growth factor signalling cascades (Reis-Filho et al, 2010). Hence, increased disease recurrence for these tumours might also be the outcome of growth factor-driven control of tumour cell proliferation, in addition to hormone-dependent cell proliferation. Treatment failure (endocrine resistance) in these tumours can thus be potentially attributed to reciprocal signalling between hormonal and growth factor signalling pathways. Oestrogen is the principal component of ER-positive tumour machinery. Therefore, for more effective tailoring of therapy, it is first crucial to understand the mechanism of action of oestrogenic response in the mammary glands, and then correlate the aetiological processes functioning downstream of the oestrogen pathway, in this subtype of breast cancer.

1.3.3 Role of oestrogen in breast carcinogenesis

Oestrogens are steroid hormones that regulate growth, proliferation and differentiation in both normal and malignant breast cells by facilitating growth of mammary epithelia, branching, alveologenesis and elongation of mammary ducts (Hovey & Trott, 2004; Sommer & Fuqua, 2001). Studies in hormone receptor mutant mouse models have suggested that for effective morphogenesis, sequential activation of hormonal signalling is imperative in the mammary epithelia (Brisken & O'Malley, 2010). Owing to hormonal milieu playing important roles in mammary gland development, any process(es) altering the sensitivity/exposure to hormones can have profound effects on the physiology of mammary tissues in the context of normal development or disease. The reason being, hormones that promote post-natal mammary morphogenesis are the ones that later drive mammary carcinogenesis.

17-β-oestradiol is the principal oestrogenic hormone. The known gene regulatory effects of oestrogen are manifested via interaction with the nuclear receptors (NRs) for oestrogen, ERα and ERβ (Griekspoor et al, 2007). One of the earliest studies establishing a relationship...
between oestrogen and breast cancer reported regression of breast tumour and prevention of tumour recurrence upon removal of the ovaries (Beatson, 1896), the major site for oestrogen production in women prior to menopause. The rationale for such observations became clear much later, with the discovery of the oestrogen-binding protein ER (Jensen et al, 1971). Studies in mouse models have suggested that the proliferative effects of oestrogen are driven by ERα whereas ERβ represses proliferation and is pro-apoptotic (Thomas & Gustafsson, 2011). Recent investigations have revealed that during the transition of normal breast to ductal in-situ carcinoma, ERα-positive cells increase significantly while ERβ-positive cells decrease (Huang et al, 2014). One possible reason for the decline in ERβ-positive cells in such breast cancers could be due to the growth advantage of ERα-positive cells, which allows for clonal expansion (Huang et al, 2014). Consistent with this view, very few breast cancer cell lines demonstrate a significant expression of ERβ. Further progression to the stage of invasive ductal carcinoma is negatively correlated to the status of ERα. Hence, the presence of ERα is regarded as a good indicator for endocrine therapy while loss of ERα is indicative of invasiveness and poor therapeutic outcomes (Herynk & Fuqua, 2007).

Molecular studies over the years have helped to elucidate hormone receptor actions. The gene encoding the ERα protein was isolated from MCF7 human breast adenocarcinoma cells in the late 1980s. Subsequent cloning and identification of the structural domains of the receptor demonstrated that ERα acts as a ligand-dependent transcription factor (TF) (Green et al, 1986; Greene et al, 1986; Walter et al, 1985). This finding brought along the obvious interpretation that since ERα is a TF, genetics must have an important role to play in determining sensitivity to oestrogen and contributing to the mammary phenotype, per se. From then, several studies have confirmed that responsiveness to oestrogen is largely dictated by genetic factors that modulate specific transcriptional programs driven by ERs in concert with additional co-regulators (Carroll & Brown, 2006; Marino et al, 2006; Wall et al, 2014b; Welboren et al, 2009a). Co-regulators in hormone receptor-positive cells were discovered two decades ago (Onate et al, 1995) and represent a class of regulatory proteins that comprise co-activators or co-repressors. Co-regulators are differentially recruited by ERs, which in turn fine-tune gene expression by activation/repression of subgroups of genes in different functional combinations. This can be achieved by modifying chromatin contexts and members of the basal transcriptional machinery. Thus, interaction of oestrogen with its cognate receptor ER can trigger an alteration in target gene activation profile to one
supporting a proliferative phenotype. This results in increased proliferation and ultimately contributes to the tumourigenic mammary phenotype.

### 1.3.4 Oestrogen Receptor

ERs (ERα and ERβ) belong to the superfamily of nuclear receptors and function as ligand-inducible TFs (Griekspoor et al, 2007). Around 50-80% of luminal breast carcinomas express ERα. Hence, understanding the molecular biology of this receptor has been a cornerstone of breast cancer research for over many years. Both ERs (α and β) are modular proteins sharing a considerably large degree of structural homology conserved through evolution, and are characterised by domains specific for DNA-, ligand-binding, dimerisation and transcriptional activation (Nilsson et al, 2001) (Figure 1.1). From amino to carboxy terminals, the principle domains of ERs are:

(i) N-terminal domain (NTD), also called A/B domain, it consists of one of the activation function domain (AF1) which partially contributes to ER’s transcriptional activity and shares minimal homology between ERα and ERβ (15% only) (Kumar et al, 2011; Tora et al, 1989).

(ii) DNA binding domain (DBD), also called C domain, shares the highest degree of homology (95%) between both ERs (Cowley et al, 1997; Kuiper et al, 1996). The DBD domain associates to oestrogen response elements (EREs) of genes either located proximally to the promoter or distally at enhancers (Tsai & O'Malley, 1994). The ERE consists of a palindromic sequence 5′AGGTCAnnnTGACCT3′ separated by three base pairs (Klein-Hitpass et al, 1989; Wood et al, 2001). The ERE guides binding of ER and also recruits co-activators (Brown & Sharp, 1990; Glass & Rosenfeld, 2000).

(iii) Hinge region (or D domain) which harbours the nuclear localisation signal, sequence homology for this domain is only around 36% between ERα and ERβ (Kuiper et al, 1996).

(iv) Ligand binding domain (LBD), also called E-(F) domain, shares a moderate 55% homology between the two ERs, contains the ligand binding pocket, the second activation function domain (AF2), receptor dimerisation interface and facilitates co-regulator recruitment through AF2.

AF1 functions in a hormone-independent manner (Gibson & Saunders, 2012; Kumar et al, 2011), whereas AF2 functionality is dependent on hormones. Hence, ligand-mediated transcriptional activity of ERs is attributed to AF2 domain of the receptors.
Figure 1.1 Domain organisation and homology of oestrogen receptors α and β

Schematic representation of ERα and ERβ showing the structural domains A-F, and the corresponding functional equivalents. Percentage of sequence homology (amino acid identity) between individual domains is indicated (numbers in brown). ERs comprise the N-terminal domain (A/B) containing the transactivation region AF1, DNA-binding domain (C), the hinge domain (D), followed by the C-terminal domain containing the ligand-binding domain and another transactivation region AF2 (E-F). Numbers (in black) represent the position of amino acid residues for individual domains. Figure is partially adapted from Kumar et al (2011).

1.3.5 Transcriptional activity of ER

There are mechanistically distinct molecular modes by which oestrogen elicits its biologic responses (Hall et al, 2001). ER-mediated transcription is the outcome of a coordinated interplay between ER, ligand oestrogen, the ERE sequence, ER interacting additional TFs/co-regulators and the chromatin context. Lipophilic oestrogen passively diffuses through the plasma membrane of target cells and signals interaction with nuclear-localised oestrogen receptors, leading to the subsequent activation of target gene transcription. The type of oestrogen signalling can vary, and can be broadly categorised as ‘genomic’ and ‘non-genomic’ events. The molecular pathways of oestrogenic action or transcriptional activity of ERs are described below:
1.3.5.1 Genomic activity of ER

1.3.5.1.1 Direct, ERE-dependent association of ER with target DNA

Pioneering work by O’Malley and group showed that ERs function as ligand-induced TFs (O’Malley, 2005). Upon binding of oestrogen to LBD of ER, ligand-bound ER binds directly to ERE(s) located proximally in target gene promoters or distally at enhancers. This leads to modification of the chromatin milieu by recruitment of transcriptional co-regulators and members of the basal transcriptional machinery such as RNA Polymerase II (Nilsson et al, 2001). The identification of a classic ERE sequence led many groups to computationally locate EREs within target gene promoter-proximal regions (Bajic et al, 2003). One such study described an important finding that only a small percentage of EREs are present in promoter-proximal regions (Bourdeau et al, 2004). From then, several genome-wide ER binding studies have mapped the majority (70-80%) of ER binding sites to either intergenic or intronic regions, with only a small percentage located proximally to gene promoters (Krum et al, 2008; Lin et al, 2007; Ross-Innes et al, 2010; Welboren et al, 2009a). In addition, many target genes harboured functional EREs that had little resemblance to the consensus ERE sequence, later named as ‘imperfect EREs’ (Curtis & Korach, 1991; Darwish et al, 1991; Slater et al, 1990). Some ERE sites only consisted of half complement of the consensus ERE sequence (AGGTCC only, instead of the palindromic pair), these were named as ‘ERE half-sites’ (Dana et al, 1994; Kato et al, 1995b; Kato et al, 1992). ERα binds with strong affinity to the classic ERE sequence (consensus sequence), however binding affinity weakens when EREs have less sequence identity (Nunez et al, 1987; Walker et al, 1984). This partly explains how ERE sequence can be a pivotal factor in determining the extent of ER-mediated target gene activation (Loven et al, 2001; Wood et al, 2001; Yi et al, 2002). Furthermore, just as ligand binding triggers conformational changes in ER, different ERE sequences are also able to exert specific allosteric effects changing the structural conformations of ERα and ERβ (Hall et al, 2002). Therefore, different ERE sequences are able to influence ER’s interaction with co-regulators and their transcriptional activity as a result (Ikeda et al, 1996).

1.3.5.1.2 Indirect, ERE-independent association of ER with target DNA

For those genes that lack EREs, other TFs can indirectly tether ER via its DBD, and influence target gene expression without direct interaction between ER and DNA. This means of indirect interaction is referred to as the non-canonical or non-classical pathway of
ER signalling. Transcription of nearly one-third (~35%) of human ER regulated genes is accomplished via an indirect association of ER to DNA (O'Lone et al, 2004).

Several TFs have been identified to engage in ‘ER-target DNA indirect binding’. Stimulating protein-1 (Sp-1) is one predominant mediator of oestrogen-sensitive gene transcription. For example, the LDL receptor gene (Li et al, 2001), CCND1 (Cyclin D1) gene (Castro-Rivera et al, 2001), retinoic acid receptor α 1 gene (Sun et al, 1998) are activated via interactions between ER and Sp-1 (Safe, 2001). Upon stimulation by oestrogen, ER signals Sp-1 to bind to its cognate sites and also facilitates recruitment of co-regulatory proteins to mediate target gene transcription (Porter et al, 1997). An example of ERα’s interaction with a protein that suppresses target gene expression is the nuclear factor-κB (NF-κB) protein. Oestrogen stimulation of ERα promotes binding of ERα to NF-κB protein, which is then blocked from binding to the promoter of interleukin-6 gene (IL-6). Thus, ERα represses IL-6 transcription by interacting with the mediator protein NF-κB (Chadwick et al, 2005). Another example of ER adopting protein-protein associations as a means to enhance target gene expression, is via activator protein-1 (AP-1) (Kushner et al, 2000). Genes that have AP-1 response elements are activated when oestrogen stimulation promotes ER binding to Jun/Fos TF complex at the AP1 binding sites (O'Lone et al, 2004). Examples of genes that are regulated via AP-1 are insulin-like growth factor 1 receptor (IGF1R), collagenase and ovalbumin (Fujimoto & Kitamura, 2004).

1.3.5.1.3 Pioneer factors

In addition to above, other factors named as ‘pioneer factors’ such as the TFs forkhead box A1 protein (FoxA1) and GATA3 (putative pioneer factor) have been discovered in recent years. Pioneer factors maintain ERα’s binding to DNA by modulating accessibility to chromatin (Jozwik & Carroll, 2012). Genome-wide TF mapping techniques have revealed that ERα binding regions in breast cancer cells are highly enriched in FoxA1 DNA-binding motifs, with FoxA1 bound to 50% of all ER-bound sites (Carroll et al, 2005). Reducing FoxA1 expression reduced ERα binding to chromatin, and suppressed expression of the ERα associated transcriptome (Carroll et al, 2005; Laganiere et al, 2005). Interestingly, binding sites for FoxA1 were not affected by oestrogen stimulation (Hurtado et al, 2011; Lupien et al, 2008), further implying that FoxA1 contributes to preparing the chromatin milieu before oestrogen-stimulated ERα is recruited to chromatin. This suggests that ERα requires pioneer factors to access distinct regions of chromatin. FoxA1 is able to independently associate and
interact with condensed chromatin, and can modulate chromatin properties in favour of target gene activation (in most occasions) (Cirillo et al, 2002; Cirillo & Zaret, 1999). However, in exceptional cases, FoxA1 also recruits additional factors such as the TLE proteins and makes the chromatin environ inaccessible, impeding transcription as a result (Sekiya & Zaret, 2007). Pioneer factors are capable of altering chromatin states by themselves. Hence, these factors limit requirement for additional chromatin modifying factors and can facilitate dynamic ER-mediated transcriptional outcomes (Hah et al, 2011).

1.3.5.1.4 Transcriptional co-regulators

ER interacts with other co-regulatory proteins, including co-activators or co-repressors, to regulate ER-mediated gene transcription (Hall & McDonnell, 2005; O'Malley & Kumar, 2009; Tremblay & Giguere, 2002). Co-activators are proteins that are recruited by DNA-binding TFs to enhance transcription, whereas co-repressors are proteins recruited by TFs to inhibit transcription. Co-regulators interact with ER in a ligand-dependent manner and allow ER to establish communication with members of the basal transcriptional apparatus, remodelling chromatin to ultimately either facilitate or impede target gene expression (Klinge, 2000). Upon binding to DNA, ERs recruit co-regulators that mostly exist in multi-protein complexes. These complexes carry out biochemical functions, such as preparation of chromatin states, initiation/elongation of transcripts, mRNA splicing, and transcription termination leading to target gene activation/repression (McKenna & O'Malley, 2002; O'Malley, 2006). Enzymatic capacities of co-regulators may include acetylation, deacetylation, ubiquitination, methylation and kinase activity.

Identification of ERα co-regulators has mostly relied on the well-characterised human oestrogen-responsive TFF1 promoter proximal region (Kumar et al, 1987; Sewack & Hansen, 1997; Stack et al, 1988). Mapping of the chromatin structure of the TFF1 promoter has revealed the presence of two nucleosomes, existing over an ERE (400 bp upstream of the TSS) and a TATAA binding site (35 bp upstream of the TSS) (Sewack & Hansen, 1997). Chromatin immunoprecipitation (ChIP) assays have provided further evidence that upon oestrogen stimulation, ERα can direct cyclic recruitment of co-regulators to the TFF1 locus (Metivier et al, 2003). Such co-regulators included histone modifiers (CBP/p300, CARM-1) and chromatin remodellers (SWI/SNF complex) that induced subtle nucleosome movements exposing the ERE and the TATAA box, for the subsequent binding by ERα and the general transcriptional apparatus (Metivier et al, 2003). Further characterisation of the TFF1 locus
have shown that the histone variant H2A.Z and the p400 nucleosome remodelling complex are also recruited to the *TFF1* promoter (Gevry et al, 2009). These changes lead to a preferential nucleosomal shift, a condition that is conducive for the recruitment of the general transcriptional apparatus, commensurate with the oestrogen-induced activation of the *TFF1* gene (Gevry et al, 2009). Furthermore, assessment of other ERα target genes also demonstrated active recruitment of H2A.Z to their promoter-proximal regions in response to oestradiol addition, widening the generality of the findings of the *TFF1* locus (Gevry et al, 2009). Taken together, these studies support the notion that co-regulators that can influence chromatin structure and composition, either directly or by the recruitment of complexes with chromatin remodelling properties, are likely to be crucial for the regulation of gene expression.

The first co-regulator of NRs, SRC-1 belongs to the p160 family of SRC (steroid receptor co-activator) and was cloned and characterised in 1995. SRC-1 was shown to not bind DNA, but rather bind NRs to subsequently mediate their transcriptional activity (Onate et al, 1995). Since then, several co-activators, such as SRC-2 (TIF2/GRIP1), SRC-3 (AIB1/RAC3/ACTR), TRAP/DRIP (Thyroid hormone receptor associated protein/vitamin D receptor-integrating protein) and co-repressors, such as NCoR (Nuclear receptor co-repressor), SMRT (Silencing mediator for retinoid and thyroid hormone receptor) have been discovered. The SRC family of co-activators also recruit other proteins such as CREB-binding protein CBP/p300 and pCAF that have histone acetylation activity (Chakravarti et al, 1996), which modify local chromatin structure to one having an open conformation. The TRAP/DRIP co-activator complex aids to connect ERs with the basal transcriptional machinery. On the contrary, co-repressors such as NCoR and SMRT recruit repressor complexes including histone deacetylases (HDACs) that enforce a repressive chromatin state (Liu & Bagchi, 2004). Other co-regulators having chromatin remodelling capacity are BRM, which is part of the SWI/SNF chromatin modelling complex and is recruited by oestrogen-stimulated ER (Belandia et al, 2002) and PRMT1, which is a member of the protein arginine methyltransferase (PRMT) family. PRMT proteins enhance the transcriptional activity of ERα by virtue of their histone methylation ability (Wang et al, 2001). These studies highlight the collective importance of co-regulators. They add an extra layer of specificity and thus are a pivotal determinant of the final transcriptional outcomes of ER’s activity.
1.3.5.2 Non-genomic activity of ER

Oestrogen-responsive gene transcription occurs over a time period of 180 minutes involving three phases of transcriptional cycling (one unproductive cycle followed by two productive cycles) marked with waves of chromatin modifications and TF associations (Metivier et al, 2003). However, some studies (Hewitt et al, 2014; Song et al, 2005; Warner & Gustafsson, 2006) reported to observe oestrogen-induced effects within a very short time frame (ranging from seconds to minutes) post oestrogen administration, which is too rapid to be resulting from transcriptional events. Hence, oestrogenic events that happen too rapidly to be as a result of transcription are thought to be non-genomic. Such non-genomic effects include activation of kinases, phosphatases and increase in membrane ion fluxes.

The non-genomic actions are frequently associated with activation of protein kinase cascades (Kahlert et al, 2000; Losel & Wehling, 2003), resulting in crosstalk with other growth factor receptors and cytoplasmic signalling pathways (Kato, 2001; Lannigan, 2003). Oestrogen activated ERα was found to directly interact and activate the IGF-1 receptor, contributing to activation of the MAPK signalling pathway. In addition, there was evidence that ERα directly interacts and activates the ErbB2 receptor (Chung et al, 2002) and the EGFR receptor as well (Razandi et al, 2002; Razandi et al, 2003). Debate continues over whether these activities involve nuclear ER (Razandi et al, 1999) or they involve a separate class of membrane-associated ER (extra-nuclear) (Pappas et al, 1995; Toran-Allerand et al, 2002). There has been speculation regarding a distinct role for ER in bridging its non-genomic to genomic actions. For example, ER by virtue of its non-genomic actions, such as by activating protein kinase cascades results in phosphorylation and activation of other TFs, and thus regulates gene transcription in this way. There are examples of several TFs such as Elk-1 (Song et al, 2002), C/EBPβ and CREB (Kousteni et al, 2003) that are phosphorylation targets of the MAP kinase pathway, once MAPK signalling cascade is activated by oestrogen. The transcriptional activity of the AP-1 TF is controlled by MAP kinase-mediated phosphorylation as well (Karin, 1995).

Growth factor signalling molecules can activate ER in a ligand-independent manner by directly phosphorylating ER (Campbell et al, 2001). This is proposed to be the mechanism for emergence of endocrine resistance in breast cancer (Dowsett et al, 2005; Osborne et al, 2005) because of the resultant loss of oestrogen dependence, giving rise to anti-oestrogen resistant tumours. Phosphorylation of serine residues located largely in the AFI domain of
ERα is attributed to the ligand-independent activities of ER (Bunone et al, 1996; Kato et al, 1995a). The list of signal transduction molecules interacting with ER is long since different molecules, and downstream targets may be active at a given time in the target cell. Hence, the collective biologic responses are likely to be diverse and unlikely to be simple.

In summary, it is evident that ER regulates expression of many genes, and such genes belong to two broad categories: those having ERE and those lacking ERE. It so appears that part of indirect regulation of gene expression by ER (via TF), is the outcome of the integration of its genomic and non-genomic activities. ER’s modulation of TF functions can occur in two distinct ways: first, by co-binding to DNA and second, by activation of cytoplasmic signal transduction cascades leading to activation of TFs (Bjornstrom & Sjoberg, 2005). The distinct modes by which ER exhibits its regulatory control is illustrated in Figure 1.2.
Figure 1.2 Gene regulatory actions mediated by ER

The canonical pathway involves binding of oestrogen to nuclear ER leading to its successive activation, followed by direct association to ERE of target genes, thereby facilitating gene expression. The non-canonical pathway can be mediated via ligand-dependent or ligand-independent ways. In the ligand-dependent pathway, upon oestrogen activation of ER, ER tethers to other DNA-bound TFs (such as Sp-1, AP-1) via protein-protein interactions to indirectly bind target genes lacking ERE sequences. Expression of non-ERE containing genes is thus mediated by ER’s indirect interaction with target DNA through binding of specific TFs. The ligand-independent pathway involves phosphorylation and successive activation of ER by protein kinases, activated via growth factor receptor-mediated signalling cascades. Both canonical and non-canonical pathways are together regarded as the genomic pathways of ER’s activity because nuclear ER is associated (directly or indirectly) with regulation of gene expression. In addition to genomic pathways, there exists a non-genomic pathway, wherein oestrogen activated extra-nuclear ER alters protein functions in the cytoplasm, such as cAMP production (Aronica et al, 1994), mobilisation of intracellular calcium (Improta-Brears et al, 1999) and also phosphorylates other TFs (such as Elk-1, C/EBPβ and CREB) by activation of protein kinases. Phosphorylated TFs then bind DNA in return and subsequently modulate expression of the associated genes. Figure is partially adapted from Heldring et al (2007).
1.3.6 ERα target gene networks

Many attempts were made to map the gene regulatory circuits controlled by ERα across several oestrogen-target tissues. In earlier years, elucidation of ERα target genes was a daunting task as expression analysis between oestrogen-induced and non-induced cells was carried out on a gene-by-gene-basis (Brown et al, 1984; Jakowlew et al, 1984). However, recent genome-wide gene expression analysis tools such as microarrays and high-throughput sequencing technologies have contributed immensely to mapping the gene networks controlled by oestrogen, not only in different oestrogen target tissues but also in disease models (including breast cancer). The rationale behind such expression profiling studies was to define the relationship between regulatory inputs (ERα), and cellular responses (breast cancer phenotype in this case) by mapping changes in gene activity.

Identification of primary gene targets of ERα along with downstream transcriptional cascades that stem from primary targets is a prerequisite to understanding oestrogen-driven breast cancer. Therefore, several groups have pursued mapping the ERα transcriptome in breast cancer cells, including breast tumour samples and cell lines such as MCF7, T47D, ZR-75-1. cDNA microarrays in particular, have been extensively used in the last one and half decades or so for profiling transcriptomic changes upon stimulation by oestrogen in breast cancer cells (Cunliffe et al, 2003; Frasor et al, 2003; Gruvberger et al, 2001; Inoue et al, 2002; Lin et al, 2004; Soulez & Parker, 2001). In one such study, Frasor and colleagues observed a general upregulation of pro-proliferative genes (and their regulators) and downregulation of anti-proliferative and pro-apoptotic genes in breast cancer cells upon stimulation by oestrogen (Frasor et al, 2003). Although large amounts of microarray data were generated by a number of studies, comparison of data across laboratories was difficult due to experimental variables such as different platforms, oestrogen treatment times and analysis strategies. However, these initial inventories of ER-responsive genes still revealed an intrinsic genetic signature representative of ERα-positive tumours of the breast (Lin et al, 2004; Zhu et al, 2006). To resolve discrepancies relating to gene expression signatures elicited by oestrogen, Ochsner and colleagues developed a web resource called GEM (Gene Expression Meta-signatures) in 2009 (Ochsner et al, 2009). This resource comprised a meta-analysis combining expression data from ten independent published datasets, which was used to generate an over-arching oestrogen-responsive gene signature representative of collective evidence from all datasets. Most of these studies were based on quantifying expression levels of genes post stimulation by oestrogen, which helped identify the
complement of genes regulated by ERα, although without distinguishing between ERα’s direct and indirect targets. A large proportion of ERα target genes lack the palindromic ERE sequence (many contain imperfect EREs, ERE half-sites) and are regulated indirectly as discussed in section 1.3.5.1.2. Therefore, combining global gene expression data with mapping of ERα-DNA binding regions was considered next for distinguishing direct versus indirect ERα targets.

1.3.7 Genome-scale analysis of ERα binding sites

Chromatin Immunoprecipitation (ChIP) assays are the cornerstone of mapping genome-wide protein-DNA associations including the binding of trans-acting factors such as ERα. Four modifications of the ChIP technique have been widely used in detecting ERα binding to chromatin in breast cancer cells. These are:

(i) ChIP combined with DNA microarray technology (ChIP-on-chip) (Carroll et al, 2005; Carroll et al, 2006; Madak-Erdogan et al, 2011; Quintin et al, 2014)

(ii) ChIP combined with massively parallel DNA-sequencing technology (ChIP-seq) (Li et al, 2013; Ross-Innes et al, 2010; Welboren et al, 2009b)

(iii) ChIP combined with high-throughput cloning and sequencing using paired-end diTags (ChIP-PET) (Lin et al, 2007)

(iv) ChIA-PET, a proximity ligation technique aimed to map long-range chromosomal interacting regions (in addition to mapping binding of ERα) combining ChIP with high-throughput paired-end tag sequencing (Fullwood et al, 2009).

Consistent with disparity observed with oestrogen-induced gene expression data, discrepancies were also reported in the number of ERα binding regions identified by these studies. Variation in ERα binding was attributed to experimental variables across studies such as the use of different technology for mapping DNA enrichments, use of different ERα antibodies, oestrogen-stimulation times/concentration used, and cell-culture conditions. Despite incomplete overlap, the ChIP studies generated the surprising finding that ERα is located mostly at regions far away from the transcription start sites (TSS) throughout the genome. As mentioned earlier in section 1.3.5.1.1, most ERα binding sites (70-80%) occur distally from promoters and only a minor fraction are confined to promoter-proximal regions (7%) (Biddie et al, 2010; Gao & Dahlman-Wright, 2011). These studies established an important concept that cis-regulatory elements can be located in gene promoters, in the introns or even at the 3’ region. A summary of published genome-scale ERα binding site analyses carried out in breast cancer cells is shown in Table 1.1. The data suggest that ERα
interacts with enhancers to mediate long-range chromosomal interactions rather than bind to promoters as previously perceived.

The identification of distal ERα binding sites was further perplexing because there was a great degree of discordance between the number of identified ERα binding sites and the number of differentially oestrogen-regulated genes. There were many more ERα binding sites than there were oestrogen-responsive genes (as identified from expression studies), implying that more than one binding site regulates the same ERα target gene. It is possible that only a subset of binding sites are functionally active under specific experimental conditions and in the cell type used.

To gain a more precise read out of oestrogen-dependent transcription, mapping of RNA Pol II occupancy in response to oestrogen treatment was used by some groups as part of their mapping studies (Carroll et al, 2006; Welboren et al, 2009b). Known to transcribe all protein-encoding genes, RNA Pol II on gene bodies had the potential to provide a direct readout of transcriptional activity improving the correlation between binding to regulation.
Table 1.1 Summary of genome-scale analyses of ERα binding in breast cancer cells

<table>
<thead>
<tr>
<th>Published study</th>
<th>Breast cancer model system</th>
<th>ERα detection technique</th>
<th>Total detected ERα binding sites</th>
<th>Topological distribution of ERα</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carroll et al (2005)</td>
<td>MCF7</td>
<td>ChIP-on-chip (Chromosome 21 and 22)</td>
<td>57 sites</td>
<td>Almost all are in non-promoter proximal regions, spanning up to &gt;100 kb from TSS</td>
</tr>
<tr>
<td>Carroll et al (2006)</td>
<td>MCF7</td>
<td>ChIP-on-chip</td>
<td>3665 sites</td>
<td>Only 4% are within 1 kb upstream of TSS</td>
</tr>
<tr>
<td>Lin et al (2007)</td>
<td>MCF7</td>
<td>ChIP-PET</td>
<td>1234 sites</td>
<td>Only 5% map to promoter proximal regions within 5 kb upstream of TSS, 38% map to intronic regions</td>
</tr>
<tr>
<td>Fullwood et al (2009)</td>
<td>MCF7</td>
<td>ChIA-PET</td>
<td>14,468 sites</td>
<td>NA</td>
</tr>
<tr>
<td>Welboren et al (2009)</td>
<td>MCF7</td>
<td>ChIP-seq</td>
<td>10,205 sites</td>
<td>Only 7% are located proximally to promoters within 5 kb upstream of TSS, 41% map to introns, 23% are located distally from genes</td>
</tr>
<tr>
<td>Madak-Erdogen et al (2010)</td>
<td>MCF-7</td>
<td>ChIP-on-chip</td>
<td>4547 sites</td>
<td>Only 5% located in proximal promoter regions within 1 kb of TSS, 50% located to intergenic regions, 30% to intronic regions, 3-7% in exons</td>
</tr>
<tr>
<td>Ross-Innes et al (2010)</td>
<td>MCF7</td>
<td>ChIP-seq</td>
<td>14,505 sites</td>
<td>NA</td>
</tr>
<tr>
<td>Li et al (2013)</td>
<td>MCF7</td>
<td>ChIP-seq</td>
<td>31,052 sites</td>
<td>Only 2.9% located proximally to promoters, 23% mapped to potential enhancers</td>
</tr>
<tr>
<td>Quintin et al (2014)</td>
<td>MCF7</td>
<td>ChIP-on-chip (in a 2 Mb genomic region including the TFF gene locus)</td>
<td>345 sites 316 sites</td>
<td>NA</td>
</tr>
</tbody>
</table>
1.3.8 Genome-wide analyses of RNA Pol II binding regions

Many groups had reported that 12-30% of genes responding to stimuli display ‘Promoter Proximal Enrichment of RNA Pol II’ (PPEP) (Core et al, 2008; Kininis et al, 2009; Muse et al, 2007). RNA Pol II mapping by Carroll and group (Carroll et al, 2006) identified that out of the unique 3,629 binding events mapped for RNA Pol II in oestrogen-stimulated MCF7 cells, 67% were enriched at promoter-proximal regions of known genes which correlated better to the number of differentially expressed genes, in contrast to the corresponding ERα binding sites. However, Welboren and group (Welboren et al, 2009b) showed that although RNA Pol II occupancy increased significantly for 596 genes (59% upregulated and 41% downregulated) in oestrogen-stimulated MCF7 cells, only 21 genes (4%) displayed PPEP. Hence, PPEP may not be a general feature of rapid oestrogen-initiated gene regulation, although it might be a feature of a small number of oestrogen-responsive genes. Despite discrepancies, assessing RNA Pol II binding is considered a good reflection of oestrogen-instigated expression of the corresponding gene, and continues to be studied (Wang et al, 2010; Zhang et al, 2014).

1.3.9 Chromosomal looping and regulation by ERα

In recent years, long range chromosomal looping that juxtaposes gene promoter and regulatory elements has emerged as a potential mechanism for ERα-mediated transcription (Deschenes et al, 2007; Fullwood et al, 2009; Pan et al, 2008; Theodorou et al, 2013). Using chromosome conformation capture (3C), it was demonstrated that multiple ERα binding sites engage in the formation of chromosomal loops (often involving cis-interactions bridging distal enhancer-promoter regions) to regulate transcription at many ERα target gene loci, such as TFF1 (Pan et al, 2008), CYP1B1, GREB1 (Deschenes et al, 2007), NRIP1 (Carroll et al, 2005), BCL2 (Perillo et al, 2008). A high-throughput genome-wide chromatin-interaction study made use of an unbiased approach called ‘ChIA-PET’ to generate the first ‘human ERα chromatin interactome map’ in oestrogen-treated MCF7 cells. In this study, high confidence distal (to gene promoters) ERα localisation sites were anchored/secured at gene promoters via long-range interactions revealing that ERα acts by extensive chromosomal looping to bring enhancers together, for orchestration of transcription. Genomic distribution of H3K4me3 (histone modification representative of active promoters), FoxA1 (cell type-specific chromatin modifier) (Meyer & Carroll, 2012) and RNA Pol II was also mapped to test the functional significance of ERα-bound chromatin interactions. Enrichment of RNA Pol II was significantly higher at ERα sites involved in promoter-proximal or promoter-distal
interactions than at idle ERα sites (sites not involved in interactions). However, H3K4me3 was enriched at promoter-proximal ERα sites only but not at ERα sites distal to gene promoters (potential enhancers), irrespective of their chromatin interaction status. Conversely, FoxA1 binding was most enriched at distal (to gene promoters) ERα interacting sites. DNase I hypersensitive sites (DHS) are representative of chromatin accessibility, and could result from the pioneering activity of FoxA1. DHS are an important measure of chromatin’s accessibility as functional regulatory elements (such as promoters, enhancers, insulators and locus control regions) exhibit hypersensitivity to nuclease. Evidence from genome-wide mapping of DHS in multiple cell lines have suggested that cell type-specific DHS are localised to transcriptional enhancers (Xi et al, 2007).

Taken together, this de-novo analysis of genome-wide ERα chromatin interactions reflected that communication with distal (to gene promoters) ERα binding sites is orchestrated by long-range looping. Such looping is guided by binding of cell type-specific chromatin modifying factors. FoxA1 being one such factor (Laganiere et al, 2005) and possibly with other uncharacterised co-factors, thereby leading to stabilised, active transcriptional machinery for co-ordinately regulating target gene expression.

1.4 Cohesin binds at ERα-anchored chromosome loops

In 2010, ChIP-seq conducted in oestrogen-stimulated MCF7 cells found a previously unrecognised connection between ERα and the chromosome cohesion protein, cohesin (Schmidt et al, 2010). Schmidt and colleagues identified multiple regions across the genome of breast cancer cells where ERα and cohesin shared overlapping binding. Cohesin depletion prevented oestrogen-stimulated re-entry into the cell cycle, demonstrating cohesin’s influence on physiological response to oestrogen in breast cancer cells. Moreover, when chromosomal interaction data from ERα ChIA-PET (Fullwood et al, 2009) was combined (discussed in section 1.3.9), a striking concordance between ERα and cohesin at oestrogen-regulated genes was observed, particularly those involving ERα-bound anchors of interacting chromatin loops. This discovery was particularly striking because it raised the possibility that alongside FoxA1, cohesin may also be involved in tethering chromatin interactions required for expression of key target genes. Previous studies have shown that cohesin is required for chromosomal interactions mediated by the 11-zinc finger chromatin binding protein CTCF (CCCTC-binding factor) (Hadjur et al, 2009; Hou et al, 2010). Hence,
cohesin’s putative involvement in mediating chromatin interactions of ERα seemed like a reasonable possibility.

Around the same time, another study carried out in *Drosophila* salivary glands, found cohesin to be directly regulating expression of a class of genes belonging to the ecdysone steroid hormone-signalling pathway. As evidence for direct regulation by cohesin, a large proportion of differentially expressed (following cohesin cleavage) ecdysone-responsive genes harboured binding sites for the Rad21 subunit of the cohesin protein. This study was further suggestive that cohesin is perhaps a general mediator of transcriptional response to steroid hormones, and that this role is conserved through evolution. Furthermore, work by our group provided additional evidence that oestrogen-responsive *MYC* oncogene, which is a known target of ERα (Butt et al, 2005) is directly and positively regulated by the cohesin protein (Rhodes et al, 2010). Taken together, these studies suggested strongly that cohesin has a so far uncharacterised role in transcriptional regulation of steroid hormone sensitive genes including ERα-responsive genes. Although the ChIP-seq and ChIA-PET studies in MCF7 cells have provided an important link between cohesin and ERα, how cohesin might regulate ERα driven gene expression signatures in breast cancer has not been investigated.

**1.5 Cohesin is a potential mediator of transcriptional response to oestrogen in breast cancer cells**

**1.5.1 The cohesin complex**

The ability to duplicate and pass a copy of its genome to daughter cells in a carefully regulated process is one of the hallmarks for continuation of life. The precise halving of the duplicated genome is primarily mediated by a multi-subunit protein complex, called cohesin, which is evolutionarily conserved from yeast to humans (Nasmyth & Haering, 2009). Cohesin protein was originally identified in yeast as a complex that holds pairs of duplicated chromosomes called sister chromatids together, from the time of their generation in S phase until their separation in anaphase (Guacci et al, 1997; Michaelis et al, 1997). The physical connection between sister chromatids (along chromosomal arms and centromeres) is vital because it helps secure the ties between sister kinetochores and microtubules of mitotic spindles to opposite spindle poles in a process called ‘chromosome biorientation’. This in turn safeguards faithful chromosomal segregation. Thus, cohesin is essential for correct sister chromatid segregation, and for maintenance of integrity of the genome.
The human cohesin complex comprises of four subunits: two ‘structural maintenance of chromosomes’ (SMC) subunits (SMC1A and SMC3) and two non-SMC subunits, called RAD21 and STAG1/2, which form a ring-like structure that holds sister chromatids during and after DNA replication (Figure 1.4A). In addition to the core cohesin complex, several auxiliary factors are required to regulate the coordinated, temporal loading, maintenance and unloading of cohesin onto chromatin (discussed in 1.5.1.1). Research over the last one and half decades or so has provided evidence that cohesin has additional roles in the interphase nuclei, such as DNA replication, DNA damage repair, organisation of interphase chromatin, chromatin remodelling and regulation of gene expression (reviewed in Losada (2014); Remeseiro et al (2013)). Depending on the nature of cohesin-mediated functions, these can be broadly categorised as ‘cohesion-dependent functions’ (pertinent to its genome duplication function) and ‘cohesion-independent functions’ (pertinent to its role in regulation of gene transcription).

1.5.1.1 Cohesion-based functions of cohesin

As mentioned in section 1.5.1, the cohesin complex encircles the newly generated sister chromatids from S phase until their physical separation in anaphase. Intriguingly, how cohesin might encircle the two sister chromatids remains controversial. One school of thought proposes that either a single or two cohesin rings topologically entrap the sister DNAs, currently known as ‘the embrace or the ring model’, without establishment of direct binding to DNA (Haering et al, 2008). In the strong ring model (one-ring model), a single cohesin ring entraps both the sister DNAs. However, in the weak ring model (two-ring model), also referred to as the ‘handcuff model’, each of the two sister DNAs is entrapped by two separate cohesin rings interconnected by interactions established between STAG1/2 and RAD21 subunits of the participating cohesin rings (Zhang et al, 2008). However, a second school of thought argues against this topological entrapment of sister DNAs model and rather proposes that more than one cohesin ring form oligomers and entrap the sister DNAs by establishing specific protein-DNA interactions, either inside or outside of the cohesin rings (Onn et al, 2008). In both the situations, integrity of the cohesin ring is considered pivotal for the prevention of premature chromatid separation. For the establishment of sister chromatid cohesion (SCC), the cohesin complex collaborates with a number of accessory and regulatory proteins. These proteins coordinate the timely establishment of SCC until metaphase, followed by its dissolution/loss at the onset of
metaphase-anaphase junction allowing effective segregation of sister chromatids in anaphase (extensively reviewed in Nasmyth & Haering (2009); Peters et al (2008)).

Cohesin is loaded onto chromatin in telophase of the preceding cell cycle by the cohesin-loading heterodimer NIPBL (Nipped-B like)-MAU2. Upon progression to G1 phase, additional factors like PDS5 (Precocious dissociation of sisters 5) and WAPL (Wings apart-like) bind to chromatin-bound cohesin and trigger cohesin unloading (Tedeschi et al, 2013). Hence, the amount of cohesin that stays on chromatin is outcome of the opposing actions of NIPBL-MAU2 and PDS5-WAPL. Until DNA is replicated, chromatin-bound cohesin is not cohesive. Cohesion is established by ESCO1/2, which acetylates two lysine residues of the SMC3 domain after passage of the DNA replication fork. Binding of another factor called Sororin to PDS5 is essential for effective cohesion establishment because Sororin is proposed to displace WAPL, thereby antagonising its cohesin-unloading action (Nishiyama et al, 2010). At the onset of mitosis, most cohesin from chromosomal arms is removed during prophase in a cleavage-independent process called ‘prophase dissociation pathway’ by the collaborative action of three kinases, namely ‘polo-like kinase 1’ (PLK1), ‘aurora B kinase’ (Aurora B) and ‘cyclin-dependent kinase 1’ (CDK1). Centromeric cohesin is however protected for allowing chromosome alignment in the metaphase plate, catalysed by the action of BUB1 through shugoshin 1 (SGO1) and protein phosphatase 2 (PP2A). During the metaphase-anaphase transition, anaphase-promoting complex/cyclosome (APC/C) becomes activated leading to degradation of Securin and activation of Separase. Separase then cleaves the RAD21 subunit of centromeric cohesin, thereby releasing sister DNA molecules via destruction of cohesin integrity. SCC is illustrated in Figure 1.3B. In addition to SCC, the cohesive form of cohesin plays important roles in repair of DNA double-strand breaks (DSBs) by homologous recombination (Cortes-Ledesma & Aguilera, 2006; Heidinger-Pauli et al, 2010), and stabilisation/restart of stalled DNA replication forks, especially at regions that are difficult to replicate such as telomeres (Remeseiro et al, 2012). Repair of DNA damage is mediated either by specific recruitment of cohesin to the chromatin region surrounding the damage, i.e., at the break site itself or the break is repaired while sister chromatids are being held as part of cohesin’s chromatid separation process (using sister chromatids as a template to repair DSBs) (Kim et al, 2002a; Strom et al, 2004; Unal et al, 2004). Thus, cohesive cohesin contributes to chromosome stability by ensuring correct chromosome biorientation for faithful chromatid segregation, prevention of collapse of replication fork machinery and facilitating its restart, and also by efficient DSB repairs.
1.5.1.2 Cohesion-independent functions of cohesin

Compelling evidence from several studies (Darwiche et al, 1999; Rollins et al, 1999) have unravelled functionally separate roles for cohesin (Monnich et al, 2009; Pauli et al, 2008; Schuldiner et al, 2008; Seitan et al, 2011). Some studies showed that the transcriptional role of cohesin in post-mitotic cells is exclusive of cohesin’s role in SCC. Subsequent genome-wide analyses of cohesin binding revealed that much of cohesin co-localizes with the insulator protein CTCF (Parelho et al, 2008; Rubio et al, 2008). More recently, both these proteins have been identified to contribute to intra- and inter-domain chromosomal organisation in the nucleus (Gibcus & Dekker, 2013; Sofueva et al, 2013; Zuin et al, 2014). Moreover, cohesin has also been shown to co-purify with the mediator complex, which itself is crucial for transcription (Kagey et al, 2010). Cohesin’s role as a chromatin architectural organiser underlies most facets of its role in genome function, whether that involves formation of chromosomal loops for long-range transcriptional regulation, control of RNA Pol II activity or alteration of local chromatin structure.
Figure 1.3

A. Cohesin protein complex

B. Facilitates removal of cohesin from chromosome arms during 'Prophase'.

- S Phase: ESCO1/2
- Binding of sororin to PDS5
- Cohesion on chromatid Chromosome in 'G1 Phase'

PCS5 retains centromeric cohesin

- Activation of cohesin for cohesion establishment between sister chromatids

PLK1, Aurora B, CDK1, WAPL

- Protection of centromeric cohesin from prophase removal pathway

SGO1, PP2A

- Activation of Separase

BUB1

- Separate cleavage of RAD21 subunit of cohesin cleaves centromeric cohesin in 'Anaphase'

APC/C

- Degradation of Securin

Faithful separation of sister chromatids
1.3 Cohesin complex and its role in SCC

(A) Architecture of the human cohesin complex. In somatic cells, cohesin is composed of SMC1 (SMC1A), SMC3, RAD21 and STAG1/2 subunits. (B) Cohesin and its auxiliary factors (associated proteins and regulators) are shown superimposed on the sister chromatid separation cycle. Non-cohesive cohesin (represented by light green ovals) is preloaded onto chromatin in telophase of the preceding cell cycle. Upon DNA replication, cohesive cohesin (represented by dark green ovals) holds replicated sister chromatids from S to metaphase, until their separation occurs in anaphase. Several proteins (highlighted in red) control the timing and location of cohesin binding to chromatin, activation of cohesiveness and dissolution of cohesion post effective separation of sister chromatids.

1.5.2 Cohesin’s implications in breast cancer

Over-, under-expression and mutations in cohesin subunits have been associated with cancer. In breast cancers, deregulation of cohesin occurs mostly via amplification/increase in gene copy number of the RAD21 subunit (near chromosomal location 8q24). However, a small number of mutations (missense, truncating and inframe), deletion as well as amplification of other cohesin subunits (SMC1A, SMC3, STAG1/2) are also observed in breast tumours (Figure 1.4). Over the years, strong links have emerged linking cohesin function to breast cancer pathophysiology and prognosis (Table 1.2). A meta-analysis of gene expression data from clinical cancer specimens demonstrated that enhanced RAD21 expression was a feature of undifferentiated cancers of the breast alongside ovarian, lung and bladder cancers (Rhodes et al, 2004). In the same year, another large scale breast cancer case-control study used clinically diagnosed breast cancer cases and identified an intronal single nucleotide polymorphism (SNP) in the RAD21 subunit of cohesin, which provided evidence that RAD21 is significantly associated with susceptibility to breast cancer (Kammerer et al, 2004). This was later confirmed by another SNP-breast cancer risk association study (Sehl et al, 2009), which identified three polymorphisms associated with breast or breast and ovarian cancers to be located in intronic regions of the RAD21 gene. Abrogation of RAD21 in breast cancer cell lines as well as in clinical breast cancer samples reduced cellular proliferation and increased sensitivity to DNA-damaging chemotherapeutic agents (Atienza et al, 2005; Xu et al, 2011b). Interestingly, RAD21 silencing was also found to result in DNA damage and enhanced apoptosis in cells overexpressing the MYC oncogene (Toyoshima et al, 2012). Overexpression of RAD21 was found to be associated with aggressive breast cancers, often leading to early relapse and delayed/poor prognosis for breast cancer patients (Oishi et al, 2007; van ’t Veer et al, 2002; Xu et al, 2011b; Yan et al, 2012). In the study carried out by Xu and co-workers, RAD21 overexpression correlated with large tumour size and shorter relapse-free survival particularly for grade 3 tumours (often characterised by a high
proliferative index). However, RAD21 expression was significantly lower in invasive tumours. These results suggest a proliferative role for RAD21 in grade 3 tumours of the breast. Additionally, RAD21 was identified as one of the five genes upregulated significantly by BRCA1 (tumour suppressor), when control MCF7 cells were compared with MCF7s ectopically expressing BRCA1 (Atalay et al, 2002). Furthermore, functional interactions between chromosome cohesion proteins and BRCA1 were identified, which could be of particular relevance because RAD21 and BRCA1 are both involved in DNA damage repair (Skibbens, 2008; Skibbens et al, 2008). This could further imply that parts of BRCA1’s role in maintenance of chromosome integrity could be mediated partially by RAD21. The newly recognised functional interactions between RAD21 and BRCA1 could thus be indicative of co-operative roles played by the two proteins in breast tumourigenesis, perhaps.

Taken together, these studies provide sufficient justification for a putative functional role of cohesin in breast tumourigenesis. Because of its role in SCC, aneuploidy and genome instability is often assumed to be the pathological driver of cancers with cohesin alterations. Indeed, this might be the case for cancers with insufficient cohesin to mediate faithful SCC. However, in breast cancers where cohesin (RAD21) is present in excess, transcriptional dysregulation of downstream genes resulting from amplified cohesin, could also be potentially contributing to tumourigenesis.
**Figure 1.4**

A. Gene alteration summary of cohesin subunits across multiple breast cancer datasets

- **RAD21**
- **SMC1A**
- **SMC3**
- **STAG1/2**

B. Oncoprints of cohesin subunits across Breast Invasive Carcinoma (TCGA Provisional) dataset

- **RAD21**
- **SMC1A**
- **SMC3**
- **STAG1/2**
Figure 1.4 Gene alteration summary for individual cohesin subunits generated using breast cancer datasets from cBioPortal (Cerami et al, 2012; Gao et al, 2013)

(A) Gene alteration summary of all cohesin subunits across available breast cancer datasets on cBioPortal. Alterations in RAD21, SMC3, SMC1A, STAG1/2 were annotated across all available breast cancer datasets on cBioPortal. The six datasets include: Breast invasive carcinoma (British Columbia, Nature, 2012) containing 65 tumour samples; Breast cancer xenografts (British Columbia, Nature, 2012) containing 29 patient xenografts; Breast invasive carcinoma (Broad, Nature 2012) containing 103 tumour samples; Breast invasive carcinoma (Sanger, Nature 2012) containing 100 tumour samples; Breast invasive carcinoma (TCGA, Provisional) containing 962 samples and Breast invasive carcinoma (TCGA, Nature 2012) containing 482 samples. (B) Oncoprints of all cohesin subunits for the largest dataset (Breast invasive carcinoma, TCGA, Provisional) comprising 962 samples showed 19.9% amplification in the RAD21 gene, followed by 2% alteration (amplification, deletion, missense, inframe and truncating mutations) in SMC1A. The dataset also displayed 1% alteration (amplification, deletion, missense and truncating mutations) in SMC3 along with 4% alteration (amplification, deletion, missense and truncating mutations) in STAG1 and STAG2 genes.
Table 1.2 Evidence of cohesin’s involvement in breast cancer

<table>
<thead>
<tr>
<th>Type of cohesin deregulation</th>
<th>Published study</th>
<th>Features of the study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upregulation of RAD21 expression in MCF7 cells with enhanced BRCA1</td>
<td>Atalay et al. 2002</td>
<td>Suppression subtractive hybridisation (SSH) was used to identify genes regulated by BRCA1 by comparing control MCF7 cells with MCF7 cells ectopically expressing BRCA1. RAD21 mRNA and protein was found to be upregulated following BRCA1 overexpression.</td>
</tr>
<tr>
<td>RAD21 mRNA overexpression in primary breast tumour tissues</td>
<td>van’t Veer et al. 2002</td>
<td>Based on expression profiling signatures from cDNA microarray analysis of 98 lymph-node-negative primary breast tumours, two tumour categories had emerged, ‘good prognosis’ group (represented by a 5 year disease-free interval prior to developing distant metastasis) and the ‘poor prognosis’ group (represented by a relatively shorter interval to developing distant metastasis). A ‘prognosis classifier’ gene signature was identified from 78 out of 98 tumours, and RAD21 mRNA was found to be significantly upregulated in the poor prognosis group.</td>
</tr>
<tr>
<td>C-to-G SNP in intron 1 of RAD21 gene in DNA samples (extracted from blood) of breast cancer patients</td>
<td>Kammerer et al. 2004</td>
<td>Genome-wide association study in 254 breast cancer patients and 268 age-matched controls, consisting of 25, 494 SNPs located within 10 kb of 15, 995 LocusLink annotated genes, identified a C-to-G SNP in intron 1 of RAD21 gene (rs1374297) as being strongly associated to susceptibility of breast cancer</td>
</tr>
<tr>
<td>RAD21 transcript overexpression in breast cancer cell lines</td>
<td>Atienza et al. 2005</td>
<td>Quantitative gene expression analysis (combination of competitive PCR and MALDI-TOF mass spectrometry) of RAD21 mRNA revealed that RAD21 level is higher in 9/11 tumourigenic breast cancer lines in comparison with normal tissue and immortalised breast cancer cell lines. siRNA abrogation of RAD21 inhibited proliferation of MCF7 and T47D cells by induction of apoptosis and sensitised MCF7 cells to anti-cancer agents, etoposide and bleomycin.</td>
</tr>
<tr>
<td>Increased expression of RAD21 mRNA in breast tumour biopsy and tissue samples</td>
<td>Oishi et al. 2007</td>
<td>cDNA microarray analysis of 31 breast tumour samples, identified two pathways ‘Wnt pathway’ and the ‘mitochondrial apoptosis pathway’ as the functional pathways associated with determining outcome of breast cancer patients having supraclavicular lymph node metastases without distant metastases. Overexpression of a group of 6 genes comprising RAD21 was correlated with poor prognostic outcomes.</td>
</tr>
<tr>
<td>SNPs in RAD21 gene in genomic DNA samples (extracted from blood) of patients with breast or breast and ovarian cancer</td>
<td>Sehl et al. 2009</td>
<td>Genotyping and whole-genome amplification assays (involving a total of 104 SNPs located in 17 genes) in genomic DNA samples of 196 breast or breast and ovarian cancer patients, and 203 familial risk-matched control patients, identified three intronic SNPs in RAD21 gene (rs16888927, rs16888997, and rs16889040) as being strongly associated with breast cancer.</td>
</tr>
</tbody>
</table>
Enhanced RAD21 expression in a subset of a cohort of familial (BRCA1, BRCA2, BRCAX) breast tumour tissues

Yan et al. 2012

Immunohistochemistry analysis of RAD21 expression in 94 breast tissue samples (28 BRCA1, 27 BRCA2 and 39 BRCAX) identified no significant differences in nuclear RAD21 expression between BRCA1, BRCA2 and BRCAX tumours. No correlation was observed between RAD21 expression and tumour size, grade and status of the lymph node. However, a significant correlation between enhanced RAD21 expression and shorter relapse-free breast-cancer specific survival was observed for BRCA2 and BRCAX tumours, (particularly for grade 3 tumours but not for grades 1 and 2 (same as Xu et al. 2012)) but not in BRCA1 cancers. Intrinsic subtype analysis further revealed worst survival for luminal breast cancers. The correlation between high RAD21 expression and poor survival was confirmed in an additional cohort of 215 breast cancers (data obtained from (Buffa et al, 2011)). An evaluation of genomic changes using previous tumour characterisation data from Waddell and co-workers (Waddell et al, 2010) (34 tumours were analysed for genomic change by SNP-CGH profiling) revealed that tumours with high RAD21 expression had a higher number of base pairs affected by genomic change, implying enhanced RAD21 levels being associated to instability of the genome. Interrogation of correlation between RAD21 expression and microRNA expression using the same cohort indicated an inverse correlation of RAD21 with mir-299-5p.
1.5.3 Cohesin’s role in transcriptional regulation

Many studies have indicated that cohesin has an emerging role in the regulation of gene transcription, which is separable from its cohesion-dependent roles. This is of particular relevance to cancers that harbour cohesin alterations. Because although aberrations in cohesin genes pose a greater chance of developing aneuploidy and genome instability, it remains to be determined whether these events are causative or merely a consequence of cancer. Hence, particularly for cancers overexpressing cohesin (such as breast cancer) where there is sufficient cohesin to ensure accurate transmission of genetic material, it remains a possibility that cohesin contributes to neoplasia via alternative mechanisms. Such mechanisms may include transcriptional modulation of key genes underlying the development of cancer. It would, therefore, be meaningful to investigate mechanisms underpinning the effect of cohesin alterations on global gene transcription, which would then help correlate its precise functional relevance to cancer.

1.5.3.1 Cohesin is a regulator of gene transcription

The very first evidence of cohesin’s role as a regulator of gene expression came from studies in yeast (Donze et al, 1999) and Drosophila (Rollins et al, 1999). In a study investigating gene silencing in *S. cerevisiae*, boundary elements that block the spread of heterochromatin were identified to flank the transcriptionally repressed *HMR* gene locus (also called ‘silent mating type loci’). Authors of this study also reported that depletion of such elements led to the spread of heterochromatin into the neighbouring regions and ectopic insertion of such elements between a ‘silencer’ and ‘promoter’ at the *HMR* gene counteracted repressive effects of the silencer on that promoter. This study was also perhaps the first evidence of cohesin’s role in boundary function delineation. Interestingly, mutations in *Smc1* and *Smc3* subunits of cohesin were found to have significant consequences on the establishment of the boundary function (Donze et al, 1999). Thus, accumulation of cohesin at silent loci was suggestive that cohesin has an important function in ‘chromatin silencing’. Another study carried out in *D. melanogaster* identified Nipped-B (functional homolog of yeast cohesin loader Scc2) as a modulator of long distance enhancer-promoter communication at the *cut* and *Ubx* genes. It was further demonstrated that cohesin binds between enhancer and promoter of the *cut* gene locus, and reduction in *Smc1* gene or mutations in *pds5* gene (cohesin accessory protein) dosage increased expression of the *cut* gene, intriguingly opposite to Nipped-B reduction (Dorsett et al, 2005; Rollins et al, 2004). The observed divergent effect was attributed to ‘biphasic effects’ mediated by cohesin, wherein based on
cohesin dosage, the same gene can increase or decrease in transcriptional response perhaps due to the formation of alternative chromatin architecture (Dorsett, 2011; Schaaf et al, 2009). Studies analysing genome-wide mapping of cohesin binding in *S. cerevisiae* showed that chromatin-bound cohesin can relocate from their initial loading sites to regions of convergent transcription (Glynn et al, 2004; Lengronne et al, 2004). This feature was later shown to be particularly important for efficient *termination of transcription at convergent gene loci* in *S. pombe* (Gullerova & Proudfoot, 2008). However, in mammalian cells, cohesin is enriched at DNase I-hypersensitive sites, but not at genes arranged in a convergent orientation (Parelho et al, 2008). In vertebrates, studies in zebrafish were the first to show evidence of gene regulatory functions controlled by cohesin where in a forward genetic screen, work by Horsfield and colleagues, identified Rad21 as a regulator of *runx1*, which is a haematopoietic TF often found to be dysregulated in human leukaemias (Horsfield et al, 2007).

Mutations in cohesin (*SMC1A, SMC3, RAD21*) and its interacting factors (*NIPBL* and *HDAC8* (histone deacetylase 8, required for deacetylation of acetylated-SMC3 post sister chromatid separation) were found to cause human developmental disorders known as ‘cohesinopathies’. The best known cohesinopathies are Cornelia de Lange Syndrome (CdLS), which is mostly caused by *NIPBL* mutations (>65%) and Roberts Syndrome (RBS) which is caused by mutations in *ESCO2* (Deardorff et al, 2007; Krantz et al, 2004; Pie et al, 2010; Vega et al, 2010; Vega et al, 2005). These syndromes are characterised by physical and mental developmental abnormalities, and are collectively termed as ‘cohesinopathies’ because the syndromes appear to originate from perturbations in proteins responsible for SCC (Bose & Gerton, 2010; Horsfield et al, 2012). Importantly, cell lines developed from CdLS patients did not display overt chromosome cohesion defects, but increased sensitivity to DNA damage and altered gene expression was apparent in CdLS cells (Deardorff et al, 2012; Dorsett & Strom, 2012; Liu et al, 2009). This suggested that transcription-based functions of cohesin are causative of the developmental defect, and pathology of CdLS is independent of cohesin’s role in SCC. Further confirmation came from animal models of CdLS that recapitulated features reminiscent of CdLS (Kawauchi et al, 2009; Muto et al, 2011; Zhang et al, 2007). It thus seems likely that transcription-based functions of cohesin are more sensitive to cohesin activity/amount (particularly during development) than its cohesion-based functions.
Cohesin also actively contributes to ‘hormone-mediated gene regulation’. The first evidence of cohesin’s role in hormonal response came from *Drosophila*, where cohesin was required for EcR (Ecdysone Receptor)-dependent axon pruning in post-mitotic neurons of the *Drosophila* mushroom body (Pauli et al, 2008; Schuldiner et al, 2008). Ablation of cohesin function in *Drosophila* salivary glands provided additional evidence that cohesin directly regulates expression of the EcR gene alongside other EcR-responsive genes (Pauli et al, 2010). The responsiveness of gene expression conditional on cohesin ablation was sufficiently rapid to imply that cohesin indisputably acts on a discrete set of EcR-responsive genes to control their expression directly. Remarkably, there is evidence of cohesin’s involvement in the transcriptional regulation of steroid hormonal response in humans as well (Geck et al, 2000; Maffini et al, 2008; Murthy et al, 2005). Interestingly, in prostate cancer cells, APRIN (unique paralog of vertebrate cohesin accessory protein PDS5) led to androgen-induced proliferative arrest, both in vitro (Geck et al, 2000) and in-vivo (Maffini et al, 2008). APRIN was also identified to possess features representative of a chromatin architectural regulator of androgen-induced chromatin changes during differentiation and cancer.

1.5.3.2 Mechanisms underpinning cohesin-mediated gene regulation

To understand the modus operandi of cohesin function, it is not only essential to figure out the timing but also how and where cohesin binds to chromatin. Genome-wide mapping of cohesin and its auxiliary factors in different model systems have elucidated the probable modes of cohesin action involving modification of local and global chromatin architecture.

1.5.3.2.1 Cohesin mediates long-distance chromosomal interactions in a ‘CTCF-dependent’ and also ‘CTCF-independent manner’

1.5.3.2.1.1 CTCF-dependent interactions

Current evidence suggests that gene regulation is mediated by bringing together enhancers and promoters. The recently recognised interplay between cohesin and CTCF is of great interest, because CTCF regulates long-range interactions in the nucleus (Parelho et al, 2008; Rubio et al, 2008; Wendt et al, 2008). CTCF is ubiquitously expressed and is cell cycle regulated peaking at the S-G2 transition (Klenova et al, 1998). CTCF binding sites were found to mostly function as transcriptional insulators (Kim et al, 2007). Insulators act as boundary elements that prevent spread of repressive heterochromatin to adjoining domains (Sun & Elgin, 1999), and can also prevent communication between gene regulatory elements
such as enhancers in a process called ‘enhancer blocking’ (Capelson & Corces, 2004). Evidence from several studies implicate CTCF and cohesin in mediating long-distance chromosomal interactions at many gene loci. For example, the imprinted IGF2/H19 locus (Nativio et al, 2009), β-globin locus (Chien et al, 2011), interferon gamma (IFN) locus (Hadjur et al, 2009), granulocyte-macrophage colony stimulating factor 2 (CSF2) (Wendt & Peters, 2009), major histocompatibility II (MHC-II) cluster (Majumder & Boss, 2011), to name a few. In mammalian cells, the total number of cohesin binding sites approximately fall in the range of 25,000 to 120,000, as mapped by ChIP-seq experiments in different cell types (Faure et al, 2012; Kagey et al, 2010; Remeseiro et al, 2012; Schmidt et al, 2010). Moreover, around 50-70% genome-wide overlap exists between cohesin and CTCF binding sites (Parelho et al, 2008; Rubio et al, 2008; Wendt et al, 2008). The substantial overlap found between cohesin and CTCF sites across human and mouse genomes suggest that binding of cohesin at CTCF sites contributes to CTCF-dependent transcriptional insulation (at least in some cases). In order to determine the functional significance of cohesin-CTCF co-localisation, CTCF and RAD21 were depleted in HeLa cells and transcriptional changes were measured for H19/IGF2. Interestingly, similar changes were observed after loss of CTCF or RAD21 (H19 transcripts decreased and IGF2 transcripts increased), suggesting that cohesin is required for CTCF’s insulatory function at H19/IGF2 (Wendt et al, 2008). CTCF is involved in the formation of DNA loops at the H19/IGF2 locus (Ling et al, 2006). However, it was the RNAi depletion of cohesin that disrupted higher order chromatin configuration of the locus providing evidence of cohesin’s requirement for correct chromatin conformation (Nativio et al, 2009). Similar results were obtained with the β-globin gene locus (Saitoh et al, 2000); CTCF’s enhancer blocking activity was reduced following depletion of CTCF or RAD21. Subsequent studies confirmed cohesin’s role in chromatin domain organisation and contribution to CTCF’s insulatory functions for additional gene loci, such as HOXA (Kim et al, 2011) and the T-cell receptor α locus (Seitan et al, 2011). Collective evidence from all these studies support the notion that cohesin regulates gene expression by orchestrating three-dimensional chromosomal interactions.

Genome-scale analyses of chromatin contacts have shown that there are frequent interactions occurring within a chromatin domain as compared to fewer interactions occurring with interspersed adjoining domains (Figure 1.5i) (Dixon et al, 2012). Domains that have frequent interactions within them are referred to as ‘topologically associating domains’ (TADs). Boundaries of such TADs are enriched for H3K4me1 (histone mark representative of
enhancers) and H3K9me3 (histone mark representative of transcriptional repression) and also with reference genes (Dixon et al, 2012). TADs can range in size from several hundred kilobases to megabases. Studies in mouse embryonic stem cells (Kagey et al, 2010; Nitzsche et al, 2011) have reported that cohesin and CTCF mediate long-range chromatin interactions defining mega-sized TADs, whereas cohesin and mediator mediate cell type-specific, short-range interactions both within and between TADs (Phillips-Cremins et al, 2013). Several studies have assessed the engagement of cohesin and CTCF in the formation of TADs (Seitan et al, 2013; Sofueva et al, 2013; Zuin et al, 2014). Across all studies, depletion of cohesin decreased intra-TAD interactions, whereas the basic integrity of TADs remained uncompromised. Cohesin reduction did not increase inter-TAD interactions in two of these studies (Seitan et al, 2013; Zuin et al, 2014), while one study reported significant enhancement in inter-TAD interactions with cohesin loss (Sofueva et al, 2013). Since TADs can be quite conserved across cell types, the reason behind the observed disparity between these studies is not very well understood as yet. Remarkably, in addition to affecting intra-TAD interactions, depletion of CTCF increased inter-TAD interaction levels as well (Seitan et al, 2013; Zuin et al, 2014).

1.5.3.2.1.2 CTCF-independent interactions

Although roughly two-thirds of cohesin binding sites coincide with CTCF, a significant proportion of cohesin binding sites are free of CTCF, and instead overlap with tissue-specific TFs. Some examples include overlap with ERα in breast cancer cells, HNF4A and CEBPA in liver carcinoma cells (Schmidt et al, 2010) and with pluripotency factors in stem cells (Nitzsche et al, 2011). In breast cancer cells, out of a total of 49,243 binding sites mapped for CTCF, 39,444 (~80%) co-localize with cohesin. 16,509 binding events for cohesin do not overlap with CTCF but rather 6,573 of these cohesin-non-CTCF sites coincide with ERα (Schmidt et al, 2010). CTCF-independent cohesin sites are enriched at enhancer elements and likely stabilize TF binding (Faure et al, 2012). In support of this viewpoint, in a recent study carried out in oestrogen-stimulated MCF7 cells, cohesin (RAD21) was found to be recruited to a subset of ERα-bound active enhancers that are enriched for the histone variant H2A.Z (Brunelle et al, 2015). Work by this group demonstrated that in oestrogen-stimulated conditions, the ERα-H2A.Z co-occupied enhancers produce enhancer RNAs (eRNAs) and subsequently recruit RNA Pol II and the RAD21 subunit of cohesin (Brunelle et al, 2015). The recruitment of cohesin at these ERα-bound enhancers strongly hint at cohesin’s contribution to ERα target gene regulation, at
least in part by stabilising eRNA-induced enhancer-promoter chromosomal looping, an observation also reported by a prior study (Li et al, 2013). Additionally, binding of cohesin to developmental stage-specific globin genes such as immunoglobulin (Degner et al, 2009) and β-globin gene loci (Chien et al, 2011) lends support to the theory that cohesin forms tissue-specific chromatin structures for effective regulation of gene expression.

1.5.3.2.2 Cohesin directly influences RNA Pol II occupancy and transition from pausing to elongation

RNA Pol II pausing and productive elongation is a key step in transcriptional regulation. Recent ChIP-chip and precision global run-on sequencing (PRO-seq) analysis of RNA Pol II in Drosophila cells from central nervous system showed a positive correlation between Rpb3 subunit and cohesin binding at active genes (Schaaf et al, 2013a). These data are consistent with prior findings that cohesin and Nipped-B preferentially bind actively transcribed regions together with RNA Pol II (Misulovin et al, 2008). Similar binding patterns were observed in mouse liver cells (Faure et al, 2012). PRO-seq data from Schaaf study (Schaaf et al, 2013a) is consistent with evidence that cohesin preferentially binds genes with promoter-proximal paused polymerase, which was based in part on genome-wide coincident binding between cohesin and the pausing factor ‘Negative Elongation Factor’ (NELF) (Fay et al, 2011). Genes with highest cohesin levels demonstrated substantially higher level of pausing, and most Pol II at the promoter was transcriptionally engaged. Furthermore, depletion of cohesin or Nipped-B resulted in increased as well as decreased pausing at cohesin-binding genes (Schaaf et al, 2013a). This is consistent with previous findings where cohesin increased expression of some genes while decreased expression of some other genes (Schaaf et al, 2009), indicating that the effect on Pol II pausing is purely gene-specific. It is proposed that cohesin might also have roles in forming interactions with insulatory elements, which consequently block transition of paused Pol II to elongation at some gene loci. Such loci include the enhancer of split and the invented-engrailed gene complexes, which are simultaneously targeted by polycomb group (PcG) of epigenetic-silencing proteins (Schaaf et al, 2009). Study carried out by Fay and colleagues (Fay et al, 2011) showed that cohesin and pausing factors (NELF, DRB sensitivity inducing factor (DSIF)) bind to a subset of actively transcribing genes (marked with H3K36me1), where they facilitate transition of paused Pol II to elongation. However, at cohesin-repressed genes, cohesin inhibits transition of paused Pol II to elongation in a NELF/DSIF independent manner (Fay et al, 2011). These results suggest that promoter-bound cohesin may promote transition to elongation for some
genes (cohesin-activated genes), while it may also block Pol II elongation at some other
genres (cohesin-repressed genes). The presence of repressor proteins (PcG) may be one of the
factors that influence cohesin’s effect on transcription elongating activity of Pol II, for
cohesin-repressed genes.

1.5.3.2.3 Cohesin interaction with epigenetic silencing proteins
Lastly, cohesin may also exert its transcriptional control via interplay with epigenetic
silencing proteins such as the PcG group of proteins, which maintain chromatin in silent
state by modifying histones. The PcG group comprises of two major protein complexes,
PRC1/2 (polycomb repressor complex 1/2) and the repressive histone mark H3K27me3
associated with polycomb silencing is generated by PRC2. Genome-wide mapping of
cohesin and Nipped-B in Drosophila cells showed that these two proteins are majorly absent
from regions enriched in H3K27me3 i.e from polycomb-silenced genes (Misulovin et al,
2008). These results suggested that cohesin antagonizes PcG silencing. Remarkably, with
depletion of cohesin or Nipped-B, genes that increased substantially in expression were
highly co-enriched for H3K27me3 and cohesin binding. This was further suggestive that
cohesin may sometimes cooperate with PcG proteins to restrain expression of some
exceptional genes (Schaaf et al, 2009). The few exceptional genes that share a rare overlap
between cohesin, PRC1 and the PRC2 generated histone mark H3K27me3 include the
inverted-engrailed gene complex, where cohesin restrains gene expression (Schaaf et al,
2009). Interestingly, such ‘cohesin-PcG restrained genes’ are not fully silenced and showed
robust increases in expression upon depletion of either PRC1 or cohesin (Schaaf et al, 2009).
Work carried out by Strubbe and co-workers identified a biochemical association between
cohesin and PRC1 in Drosophila embryos using an inducible biotinylation-tagging approach
(Strubbe et al, 2011). Schaaf and colleagues (Schaaf et al, 2013b) provided additional
evidence that cohesin controls PRC1 binding to many genes. An unexpected finding was that
PRC1 was identified to preferentially bind a large proportion of active cohesin-binding
genres, contrary to conventional thinking that would expect PRC1 to associate with PcG-
silenced genes only. At ‘cohesin-PRC1 active genes’, cohesin facilitates PRC1 binding.
Conversely, cohesin and PRC1 binding was mutually antagonistic at ‘PcG-silenced genes’
(marked by H3K27me3, having both PRC1 and PRC2 binding) (Schaaf et al, 2013b). At
PcG-silenced genes, reduction in cohesin increased silencing, while reduction in PRC1
decreased silencing. Thus, at active genes, cohesin aids binding of PRC1 while at silenced
genes, cohesin antagonizes binding and activity of PRC1 with the exception of the ‘cohesin-PcG restrained genes’ where cohesin and PRC1 together promote gene silencing.

Taken collectively, there is a high possibility that cohesin dysfunction leads to neoplasia by differential regulation of tissue-specific gene expression via altering both local and global chromatin contexts.
Figure 1.5 Transcription oriented functions of cohesin

(i) Cohesin together with CTCF is implicated in higher-order genome organisation by formation of discrete functional units called TADs. TADs possibly contribute to transcriptional regulation by restricting the number and types of chromatin interactions that can occur between gene promoters and non-protein-coding regulatory elements within and between TADs (intra- and inter-TAD interactions). TAD boundaries, in some cases, can demarcate functionally distinct regions as well, an example is the *Hoxa* locus. (ii) Cohesin mediates long-distance chromatin interactions between distal enhancers and proximal promoters via entrapping two distal DNA segments forming a loop. TADs are also thought to result from chromatin looping between near by and
interspersed genomic regions. (iii) Cohesin also mediates long-range spatial interactions for the formation of distinct gene clusters together with CTCF, where the CTCF binding sites serve as a barrier, preventing interactions with regulatory elements of adjoining gene clusters. Such clusters can isolate (or insulate) a genomic region and to ensure its independent function, can be regulated by a common set of factors. (iv) Cohesin’s influence on transcriptional elongation. Cohesin may facilitate or inhibit the transition of paused RNA Pol II to elongation as evidenced in *Drosophila* and mouse liver cells in a gene-context dependent manner, independent of pausing factors (NELF/DSIF) for cohesin-repressed genes. (v) Cohesin’s interaction and modulation of PcG activity. Cohesin-PcG (PRC1) interaction at active genes facilitates transcription, however cohesin-PcG (PRC1/2) interaction at PcG-silenced genes represses transcription. Cohesin aids in binding of PRC1 at active genes whereas cohesin antagonizes PRC1 binding at PcG-repressed genes.

1.6 Overall aim of this thesis

SNPs in the *RAD21* gene are strongly associated with breast cancer risk, and RAD21 overexpression confers poor prognosis and treatment resistance in luminal breast cancers. Despite frequent amplification and overexpression of RAD21 in breast cancer, it is not known how this contributes to cancer progression, and the consequent molecular mechanisms are not very well understood. Genome-wide analyses in ER-positive breast cancer cells have indicated that ERα co-localizes with the RAD21 subunit of cohesin at multiple ERα-anchored chromatin loops, mostly at oestrogen-regulated genes (Schmidt et al, 2010). This study raised the possibility that cohesin and ERα cooperate to drive expression of a network of oestrogen-responsive genes. Transcriptional dependence on cohesin for oestrogenic response at the *MYC* gene (McEwan et al, 2012) led me to hypothesize that cohesin may also be involved in regulating oestrogenic responses of other ER target genes. The goal of this study was to determine if, and how cohesin regulates expression of ER target genes using the oestrogen-sensitive MCF7 breast cancer cells as a model.

A microarray expression analysis in control and RAD21-depleted MCF7 cells was carried out to identify additional genes in breast cancer that depend on cohesin for their transcriptional regulation by oestrogen. In an investigation of mechanism, dependence on cohesin for ERα binding to its cognate sites at oestrogen-responsive genes was tested for select candidate genes. The aim was to investigate whether a direct regulatory, cause-and-effect relationship exists between cohesin and cohesin-dependent ER target genes. Next, the cohesin-targeting efficacy of a HDAC8 inhibitor PCI-34051 was tested to determine whether the chemical inhibitor could target cohesin’s function pertaining to transcriptional regulation of ER target genes. This was aimed to determine if the cohesin-inhibitor PCI-34051 has the potential to specifically target ER-positive cancers through the cohesin-oestrogen pathway.
Also, whether cohesin could be a new therapeutic target for treatment of ER-positive breast cancers.

Delineation of cohesin-dependent oestrogen-responsive gene signature is necessary to understand cohesin’s precise role in ER-positive breast cancers. Identification of the affected oncogenic pathways and understanding of cohesin’s gene regulatory mechanisms is further essential to estimate how cohesin disruption would interact/integrate with existing anti-oestrogen therapies. Overall, this study was aimed to improve current understanding of cohesin’s role in oestrogen-sensitive breast cancer pathology and elucidate cohesin’s potential as a target for ER-positive breast cancer therapy.
CHAPTER 2

Materials and Methods

2.1 Materials

All standard chemicals and solvents used were purchased from reputable manufacturers/suppliers and were of analytical grade unless otherwise specified. The details of all chemicals/reagents, consumables, solutions/buffers, antibodies, enzymes, and kits used in this study are chronologically listed in sub-sections 2.1.1 to 2.1.5. Primer sequences (gene expression and ChIP) are included in section 2.1.6. All primers were synthesised by Integrated DNA Technologies (IDT, Custom Science, NZ) unless otherwise specified.

2.1.1 Chemicals and Reagents

Table 2.1 List of common chemicals and reagents

<table>
<thead>
<tr>
<th>Chemicals/Reagents</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-Mercaptoethanol</td>
<td>BIORAD, USA</td>
</tr>
<tr>
<td>Ammonium persulfate (APS)</td>
<td>Acros Organics, USA</td>
</tr>
<tr>
<td>D-(-)-Glucose</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>D-(-)-Glucose (ultrapure)</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Deoxycholic acid/Sodium deoxycholate</td>
<td>AppliChem, Germany</td>
</tr>
<tr>
<td>di-Sodium hydrogen phosphate dihydrate (Na₂HPO₄·2H₂O)</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Dimethyl sulfoxide (DMSO)</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Ethanol</td>
<td>Scharlab S.L., Spain</td>
</tr>
<tr>
<td>Ethidium bromide (EtBr)</td>
<td>Merck, USA</td>
</tr>
<tr>
<td>Ethylene diamine tetraacetic acid (EDTA)</td>
<td>Fisher Scientific, USA</td>
</tr>
<tr>
<td>Ethylene glycol tetraacetic acid (EGTA)</td>
<td>CalBiochem, USA</td>
</tr>
<tr>
<td>Ficoll 400</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Sodium hydroxide (NaOH)</td>
<td>VWR International Limited, USA</td>
</tr>
<tr>
<td>Formaldehyde</td>
<td>Fisher Chemical, USA</td>
</tr>
<tr>
<td>Glycine</td>
<td>Thermo Scientific, USA</td>
</tr>
<tr>
<td>Hydrochloric acid (HCl)</td>
<td>BDH, UK</td>
</tr>
<tr>
<td>Lithium chloride (LiCl)</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Methanol</td>
<td>Fisher Chemical, USA</td>
</tr>
<tr>
<td>Monopotassium phosphate (KH₂PO₄)</td>
<td>BDH, UK</td>
</tr>
<tr>
<td>N-(2-hydroxyethyl)piperazine-N-(2-ethanesulfonicacid) (HEPES)</td>
<td>Acros Organics, USA</td>
</tr>
<tr>
<td>Nonidet P-40</td>
<td>Fluka Biochemica, USA</td>
</tr>
<tr>
<td>Paraformaldehyde (PFA)</td>
<td>Fisher Chemical, USA</td>
</tr>
<tr>
<td>Phosphate buffer saline (PBS)</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Potassium chloride (KCl)</td>
<td>BDH, UK</td>
</tr>
<tr>
<td>Potassium hydroxide (KOH)</td>
<td>BDH, UK</td>
</tr>
<tr>
<td>Sodium acetate (NaC₂H₃O₂)</td>
<td>BDH, UK</td>
</tr>
<tr>
<td>Sodium bicarbonate (NaHCO₃)</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Chemical Name</td>
<td>Supplier</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Sodium chloride (NaCl)</td>
<td>VWR, Prolab, Turkey</td>
</tr>
<tr>
<td>Sodium dodecyl sulfate (SDS)</td>
<td>Fisher Scientific, USA</td>
</tr>
<tr>
<td>Sodium pyruvate (powder)</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Tetramethylethylenediamine (TEMED)</td>
<td>BIORAD, USA</td>
</tr>
<tr>
<td>Trizol reagent</td>
<td>Life Technology, USA</td>
</tr>
<tr>
<td>Dithiothreitol (DTT)</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Tris base</td>
<td>J.T Baker, USA</td>
</tr>
<tr>
<td>Triton X-100</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Tween 20</td>
<td>Sigma, USA</td>
</tr>
</tbody>
</table>
## 2.1.2 Consumables

### Table 2.2 List of consumables

<table>
<thead>
<tr>
<th>Consumable</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novex sharp pre-stained protein standard</td>
<td>Novex, Life Technology, USA</td>
</tr>
<tr>
<td>Odyssey protein weight marker</td>
<td>LI-COR Biosciences, USA</td>
</tr>
<tr>
<td>Odyssey blocking buffer</td>
<td>LI-COR Biosciences, USA</td>
</tr>
<tr>
<td>MagicMark XP Western protein standard</td>
<td>Novex, Life Technology, USA</td>
</tr>
<tr>
<td>Precision plus protein standards</td>
<td>BIORAD, USA</td>
</tr>
<tr>
<td>1 kb plus DNA ladder</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>Salmon sperm DNA/Protein A agarose 50% slurry</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>17-β-Oestradiol</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Lipofectamine RNAimax</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>OPTI-MEM reduced serum media</td>
<td>Gibco, Life Technology, USA</td>
</tr>
<tr>
<td>Fetal bovine serum (FBS), NZ origin</td>
<td>Gibco, Life Technology, USA</td>
</tr>
<tr>
<td>Dulbecco’s Modified Eagle Medium (DMEM)</td>
<td>Gibco, Life Technology, USA</td>
</tr>
<tr>
<td>DMEM (Powder)</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Trypsin-EDTA (0.5%), No phenol red</td>
<td>Gibco, Life Technology, USA</td>
</tr>
<tr>
<td>Dextran coated charcoal</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>YOYO-1 stain*</td>
<td>Life Technology, USA</td>
</tr>
<tr>
<td>Phosphate buffer saline (PBS) tablets</td>
<td>Oxoid, UK</td>
</tr>
<tr>
<td>Bovine serum albumin (BSA)</td>
<td>MoreGate Biotech, Australia</td>
</tr>
<tr>
<td>Sodium hydroxide pellets</td>
<td>BDH, UK</td>
</tr>
<tr>
<td>Trypan blue stain</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Ponceau stain</td>
<td>BIORAD, USA</td>
</tr>
<tr>
<td>Coomassie brilliant blue R-250</td>
<td>BIORAD, USA</td>
</tr>
<tr>
<td>Bromophenol blue</td>
<td>BDH, UK</td>
</tr>
<tr>
<td>HDAC8 inhibitor PCI-34051</td>
<td>Cayman Chemical, USA</td>
</tr>
<tr>
<td>Thymidine</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Propidium iodide (P.I)</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT)</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Random primers</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>Hoechst stain (H33342)</td>
<td>Life Technology, USA</td>
</tr>
<tr>
<td>Orange G dye</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>ProLong Gold anti-fade mountant</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>Formaldehyde solution, 37-41%</td>
<td>Fisher Scientific, UK</td>
</tr>
<tr>
<td>Phenol:chloroform:isooamylalcohol</td>
<td>Sigma, USA</td>
</tr>
<tr>
<td>Protease inhibitor tablets</td>
<td>Roche Applied Science, Germany</td>
</tr>
<tr>
<td>10 mM dNTPs</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>RNase OUT Recombinant RNase inhibitor</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>UltraPure™ DNase-RNase free distilled water</td>
<td>Life Technology, USA</td>
</tr>
</tbody>
</table>

*Kindly provided by Dr. Adele Woolley, Senior lecturer, Department of Pathology, University of Otago, Dunedin*
2.1.3 Solutions and buffers

Table 2.3 List of general solutions and buffers

<table>
<thead>
<tr>
<th>Solution</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>70% Ethanol</td>
<td>Absolute ethanol dissolved in ddH₂O at a final concentration of 70% (v/v)</td>
</tr>
<tr>
<td>DEPC-H₂O</td>
<td>0.1% DEPC dissolved in MiliQ-H₂O, incubated overnight at room temperature, followed by sterilisation by autoclaving</td>
</tr>
<tr>
<td>10x PBS</td>
<td>NaCl, Monopotassium phosphate (KH₂PO₄), di-Sodium hydrogen phosphate dihydrate (Na₂HPO₄·2H₂O), Potassium chloride (KCl) dissolved in ddH₂O at a final concentration of 1.5 M, 0.02 M, 0.1 M and 0.03 M, respectively. pH adjusted to 7.2 using HCl, sterilised by autoclaving</td>
</tr>
<tr>
<td>1x RIPA buffer</td>
<td>NaCl, NP-40/Triton X-100, Deoxycholic acid, 10% SDS, 1.5 M TRIS-HCl (pH-8) dissolved in ddH₂O at a final concentration of 150 mM, 1% (v/v), 0.5 % (w/v), 0.1% (v/v), and 50 mM, respectively. Stored at 4 °C. Protease inhibitors added prior to use</td>
</tr>
<tr>
<td>MTT (5 mg/ml)</td>
<td>5 mg of MTT dissolved in 1 ml of PBS for a working stock of 5 mg/ml</td>
</tr>
<tr>
<td>150 mM Thymidine</td>
<td>363.345 mg of thymidine dissolved in 10 ml of ddH₂O</td>
</tr>
<tr>
<td>0.5M TRIS-HCl (pH 6.8)</td>
<td>60.6 g TRIS base dissolved in 1 L of ddH₂O. pH adjusted to 6.8 using HCl</td>
</tr>
<tr>
<td>1.5M TRIS-HCl (pH 8.8)</td>
<td>181.5 g TRIS base dissolved in 1 L of ddH₂O. pH adjusted to 8.8 using HCl</td>
</tr>
</tbody>
</table>

Table 2.4 List of solutions/buffers used for DNA techniques

<table>
<thead>
<tr>
<th>Solution</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1x PBS</td>
<td>1 PBS tablet dissolved per 100 ml of ddH₂O</td>
</tr>
<tr>
<td>1x Tris-Acetate-EDTA (TAE) buffer</td>
<td>1 M TRIS (pH 8), Glacial acetic acid and 0.5 M EDTA (pH 8) added to ddH₂O at a final concentration of 40 mM, 20 mM and 1 mM, respectively</td>
</tr>
<tr>
<td>4x DNA loading dye</td>
<td>Orange G (DNA tracking dye) and Ficoll 400 added to 1x TAE buffer at a final concentration of 0.2% (w/v) and 15% (v/v), respectively</td>
</tr>
<tr>
<td>1x DNA lysis buffer</td>
<td>5 M NaCl, 1 M TRIS-HCl (pH 8), 0.5 M EDTA (pH 8) and 10% SDS added to ddH₂O at a final concentration of 100 mM, 50 mM, 25 mM and 0.1% (v/v), respectively</td>
</tr>
</tbody>
</table>
Table 2.5 List of solutions/buffers used for western blotting

<table>
<thead>
<tr>
<th>Solution</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>10% Resolving gel</td>
<td>5 ml 40% acrylamide, 5 ml 1.5 M TRIS-HCl (pH 8.8), 200 µL 10% SDS, 200 µL 10% APS, 8 µL TEMED dissolved in 9.6 ml of ddH₂O</td>
</tr>
<tr>
<td>4% Stacking gel</td>
<td>2 ml 40% acrylamide, 2.5 ml 0.5 M TRIS-HCl (pH 6.8), 100 µL 10% SDS, 50 µL 10% APS, 10 µL TEMED dissolved in 5.3 ml of ddH₂O</td>
</tr>
<tr>
<td>4x SDS Protein loading dye</td>
<td>100% Glycerol, 0.5 M TRIS (pH 6.8), SDS, bromophenol blue and 100% β-mercaptoethanol added to ddH₂O at a final concentration of 40% (v/v), 240 mM, 8% (w/v), 0.04% (w/v) and 5% (v/v), respectively</td>
</tr>
<tr>
<td>10x Protein electrophoretic buffer (PEB)</td>
<td>30.3 g TRIS base, 144 g glycine, 10 g SDS dissolved in 1 L of ddH₂O. Diluted 1:10 for 1x PEB</td>
</tr>
<tr>
<td>1x Protein transfer buffer</td>
<td>10% (v/v) Methanol added to 1x PEB</td>
</tr>
<tr>
<td>0.1% Ponceau stain</td>
<td>0.1 g Ponceau dissolved in 100 ml of ddH₂O</td>
</tr>
<tr>
<td>0.25% Coomassie brilliant blue solution</td>
<td>2.5 g Coomassie brilliant blue R-250, 500 ml methanol (50%, v/v) and 100 ml of acetic acid (10%, v/v) dissolved in 1 L of ddH₂O</td>
</tr>
<tr>
<td>Destaining solution</td>
<td>500 ml Methanol (50%, v/v), 100 ml acetic acid (10%, v/v) dissolved in 1 L of ddH₂O</td>
</tr>
<tr>
<td>Wash Buffer (PBS-Tween)</td>
<td>0.1% (v/v) tween dissolved in 1x PBS</td>
</tr>
</tbody>
</table>

Table 2.6 List of solutions used for flow cytometry

<table>
<thead>
<tr>
<th>Solution</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1x PBS - 0.1% EDTA</td>
<td>500 µl 0.5 M EDTA (pH 8) added to 500 ml of 1x PBS</td>
</tr>
<tr>
<td>1x PBS - 1% FBS</td>
<td>5 ml FBS added to 500 ml of 1x PBS</td>
</tr>
<tr>
<td>Propidium iodide (PI) - RNase A</td>
<td>150 µl of 1 mg/ml of PI and 187.5 µl of 20 mg/ml of RNase A made up to 15 ml using 1x PBS - 1% FBS solution. Final concentration of PI used was 10 µg/ml and that of RNase A was 250 µg/ml</td>
</tr>
</tbody>
</table>
Table 2.7 List of solutions used for ChIP

<table>
<thead>
<tr>
<th>Solution</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1 M TRIS-HCl (pH 8.1)</strong></td>
<td>121.4 g TRIS dissolved in 1 L of ddH₂O, pH adjusted to 8.1 using HCl, sterilised by autoclaving</td>
</tr>
<tr>
<td><strong>4 M LiCl</strong></td>
<td>16.96 g LiCl slowly dissolved in 100 ml of ddH₂O, sterilised by autoclaving</td>
</tr>
<tr>
<td><strong>3 M Sodium acetate (pH 5.2)</strong></td>
<td>24.6 g sodium acetate dissolved in 100 ml of ddH₂O, pH adjusted to 5.2 using glacial acetic acid. Sterilised by autoclaving</td>
</tr>
<tr>
<td><strong>1 M NaHCO₃</strong></td>
<td>8.4007 g sodium bicarbonate dissolved in ddH₂O, filter sterilised</td>
</tr>
<tr>
<td><strong>10% Sodium deoxycholate</strong></td>
<td>5 g sodium deoxycholate dissolved in 50 ml of ddH₂O, filter sterilised and kept in dark (protected from light)</td>
</tr>
<tr>
<td><strong>1 M HEPES-KOH (pH 7.5)</strong></td>
<td>47.7 g of HEPES dissolved in 200 ml of ddH₂O, pH adjusted to 7.5 using KOH, filter sterilised</td>
</tr>
<tr>
<td><strong>0.5 M EDTA (pH 8)</strong></td>
<td>18.61 g EDTA dissolved in 100 ml of ddH₂O, pH adjusted to 8 using NaOH pellets, sterilised by autoclaving</td>
</tr>
<tr>
<td><strong>0.5 M EGTA (pH 8)</strong></td>
<td>19.02 g EGTA dissolved in 100 ml of ddH₂O, pH adjusted to 8 using NaOH pellets, sterilised by autoclaving</td>
</tr>
<tr>
<td><strong>2.5 M Glycine</strong></td>
<td>18.76 g glycine dissolved in 100 ml of ddH₂O by brief heating in the microwave for up to 1-2 minutes, filter sterilised</td>
</tr>
<tr>
<td><strong>5 M NaCl</strong></td>
<td>146.1 g NaCl dissolved in 500 ml of ddH₂O by brief heating in the microwave (up to 5 minutes with occasional breaks)</td>
</tr>
<tr>
<td><strong>11% Formaldehyde solution</strong></td>
<td>500 µl 1 M HEPES-KOH (pH 7.5), 200 µL 5 M NaCl, 20 µl 0.5 M EDTA (pH 8), 10 µL 0.5 M EGTA (pH 8) and 2.98 ml 37% formaldehyde solution dissolved in 6.29 ml of ddH₂O. Final concentration of the constituents were: HEPES - 50 mM, NaCl - 100 mM, EDTA - 1 mM, EGTA - 0.5 mM and formaldehyde - 11% respectively</td>
</tr>
</tbody>
</table>
### Table 2.8 List of buffers used for ChIP

<table>
<thead>
<tr>
<th>Buffer</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lysis buffer</strong></td>
<td>1 M TRIS-HCl (pH 8.1), 10% SDS, 0.5 M EDTA (pH 8) added to ddH$_2$O at a final concentration of 50 mM, 1% (v/v) and 10 mM, respectively. Protease inhibitor tablets were added prior to use</td>
</tr>
<tr>
<td><strong>ChIP dilution buffer</strong></td>
<td>1 M TRIS-HCl (pH 8.1), 10% SDS, triton X-100, 0.5 M EDTA (pH 8) and 5 M NaCl added to ddH$_2$O, at a final concentration of 16.7 mM, 0.01% (v/v), 1.1% (v/v), 1.2 mM and 16.7 mM, respectively. Protease inhibitor tablets were added prior to use</td>
</tr>
<tr>
<td><strong>Low-salt wash buffer</strong></td>
<td>1 M TRIS-HCl (pH 8.1), 10% SDS, triton X-100, 0.5 M EDTA (pH 8) and 5 M NaCl added to ddH$_2$O, at a final concentration of 20 mM, 0.1% (w/v), 1% (v/v), 2 mM and 150 mM, respectively</td>
</tr>
<tr>
<td><strong>High-salt wash buffer</strong></td>
<td>1 M TRIS-HCl (pH 8.1), 10% SDS, triton X-100, 0.5 M EDTA (pH 8) and 5 M NaCl added to ddH$_2$O, at a final concentration of 20 mM, 0.1% (v/v), 1% (v/v), 2 mM and 500 mM, respectively</td>
</tr>
<tr>
<td><strong>LiCl wash buffer</strong></td>
<td>1 M TRIS-HCl (pH 8.1), 4 M LiCl, Nonidet P-40, 10% sodium deoxycholate and 0.5 M EDTA (pH 8) added to ddH$_2$O, at a final concentration of 10 mM, 0.25 mM, 1% (v/v), 1% (w/v) and 1 mM, respectively</td>
</tr>
<tr>
<td><strong>TE buffer</strong></td>
<td>1 M TRIS-HCl (pH 8.1) and 0.5 M EDTA (pH 8) added to ddH$_2$O, at a final concentration of 10 mM and 1 mM, respectively</td>
</tr>
<tr>
<td><strong>Elution buffer</strong></td>
<td>10% SDS and 1 M NaHCO$_3$ added to ddH$_2$O, at a final concentration of 1% (v/v) and 100 mM, respectively (made fresh every time)</td>
</tr>
</tbody>
</table>

### Table 2.9 List of solutions used for immunofluorescence

<table>
<thead>
<tr>
<th>Solution</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>4% PFA</strong></td>
<td>4% (w/v) paraformaldehyde dissolved in PBS by incubating at 65 °C in a water bath (made fresh every time)</td>
</tr>
<tr>
<td><strong>10% Triton-X</strong></td>
<td>10% (v/v) triton-X diluted in 1x PBS</td>
</tr>
<tr>
<td><strong>Hoechst (H33342)</strong></td>
<td>Working stock of 1 mg/ml stock was prepared in ddH$_2$O</td>
</tr>
<tr>
<td><strong>Blocking buffer</strong></td>
<td>10% (v/v) Fetal bovine serum (FBS) dissolved in 1x PBS</td>
</tr>
</tbody>
</table>
2.1.4 Antibodies

2.1.4.1 Antibodies used for western blotting

Primary antibodies used for western blotting:
(i) Rabbit polyclonal anti-RAD21 (Ab992, Abcam, UK) in 1:1000 dilution
(ii) Rabbit polyclonal anti-ERα ((HC-20, sc-543) Santa Cruz Biotech, USA) in 1:1000 dilution
(iii) Rabbit monoclonal anti-Total SMC3 (D47B5, Cell Signaling technology, USA) in 1:1000 dilution
(iv) Mouse monoclonal anti-acetyl SMC3 (Custom antibody kindly provided by Dr. Katsuhiko Shirahige, Institute of Molecular and Cellular Biosciences, University of Tokyo) in 1:500 dilution
(v) Mouse polyclonal anti-PARP (556362, BD Pharmeden, USA) in 1:2000 dilution
(vi) Mouse anti-Y-Tubulin (T5326, Sigma, USA) in 1:5000 dilution

Secondary antibodies used:
(i) IRDye 800CW goat polyclonal anti-mouse IgG (926-32210, LI-COR Biosciences, USA) in 1:15,000 dilution
(ii) IRDye 680CW goat polyclonal anti-rabbit IgG (926-32221, LI-COR Biosciences, USA) in 1:15,000 dilution

2.1.4.2 Antibodies used for ChIP

In addition to RAD21 and ERα antibodies, rabbit polyclonal anti-RNA Polymerase II-serine-5-phosphorylation (RNAPIIser5-P) (Ab5131, Abcam, UK) antibody was used for ChIP experiments.

2.1.4.3 Antibodies used for immunofluorescence

In addition to total-SMC3 (used in 1:200 dilution) and acetyl-SMC3 (used in 1:100 dilution), following antibodies were used for immunofluorescence experiments:
(i) Mouse monoclonal anti-α-Tubulin (T6199, Sigma, USA) in 1:500 dilution

Secondary antibodies:
(ii) Goat polyclonal anti-mouse antibody, Alexa fluor 488 (A-11001, Life technology) in 1:2000 dilution
(iii) Goat polyclonal anti-rabbit antibody, Alexa fluor 568 (A-11011, Life technology, USA) in 1:2000 dilution
### 2.1.5 Enzymes used in molecular techniques

**Table 2.10 List of enzymes**

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteinase K</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>RNAse A</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>RNAse H</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>Superscript III Reverse transcriptase</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>Platinum SYBR Taq DNA polymerase</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>(with Rox)</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>Platinum Taq DNA polymerase</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>SYBR Premix Ex Taq</td>
<td>TaKaRa BIO INC., Japan</td>
</tr>
</tbody>
</table>

### 2.1.6 Kits used in molecular techniques

**Table 2.11 List of kits**

<table>
<thead>
<tr>
<th>Kits</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleospin Total RNA isolation kit</td>
<td>Macherey-Nagel, Germany</td>
</tr>
<tr>
<td>RNA 6000 Nano kit</td>
<td>Agilent Technologies, USA</td>
</tr>
<tr>
<td>Superscript III Reverse transcriptase kit</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>Platinum SYBR Green qPCR SuperMix-UDG with Rox</td>
<td>Invitrogen, Life Technology, USA</td>
</tr>
<tr>
<td>SYBR Premix Ex Taq ™ (Tli RNase H Plus), Bulk</td>
<td>TaKaRa BIO INC., Japan</td>
</tr>
<tr>
<td>BCA Protein assay kit</td>
<td>Thermo Scientific, USA</td>
</tr>
<tr>
<td>QIAquick® PCR purification kit</td>
<td>Qiagen, Netherlands</td>
</tr>
</tbody>
</table>
### 2.1.7 Primers used

#### Table 2.12 List of primers

<table>
<thead>
<tr>
<th>Amplicon</th>
<th>Forward primer (F)</th>
<th>Reverse primer (R)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gene expression primers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPIA*</td>
<td>ACGGCGAGCCCTTGG</td>
<td>TTTCTGCTGTCTTTGGGACCT</td>
</tr>
<tr>
<td>GAPDH*</td>
<td>TGCAACCACAAACTGTCTTAGC</td>
<td>GGCATGGAAGTGTGGTGTCATGAG</td>
</tr>
<tr>
<td>RAD21*</td>
<td>CAATGCCAACCATGACTCAT</td>
<td>CGGTGTAAGACACGGTGTTAAG</td>
</tr>
<tr>
<td>MYC*</td>
<td>TCGGGGTAGTGGAAACCCAG</td>
<td>CAGCAGCTCAGATTCTTTTCC</td>
</tr>
<tr>
<td>GREB1*</td>
<td>ATCAACTGCTCGGACTTTCAG</td>
<td>TGAAGCTCCGTCCTGACAGATG</td>
</tr>
<tr>
<td>ESR1*</td>
<td>CAGGATCTTCAAGCCAGGCAC</td>
<td>ATGCTCAACTGGGCGGAGAG</td>
</tr>
<tr>
<td>SOX4</td>
<td>CTGGTCCGTGTCCGGGCTC</td>
<td>AGCTGTAATGGGCAAGGAGTT</td>
</tr>
<tr>
<td>IL20</td>
<td>CAAGACAACAGGCTCGCAAG</td>
<td>TGTCAGGGGTCTGCTAGTT</td>
</tr>
<tr>
<td>BMPR2</td>
<td>ATCGTGAGCCACACAGTCAATCCA</td>
<td>GACCAATTGGTGACACGCCTA</td>
</tr>
<tr>
<td>DKK1</td>
<td>CATCAAGAAACCTTGCCCCCA</td>
<td>CGGTGACTGGGTATCTGCTAGT</td>
</tr>
<tr>
<td>PIM2</td>
<td>TGACTTTAGTTGGAGCAAGGAG</td>
<td>GGATGCTCCCCACACACCAT</td>
</tr>
<tr>
<td>THBS1</td>
<td>AACAACCCCAACACCCCAGTT</td>
<td>TTGAGCGAGCGACGTACAGTC</td>
</tr>
<tr>
<td>BAG1</td>
<td>GAAGCGACCTGGAAGTTGAAGA</td>
<td>TTCTTGGTTGCTGGAGAGA</td>
</tr>
<tr>
<td>CXCL12</td>
<td>CTCTGAGGGGATGTGAATGG</td>
<td>GCCTCCATGGGACATACATAG</td>
</tr>
<tr>
<td>IRS2</td>
<td>CCACCATCGTGAAGAGTGAAG</td>
<td>TTGCTTTGGTGTCGTGCCCTA</td>
</tr>
<tr>
<td>CCNG2</td>
<td>CTCCGGCAGATGAAGGAGTATT</td>
<td>ATCAATTCTCCGGGATGCTAGCT</td>
</tr>
<tr>
<td>TFF2</td>
<td>GGGTCCCCCTGGTTTTCC</td>
<td>GAGACCTCCATGACCACCTGA</td>
</tr>
<tr>
<td>BAX**</td>
<td>GACAGGGGCGCTTTTTTGTCA</td>
<td>TGTCACGTCGCAATCATC</td>
</tr>
<tr>
<td>PUMA**</td>
<td>GGTGGCGGAGCGACACCTC</td>
<td>CGGCGAGGCTGGGAGT</td>
</tr>
<tr>
<td>BCL2**</td>
<td>GGATGCCTTTGTTGGAACTGTG</td>
<td>GGTCCTTGGGAAATTGATGTT</td>
</tr>
<tr>
<td><strong>ChIP primers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOX4 P1</td>
<td>CGCTAGGAGAAATGACCCCGAGA</td>
<td>TTCAGTTGACCGTGAAACC</td>
</tr>
<tr>
<td>SOX4 P2</td>
<td>CAGCAACCAGCATTCGAGA</td>
<td>CCTCTCTCGTCTCCTCACC</td>
</tr>
<tr>
<td>SOX4 P3</td>
<td>ATCCCAATATATTTTCTCCCTG</td>
<td>GCTTAAGCCTCCAGCAGAAGGATT</td>
</tr>
<tr>
<td>SOX4 P4</td>
<td>GGTCTTGAACCTGCTGGCTTAG</td>
<td>CCCGTGTTGGCAATTACAGTT</td>
</tr>
<tr>
<td>IL20 P1</td>
<td>CACCCAGGGTGCCTGACTA</td>
<td>GCAAGACGTGATGGGCAC</td>
</tr>
</tbody>
</table>

* Primers synthesised by Invitrogen, pre-available in the Horsfield Lab
** Primers were kindly provided by Dr. Adele Woolley, Department of Pathology, University of Otago, Dunedin
2.1.8 Model system used
All of this work is carried out in the ERα-positive human breast adenocarcinoma cell line MCF7 (ATCC-HTB-22).

2.2 Cell culture maintenance and manipulation

2.2.1 Reviving cells from liquid nitrogen stock/Thawing frozen cells
To revive cells from liquid nitrogen storage, a vial of cryopreserved working stock of cells was removed, immediately placed on ice and quickly thawed (<1 minute) in the water bath set at 37 °C by gently swirling the vial. In the laminar flow hood, the exterior of the vial was first wiped with 70% ethanol, and thawed cells were then transferred slowly (drop wise) to a 15 ml tube containing pre-warmed growth medium (pre-warmed to 37 °C). The diluted cell suspension was then centrifuged at 250 × g for 5 minutes. The supernatant was aspirated off (using a glass Pasteur pipette) and the cell pellet was resuspended in 7.5 ml of culture medium and transferred to a T25 cell culture flask. Cell growth was henceforth maintained as described in the following sections.

2.2.2 Maintenance of cell line
MCF7 cells used in this study were maintained in Dulbecco’s Modified Eagle Medium (DMEM, Life Technology or Sigma, USA) culture medium supplemented with 10% fetal bovine serum (FBS, Gibco, Life Technology, USA) in a 37 °C humidified cell culture incubator (CellStar, Sigma, USA) set at 5% CO₂.

For optimal growth, culture medium was changed every alternate day, and when cell confluence was nearly 80-85% with minimal room for expansion, cells were sub-cultured or passaged to stimulate continued proliferation. Cells were sub-cultured prior to reaching maximum confluence. Although transformed cells can continue to proliferate despite culture flasks becoming fully confluent, however, due to constricted flask space, contact inhibition leads to deterioration of proliferative capacity of the cells. Hence, adherent cultures of MCF7 cells were sub-cultured while still in log phase.

2.2.2.1 Sub-culturing/Splitting/Passaging
For sub-culturing, exhausted culture medium was aspirated off and cells were rinsed with PBS prior to enzymatic dissociation by trypsin, without disturbing the cell layer. This step
washes off any traces of serum, calcium, magnesium that could interfere with the action of the dissociation agent. Following the wash step, cells were detached from the surface of the culture vessel by adding an appropriate volume of pre-warmed enzymatic dissociation reagent 0.05% trypsin-EDTA (phenol-red free, Gibco, Life Technology, USA). Cells were completely covered with trypsin by gently rocking the culture vessel. The vessel was then transferred to the cell culture incubator for 3-5 minutes (set at 37 °C) to expedite cell detachment. Upon cell detachment (after >90% of the cells had detached), pre-warmed culture medium (three times the volume of trypsin) was added to inactivate the trypsin, and the medium was evenly dispersed by pipetting several times. Next, the cell suspension was transferred to a 15 ml tube and centrifuged at 250 × g for 5 minutes. The supernatant was aspirated off and the cell pellet was then resuspended in a minimal volume of culture medium. A small volume from this was removed for cell counting. For appropriate cell seeding density, trypan blue exclusion test was performed and cell viability was measured using a haemocytometer and cell-counter (discussed below). The cell suspension was then diluted as per the seeding density requirements and plated in a new culture vessel to be returned to the incubator.

2.2.2.2 Trypan blue cell counts using a haemocytometer

The cell counting chamber of the haemocytometer was wiped with 70% ethanol, and a coverslip was appropriately placed in position. 0.4% working stock solution of sterile-filtered trypan blue was prepared in PBS, and 100 µl of this solution was added to 100 µl of the diluted cell suspension creating a dilution factor of 2. 10-20 µl of this solution was loaded to the counting chamber of the haemocytometer. The haemocytometer was placed under the 10x objective of an inverted phase-contrast microscope (LH50A Olympus, Japan), and cell numbers in gridded squares of the counting chamber were counted. Cell counts were performed for the four corner squares of one chamber and the average count per square was determined. The number of viable cells per 1 ml of the cell suspension was calculated by correcting for the dilution factor using the formula:

\[
\text{Average number of viable cells per square} \times 10^4 \times \text{dilution factor} = \text{cells/ml of cell suspension}
\]

Percentage viability was subsequently determined using the formula:

\[
\% \text{ viable cells} = \left( \frac{\text{Number of unstained cells}}{\text{total number of cells}} \right) \times 100
\]
2.2.3 Freezing cells for liquid nitrogen storage/Cryopreservation
Surplus of cells available from sub-culturing were stored in liquid nitrogen, by
cryopreserving them in complete growth medium (supplemented with FBS) in the presence
of the cryoprotective agent dimethyl sulfoxide (DMSO, Sigma, USA). DMSO helps in
reducing the freezing point of the medium and facilitates effective preservation of cells by
ensuring reduced risk of cell damage caused due to ice crystal formation at low
temperatures. Freezing medium (60% culture medium, 20% FBS and 20% DMSO) was
prepared and stored at 4 °C. Cells of low passage number and high density were harvested
by trypsinisation. Cell pellet was obtained by spinning at 250 × g for 5 minutes, following
resuspension in culture medium (half the volume of the required final medium). A similar
volume (as that of the cell suspension creating a dilution of 1:1) of ice-cold freezing medium
was added drop wise to the resuspended cell suspension, and after gentle homogeneous
mixing, aliquots were dispensed in cryovials. Next, cryovials were placed in an isopropanol
freezing container (Invitrogen, Life Technology, USA) and stored overnight at -80 °C, and
transferred to the liquid nitrogen dewar, the following day.

2.2.4 Making of culture medium
Phenol red free, powdered DMEM media (Sigma, USA) containing L-Glutamine and 1 g/L
of glucose was reconstituted in 900 ml of sterile ddH₂O. To this, 3.5 g of ultrapure glucose
(Sigma, USA) (to make the final concentration 4.5 g/L), 10 ml of sodium pyruvate (Sigma,
USA) and 3.7 g of sodium bicarbonate (Sigma, USA) was added. The final volume was
made up to 1 L by adding more ddH₂O. pH of the media was adjusted to ~6.7 using 1 N HCl
or 1 N NaOH. Media was immediately filter-sterilised using ‘Corning® bottle-top filter
vacuum system’ (Sigma, USA) which has a pore size of 0.22 microns. Media was aseptically
dispensed in sterile containers and stored at 4 °C for future use (typically used within a
month).

2.2.5 Charcoal-dextran stripping of FBS for oestrogen-free culture
medium
For assessing oestrogen’s influence on genome-wide gene expression, culture medium
(DMEM) was deprived of oestrogen by supplementing with charcoal-dextran stripped FBS.
Dextran-treated charcoal selectively removes hormones without non-specific loss of
essential serum components. For hormone stripping, 2 g of dextran-coated charcoal (Sigma,
USA) per 100 ml of FBS was added to a glass bottle, followed by incubation at 55 °C in the
water bath for 30 minutes. After incubation, the FBS-charcoal mix was transferred to 50 ml tubes and centrifuged at maximum speed (3220 × g) for 10 minutes. The supernatant was carefully transferred (without agitating charcoal settled at the bottom) to a new bottle. For every 100 ml of FBS, 2 g of additional dextran-coated was added and re-incubated at 55 °C as before. Previous steps were repeated, and the final recovered supernatant was filter-sterilised and aliquoted in 50 ml tubes and stored at -20 °C for future use.

2.2.6 siRNA Transfection

RAD21 gene silencing was achieved using small interfering RNAs (siRNAs). The siRNAs used in this study are listed in table 2.13. For RNA interference (RNAi) analysis, lipofectamine RNAimax reagent (Invitrogen, Life Technology, USA) was used for efficient delivery of siRNAs to MCF7 cells using the reverse transfection technique (suitable for MCF7 cells). siRNAs were reconstituted aseptically as per the manufacturer’s guidelines. The reverse transfection mix consisted of the respective siRNA (one at a time) and lipofectamine (5 µl per well), diluted using the reduced serum medium OPTI-MEM (500 µl per well) (Gibco, Life Technology, USA). In reverse transfection, the transfecting complex is prepared inside the well of a culture plate, to which cells and culture medium are added. For high throughput transfection, the transfection mix (siRNA + lipofectamine RNAimax + OPTI-MEM) was prepared, followed by harvesting cells by trypsinisation. The transfection mix was left to incubate at room temperature for 15-20 minutes for homogeneous mixing prior to performing the transfection procedure. During the incubation time, viable cells were counted (by trypan blue staining described in section 2.2.2.2) and the cell suspension was prepared by diluting with culture medium to achieve appropriate cell numbers. The transfection mix was added to each well of the assay plate (6 well plate in this case), followed by drop wise addition of the cell suspension. 300,000/400,000 cells were plated in each well of a 6 well plate. The final concentration of siRNA used was 5-10 nM, 5 nM for microarray experiments and 10 nM for the ChIP experiments. The assay plate was swirled around, gently tapped on each side to ensure that the cell suspension was evenly distributed. At the end of the transfection procedure, the assay plates were carefully transferred to the cell culture incubator.
Table 2.13 siRNAs used in this study

<table>
<thead>
<tr>
<th>siRNA target</th>
<th>Name</th>
<th>siRNA sequence</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-targeting control</td>
<td>ON-TARGET plus Non-Targeting siRNA #1, D-001810-01</td>
<td>UGGUUUACAUGUCGACUAA</td>
<td>Dharmacon, Thermo Scientific</td>
</tr>
<tr>
<td>RAD21</td>
<td>On target plus siRNA, Human RAD21, J-006832-06</td>
<td>GCUCAGGCCCCUGGGAAUA</td>
<td>Dharmacon, Thermo Scientific</td>
</tr>
<tr>
<td>RAD21</td>
<td>On target plus siRNA, Human RAD21, J-006832-07</td>
<td>GGGAGUAGUUCGAUCAUA</td>
<td>Dharmacon, Thermo Scientific</td>
</tr>
<tr>
<td>RAD21</td>
<td>On target plus siRNA, Human RAD21, J-006832-08</td>
<td>GACCAAGGUUCAUUAUA</td>
<td>Dharmacon, Thermo Scientific</td>
</tr>
</tbody>
</table>

2.3 Gene expression studies

2.3.1 Total RNA isolation

For gene expression analyses, total RNA was isolated from cultured MCF7 cells seeded at a density of $3 \times 10^5$ cells per well of a 6 well plate (BD falcon, BD Biosciences, USA) using ‘Nucleospin RNA Isolation kit’ according to manufacturer’s instructions (Machery Nagel, Germany). For RNA isolation, used media from the wells of the culture plate was aseptically aspirated off at specified time points, followed by addition of 350 µl of lysis buffer (Buffer RA1, supplied with the kit). The cell-lysis buffer mix was thoroughly homogenised by pipetting throughout the surface of the wells, followed by storage of the plate at -80 °C. Plate edges were sealed using a moisture-resistant, self-sealing film (Parafilm M®, Bemis Flexible packaging, USA) prior to storage. RNA extraction was continued at suitable times using manufacturer’s guidelines afterward.

2.3.1.1 Determination of RNA concentration and purity

After RNA isolation, concentration and purity was assessed by measuring absorbance at 230 ($A_{230}$), 260 ($A_{260}$) and 280 ($A_{280}$) nm, using the spectrophotometer (NanoDrop ND-1000, Nanodrop Technologies Inc., USA). RNA samples with $A_{260}/A_{280}$ and $A_{260}/A_{230}$ ratios ranging in between 1.8-2.1 were rated to be of optimal quality to proceed for expression analyses.
2.3.1.2 RNA integrity analysis by Bioanalyzer profiling

For cDNA microarray analysis, RNA samples from the four biological replicates (experiment design described in section 3.4, Chapter 3) were additionally analysed for integrity. RNA integrity analysis was carried out using the RNA 6000 Nano kit (Agilent Technology, USA) run on a 2100 Bioanalyzer system (Agilent Technology, USA) as per the manufacturer’s instructions. RNA samples with RNA integrity numbers (RINs) ranging between 7.9-10 were considered of high quality to proceed for microarray expression analysis.

2.3.2 cDNA synthesis

Quality controlled RNAs were used to synthesize random-primed cDNA using the Superscript III reverse transcriptase kit (Invitrogen, Life Technology, USA). 500 ng of total RNA was added to 1 µl of 10 mM dNTP and 0.83 µl of 300 ng/µl working stock or 250 ng of random primers (Invitrogen, Life Technology, USA). This was followed by incubation in the thermal cycler (2720 Thermal Cycler, Applied Biosystems, Life Technology, USA) at 65 °C for 5 minutes. This step denatures the RNA secondary structure before quick chilling on ice (for a minute or two) to allow random primers to anneal to the RNA templates. 4 µl of 5x First strand buffer (FS Buffer) (comes with the kit), 1 µl each of 0.1 M DTT (comes with the kit), reverse transcriptase Superscript III and the recombinant ribonuclease inhibitor RNase OUT (Invitrogen, Life Technology, USA) were added to samples. The tubes were briefly spun for homogeneous mixing of the reaction constituents. Reverse transcription reaction was extended by incubating the tubes at 25 °C for 5 minutes, 50 °C for 60 minutes and 70 °C for 15 minutes. 1 µl of the endoribonuclease RNase H (Invitrogen, Life Technology, USA) was added at the end, and the tubes were incubated at 37 °C for a further 20 minutes. This step removes complementary RNA from the RNA-cDNA hybrid to yield single-stranded DNA. The newly synthesised cDNAs (25 ng/µl) were diluted to a final concentration of 5 ng/µl using ultra-pure distilled water (Invitrogen, Life Technology, USA) prior to use in qRT-PCR reactions.

2.3.3 Quantitative reverse transcriptase PCR (qRT-PCR)

Quantitative RT-PCR was used to detect and quantify the amount of target genes expressed in treated samples. Real-time qPCR was performed in technical duplicates either using Platinum SYBR Green qPCR Super-Mix-UDG with ROX (Invitrogen, Life Technology, USA) or TaKaRa SYBR Premix Ex Taq™ (TaKaRa BIO INC., Japan) on ABI 7300 Real
Time PCR System (Applied Biosystems, USA), or LightCycler 480 System (Roche Life Sciences, USA). qPCR reactions were performed in 20 µl volumes containing 1 µl of 5 ng/µl of template cDNA, 1 µL each of 3 µM forward and reverse primers, 7 µl of ultrapure distilled water and 10 µl of platinum SYBR green or TaKaRa SYBR premix. A non-template control was included for every gene/primer set tested.

qRT-PCR settings for the ABI 7300 system was: 1 cycle of 50 °C for 2 minutes, 1 cycle of 95 °C for 2 minutes, 40 cycles of 95 °C for 15 seconds and 60 °C for 30 seconds. This is followed by a dissociation stage comprising 1 cycle of 95 °C for 15 seconds, 60 °C for 30 seconds and 95 °C for 15 seconds.

qRT-PCR settings for LightCycler 480 was: 1 cycle of 95 °C for 30 seconds, 40 cycles of 95 °C for 5 seconds and 60 °C for 30 seconds. This is followed by a dissociation stage comprising 1 cycle of 95 °C for 5 seconds, 60 °C for 1 minute and 95 °C (acquisition mode: continuous), followed by a cycle of cooling at 50 °C for 30 seconds.

Data was extracted using SDS software version 1.4 (Applied Biosystems) or LightCycler 480 Software version 1.5 (Roche Life Sciences). Data was imported into the qBasePLUS software (Biogazelle, Belgium) for calculation of cycle threshold (Ct) values and for relative quantification by normalising to expression levels of reference genes post correction for real-time PCR efficiency. Melting/dissociation curves were collected and assessed for primer-dimer formation or any non-specific product amplification. The expression levels of two reference genes PPIA (Peptidylprolyl isomerase A, also referred to as PPIA or Cyclophilin A) and GAPDH (Glyceraldehyde 3-phosphate dehydrogenase) were found to vary minimally in response to siRNA depletion of RAD21 in MCF7 cells in our lab’s previous study. Hence, these two genes were deemed suitable for use as reference genes for this study as well. Relative quantification was performed to determine changes in expression levels of steady state mRNA levels of a target gene across control/treated samples and expressed relative to the expression levels of endogenous reference genes.

2.3.3.1 Primer design

Expression primers used in this study are listed in Table 2.12. Gene-specific primers were designed either using NCBI’s primer designing tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) or using Primer3 (version 0.4.0) (http://biotools.umassmed.edu/bioapps/primer3 www.cgi). To avoid amplification of contaminating genomic DNA, primers were designed spanning exon-exon junctions. For reliable and efficient amplification of target DNA, amplicon sizes were kept in the range of
90-150 bp. GC content was kept in the range of 40-60% for maximum product stability, melting temperature was kept in between 57-60 °C. To lower the probability of potential hairpin/primer-dimer formation, self-complementarity was kept as low as possible relative to the nature of the sequence of interest. After selection of unique primer sequences, these were also tested for non-specific amplification using the in-silico PCR tool available on UCSC genome browser (https://genome.ucsc.edu/cgi-bin/hgPcr). This was done for prior indication of any possible off-target DNA templates that the designed primers could hybridise to, for decreasing chances of non-specific amplification.

2.3.3.2 Primer validation

All primers were validated prior to use. The amplification efficiency and sensitivity of the qPCR assay was determined by performing ten fold dilutions (three to five) using template cDNA of known concentration. A standard curve was generated using Sequence Detection 7300 software (SDS), version 1.4 for the ABI 7300 machine or Light Cycler 480’s built-in software for each primer pair, to measure the performance efficiency of the primers. For a 10-fold dilution series, a perfectly efficient (100%) PCR has a slope of -3.32. However, slopes in the range of -3.0 to -3.9 (80-100% efficiency) are considered acceptable to proceed with analysis. Slopes outside of this range are suggestive of inefficient amplification or of reactions amplifying sequences non-specifically, both of which are detrimental to attaining reliable quantitative results. Those primer pairs that had slopes in the acceptable range were chosen to proceed with further analysis. Next, to verify that the amplified product is the desired product, all PCR reactions were subjected to melt curve/dissociation analysis. Target-specific primers were expected to produce dissociation curves with single peaks indicating target specificity. Those primer pairs that generated single peaked dissociation curves were selected for further validation. Primer specificity was also simultaneously evaluated by running the PCR amplified products on ethidium bromide (EtBr) stained agarose gels (described next). Target-specific primers were expected to generate amplicons of the correct size corresponding to the predicted amplicon length when run on agarose gels. After confirming a single amplified product of the desired size, the remaining PCR product was purified using ‘PCR purification kit’ (QIAquick® PCR purification kit, Qiagen, Netherlands) and sent for full Sanger sequencing at the Genetic Analysis Services of University of Otago. For sequencing, purified PCR products were sent as part of a mixture containing 1 ng/100 bp of purified DNA, 3.2 pmol of forward/reverse primer, made up to a total volume of 5 µl using ultra-pure distilled water (Life Technology, USA).
2.3.3.2.1 Agarose gel electrophoresis
Agarose powder (1.5%, w/v) was dissolved in 1x TAE buffer (Table 2.4) and the gel was cast in a gel-casting tray (BIORAD, USA) with appropriately placed combs. DNA was mixed with 4x DNA loading dye (at a final concentration of 1x, v/v) and loaded into designated wells, alongside loading 1 µl of 1 kb plus ladder (Invitrogen, Life Technology, USA) mixed with 4x loading dye. Electrophoresis was carried out at 100 volts for 1.5 hours, followed by EtBr staining of the gel by dissolving the gel in EtBr bath for 15 minutes. The stained gel was then visualised under ultraviolet (UV) light on a Gel Doc XR system (BIORAD, USA).

2.3.4 Microarray expression analysis
2.3.4.1 Sample preparation
For preparing RNA samples for microarray analysis, MCF7 cells were grown at a density of $3 \times 10^5$ cells per well of a 6 well plate, in hormone-free culture medium (DMEM supplemented with 10% charcoal-dextran treated FBS) for 24 hours. This was followed by reverse transfection (described in section 2.2.6) with control or RAD21 siRNAs (5 nM for 48 hours). After 48 hours, cells were subsequently stimulated with 100 nM of 17-β-oestradiol or vehicle (absolute ethanol) for 3 and 6 hours. Four independent biological replicates were generated. RNA from these samples was extracted as described in section 2.3.1, and RNA quality and integrity was examined as described in sections 2.3.1.1 and 2.3.1.2. Cells from the specified time points were also harvested for protein. RAD21 knockdown was confirmed at the transcript level by qRT-PCR (described in section 2.3.3) as well as at the protein level by western blotting (described in section 2.4.2). Based on quality control analysis, RNA samples from three replicates (RNA samples with RIN in the range of 7.9-10) were sent for microarray expression analysis at New Zealand Genomics Limited (NZGL), Auckland. Microarray procedure was carried out by Liam Williams, using ‘Affymetrix PrimeView Human Gene Expression arrays’ and ‘GeneChIP 3’ IVT Express Kit’ as per the manufacturer’s guidelines.

2.3.4.2 Data processing and analysis
Affymetrix PrimeView data (original CEL files) was pre-processed and normalised using robust multiarray average (RMA) algorithm. Gene expression analysis was performed using the statistical software R, and a number of R and Bioconductor packages. R and the R package limma were used to calculate potentially differentially expressed transcripts
between the treatment groups (control versus test, described in section 3.6.1). A linear model was used to estimate the fold changes for each gene and an empirical Bayes method (*eBayes* method in the Bioconductor package limma) was used to moderate the standard deviations of the estimated log-fold changes. Benjamini-Hochberg method was used to correct the p-values for multiple testing, and the corrected p-value was termed as the adjusted p-value. The Affymetrix probe IDs were annotated using the Bioconductor package hgu219.db_2.7.1. Above mentioned analyses were carried out by NZGL’s research programmers Vicky Fan and Louis Ranjard of the University of Auckland. Gene functions were annotated with Gene Ontology (GO) and NIH-DAVID (National Institute of Health-The Database for Annotation and Visualisation and Integrated Discovery, version 6.7), and pathway analysis was carried out using IPA (Ingenuity Pathway Analysis) and WebGestalt (WEB based GEne SeT AnaLysis Toolkit).

### 2.4 Protein expression studies

#### 2.4.1 Total cell lysate extraction

Used media from wells of the 6 well plate was aspirated off, and cells were washed with ice-cold PBS. After rinsing off the PBS, cells were lysed in ice-cold Radio-Immunoprecipitation Assay buffer (RIPA) (protease inhibitor tablet (Roche, USA) was added prior to use) by adding 60-100 µl of the RIPA buffer (volume of buffer was relative to confluence of the well). Cells were collected by scraping with a plastic cell scraper (Sigma, USA) and transferred to a labelled 1.5 ml eppendorf tube. Next, the cell suspension was homogenised by vigorous pipetting. The tube was placed on ice for 10 minutes and spun down at 13,000 rpm for 15 minutes at 4 °C. The supernatant (containing proteins) was transferred to another freshly labelled tube and stored at -20 °C, with an aliquot kept aside on ice for protein quantitation.

#### 2.4.2 Protein quantitation, SDS-PAGE and Western blotting

##### 2.4.2.1 Protein quantitation - BCA assay

Protein concentration was quantitated using the colourimetric ‘Bicinchoninic Acid’ kit (BCA Assay kit, Pierce, Thermo Scientific, USA) using the manufacturer’s guidelines. Frozen protein lysates were thawed on ice while the assay working reagent mastermix (50 parts reagent A + 1 part reagent B) was being prepared. Serial dilutions of albumin standards (BSA standards, supplied with the BCA kit) were also simultaneously prepared using 1x RIPA as a diluent. Reactions were prepared in duplicate on a 96 well plate format. Protein
samples/BSA standards and the freshly prepared detection reagent mastermix were added in a 1:20 ratio for each BCA reaction. Following plate set-up, reaction constituents were mixed by shaking the plate in a plate shaker (MS1 Minishaker, IKA-Works Inc., USA) for 30 seconds, followed by incubating the plate at 37 °C for 30 minutes in a non-sterile incubator (Sanyo, UK). At the end of the incubation period, absorbance was measured at 570 nm using a microplate reader (BIORAD, USA). Protein concentration of samples was subsequently determined by generating a standard curve using the series of albumin dilution standards, resulting in a linear relationship between concentration and absorbance.

2.4.2.2 SDS-PAGE and Western blotting
The volume equal to 30-60 µg of total protein lysate was added to 4x protein loading dye (to give 1x concentration, Table 2.5) and mixed by quick spinning. The mixture was denatured by boiling at 95 °C for 10 minutes. A 10% resolving gel and a 4% stacking gel was assembled, and protein samples were subsequently separated by polyacrylamide gel electrophoresis (PAGE) using a Mini-PROTEAN 3 electrophoresis system (BIORAD, USA). To determine electrophoretic speed and protein size, 10 µl of Novex sharp pre-stained protein standard (Life Technology, USA) and 2.5 µl of Odyssey® protein molecular weight marker (LI-COR Biosciences, USA) were loaded in separate wells alongside samples in appropriate wells. Electrophoresis was carried out at 20 mA per gel at room temperature using 1x protein electrophoretic buffer (Table 2.5) for 2-2.5 hours or until the desired separation was achieved. Following protein separation, the stacking gel was discarded, and proteins were transferred to Nitrocellulose (NC) membrane (Hybond-C Extra, GE Healthcare Life Sciences, NZ) using the Mini-Trans-Blot cell system (BIORAD, USA). Protein transfer was run overnight using a transfer cassette (comprising the NC membrane sandwiched between transfer buffer soaked layers of sponge and Whatman papers) at 12 volts in ice-cold 1x protein transfer buffer (Table 2.5). To reduce heat generated during protein transfer, an icepack was placed in the transfer cell alongside surrounding the entire transfer apparatus with ice. Following protein transfer, the membrane was removed from the transfer cassette and washed in PBS prior to blocking for a minimum of 1 hour at room temperature in Odyssey® Blocking buffer (LI-COR Biosciences, USA). The membrane was then incubated in the appropriate primary (1°) antibody diluted in the blocking buffer for 1.5 hours at room temperature. The membrane was washed 5 × 5 minutes in PBS-T (Table 2.5). After washing off excess of 1° antibody, the membrane was incubated with the appropriate fluorescent IRDye secondary (2°) antibodies (section 2.1.4.1) diluted in blocking buffer and
incubated at room temperature for 1.5 hours, protected from light (as 2° antibodies used were light sensitive). After completion of 2° antibody incubation, the membrane was again washed $5 \times 5$ minutes in PBS-T, followed by $2 \times 5$ minutes washes in PBS. After the final wash, excess wash solution was removed by gently tapping the edge of the membrane on a Whatman paper and complete drying of the membrane was achieved by sandwiching between two Whatman papers. The membrane was light protected by wrapping with aluminium foil, and subsequently visualised on the Odyssey Infrared detection and imaging system and detected bands were quantified using the Image Studio Lite software (LI-COR Biosciences, USA).

### 2.5 Chromatin Immunoprecipitation (ChIP) analysis

The ChIP technique makes use of the standard principles of immunoprecipitation to determine if a given protein binds to a specific DNA sequence and is extensively used to study differential recruitment of proteins to genomic loci of interest.

#### 2.5.1 Principle of ChIP

Firstly, the dynamic protein-DNA interactions between regions of close proximity on the chromatin template, at times of interest (45 minutes post oestradiol stimulation following 48 hours of RAD21 depletion, in this case) are chemically preserved by cross-linking for examination in varied functional states (in the presence and absence of cohesin, in this case). Cross-linked chromatin is then isolated from the sample and fragmented (creating random fragments) by sonication, to keep the average fragment size in the range of 200-800 bp. The lysate is pre-cleared by sedimentation and protein-DNA complexes are immunoprecipitated using a high-quality antibody, specific to the protein of interest (ERα, RAD21 and RNAPIIser5-P antibodies in this case). Following immunoprecipitation, magnetic or agarose beads (Protein A agarose/Salmon sperm DNA beads in this case) that bind to the antibody are used to collect antibody-bound protein-DNA complexes. This step is followed by stringent washes to remove any non-specifically bound chromatin before protein-DNA complexes are eluted off the beads. Next, the protein-DNA cross-links are reversed, proteins are digested and immunoprecipitated DNA is recovered and purified. The retrieved DNA is analysed by qPCR amplification using site-specific primers, to detect interaction of the targeted DNA molecule with proteins of interest. Enrichment of protein binding is quantified by measuring changes in fluorescence relative to abundance of the target sequence present in input chromatin.
2.5.2 ChIP Protocol

MCF7 cells were grown in large petridishes (BD Falcon, BD Biosciences, USA), used media was aspirated off, and cells were rinsed with PBS, followed by fixing with 1% formaldehyde solution (Table 2.7). Formaldehyde was quenched with 125 mM glycine, and fixed cells were washed thrice with PBS to wash off the formaldehyde-glycine solution. After washing, an additional 10 ml of PBS was added, and cells were scraped off the petridish using a cell-scaper (Sigma, USA). PBS containing scraped cells was transferred to a 50 ml tube and remaining cells in the petridish were collected by rinsing twice with PBS and subsequently added to the 50 ml tube, followed by pelleting of cells by spinning at 1350 × g for 5 minutes. The pellet of cross-linked cells was frozen at -80 °C until use. Fixed cells were lysed in ChIP lysis buffer containing protease inhibitor (Table 2.8) and incubated for 5 minutes on ice, prior to chromatin fragmentation by sonication. Equal volumes of lysed cells were aliquoted in 1.5 ml eppendorf tubes and sonicated for 5 × 10 second pulses at 25% amplitude using the Sonics Vibra-Cell sonicator (Model VCX 130, Sonics and materials Inc., USA). Optimisation of sonication conditions is discussed in section 2.5.2.1. Following sonication, samples were centrifuged at 10,000 × g for 10 minutes at 4 °C to solubilize sonicated chromatin. For each sample, the supernatant was pooled from the individual tubes and concentration of chromatin was estimated by spectrophotometric reading (NanoDrop ND-1000, Nanodrop Technologies Inc., USA). Nanodrop readings were used to normalize the amount of chromatin amongst control and test samples (Control siRNA versus RAD21 siRNA treatments, in this case) going into each immunoprecipitation reaction. Chromatin was normalised and diluted using ChIP dilution buffer (Table 2.8). Equal amounts of diluted chromatin were pre-cleared using Salmon sperm DNA/Proteinase A Agarose 50% slurry (Millipore, USA) for 30 minutes, with rotation at 4 °C. Following incubation, agarose beads were pelleted by centrifugation at 4000 rpm at 4 °C for 2 minutes. 5% of the pre-cleared chromatin was kept aside as ‘Input chromatin’ while the rest was divided into equal volumes and incubated overnight at 4 °C with specific antibodies. For every ChIP, a no-antibody control was also performed. Next day, immuno-complex (antibody-bound DNA-protein complex) was collected by adding additional volume of Salmon sperm DNA/Proteinase A Agarose slurry and incubated for 3 hours at 4 °C. Beads were collected by gentle centrifugation at 1000 rpm at 4 °C for 1 minute. Supernatant was discarded, and beads containing the immuno-complex were washed for five minutes, sequentially with 1 ml of low-salt, high-salt, and LiCl wash buffers (Table 2.8). Samples were kept on ice and beads were washed for 5 minutes with occasional inverting, followed by pelleting by centrifugation
at 1000 rpm at 4 °C for 1 minute, and discarding supernatant before the next wash. Beads were washed twice with 1 ml of TE buffer (Table 2.8), incubated on ice with occasional inverting, and pelleted by spinning at 1000 rpm at 4 °C for 1 minute. Bound protein-DNA complexes were eluted off the beads by incubating in elution buffer (Table 2.8) at 65 °C for 15 minutes with occasional inverting, followed by centrifugation at 4000 rpm for 2 minutes. After centrifugation, the eluate was collected, and the elution was repeated. For each sample, the eluates were combined and cross-linking was reversed by adding 5 M NaCl at a final concentration of 200 mM and RNase A (Invitrogen, Life Technology, USA) followed by overnight incubation at 65 °C. Next day, remaining proteins in the samples were digested by adding Proteinase K (Invitrogen, Life Technology, USA) at a final concentration of 40 µg/ml and incubated at 45 °C for 60 minutes. DNA was recovered by phenol:chloroform:isoamyl alcohol (Sigma, USA) extraction and precipitated with 1/10th volume of 3 M sodium acetate and 2.5 volumes of ethanol. Extracted DNA was resuspended in 22 µl of ultrapure distilled water (Invitrogen, Life Technology, USA). 1 µl of resuspended pre-cleared or immunoprecipitated chromatin was used for qPCR analyses.

2.5.2.1 Sonication characterisation

For accurate quantification of ChIP-enriched DNA by qPCR, the recommended optimal chromatin range is between 100-1000 bp (Aparicio et al, 2005). There are two standard approaches employed to shear chromatin, one is the enzymatic digestion approach (micrococcal nuclease digestion) and the second is mechanical shearing (ultrasonicator systems using ultrasonic sound waves). Out of the two approaches, mechanical shearing was chosen because nuclease fragmentation of chromatin would have been incompatible with formaldehyde cross-linking (previous step) of protein-DNA interactions (Johnson & Bresnick, 2002).

For sonication optimisation, the cross-linked cell pellet (of each sample) resuspended in lysis buffer was aliqoted into 1.5 ml eppendorf tubes. Sonication was performed using Sonics Vibra cell VCX 130 ultrasonic processor with 5 pulses of 8 or 10 seconds pulses at 20% or 25% amplitude frequencies, respectively. The degree of DNA fragmentation was examined by running the purified ChIP-enriched DNA on 1.5% agarose gels. During sonication optimisation, two variables were tested (time and amplitude, Figure 2.1A) to optimise yield of the ideal range of chromatin fragments. 5 × 10 second pulses at 25% amplitude (Figure 2.1A, lane 1) was found to generate fragments ranging between 250-850 bp (represented by
DNA smears apparent in the size range) which was ideal for the intended downstream ChIP application (Figure 2.1B). Hence, this sonication condition was used across all my ChIP experiments. To ensure consistent DNA fragmentation across all ChIP biological replicates, purified DNA from each replicate was size checked by agarose gel electrophoresis prior to protein enrichment analysis by qPCR.

**Figure 2.1**

**Figure 2.1 Sonication characterisation in MCF7 cells**

To determine correct sonication conditions for generating chromatin fragments of optimal size range (100-1000 bp), different pulse durations and amplitude frequencies were tested using the Sonics Vibra-Cell sonicator (Model VCX 130, Sonics and materials Inc., USA). (A) Shows purified sonicated DNA fragments run on a 1.5% agarose gel, obtained from the optimisation experiment. Lane 1: 5 sets of 10 seconds pulses at 25% amplitude; Lane 2: 5 sets of 8 seconds pulses at 25% amplitude; Lane 3: 5 sets of 8 second pulses at 20% amplitude; Lane 4: 5 sets of 10 second pulses at 20% amplitude. Lane 1 yielded the correct sized fragments (red arrows are indicative of the fragment size ranging between 250 to 850 bp). Hence, 5 sets of 10 seconds pulses set at 25% amplitude condition was chosen for prospective ChIP experiments. (B) Shows a representative agarose gel electrophoresis image showing correctly sized (250-850 bp) chromatin fragments from one of the ChIP biological replicates. Lane 1: Control siRNA + vehicle; Lane 2: Control siRNA + oestradiol; Lane 3: RAD21 siRNA + vehicle; Lane 4: RAD21 siRNA + oestradiol. Red arrows are indicative of the 250-850 bp fragment size range.
2.5.3 Detection and quantification of ChIP-ed DNA

2.5.3.1 ChIP primer design and validation

ChIP primers used in this study are listed in Table 2.12. ChIP-qPCR experiments were carried out to firstly validate previously identified RAD21 and ERα binding at genes of interest and investigate changes in ERα and RNA Pol II binding upon depletion of RAD21 in MCF7 cells. For identifying sites to design primers to, publicly available ChIP-seq tracks (.GFF files) from Schmidt et al (2010) were incorporated into ‘Geneious bioinformatic software’ (version R7) and primers were designed within previously identified peaks by primer3 (http://biotools.umassmed.edu/bioapps/primer3 www.cgi). Primers were designed as previously discussed in section 2.3.3.1. Primer specificity and efficiency was also assessed as mentioned in section 2.3.3.2. The only difference being, for generation of standard curves for ChIP primers, purified sonicated DNA was used as a template DNA instead of cDNA.

2.5.3.2 Standard curve based detection and quantification of protein binding at sites of interest

To quantify the amount of target DNA immunoprecipitated in a given ChIP sample, the pre-established standard curve (of the corresponding primer pair) consisting of serial dilutions of a known concentration of template DNA was employed. General reaction conditions were similar to expression primers as outlined in section 2.3.3 where 1 µl of resuspended ChIP DNA was used in a 20 µl qPCR reaction. For generation of standard curve for a primer set, Ct values of each sample were plotted against the log of the dilution of that sample, and the line of best fit was calculated using the equation, \( y = mx + c \). In this equation, \( y \) represents the Ct value, \( m \) represents slope of the line, \( x \) represents log of the amount of diluted DNA and \( c \) represents the Y-intercept. The unknown amount of target DNA (\( x \)) immunoprecipitated in each ChIP sample was calculated by rearranging the equation, \( x = (y-c)/m \). \( c \) (Y-intercept) and \( m \) (slope) of the line were known from establishment of the standard curve for every primer set and \( y \) (Ct value) was determined by qPCR analysis of ChIP samples. qPCR reactions were performed in duplicate for each ChIP replicate. \( x \) was calculated by subtracting the mean Ct value from the Y-intercept of the standard curve and dividing by the slope of the line, \( x = (\text{mean Ct value} - \text{Y-intercept})/\text{slope} \). \( x \) was then converted from the log scale by \( 10^x \).

Subtracting the amount of DNA in the ‘no-antibody sample’ from the amount of DNA in the ‘immunoprecipitated with antibody sample’ and dividing by the amount of DNA in the
corresponding ‘input sample’ then calculated the immunoprecipitation efficiency for each sample.

\[
\text{Immunoprecipitation efficiency of a sample} = \frac{\text{DNA in ‘antibody sample’} - \text{DNA in ‘no-antibody sample’}}{\text{DNA in ‘input sample’}}.
\]

The immunoprecipitation efficiency was expressed as % input, obtained by multiplying immunoprecipitation efficiency by 100. This signified the biological relevance of protein binding at a particular site of interest. The formula was used to validate the previously identified associations of ER\(\alpha\) and RAD21 proteins at the \(SOX4\) and \(IL20\) gene loci in MCF7 cells. The same formula was used to determine alterations in the fraction of ER\(\alpha\) and RNAPII\text{ser5-P} bindings, upon loss of RAD21 binding at these gene loci. Data from each ChIP replicate was subjected to mean centring and results were presented as an average of four biological replicates.

### 2.6 Immunofluorescence analysis

For immunofluorescence analysis, \(6.25 \times 10^4\) cells were seeded onto sterile coverslips in a 24 well plate format and treated as per the requirements of the experiment. Composition of the solutions used is listed in Table 2.9, and antibodies and their dilutions used are listed in section 2.1.4.3. Immunofluorescence stainings were carried as follows:

**Specimen preparation (PFA fixation of cells):**

i. After the specified treatment, used media from the wells of the 24 well plate was aspirated off, and 500 \(\mu\)l of thawed 4% PFA (thawed at 65 °C) was added to each well.

ii. Cells were fixed for 10 minutes at RT.

iii. The fixative (4% PFA) was aspirated off and cells were rinsed 3 × 2 minutes with 1 ml of PBS.

iv. Following rinsing, 1 ml of PBS was added to the wells and edges of the plate was sealed with a thermoplastic sealing film (Parafilm, Bemis Flexible Packaging, USA) and kept at 4 °C until immunostained (for a week or two).

**Immunostaining:**

iv. PBS was aspirated off and cells were permeabilised by adding 500 \(\mu\)l of 0.3 - 0.5% triton X-100 (10% Triton X-100 stock was diluted in 1x PBS) for 10 minutes at RT.

v. Triton was aspirated off and cells were rinsed 3 × 2 minutes with 1 ml of PBS, followed by adding 500 \(\mu\)l of blocking buffer.

vi. Cells were blocked at RT for 1.5 hours.
vii. Meanwhile primary antibodies were diluted in blocking buffer at the appropriate concentrations and added to the cells. This was followed by overnight incubation at 4 °C for immunostaining with anti-total and acetyl-SMC3 antibodies. However, for α-tubulin immunostainings, cells were incubated for an hour at RT.

vii. After incubation with primary antibody, cells were rinsed 3 × 2 minutes with PBS. 1 µg/ml of Hoechst (H33342) and appropriate fluorescent secondary antibodies were diluted in blocking buffer at the appropriate concentrations and added to the cells and incubated for 1.5 hour at RT. Coverslips were kept protected from light henceforth.

viii. Cells were rinsed 3 × 2 minutes with PBS.

ix. Coverslips were mounted onto frosted glass slides (LabServ, Thermo Scientific, USA) using the mounting medium ‘ProLong Gold Anti-fade mountant’ (Invitrogen, Life Technology, USA). Mounting medium was thawed at RT for an hour prior to use.

x. Slides were then allowed to dry by keeping in dark (protected from light) for 5-6 hours or until dried before imaging.

Imaging:

xi. Slides were viewed at 10x, 20x and 60x magnification on Nikon C2 confocal microscope using the appropriate excitation lasers, laser 488 for green (anti-rabbit antibody), laser 561 for red (anti-mouse antibody) and laser 405 for blue fluorescent light (Hoechst) and images were captured using Z-series feature of the NIS elements software. Exposure time and camera settings between control and treated cells were kept constant. Images were Z-stacked and processed using the FIJI software.

2.7 Cell cycle analysis

Analysis of cell cycle progression was accomplished by quantitating cellular DNA by staining with a fluorescent DNA binding dye propidium iodide (PI) (Sigma, USA). Nuclei of cells were labelled with PI and population of cells in each phase of the cell cycle was quantitated based on differences in fluorescence intensities. Quiescent (cells in G_0) and cells in G_1 phase have one copy of DNA and thus have 1x fluorescence intensity. Cells in S phase are synthesising DNA; hence, these have fluorescence values between 1x and 2x. However, cells in G_2-M have doubled their DNA; hence, these are represented by 2x fluorescence intensity. Solutions used in flow cytometric analyses are listed in Table 2.6.
2.7.1 Asynchronous cell cycle analysis

Ethanol fixation of cells:
For cell cycle analysis, $3 \times 10^5$ cells were seeded in a 6 well plate format in the appropriate media and treated as per the requirements of the experiment. After the specified treatment, used media containing floating cells were collected in a 15 ml tube and 1 ml of ‘1x PBS – 0.1% EDTA’ solution was added to the well and incubated at 37 °C in a non-sterile incubator for 5 minutes. Following cell detachment, cell solution was transferred to the appropriate tube containing floating cells from the same well. The tube was placed on ice, next. Cells were then pelleted by centrifugation at 2000 rpm for 5 minutes at 4 °C. Supernatant was aspirated off and cell pellet was washed in 5 ml of ‘1x PBS – 1% FBS’, followed by re-pelleting at 2000 rpm for 5 minutes at 4 °C. Cell pellet was resuspended in 500 µl of ice-cold PBS and fixed in 7 ml of pre-chilled (at -20 °C) 96% ethanol under slow vortexing (ethanol was added drop wise). Ethanol fixes and also simultaneously permeabilises the cell membrane. Following ethanol fixation, samples were stored at 4 °C overnight or up to 2 weeks until analysed.

Propidium iodide (PI) staining:
Following ethanol fixation, cells were pelleted by centrifugation at 2000 rpm for 5 minutes at 4 °C and ethanol was aspirated off. Cell pellet was washed in 5 ml of ‘1x PBS - 1% FBS’, followed by re-pelleting at 2000 rpm for 5 minutes at 4 °C. Cell pellet was next resuspended in 400 µl of ‘PI – RNase A’ solution having a final concentration of 10 µg/ml of PI and 250 µg/ml of RNase A (Invitrogen, Life Technology, USA) (to remove cellular RNA). The cell pellet was homogeneously mixed to obtain a single cell suspension and transferred to FACS tube. Tubes were incubated at 37 °C in a non-sterile incubator (protected from light) for 30 minutes to expedite PI binding to DNA. After incubation, FACS tubes were protected from light and placed on ice for flow cytometric analyses.

Flow cytometric analyses:
Cell cycle profiles were analysed either on a BD FACSCalibur Flow Cytometer (BD Biosciences, USA) or on a Beckman Coulter Gallios Flow Cytometer (Beckman Coulter, USA). The interface software was ‘CellQuest Pro Software’ (BD Biosciences) for the FACSCalibur. ‘Acquisition software’ was the interface software for Gallios Flow Cytometer. PI was detected in FL2 channel of the FACSCalibur, and FL3 channel of the Gallios Flow Cytometer. All samples from individual experiments were analysed at the same time, and control samples were used to calibrate instrument settings (voltage and current) prior to analysing the test samples. Using area (A), height (H) and width (W) parameters,
cells were gated to exclude cell aggregates, debris, and doublets. Typically a minimum of 10,000 events from each sample was acquired and analysed. The sample files were further processed and presented as histograms (with fluorescence intensity on the X-axis and cell counts on the Y-axis) using the Flow Jo software (version 9.7, Tree Star, USA).

2.7.2 Synchronised cell cycle analysis

Cell synchronisation by double thymidine block

Principle:
Cells were synchronised at G1-S phase by using two sequential exposures of the chemical inhibitor of DNA synthesis, thymidine (Sigma, USA). Each exposure was separated by a time interval marked with recovery from the thymidine block along with simultaneous cell growth. This method of cell synchronisation is proposed to work on the principles of excess thymidine-induced feedback inhibition of DNA replication and is suitable for investigating events occurring in S and G2-M phases of the cell cycle.

Procedure:
3 × 10^5 cells were seeded in a 6 well plate format in the appropriate media and allowed to grow to 25-30% confluence. 2 mM thymidine was added to the media for 18 hours (1st block). After 18 hours, the first block was released by aspirating used media, followed by 2 × 2 minutes washes with PBS (pre-warmed to RT). Cells were re-fed with fresh media and allowed to grow for the next 9 hours, followed by re-treatment with 2 mM thymidine for another 17 hours (2nd block). The second block was released at the end of 17 hours as before, synchronised cells were then allowed to cycle and treated as per the requirements of the experiment.

Cell cycle analysis:
After the specified period of treatment, cells were harvested, fixed, PI stained and analysed by flow cytometry on a Gallios Flow Cytometer (Beckman Coulter, USA) as described in section 2.7.1.

2.8 Functional assays

2.8.1 Assessment of cell proliferation

Cell proliferation was measured by the end-point based colourimetric MTT assay, and also by confluence-based, continuous kinetic monitoring of cell growth in real-time, using the Essen IncuCyte FLR digital imaging system (Essen Bioscience, USA).
2.8.1.1 MTT assay

Principle:
MTT assay is used to quantify cell proliferation based on the ability of metabolically active cells to reduce the yellow tetrazolium dye, MTT into a purple coloured formazan product. Oxidoreductase enzymes catalyse the reaction. The formazan product is solubilised and quantified via spectrophotometry (by reading absorbance at 570 nm) allowing quantification of changes in the rate of cell proliferation. The absorbance readings serve as a guide to measuring the proportion of metabolically active and hence proliferating cells, in the total cell population.

Procedure:
For the assay, appropriate numbers of cells were seeded (in 200 µl of culture media) in a 96 well plate format and treated as per the requirements of the experiment. For each treatment, samples were prepared in quadruplets. At the end of the treatment period, 10 µl of 5 mg/ml stock of the MTT reagent (Sigma, USA) was added to each well (at a final concentration of 0.25 mg/ml) and the plate was incubated at 37 °C in the cell culture incubator for 3.5-4 hours (or until the purple precipitate was visible). At the end of the incubation period, culture medium containing MTT reagent was aspirated off and 100 µl of DMSO was added to each well and also to four blank wells (DMSO only wells). Using the plate shaker, DMSO added to the wells was homogeneously mixed by brief shaking (10-20 seconds), and absorbance was read at 570 nm by labelling the four DMSO only wells as blank, using the microtitre plate reader (BIORAD, USA). For each treatment type, average values from quadruplet readings were determined (by subtracting the average reading of DMSO only wells) and cell proliferation was quantitated relative to control wells.

2.8.2.2 IncuCyte assessment of cell proliferation

IncuCyte (Essen Bioscience, USA) is an automated microscopy gantry that allows for monitoring of cell proliferation based on area (confluence) metrics, by temporal gathering and archiving of kinetic assay readouts (confluence-based measures and morphological imaging) in the native environment of a cell culture incubator. Using proprietary algorithms of monolayer confluence, cell proliferation is measured in real-time, in a label-free, non-invasive way, and proliferation is quantified by generation of a linear relationship between cell number and confluence.
Procedure:
To cross-verify MTT proliferation assessments, for every time point and treatment type, an additional cell culture plate (with similar cell seeding densities) was placed in the IncuCyte microscopy platform, and monitored in parallel. The cell plate was placed within one of the six microplate trays of the IncuCyte FLR, and the scan type was set to ‘Phase-contrast’ for acquiring images every 2-3 hours. At the end of the specified treatment periods, kinetic growth curves were generated, and cell growth was quantified and assessed.

2.8.2 Assessment of cell death
Quantitative assessment of cell death was carried out by YOYO-1 nuclear staining of DNA, using the IncuCyte FLR live-imaging system and also by the conventional trypan blue exclusion assay. In addition, cell death was also qualitatively assessed by agarose gel electrophoretic assessment of fragmentation of genomic DNA (gDNA) (assessment of chromatin cleavage) and by confocal microscopic evaluation of Hoechst (H33342) stained cells (morphologic analysis of fragmented/condensed nuclei).

2.8.2.1 YOYO-1® staining of DNA
Principle:
YOYO-1®* (Life Technology, USA) is a cell impermeant cyanine dimeric nucleic acid stain. When added to the culture medium, YOYO-1 binds only to dsDNA of cells that have lost plasma membrane integrity, and thus allows for evaluation of death-inducing ability of any cytotoxic agent when analysed on an IncuCyte FLR digital imaging system. At the end of the specified treatment periods, YOYO-1 fluorescently (green fluorescence) stains nuclear DNA of damaged cells, which can be morphologically monitored and quantified using the IncuCyte FLR object counting algorithm.

Procedure:
3 × 10^5 cells were seeded in the appropriate culture medium in triplicate in a 6 well plate format and were allowed to grow overnight. Following day, together with appropriate concentration of the cytotoxic agent, YOYO-1 reagent was added to each well at a final concentration of 100 nM (1:10,000 dilution of stock reagent). The cell plate was placed within a microplate tray of the IncuCyte FLR. Scan type was set to ‘Fluorescence & Phase-contrast’ and cell death was assessed using the built-in ‘object confluence’ and ‘object counting algorithm’ of the IncuCyte software. After the final scan, data were exported, and cell death was quantified relative to control wells.
2.8.2.2 Trypan blue exclusion assay

Cell death was also quantitatively assessed by one of the most conventional membrane integrity assays, the trypan blue exclusion assay as described in section 2.2.2.2. Based on the same principle of viable cells with intact cell membranes prevent uptake of the trypan blue stain. However, due to the loss of membrane integrity in dead or dying cells, trypan blue stain is taken up by cells with disrupted membranes. Percentage of dead cells was calculated as per the following:

\[
\% \text{ non-viability} = \left( \frac{\text{Total number of stained cells}}{\text{Total number of cells}} \right) \times 100
\]

2.8.2.3 Assessment of DNA fragmentation

Chromatin cleavage is a hallmark of cell death. Hence, separating purified gDNA on a conventional 1.8% agarose gel qualitatively assessed internucleosomal DNA fragmentation. gDNA was extracted sequentially by proteinase K digestion, followed by phenol:chloroform:isoamyl alcohol extraction and subsequent ethanol precipitation.

gDNA Extraction Protocol:

i. 3 × 10^5 cells were plated in a 6 well plate format and treated as per the requirements of the experiment. At the end of the treatment period, used media (containing floating cells) was collected in a 15 ml tube. Cells were washed with 1 ml of PBS, and harvested by trypsination and collected in the same tube containing floating cells of the respective well. This was followed by centrifugation at 10,000 rpm for 5 minutes.

ii. Supernatant was discarded and 200 µl of DNA lysis buffer (Table 2.4) was added to each cell pellet, to which proteinase K was added at a final concentration of 1 mg/ml (to digest contaminating proteins).

iii. The samples were incubated at 55 °C for 1 hour on a thermomixer (Thermomixer Comfort, Eppendorf, Germany) set at 300 rpm.

iv. The total volume for each sample was made up to 500 µl with ddH₂O, followed by incubating the samples at 95 °C (water bath) for 5 minutes for inactivating the proteinase K.

v. Equal volumes of phenol:chloroform:isoamyl alcohol (Sigma, USA) was added to each sample, followed by centrifugation at 16,000 × g for 5 minutes at RT.

vi. The aqueous phase was transferred to a freshly labelled 2 ml tube.

vii. Next, 1/10th volume of 5 M NaCl and 2.5 volumes of 96% ethanol were added to each tube containing 500 µl of the aqueous phase. The tubes were placed in the -80 °C freezer for 1.5-2 hours for precipitation of DNA.

viii. The tubes were centrifuged at 16,000 × g at 4 °C for 30 minutes to pellet the DNA.
ix. The supernatant was carefully discarded without disturbing the cell pellet.

x. 500 µl of ice-cold 70% ethanol was added to each sample, followed by spinning at 16,000 × g at 4 °C for 5 minutes.

xi. Ethanol was discarded, and the pellet was dried at RT for 10-15 minutes.

xii. The DNA pellet was resuspended in 100 µl of ddH$_2$O and DNA concentration was determined by nanodrop measurement (NanoDrop ND-1000, Nanodrop Technologies Inc., USA), followed by storage at -20 °C until use.

xiii. 1-2 µg of purified gDNA was separated on 1.8% agarose gel (as described in section 2.3.3.2.1) and assessed for fragmentation of DNA.

### 2.8.2.4 Hoechst staining of cells

In addition to assessment of gDNA fragmentation, changes to nuclear morphology (nuclei fragmentation or condensation) were also assessed by Hoechst staining of cells. Hoechst (H33342) is a fluorescent bisbenzamide dye that fluoresces blue upon contact with DNA and once bound, allows easy visualisation of nuclear DNA, when seen on a fluorescence microscope with appropriate excitation.

**Procedure:**

For this, 6.25 × 10$^4$ cells were seeded on sterile coverslips placed in a 24 well plate format and treated as per the requirements of the experiment. After the specified treatment, used media from the wells of the 24 well plate was aspirated off, and 500 µl of thawed 4% PFA (thawed at 65 °C) was added to each well.

ii. Cells were fixed for 10 minutes at RT.

iii. The fixative (4% PFA) was aspirated off and cells were rinsed 3 × 2 minutes with 1 ml of PBS.

iv. Cells were permeabilised by adding 500 µl of 0.2% triton X-100 (10% Triton X-100 stock was diluted in 1x PBS) for 10 minutes at RT.

v. Triton was aspirated off and cells were rinsed 3 × 2 minutes with 1 ml of PBS, followed by adding of Hoechst labelling solution (H33342, at a final concentration of 1 µg/ml) to the cells and incubating for 30 minutes at RT. Coverslips were kept protected from light henceforth.

vi. Hoechst labelling solution was aspirated off, and cells were rinsed 3 × 2 minutes with PBS.
vii. Coverslips were mounted onto frosted glass slides (LabServ, Themo Scientific, USA) using the mounting medium ‘ProLong Gold Anti-fade mountant’ (Invitrogen, Life Technology, USA). Mounting medium was thawed at RT for an hour prior to use.

viii. Slides were allowed to dry by keeping in dark (protected from light) for 1-2 hours.

Imaging:

ix. Slides were viewed at 10x, 20x and 60x magnification on Nikon C2 confocal microscope using the appropriate excitation laser (laser 405 for blue fluorescent light in this case) and images were captured using Z-series feature of the NIS elements software. Exposure time and camera settings between control and treated cells were kept constant. Images were Z-stacked and processed using the FIJI software.

2.9 Statistical analysis

GraphPad Prism (versions 5 and 6, GraphPad software Inc., USA) was used to graph and analyze most of the data, except for the microarray analyses. Details of specific statistical tests are included in respective figure legends.
CHAPTER 3

Identifying cohesin-dependent oestrogen-responsive genes in oestrogen receptor-positive MCF7 breast cancer cells

3.1 Introduction

Cohesin may affect transcription independent of its role in chromatid cohesion (Monnich et al, 2009; Pauli et al, 2010) in a remarkably tissue-specific manner (Horsfield et al, 2007) suggesting it has a pivotal but as yet poorly understood role in tissue-specific transcription.

As mentioned previously (in section 1.4), in breast cancer cells, cohesin binding sites frequently coincide with ERα binding sites and the cellular response to oestrogen changes with reduction in cohesin (Schmidt et al, 2010). The coincident genomic binding of cohesin and ERα underpinned the possibility of there existing a unique regulatory module (mutual dependence) for driving transcription of ER target genes.

Despite cohesin’s established role in regulating transcription across the genome (Kagey et al, 2010; Liu et al, 2009), there is uncertainty surrounding this role for individual ER target genes. In this regard, a clear relationship exists between cohesin and ERα in regulating expression of the MYC oncogene. However, whether there is a general transcriptional dependence on cohesin for oestrogen-responsive genes remains elusive. Therefore, to characterize cohesin-dependent genetic variation in oestrogen-sensitive breast cancer and to identify oestrogen-responsive genes potentially regulated by cohesin, a loss-of-function approach was undertaken. The rationale was to investigate the impact of cohesin removal on the expression of downstream genes, to unravel important functional connections and identify consequences on genome-wide transcription. Although RAD21 overexpression seems to be predominant in breast cancers (Atienza et al, 2005; Yan et al, 2012), due to technical challenges associated with stoichiometric overexpression of individual subunits of the multi-subunit protein, a gain-of-function or an overexpression model was not considered for this study. Moreover, enough is known about the consequences resulting from overexpression of RAD21 (Xu et al, 2011b) but the phenotypes do not necessarily reflect actual function of the gene. In this regard, it remains to be addressed whether transcriptional dysregulation caused by cohesin (RAD21) overexpression is implicated in causing/driving breast tumourigenesis. Hence, to understand the functional relevance of cohesin levels in ER-positive cells, a loss-of-function path was deemed suitable for this study. This approach
was expected to identify ER-responsive genes that are dependent on cohesin for transcriptional activity.

3.2 Hypothesis
Cohesin facilitates oestrogen-dependent transcription and engages in the regulation of multiple ER target genes in breast cancer.

3.3 Aim
The first aim of the project was to identify additional genes in breast cancer that are both direct targets of ERα and have altered expression upon removal of cohesin. The second aim was to identify the molecular pathways in which such cohesin-dependent oestrogen-responsive genes are operating.

3.4 Experiment Design
To uncover comprehensive features of cohesin’s participation in oestrogen-mediated transcription, I analysed gene expression in MCF7 cells depleted of the RAD21 subunit of cohesin. The MCF7 cell line is a well-characterised model for ER-positive breast cancers as it demonstrates exquisite oestrogen sensitivity through constitutive expression of ERα. The absence of mutations in cohesin subunits and its regulators (https://cansar.icr.ac.uk/cansar/cell-lines/MCF-7/mutations/), coupled with an overexpression of RAD21 (Atienza et al, 2005) makes the MCF7 cell line a biologically relevant model to investigate the functional relevance of cohesin in ER-positive breast cancer. For assessing responsiveness to exogenously added oestrogen, endogenous oestrogen was stripped out of the growth medium by supplementing the medium with charcoal dextran-treated serum (section 2.2.5). MCF7 cells were deprived of the steroid hormone oestrogen for a day, followed by RNAi-mediated depletion of the RAD21 subunit of cohesin. Forty eight hours post-siRNA treatment, a time point at which robust knockdown of RAD21 is achieved without discernible interference with the progression of cell cycle, cells were stimulated with 100 nM of 17-β-oestradiol (oestrogen) or treated with equivalent volumes of vehicle (ethanol). Cells were harvested at 3, 6, and 24 hours post-stimulation for analysis. Four independent biological replicates were generated. The oestradiol concentration used in my experiments was primarily derived from the oestradiol dose used in the Schmidt study (Schmidt et al, 2010), as I intended to carry out my transcriptome analysis in similarly treated MCF7 cells. The appropriateness of the used oestradiol dose could be
additionally justified by the fact, that a number of other research groups have used a similar concentration in their studies (Carroll et al, 2006; Fullwood et al, 2009; Kininis et al, 2009; Li et al, 2013; Theodorou et al, 2013). Although the plasma concentration of oestradiol in pre- and post-menopausal women is approximately hundred to thousand times lower than the used dose (Folkerd et al, 2014), the methodology followed in my experiments is comparable with previously published work. Furthermore, it is interesting to add that gene expression analysis carried out in our lab using lower doses of oestradiol (1 and 10 nM) (Jisha Antony, unpublished data), produced similar changes as that of 100 nM used across my study. Gene knockdown was confirmed both at the transcript and protein levels by qRT-PCR and immunoblot for all the four biological replicates. Given cohesin’s essential role in cell division, cell cycle progression was assessed by propidium iodide staining of DNA using flow cytometry. Phenotypic changes upon RAD21 knockdown were also assessed by IncuCyte FLR based growth monitoring, imaging and proliferation assays. The critical first step in obtaining meaningful microarray data is the use of high quality intact RNA. Hence, RNA samples from all the four replicates were assessed for integrity and quality by Bioanalyzer profiling. Upon ligand addition, oestrogen-dependent gene transcription occurs over a time period of 180 minutes (Metivier et al, 2003). Therefore, in order to attribute a direct role for cohesin in regulating oestrogen-responsive genes, it was essential to capture alterations in transcription as soon as oestradiol was added. Hence, quality controlled RNAs from three best of the four replicates from the 3 and 6 hours early response time points were analysed by cDNA microarray using Affymetrix GeneChip PrimeView Human Gene Expression Array. Global changes in gene expression were analysed to determine the regulatory network that cohesin is functioning in this class of breast cancer where cell proliferation is fuelled by oestrogen.
Figure 3.1 Schematic illustration of the experiment design used for sample preparation for cDNA microarray analysis

This figure illustrates the experimental steps followed, prior to the analysis of cohesin-dependent changes in expression of genes responding early to oestrogen by cDNA microarray.

3.5 Generation of independent biological replicates

3.5.1 Optimisation of RAD21 knockdown by RNAi-mediated gene silencing

RNAi is a sequence-specific gene silencing technique, which has been extensively used for genome-wide loss-of-function studies (Hannon, 2002). On-target plus RAD21 siRNA (Dharmacon, Thermo Scientific, USA) was used in this study to selectively and robustly deplete the cohesin protein in MCF7 cells. The on-target plus siRNAs allow guaranteed target gene knockdown with minimal off-target effects, as the siRNA design filters eliminate common seed regions and also incorporate a dual-strand modification pattern. The seed
region is the hexamer or heptamer sequence of nucleotides (positions 2-7 or 2-8) present in the anti-sense (guide) strand of a siRNA, which often induces off-target gene modulation due to its sequence complementarity with the 3’ UTRs. As a majority of unintended gene modulation is associated with siRNA seed region-related off-targeting, the on-target plus siRNAs are modified in the seed regions of their guide strands to prevent micro RNA-like off-targeting. These siRNAs additionally demonstrate a dual strand chemical modification wherein the sense (passenger) strand is blocked from RISC (RNA-induced silencing complex) uptake, thereby facilitating guide strand loading and simultaneously reducing passenger strand-mediated off-targets.

Three effective siRNAs from a Dharmacon smartpool, targeting different regions of the RAD21 gene were assessed for their cohesin-silencing effects (Figure 3.2). Gene knockdown was optimised and RAD21 siRNA DHA-J-006832-06 (GE Dharmacon, Thermo Scientific, USA) was chosen for further experiments. Although, all three siRNAs were effective in target gene silencing, however, the statistical significance of target mRNA/protein suppression was consistently higher (p-value <0.0001) for siRNA no. 6, for up to 72 hours, in comparison to RAD21 suppression by other siRNAs tested.
Figure 3.2 Optimisation of cohesin knockdown in MCF7 cells

The first step to identifying cohesin-dependent events in oestrogen-regulated transcription was the efficient knockdown of the RAD21 subunit of cohesin. Gene silencing efficacy of three effective siRNAs targeting distinct regions of the *RAD21* gene were tested. Oestrogen-deprived MCF7 cells were transfected with 5 nM of siRNA targeting *RAD21* (or non-targeting control) for 48 and 72 hours, and cells from these time points were harvested for RNA and protein. (A) Transcript levels of RAD21. RNA levels were determined using qRT-PCR and are shown relative to the mean of reference genes, *PPIA* and *GAPDH*. Error bars represent the SEM from 3 independent experiments. (B) Representative immunoblots showing decrease in RAD21 protein after transfection with RAD21 siRNA for the indicated time points. The bar graph depicts quantification of total RAD21 protein levels, normalised to γ-Tubulin. All three siRNAs were significantly effective in suppressing target mRNA (p-value <0.0001 for oligo no. 6 and 7; and p-value <0.001 for oligo no. 8) and protein expression (p-value <0.0001 for oligo no. 6, 7 and 8) post 48 hours of siRNA treatment. However, oligo number 6 was consistent in continuing to suppress expression of target mRNA and protein, up to 72 hours of
siRNA transfection (p-value <0.0001). Hence, siRNA no. 6 was chosen for use in all future experiments. Statistical significance was determined using 2-way ANOVA.

3.5.2 Profiling alterations in transcriptional responses to oestrogen with loss of cohesin required a robust knockdown of cohesin

Cohesin has a diverse spectrum of cellular roles, the best known of which is to ensure error-free chromosome segregation. Because of its role in the cell cycle, it was necessary to knockdown cohesin to a level that allowed for disruption of transcription without deleterious interference with cell division. Hence, the level of decrease in RAD21 transcript and protein was assessed for all four biological replicates. There was a significant reduction in the transcript level of RAD21 and approximately 75% decrease in protein levels, across all experiments (Figure 3.3). This level of knockdown was suitable to analyse the transcriptional role of cohesin as RAD21-depleted cells remained viable (Figure 3.6) over the period of study. Moreover gene expression was expected to be sensitive to moderate changes in cohesin levels, as previously reported in *Drosophila* (Dorsett, 2011), and also in CdLS patients wherein mild to moderate cohesin deficits altered gene expression of thousands of genes with no significant effects on chromosome cohesion or segregation (Liu et al, 2009). Hence, the amount of knockdown achieved was deemed suitable for the intended purposes of this study.
Figure 3.3

Oestrogen-deprived MCF7 cells were transfected with 5 nM of non-targeting control or RAD21 siRNA for 48 hours, then treated with vehicle or 100 nM 17-β-oestradiol for the indicated time periods. Knockdown of RAD21 was verified at both the transcript and protein levels at the indicated time points. (A) Transcript levels of RAD21 in MCF7 cells. RNA levels were determined using qRT-PCR and are shown relative to the mean of reference genes, *PPIA* and *GAPDH* (Appendix A). Error bars represent the SEM from 3 independent experiments. Significance was determined using two-way ANOVA. (B) Representative immunoblots showing decrease in RAD21 protein upon transfection with RAD21 siRNA at the indicated time points. The bar graph depicts quantification of total RAD21 protein levels, normalised to γ-tubulin, in MCF7 cells transfected with control or RAD21 siRNA. Total RAD21 protein levels decreased by 75% in RAD21 siRNA-treated conditions. Error bars represent SEM from 3 independent experiments. Significance was determined by paired ratio t-test.
3.5.3 Verification of oestrogen induction by assessing expression of well-characterised oestrogen-regulated genes

In order to correlate impacts of cohesin-dependent transcriptional dysregulation to oestrogenic response, it was important to verify if MCF7 cells were oestrogen-induced to a sufficient level, to evoke a robust response. To do so, I examined the expression levels of two well-characterised oestrogen-regulated genes MYC and GREB1 (Pellegrini et al, 2012), which also happen to be key proliferation genes in breast cancer. As mentioned before, MYC was previously identified by our group, to be directly dependent on cohesin for its response to oestrogen and is a validated model for verification of oestrogen stimulation (McEwan et al, 2012). In addition to MYC, I examined the level of another well-known oestrogen-responsive gene, GREB1 that belongs to a family of genes discovered in the year 2000 that were given the name GREB (Genes Regulated by Oestrogen in Breast Cancer) for their roles in oestrogen-induced tumour cell proliferation. GREB1 levels were shown to strongly correlate with levels of ERα in breast cancer cells (Rae et al, 2005). Hence, I included this gene to confirm that the oestrogen stimulatory conditions used in the experimental setup were optimal for the purposes of my study. Both genes demonstrated significant induction by oestrogen across all biological replicates (Figure 3.4). Oestrogen-induced expression levels of MYC confirmed that in the absence of RAD21, the expression level of MYC was significantly downregulated, implying RAD21 is required to maintain its response to oestrogen. On the contrary, oestrogen induction of GREB1 remained unaffected (statistically non-significant) despite depletion of RAD21. These results highlighted that not all oestrogen-responsive genes require cohesin for their expression.
Figure 3.4

**Figure 3.4 Verification of oestrogen induction by assessing levels of MYC and GREB1, two well-characterised oestrogen-regulated genes**

Oestrogen-deprived MCF7 cells were transfected with 5 nM of non-targeting control or RAD21 siRNA for 48 hours, then treated with vehicle or 100 nM 17-β-oestradiol for the indicated time periods. Total RNA was harvested at 3, 6, and 24 hours post vehicle/oestradiol treatments. (A) Transcript levels of MYC (B) Transcript levels of GREB1. RNA levels were determined using qRT-PCR and are shown relative to the mean of reference genes, PPIA and GAPDH. Error bars represent the SEM from 3 independent experiments. Significance was determined using two-way ANOVA.

3.5.4 Assessment of RNA quality

Crucial to any microarray analysis is the use of high quality, intact RNA starting material. As RNA quality is of paramount importance for generating biologically meaningful gene expression data, total RNA of all biological replicates of the 3 and 6 hour time points, post extraction from siRNA-treated cells were assessed by Bioanalyzer profiling. Integrity was
evaluated by Agilent Technology’s RNA integrity number (RIN) software that assigns a numerical integrity value on a scale of 1 to 10 (10 being most intact and 1 being most degraded) based on the complete electrophoretic trace of the RNA sample being run. This technique of microcapillary-electrophoretic-RNA separation-based integrity analysis is regarded more accurate than the traditional evaluation of integrity by measuring the ratio of 28s/18s ribosomal bands only. The RIN number of the samples for microarray ranged between 7.9-10 across both the time points and were considered suitable for the intended downstream application. An overall summary of quality assessments of RNAs from the 3 and 6 hour time points, is presented in Table 3.1. The Bioanalyzer profiles of RNAs from the 3 hour time point are shown below and that from the 6 hour time point are included in Appendix B.
Figure 3.5 Characterisation of RNA intactness by Bioanalyzer profiling

RNAs extracted from vehicle-treated/oestrogen-stimulated, control/RAD21-depleted MCF7 cells were assessed for quality and integrity by Bioanalyzer profiling. (A) Electropherograms of the 12 samples of the 3 hour time point. (B) Electrophoretic gel image of the run samples of the 3 hour time point. RNAs were size separated by voltage and detected by laser-induced fluorescence. The RIN algorithm calculated integrity by taking into account, the total electrophoretic separation of the RNA samples. High-quality intact RNAs were represented by distinct peaks of 18s and 28s ribosomal components and also by a 28s/18s ratio of <1.8.
Table 3.1 Summary of RNA integrity assessment of samples from the 3 and 6 hour time points

<table>
<thead>
<tr>
<th>Sample name</th>
<th>rRNA ratio (28s/18s)</th>
<th>RNA integrity number (RIN)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>3 hour time point</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control siRNA + oestradiol, Replicate 1</td>
<td>2.1</td>
<td>9.9</td>
</tr>
<tr>
<td>Control siRNA + vehicle, Replicate 1</td>
<td>2.2</td>
<td>9.7</td>
</tr>
<tr>
<td>RAD21 siRNA + oestradiol, Replicate 1</td>
<td>1.8</td>
<td>9.9</td>
</tr>
<tr>
<td>RAD21 siRNA + vehicle, Replicate 1</td>
<td>2.0</td>
<td>9.9</td>
</tr>
<tr>
<td>Control siRNA + oestradiol, Replicate 2</td>
<td>2.0</td>
<td>8.6</td>
</tr>
<tr>
<td>Control siRNA + vehicle, Replicate 2</td>
<td>2.0</td>
<td>8.2</td>
</tr>
<tr>
<td>RAD21 siRNA + oestradiol, Replicate 2</td>
<td>2.1</td>
<td>10</td>
</tr>
<tr>
<td>RAD21 siRNA + vehicle, Replicate 2</td>
<td>2.6</td>
<td>9.5</td>
</tr>
<tr>
<td>Control siRNA + oestradiol, Replicate 3</td>
<td>2.1</td>
<td>9.7</td>
</tr>
<tr>
<td>Control siRNA + vehicle, Replicate 3</td>
<td>2.4</td>
<td>8.8</td>
</tr>
<tr>
<td>RAD21 siRNA + oestradiol, Replicate 3</td>
<td>2.1</td>
<td>8.1</td>
</tr>
<tr>
<td>RAD21 siRNA + vehicle, Replicate 3</td>
<td>1.9</td>
<td>7.9</td>
</tr>
<tr>
<td><strong>6 hour time point</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control siRNA + oestradiol, Replicate 1</td>
<td>2.6</td>
<td>10</td>
</tr>
<tr>
<td>Control siRNA + vehicle, Replicate 1</td>
<td>2.7</td>
<td>10</td>
</tr>
<tr>
<td>RAD21 siRNA + oestradiol, Replicate 1</td>
<td>2.5</td>
<td>10</td>
</tr>
<tr>
<td>RAD21 siRNA + vehicle, Replicate 1</td>
<td>2.8</td>
<td>8.4</td>
</tr>
<tr>
<td>Control siRNA + oestradiol, Replicate 2</td>
<td>2.4</td>
<td>9.9</td>
</tr>
<tr>
<td>Control siRNA + vehicle, Replicate 2</td>
<td>2.5</td>
<td>10</td>
</tr>
<tr>
<td>RAD21 siRNA + oestradiol, Replicate 2</td>
<td>2.8</td>
<td>9.9</td>
</tr>
<tr>
<td>RAD21 siRNA + vehicle, Replicate 2</td>
<td>2.7</td>
<td>10</td>
</tr>
<tr>
<td>Control siRNA + oestradiol, Replicate 3</td>
<td>3.1</td>
<td>9.7</td>
</tr>
<tr>
<td>Control siRNA + vehicle, Replicate 3</td>
<td>2.8</td>
<td>10</td>
</tr>
<tr>
<td>RAD21 siRNA + oestradiol, Replicate 3</td>
<td>2.6</td>
<td>10</td>
</tr>
<tr>
<td>RAD21 siRNA + vehicle, Replicate 3</td>
<td>2.8</td>
<td>10</td>
</tr>
</tbody>
</table>

This table summaries the 28s/18s ratios and the RNA integrity numbers of samples of the 3 and 6 hour time points. For all samples tested, the ratios of 28s:18s ribosomal subunits were greater than 1.8 and the RIN numbers ranged between 7.9-10. Hence, these RNAs were considered to be of high quality and suitable to proceed with analysis by cDNA microarray.
3.5.5 Depletion of cohesin subunit RAD21 slowed cell cycle progression of MCF7 cells, without markedly affecting cellular viability

The consequences of cohesin knockdown on cell cycle progression were evaluated, to be sure that the transcriptional dysregulations observed are separable from cell cycle disruption. Following 24 hours of vehicle-treatment/oestrogen-stimulation and a total of 72 hours of cohesin depletion, cell cycle progression was analysed by propidium iodide staining of DNA followed by flow cytometry (Figure 3.6). DNA content was used as a means to identify anomalies or distinguish between cell populations of control versus RAD21-depleted cells. The proportion of cells in each cell cycle phase was analysed using the Flow Jo software (version 9.7). A significant reduction in the transition of MCF7 cells from G1 to S/G2-M phase was observed in both basal and oestrogen-induced progression of the cell cycle. However, RAD21-depleted cells showed no evidence of a sub-G1 population and retained a slight growth response to oestrogen, which indicated that these cells remained viable. Although, there was a cell cycle delay, there was no evidence of cell populations with < 2N or > 4N DNA content. The absence of abnormal DNA complement (< 2N or > 4N) in cell populations from the 24 hours time point indicated that analysis of transcriptional responses at the earlier time points (3 and 6 hours) are likely to reflect direct, primary transcriptional responses to cohesin-depletion.
Figure 3.6 Depletion of cohesin subunit RAD21 delayed cell cycle progression of MCF7 cells

Oestrogen-deprived MCF7 cells were transfected with 5 nM of non-targeting control or RAD21 siRNA for 48 hours, then treated with vehicle or 100 nM 17-β-oestradiol for 24 hours. Cells were then fixed, and stained with propidium iodide. Cell cycle profile analyses were carried out on BD FACSCalibur and processed using the FlowJo software (version 9.7). (A) Representative histograms of the flow cytometry analyses showing intensity of fluorescence on the X-axis and cell number on the Y-axis. (B) Bar graphs showing the percentage of cells quantified in the G0/G1, S and G2/M phases of the cell cycle in control and RAD21-depleted, vehicle-treated/oestadiol-stimulated conditions. Error bars represent the SEM from two independent experiments. Significance was determined by two-way ANOVA. In vehicle- and oestrogen-treated cells, RAD21 depletion reduced the progression of cells from G0/G1 to S, G2/M phase.
3.5.6 RAD21 depletion did not affect cellular growth and morphology

Cohesin can directly regulate genes involved in cell proliferation (Dorsett, 2011). MYC is one of the key proliferation genes and its binding and regulation by cohesin is universally conserved in humans and also across many other species (Kawauchi et al, 2009; Rhodes et al, 2010; Schaaf et al, 2009). To determine if the level of cohesin-depletion (up to 75% in the MCF7 cells) triggered any gross physiological stress that could have affected cell growth during the window of study, I examined the impact of RAD21 knockdown and its effect on oestrogen responsiveness, morphology and kinetic proliferation of MCF7 cells. The growth of siRNA-transfected oestrogen-stimulated MCF7 cells was periodically monitored using Essen IncuCyte FLR digital imaging system (Essen Bioscience). Kinetic growth curves were obtained by using IncuCyte’s algorithm for confluence. A comparison of confluence over time in control versus RAD21-depleted cells indicated that cells were still able to grow in the absence of RAD21 for up to 72 hours post knockdown with 24 hours of oestrogen stimulation (Figure 3.7A). The effect of RAD21 depletion on cell proliferation was also investigated using the MTT assay. Consistent with the IncuCyte results, the MTT assay also showed that depletion of RAD21 does not inhibit oestrogen-responsive cell proliferation up to 72 hours post siRNA-treatment (Figure 3.7B). To analyse cellular morphology, upon RAD21 depletion, phase-contrast images of control- and RAD21 siRNA-transfected cells were taken at 48 and 72 hours post-transfection. The morphology of RAD21-depleted cells looked similar to control transfected cells (Figure 3.7C). Although RAD21 depletion for up to 72 hours led to a cell cycle delay (Figure 3.6), these data suggest that there were no significant changes in growth profiles, cell viability or phenotype.
Figure 3.7. Growth profiles of MCF7 cells depleted of the RAD21 subunit of cohesin

Oestrogen-deprived MCF7 cells were transfected with 5 nM of non-targeting control or RAD21 siRNA for 48 hours, then treated with vehicle or 100 nM 17-β-oestradiol for 0 or 24 hours. (A) Kinetic growth curve assessment by IncuCyte FLR digital imaging. IncuCyte growth profiles of MCF7 cells transfected with control or RAD21 siRNA. Oestrogen-deprived MCF7 cells, transfected with control or RAD21 siRNA were seeded onto 96 well plates and 48 hours later were stimulated with 100 nM oestrogen. Growth profiles of the cells were periodically monitored for up to 24 hours post oestrogen treatment. Data points were collected and quantitated using monolayer confluence. The confluence versus time graph indicates that following depletion of RAD21, growth of these cells was not compromised for the time period studied. Error bars represent the
SEM from two repeats. (B) Assessment of cell proliferation by MTT assay. The impact of RAD21 depletion on the proliferation of MCF7 cells was determined by MTT assay in control or RAD21-depleted cells, following 0 and 24 hours of stimulation with oestrogen. MTT labelling solution was added at a final concentration of 0.25 mg/ml at the end of each treatment period. Formed formazan crystals were then solubilised and absorbance was measured at 570 nm using a spectrophotometer. No significant growth suppression was observed following RAD21 depletion. Error bars are SEM from two repeats. (C) Phenotypic assessment of siRNA-treated cells by visualisation of cell morphology. Cells were visualised on an IX71 Olympus inverted microscope. Phase-contrast morphological images were taken at a magnification of 20x using the Olympus DP71 camera and DC software, of control and RAD21-depleted cells treated with 0 and 24 hours of 100 nM oestrogen. Images were digitally zoomed and processed using the ImageJ software. The phenotype of MCF7 cells remained unaltered following RAD21 depletion.

### 3.5.7 RAD21 depletion did not affect ERα transcript and protein levels

As discussed in Chapter 1, the cell type-specific actions of the ligand oestrogen are mediated by oestrogen receptors α and β. So, in the event of ligand-modulated transcriptional dysregulation, one reasonably obvious question that emerges is whether levels of the ligand receptors remain stable. One example of cohesin influencing hormone receptor levels came from *Drosophila*, where cohesin removal was shown to transcriptionally repress the EcR (Ecdysone) receptor in *Drosophila* salivary glands (Pauli et al, 2010). However, in contrast, no alteration in the transcript levels of oestrogen receptor was observed in the microarray analysis carried out in zebrafish depleted for Rad21 (Rhodes et al, 2010). In support of the observation that cohesin depletion does not affect hormone receptor levels in vertebrates, other studies (Quintin et al, 2014; Schmidt et al, 2010) showed unaltered ERα protein levels in RAD21-depleted ER-positive cells. Moreover, consistent with my results (Figure 3.6), oestrogen-stimulated entry into the cell cycle was impeded owing to depletion of RAD21 (Schmidt et al, 2010).

Assuming there is a general genetic dependence on cohesin for expression of a subset of oestrogen-responsive genes, I examined the correlation between cohesin levels and expression of ERα. Data mining on cBioPortal for cancer genomics (Cerami et al, 2012; Gao et al, 2013) revealed ERα protein to be the top 11th protein that underwent a change in expression (p-value <0.013) in a cohort of breast invasive carcinoma (including 962 tumour samples, TCGA Provisional study), when there is an alteration in *RAD21* (mostly amplification). To further determine if cohesin depletion alters hormone receptor levels, I examined the effect of RAD21 knockdown on the expression of ERα, both at the transcript and protein levels.
Consistent with previously published data (Quintin et al, 2014; Schmidt et al, 2010) my results indicated that ERα mRNA (Figure 3.8A) and protein (Figure 3.8B) levels remained unperturbed with knockdown of RAD21 in the presence of oestrogen. This was not surprising considering the reported long half-life of ERα. Hence, with up to 72 hours of RAD21 removal and 24 hours of oestrogen stimulation, there was no alteration in the expression levels of the ERα protein. Notably, the basal expression of ERα transcript and protein were increased significantly, despite an absence of the cognate ligand oestrogen. Owing to these results, it seems likely that cohesin has a basal repressive effect on the expression of ESR1, considering expression of ERα goes up significantly when breast cancer cells are deprived of cohesin. These results hint towards the mechanisms governing dependence on cohesin for modulation of oestrogen responsiveness of ER target genes (in MCF7 cells), involving a strategy other than reducing levels of ERα by decreasing expression of the ESR1 gene.
Figure 3.8

RAD21 depletion did not affect oestradiol-induced transcript and protein expression of ERα

Oestrogen-deprived MCF7 cells were transfected with 5 nM of non-targeting control or RAD21 siRNA for 48 hours. Cells were then treated with vehicle or 100 nM 17-β-oestradiol for the indicated time periods. Total RNA and protein were harvested at 6 and 24 hours post vehicle/oestradiol treatments. (A) Transcript levels of ERα in MCF7 cells. RAD21 depletion did not alter oestrogen-induced mRNA levels of ERα. Notably, there was a significant increase in basal expression of ERα in vehicle-treated conditions, conditional on RAD21 knockdown. RNA levels were determined using qRT-PCR and are shown relative to the mean of reference genes, PPIA and GAPDH. Error bars represent the SEM from three independent experiments. Significance was determined using two-way ANOVA. (B) Representative immunoblots showing a significant increase in basal levels of ERα protein (in vehicle-treated condition) upon transfection with RAD21 siRNA, at the indicated time points. The bar graph depicts quantification of total ERα protein levels, normalised to γ-tubulin, in MCF7 cells transfected with control or RAD21 siRNA. Total ERα protein levels were increased despite an absence of oestrogen. Error bars represent SEM from three independent experiments. Significance was determined by two-way ANOVA.
3.6 Identification of oestrogen-responsive genes with altered expression upon cohesin depletion

To identify genes that undergo changes in expression with cohesin removal, differential gene expression of RAD21-depleted oestrogen-deprived MCF7 cells was profiled by cDNA microarray.

Global changes in gene expression in RAD21-depleted cells in the presence and absence of oestrogen were analysed using Affymetrix PrimeView Human Gene Expression arrays using the GeneChip 3’IVT Express Kit at NZGL, Auckland. Detailed experiment design is described in section 2.3.4.1. The array comprised a total of 48,658 reference sequence probe sets (11 probes per set for well annotated sequences and 9 probes per set for others; and each gene represented by one or more probesets) covering approximately 20,000 well annotated genes. Two early response time points (3 and 6 hours post stimulation with oestrogen) were chosen to profile gene transcription. Quality controlled RNAs from three independent biological replicates (Figure 3.5, Appendix B) were hybridised to the arrays by Liam Williams of University of Auckland. The biological replicates tightly clustered together and showed reasonably similar expression values for the specific treatments. Results from these analyses are, therefore, more likely to be real and less likely to be due to technical artefacts. This section presents results from the bioinformatic analyses of the cDNA microarray.

3.6.1 Microarray data analysis procedures

Data was processed and analysed as described in section 2.3.4.2. NZGL’s research programmers, Vicky Fan and Louis Ranjard of University of Auckland, used the statistical software R, and various Bioconductor packages (limma, affy, affyPLM, gplots, AnnotationDbi) to carry out this analysis. Analyses were performed to identify potentially differentially expressed transcripts for:

i. Differences between ‘vehicle-treated Control’ and ‘oestradiol-stimulated Control’:

(Control siRNA + oestradiol) – (Control siRNA + vehicle)

This comparison (Control.V vs Control.E) was undertaken to identify oestrogen-responsive genes.

ii. Differences between ‘Control’ and ‘RAD21 knockdown’ in the presence of vehicle:

(RAD21 siRNA + vehicle) – (Control siRNA + vehicle)
This comparison (Control.V vs RAD21.V) was undertaken to identify genes that change in expression with cohesin depletion, in the absence of oestrogen. These are cohesin-dependent genes whose basal expression becomes altered with the removal of cohesin.

iii. Differences between ‘oestradiol-stimulated Control’ versus ‘oestradiol-stimulated RAD21 knockdown’:

**(RAD21 siRNA + oestradiol) – (Control siRNA + oestradiol)**

Differentially expressed genes emerging from this analysis (Control.E vs RAD21.E) are potentially co-regulated by cohesin and ERα. This list of genes was used for most of the expression analysis.

Gene functions were annotated using Gene Ontology (GO). To get clues about the molecular/biological processes involving cohesin-ER network genes, functional annotation clustering was undertaken using NIH-DAVID, version 6.7 (Huang da et al., 2009). For insights into the regulatory mechanisms and associated biochemical pathways, WebGestalt and IPA softwares were used. Data was submitted at GEO under accession number GSE59908.

### 3.6.2 RAD21 depletion altered oestrogen responsiveness of a discrete subset of genes in MCF7 cells

As mentioned before, RAD21 is frequently overexpressed in breast cancers, and therefore it may be important to identify oestrogen target genes whose expression is influenced by RAD21/cohesin levels. Global changes in gene expression were analysed, and several genes were identified as differentially regulated upon RAD21 removal at the two early response time points. (Figure 3.9, Appendices C, D and E).

The heat maps (Figure 3.9A, Appendix C) show the contrasting gene expression patterns in ‘control’ and ‘cohesin knockdown (RAD21)’ cells after 3 and 6 hours of oestradiol stimulation. A number of transcripts had evidence for significant differential expression after adjustment for multiple testing (Benjamini-Hochberg false discovery rate (FDR) correction). To obtain a first perspective on the number of differentially expressed transcripts, four different adjusted p-value thresholds (<0.05, <0.005, <0.001, <0.0005) were used. However, the moderately significant amongst the four cut-offs (<0.005) was chosen for final data presentation, as with stringent filtering, there remains a possibility to miss real biological effects. Further hierarchical clustering uncovered transcripts that were more abundant (green colour) or less abundant (red colour) upon cohesin removal in oestrogen-stimulated MCF7.
cells. The heat map is more distinct for the 3 hour time point as compared to the 6 hours one suggestive of the earlier transcriptional events being more robustly differentiated and the response to oestradiol changing as the response develops over time. More transcripts were differentially regulated at 3 hours (218 probes representing 153 genes, adjusted p-value <0.005) (Appendix C) than at 6 hours (110 probes representing 77 genes, adjusted p-value <0.005) (Figure 3.9B), indicating statistically significant differential regulation of genes that respond transcriptionally within the first few hours of oestrogen stimulation. For convenience of presentation, the heat map of transcripts with adjusted p-values less than 0.001 from the 3 hour time point is included in Figure 3.9, as the total number of transcripts with adjusted p-values less than 0.005 (218 transcripts) could not be accommodated in a single heat map. Hence, these have been presented as part of four heat maps in Appendix C. 74 probes (representing 53 genes) were similarly differentially regulated at both 3 and 6 hours post oestrogen stimulation, indicating consistency of the early transcriptional response to oestrogen in RAD21-depleted cells. Out of the 153 genes that were differentially expressed at the 3 hour time point, 60 genes were upregulated, and 93 genes were downregulated in oestrogen-stimulated RAD21-depleted cells in comparison to vehicle versus oestrogen-stimulated control cells (Control vs Control.E). Similarly, for the 6 hour time point, 32 genes were upregulated and 45 genes were downregulated. This suggests that cohesin is involved in bidirectional regulation of oestrogen-responsive genes, and cohesin is not merely a facilitator but also a repressor of ER-mediated transcription. 12 (16%) out of the 77 genes dysregulated at the 6 hour time point and an additional 3 genes (identified from the adjusted p-value cut-off of 0.05 for being important in the aetiology of breast cancer) were selected for experimental validation (qRT-PCR), which will be discussed in detail in Chapter 4.

The fraction of differentially regulated genes in RAD21-depleted cells was 22% of the total number of oestradiol-responsive transcripts at 3 hours post-stimulation (n=72, adjusted p-value <0.005), and 12% at 6 hours post-stimulation (n=169, adjusted p-value <0.005). If the p-value is relaxed a bit further (adjusted p-value <0.05), the fraction of differentially regulated genes in RAD21-depleted cells of all the oestrogen-responsive transcripts, at 3 hours was 33% (n=270), and 23% (n=476) for the 6 hour time point. A summary of this data is presented in Table 3.2. This indicates that not all oestrogen-sensitive genes depend on cohesin for their oestrogenic response.

Scatterplots (Figure 3.10) were generated with vehicle versus oestrogen-dependent fold changes in control cells plotted on the Y-axis, and oestrogen-dependent fold changes in
control versus RAD21-depleted cells plotted on the X-axis, for the genes significantly regulated by RAD21 at both the 3 and 6 hour time points. Genes with transcription influenced by cohesin depletion appeared to cluster at least in four different regulatory categories: (i) genes that are normally upregulated by oestrogen but repressed by removal of cohesin, (ii) genes that are repressed by oestrogen but de-repressed by removal of cohesin (iii) genes that are upregulated by oestrogen and further upregulated by removal of cohesin, (iv) genes that are repressed by oestrogen and further repressed by removal of cohesin. The complete lists of these genes are included in Appendices D (3 hour time point) and E (6 hour time point).

Some degree of variation has been reported in between previously published genome-wide oestrogen-responsive microarray studies (Cheung & Kraus, 2010; Jagannathan & Robinson-Rechavi, 2011). To find out how well the identified gene signatures from my microarray overlapped with those of others, data were compared to gene expression signatures of previously published oestrogen-regulated expression studies. Publicly available datasets deposited in GEO were accessed, and the statistical test used for p-value correction applied across my dataset (Benjamini & Hochberg FDR correction) was applied to the GEO accessed studies. Appropriate adjusted p-value filters were used, and comparisons were drawn between my dataset and many of GEO-accessed expression studies. Comparison with these datasets revealed that one-fourth to two-thirds of my gene lists overlapped with datasets represented in the GEO submitted studies GSE9936, GSE4006, GSE8597, GSE5840, and GSE11324. The observed non-overlapping genes might be explained by one or a combination of the following factors: the source from where the cell lines were obtained, differences in growth conditions, duration/concentration of oestrogen treatment, the complexity of experiment design, the cDNA array platforms used, or the programs and baseline statistics (such as data normalisation, processing, computation parameters) used for the bioinformatic analysis of the generated data. A limitation to comparing microarray data from multiple studies is the unavailability of sufficient description of the procedures and analytical options used for the data analysis (Ioannidis et al, 2009). Incomplete data annotation leads to the generation of considerable discrepancies that make cross-comparison of results from different datasets challenging, ambiguous and inconclusive. However, considering a reasonable degree of overlap (one-fourth to two-thirds) existed between genes identified in my dataset and previous studies, my array data are likely to be representative.
Figure 3.9 Transcriptional changes associated with cohesin knockdown in oestradiol-stimulated MCF7 cells

Oestrogen-deprived MCF7 cells were transfected with 5 nM of siRNA targeting RAD21 (or non-targeting control) for 48 hours. Oestradiol (100 nM) was added, and cells were harvested at 3, 6 and 24 hours post-stimulation. RNAs from the 3 and 6 hour time points were analysed by microarray. (A) Heat map showing the contrast in gene expression patterns in ‘control’ and ‘cohesin (RAD21) knockdown’ MCF7 cells post 3 hours
(left) and 6 hours (right) of oestradiol stimulation. After correcting the p-values for multiple testing (by Benjamini-Hochberg method), differentially expressed transcripts with adjusted p-value <0.001 (3 hour time point) and <0.005 (6 hour time point) were used to generate the heat maps. Additional heat maps with adjusted p-value <0.005 covering 218 differentially expressed transcripts from the 3 hour time point are available in Appendix C. The columns (Y-axis) in the heat map represent biological replicates of samples while the rows (X-axis) represent genes classified into clusters based on most similar expression patterns. Expression patterns were assessed based on z scores of expression levels where red indicates low expression, black indicates intermediate and green indicates high expression. (B) Venn diagrams depicting differentially expressed transcripts. Venn diagrams were obtained from comparing the hybridisation intensities of two treatment groups (‘Control versus RAD21 knockdown’ and oestradiol-stimulated ‘Control versus oestradiol-stimulated RAD21 knockdown’) with an adjusted p-value threshold of <0.005; Left, 3 hour time point (3 hours of oestradiol exposure); right, 6 hour time point (6 hours of oestradiol exposure); center, comparison between oestradiol-stimulated transcripts at 3 and 6 hour time points.
Table 3.2 Summary of significantly differentially expressed transcripts from the 3 and 6 hour array

<table>
<thead>
<tr>
<th></th>
<th>3 hour array</th>
<th>6 hour array</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of transcripts</td>
<td>Number of transcripts</td>
</tr>
<tr>
<td>Adjusted p-value &lt;0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Control.Vehicle vs. Control.Oestrogen</strong></td>
<td>270</td>
<td>476</td>
</tr>
<tr>
<td><strong>Control.Oestrogen vs. RAD21.Oestrogen</strong></td>
<td>90</td>
<td>107</td>
</tr>
<tr>
<td>% of oestrogen-responsive genes affected by RAD21 knockdown</td>
<td>33</td>
<td>22</td>
</tr>
<tr>
<td><strong>Control.Vehicle vs. Control.Oestrogen</strong></td>
<td>72</td>
<td>169</td>
</tr>
<tr>
<td><strong>Control.Oestrogen vs. RAD21.Oestrogen</strong></td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>% of oestrogen-responsive genes affected by RAD21 knockdown</td>
<td>22</td>
<td>12</td>
</tr>
</tbody>
</table>

*Control.Vehicle is control siRNA transfected MCF7 cells treated with vehicle
Control.Oestrogen is control siRNA transfected MCF7 cells stimulated by oestrogen
RAD21.Oestrogen is RAD21 siRNA transfected MCF7 cells stimulated by oestrogen

This table shows the percentage of oestrogen-responsive genes at significance levels (adjusted p-value cut-offs <0.05 and <0.005) affected by RAD21 knockdown post 3 and 6 hours of oestradiol treatment in MCF7 cells.
Figure 3.10
3.6.3 Functional annotation of the differentially regulated genes

To determine the biological processes involving RAD21-regulated oestrogen-responsive genes, functional annotation clustering was undertaken using NIH-DAVID (Table 3.3). Up- and downregulated genes were clustered separately. For functional annotation enrichment analysis, the tool ‘functional annotation clustering’ was chosen over ‘functional annotation chart’ to avoid redundant categories coming up in the final analysis. The clustering tool uses a grouping algorithm that puts similar annotations together based on the degrees of co-association of genes. An EASE score threshold, a modified Fisher Exact P-value cut-off of 0.05 was used for gene-enrichment analyses, the smaller the resulting EASE score for a cluster, the more enriched it is. These scores are included as significance in Table 3.3 (3 hour time point) and Appendix F (6 hour time point).

Functional analysis indicated that significantly differentially expressed genes were not clustered in the genome. However, the range of biological processes that these genes are involved in are quite distinguishable from one another. The upregulated genes clustered in biological roles governing cell-cell adhesion, constitution of cellular plasma membrane, cytoplasm and the cell junctions. Upregulated genes also have roles in ion binding and developmental processes (Appendix F). Downregulated genes were related to positive regulation of developmental processes, modulation of catalysis of phosphate groups, cell migration, glycoproteins and regulation of protein metabolic activities. Repressed genes were also involved in activities regulating the cell cycle, cell death and maintenance of cellular homeostasis. Another functional cluster that was found enriched for the repressed genes of the 6 hour time point was positive regulation of leukocyte activation. The functional categorisation of the dysregulated genes revealed that cohesin-dependent genes of the oestrogen response pathway have roles in a variety of biological processes ranging from modulation of kinase activity to the regulation of developmental processes.
Table 3.3 Functional annotations of dysregulated genes in oestrogen-stimulated RAD21-depleted MCF7 cells at the 3 hour time point

Upregulated genes:

<table>
<thead>
<tr>
<th>DAVID functional annotation term</th>
<th>Genes</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell-cell adhesion</td>
<td>FREM2, B4GALT1, CELSR2, PCDHB13, PCDHB14, PCDHB16, PCDHB3, TLN2, VWF</td>
<td>3.4E-03</td>
</tr>
<tr>
<td>Plasma Membrane</td>
<td>FREM2, RGNEF, SWAP70, B4GALT1, APOA1, CELSR2, CACNA1I, DOK7, FMN1, IGSF5, IL17RB, MR1, KCNK6, PSKH1, PCDHB13, PCDHB14, PCDHB16, PCDHB3, SLC22A5, SLC6A6, SYTL2, TLN2, UTRN, VWF</td>
<td>7.5E-03</td>
</tr>
<tr>
<td>Topological domain: Cytoplasmic</td>
<td>FREM2, SIDT2, B4GALT1, CELSR2, CACNA1I, CHST8, CLCN4, GDPD1, IGSF5, IL17RB, MR1, PAM, KCNK6, PCDHB13, PCDHB14, PCDHB16, PCDHB3, SLC22A5, SLC6A6, TMPRSS3, VASN</td>
<td>3.8E-03</td>
</tr>
<tr>
<td>Ion binding</td>
<td>FREM2, KLF9, RGNEF, B4GALT1, CELSR2, CACNA1I, CLCN4, C9orf95, CSRP2, CYP1B1, IDH1, MSR3, PAM, KCNK6, PCDHB13, PCDHB14, PCDHB16, PCDHB3, SLC22A5, SYTL2, SYTL5, UTRN, ZNF185, ZFP36L1</td>
<td>5.4E-02</td>
</tr>
<tr>
<td>Cell junction</td>
<td>B4GALT1, DOK7, FMN1, IGSF5, TLN2, UTRN</td>
<td>4.7E-02</td>
</tr>
</tbody>
</table>

Downregulated genes:

<table>
<thead>
<tr>
<th>DAVID’s functional annotation term</th>
<th>Genes</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive regulation of developmental process</td>
<td>KITLG, SMAD2, IL20, IRX3, NRCAM, NUMBL, THBS1</td>
<td>1.8E-03</td>
</tr>
<tr>
<td>Regulation of kinase activity</td>
<td>KITLG, ADRA2A, CXCR4, CCNG1, CDKN3, IRS1, PDCD4, THBS1</td>
<td>1.3E-03</td>
</tr>
<tr>
<td>Regulation of cell migration</td>
<td>CXCR4, IRS1, MMP9, RTN4, THBS1</td>
<td>7.9E-03</td>
</tr>
<tr>
<td>Glycoprotein</td>
<td>ABCC1, DENND5B, EFEMP1, KITLG, MYC ST8SIA4, B4GALT6, AMIGO2, ADRA2A, CXCR4, DKK1, FAM171B, FAM20C, MMP9, NRXN3, NRCAM, PVRL2, PRTG, RPRM, SGC, SPCS3, S1PR3, STG1, SUMF1, SUSD4, THBS1, TPST2</td>
<td>3.9E-02</td>
</tr>
<tr>
<td>Regulation of cellular protein metabolic process</td>
<td>KITLG, FEM1A, IL20, PDCD4, SELT, TRMT6, THBS1</td>
<td>2.3E-02</td>
</tr>
<tr>
<td>Cell cycle</td>
<td>AKAP8, DBF4, RAD21, CCNG1, CDKN3, RPRM, THBS1, TACC1, MYC</td>
<td>2.7E-02</td>
</tr>
<tr>
<td>Cell death</td>
<td>RAD21, CXCR4, MMD, PDCD4, RTN4, SPG20, SGMS1, THBS1, MYC</td>
<td>1.8E-02</td>
</tr>
<tr>
<td>Cellular homeostasis</td>
<td>CXCR4, CLDN1, SELT, S1PR3, STC1, TXNRD1, MYC</td>
<td>2.1E-02</td>
</tr>
</tbody>
</table>
Classification of dysregulated genes (Control.E vs RAD21.E, p-value <0.005) based on their molecular and biological functions. Functional clustering tool of NIH-DAVID (version 6.7) was used to identify enrichment of functionally related group of genes (p-value <0.005) in order to understand the biological, molecular processes regulated by RAD21 in oestrogen-stimulated MCF7 cells at the 3 hour time point.

3.6.4 Functional network and pathway analysis of differentially regulated genes

To further characterize how cohesin-dependent oestrogen-responsive genes are networked in global regulatory modules, IPA (Ingenuity Pathway Analysis) tool was used. IPA analyses transcriptomic data using ‘Ingenuity Knowledge Base (IKB)’ which is a large repository of biological/molecular/chemical findings compiled using different sources, to identify the networks most significantly perturbed relative to one’s dataset. Genes from the 3 and 6 hour time points that met the adjusted p-value significance cut-off of <0.005 were used to identify the ‘Focus genes’. These genes were then used as starting points for the generation of functional networks using IKB. The networks were assigned scores based on the algorithms for significance of association. Networks with scores exceeding three were considered to be statistically significant (Calvano et al, 2005). A Fisher’s exact test was also employed to calculate p-values for determining that the functions assigned to each subset of genes (for a function in particular) is assigned to it, due to chance alone.

153 genes from the 3 hour time point associated with ten high-scoring networks. Significantly affected networks are presented in Appendix G. While analysing disease association of the dataset, cancer emerged as the most strongly associated disease (p-value range: 0.00736 - 0.000421) involving dysregulations in 121 genes. Further sub-classification of the cancer-associated genes revealed correlations of many of these genes to mammary tumours, ductal carcinoma, breast or ovarian cancers and cancer-related processes, namely cell viability of cancer cells, metastasis, differentiation and proliferation of cancer cells. These are presented in Figure 3.11. Perturbed molecular and cellular functions involved genes associated to cell-to-cell signalling and interaction (32 genes, p-value range: 0.00736 – 0.00000173); cellular assembly and organisation (32 gene, p-value range: 0.00736 – 0.00000173); cell function and maintenance (35 genes, p-value range: 0.00736 – 0.00000173); cellular growth and proliferation (59 genes, p-value range: 0.00736 – 0.00000411) and lipid metabolism (21 genes, p-value range: 0.00736 – 0.00000744). These functional categories are consistent with results obtained from NIH-DAVID (Table 3.3) substantiating strength of the gene family-functional associations.
77 genes of the 6 hour time point mapped to five high-scoring functional networks. A summary of this analysis is included in Appendix G. In addition to cancer, endocrine system disorders were most strongly associated to the dysregulated genes of the 6 hours dataset. Cancer was associated with 58 genes having a p-value range of 0.0171 – 0.00049 followed by endocrine system disorder involving 12 genes with p-value ranging between 0.0165 – 0.000560. Affected molecular and cellular functions involved genes associated to similar functional categories as that observed for the 3 hour time point. These include cell-to-cell signalling and interaction (19 genes, p-value range: 0.0171 – 0.000000213); cellular assembly and organisation (19 genes, p-value range: 0.0171 – 0.000000213); cell function and maintenance (19 genes, p-value range: 0.0171 – 0.000000213); cell movement (25 genes, p-value range: 0.0171 – 0.00000408) and cell death and survival (26 genes, p-value range: 0.0171 – 0.000101). An important observation of the upstream regulator analysis feature of the IPA tool was that ESR1 followed by HRAS were identified as the two top upstream regulators common to both the 3 and 6 hours datasets. It is expected that genes of the oestrogen response pathway would be regulated by ESR1. However, since more than one upstream regulator often modulates genes, it is plausible that some of the identified genes would well be modulated by additional co-activators and other co-regulatory molecules as well.

Next, to decipher the principal canonical networks affected in my dataset, the gene lists were subjected to molecular pathway analysis using another bioinformatics tool called WebGestalt. Pathway common analysis identified β1 integrin, ErbB1, EGFR and P13K/mTOR pathways as few of the major molecular pathways affected by depletion of cohesin. All of these pathways have pre-established associations with cell proliferation and cancer. Results from the 3 hour time point are included in Table 3.4 and those from the 6 hour time point are included in Appendix H.
Figure 3.11 Correlations of the significantly perturbed cancer-associated genes of the 3 hour time point to the underlying biological functions

Gene products and their associations to breast or ovarian cancers (i, ii) and the different biological processes implicated in cancer (iii, iv), as identified by Ingenuity knowledge base. The prediction legend summarises the correlations based on comparisons drawn between the expression level of a gene(s) as observed in the current dataset with known function of the gene(s) extracted from the knowledge base. Genes with green nodes are downregulated, and those with red nodes are upregulated in the dataset. An orange edge connector implies activation of the biological process, blue indicates inhibition, and yellow if the findings appear inconsistent.
### Table 3.4 Molecular pathway enrichment analysis of dysregulated genes in oestrogen-stimulated, RAD21-depleted MCF7 cells

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Number of genes</th>
<th>Total number of reference genes in the pathway</th>
<th>Ratio of enrichment</th>
<th>Raw p-value</th>
<th>Adjusted p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>β1 integrin cell surface interactions</td>
<td>19</td>
<td>1351</td>
<td>4.15</td>
<td>1.73E-07</td>
<td>8.23E-07</td>
</tr>
<tr>
<td>ErbB1 downstream signalling</td>
<td>18</td>
<td>1288</td>
<td>4.13</td>
<td>4.11E-07</td>
<td>8.23E-07</td>
</tr>
<tr>
<td>EGFR-dependent endothelin signalling events</td>
<td>18</td>
<td>1289</td>
<td>4.12</td>
<td>4.15E-07</td>
<td>8.23E-07</td>
</tr>
<tr>
<td>Class I P13K signalling events</td>
<td>18</td>
<td>1288</td>
<td>4.13</td>
<td>4.11E-07</td>
<td>8.23E-07</td>
</tr>
<tr>
<td>mTOR signalling pathway</td>
<td>18</td>
<td>1288</td>
<td>4.13</td>
<td>4.11E-07</td>
<td>8.23E-07</td>
</tr>
<tr>
<td>Insulin pathway</td>
<td>18</td>
<td>1288</td>
<td>4.13</td>
<td>4.11E-07</td>
<td>8.23E-07</td>
</tr>
<tr>
<td>IGF1 pathway</td>
<td>18</td>
<td>1291</td>
<td>4.12</td>
<td>4.25E-07</td>
<td>8.23E-07</td>
</tr>
<tr>
<td>IL5-mediated signalling events</td>
<td>18</td>
<td>1292</td>
<td>4.12</td>
<td>4.30E-07</td>
<td>8.23E-07</td>
</tr>
<tr>
<td>PDGF receptor signalling network</td>
<td>18</td>
<td>1293</td>
<td>4.11</td>
<td>4.34E-07</td>
<td>8.23E-07</td>
</tr>
<tr>
<td>IFN-gamma pathway</td>
<td>18</td>
<td>1296</td>
<td>4.10</td>
<td>4.49E-07</td>
<td>8.23E-07</td>
</tr>
<tr>
<td>Plasma membrane oestrogen receptor signalling</td>
<td>18</td>
<td>1301</td>
<td>4.09</td>
<td>4.75E-07</td>
<td>8.23E-07</td>
</tr>
<tr>
<td>ErbB receptor signalling network</td>
<td>18</td>
<td>1312</td>
<td>4.05</td>
<td>5.36E-07</td>
<td>8.38E-07</td>
</tr>
<tr>
<td>TRAIL signalling pathway</td>
<td>19</td>
<td>1328</td>
<td>4.23</td>
<td>1.33E-07</td>
<td>8.23E-07</td>
</tr>
</tbody>
</table>

Signalling cascades in which cohesin-regulated oestrogen-responsive genes are over-represented. WebGestalt was used to identify the enriched pathways of dysregulated genes at the 3 hour time point. Statistically significant pathway enrichment was determined using a hypergeometric test with a significance p-value cut-off <0.001. Benjamini-Hochberg method was used to correct the p-values for multiple testing and the corrected p-value was termed as the adjusted p-value.

### 3.7 Discussion

Recent developments in the field of cohesin biology have highlighted that the transcriptional regulatory role of cohesin is highly tissue-specific. Cohesin’s coincident binding with cell type-specific transcription factors, such as ERα in MCF7 cells and HNF4A and CEBPA in
HepG2 hepatocellular carcinoma cells (Schmidt et al, 2010) have supported the idea that cohesin is an important element of the transcriptional regulatory modules. Also, cohesin functions in a tissue- and context-dependent way.

There is strong support for the hypothesis that cohesin could influence transcription of steroid hormone-responsive genes from many prior studies (Li et al, 2013; Pauli et al, 2010; Prenzel et al, 2012; Quintin et al, 2014; Schmidt et al, 2010). This study set out to analyse how cohesin depletion affects the biology of oestrogen-sensitive breast cancer cells, and identify genes that are transcriptionally dependent on cohesin for their cellular response to oestrogen by using the MCF7 model system. A signature of cohesin-dependent oestrogen-responsive genes was identified generating some interesting findings, and the consequences on downstream biological and disease processes were explored.

A loss-of-function approach (siRNA-mediated silencing of the RAD21 subunit of cohesin) was undertaken to identify such genes. I aimed to profile gene expression changes that occurred relatively early (after 3 and 6 hours of oestrogen addition) following depletion of the RAD21 protein. With increasing time, the probability of gene expression perturbations occurring due to other secondary or tertiary factors (such as changes in cell physiology, defective chromosome segregation or DNA repair) becomes greater. Robust knockdown of the RAD21 subunit of the cohesin protein enabled assessment of immediate (3 and 6 hours post oestrogen addition) and late (24 hours post oestrogen addition) effects on oestrogen-responsive gene expression following cohesin depletion. Knockdown of the cohesin subunit RAD21 (to a level of 75% across all experiments) was accomplished. There is considerable technical difficulty in investigating the non-mitotic functions of proteins essential for cell proliferation, as there always remains an increased risk of compromising cell proliferation. Therefore, to rule out possibilities of indirect effects on transcription caused by cohesin’s role in chromosome separation, cell cycle analysis was performed 24 hours post-oestrogen addition (72 hours post RAD21 depletion, Figure 3.6). Cell cycle analysis confirmed that up to 72 hours of cohesin disruption, the transition of breast cancer cells from G\textsubscript{1} to S/G\textsubscript{2}-M was reduced to a significant extent both in the presence and absence of oestrogen. However, despite a slower transit through the cell cycle, no sub-G\textsubscript{1} population was observed implying that the cells remained viable. Cells with post-mitotic chromosome complement of greater or less than 4N were also not observed suggesting that these cells did not undergo missegregation of chromatids (with no evidence of aneuploidy) for the time period studied. It is notable in this regard that to additionally rule out the prospect of confounding cell cycle
regulated genes affecting the transcriptome, expression data from both the 3 and 6 hour time points were screened for the presence of dysregulated expression of key cell cycle regulators. It is interesting to note that only three (CDKN3, CCNG1 and CCNG2) out of eighty four key cell cycle regulating genes had evidence of differential expression upon cohesin depletion in my dataset. Hence, I contend that upto 75% of cohesin knockdown was suitable for teasing out the transcriptional regulatory function of cohesin without adversely affecting its cell cycle centric functions.

Keeping aside the siRNA seed region-related off-targeting (discussed in section 3.5.1), some degree of unintended transcript modulation is also associated with the induction of an immune stimulatory response due to the siRNA-induced activation of Toll-like receptors (TLRs). However, key members of the TLR family (such as TLR3, TLR7, TLR8), interferons, TNFα (Tumour necrosis factor α) and interleukins (such as IL-6) that are principal mediators of the TLR-induced inflammatory response were not identified in my dataset. Hence, it could be regarded that the contribution of RNAi-mediated artifacts to the observed transcriptome is marginally less (if at all any).

Phenotypic effects of cohesin disruption on cell proliferation and morphology were also assessed to ensure effects on transcription were not occurring due to physiological stress caused by depletion of cohesin. As measured by the MTT assay and assessed by IncuCyte growth monitoring, RAD21 depletion did not have a detrimental impact on the proliferation of MCF7 cells up to 24 hours of oestrogen stimulation (Figure 3.7). Phase-contrast microscopy revealed no discernible phenotypic alterations in RAD21-deplete cells when compared to control MCF7 cells, suggesting an absence of physiological stress.

I did not observe a significant alteration in the transcript or protein levels of ERα in MCF7 cells (Figure 3.8) following 24 hours of oestrogen addition. Intriguingly, a study published in 2012 (Prenzel et al, 2012) reported that SMC3-mediated siRNA depletion of cohesin downregulated levels of the ERα protein as early as 12 hours after transfection with SMC3 siRNA. Perhaps the observed downregulation of ERα could be attributed to siRNA-transfected cells becoming completely quiescent (non-cycling) over the period of study. It was noteworthy that in the Drosophila model system, although there was a decline in EcR-B1 mRNA within 4 hours of cohesin cleavage, the change in protein levels of the EcR protein happened much later (post 8 hours). Therefore, direct transcriptional regulation of
the EcR gene by cohesin could not be solely accounted for the early affected ecdysone-responsive genes (within the first 4 hours).

ERα protein levels remaining unaltered by cohesin depletion in ER-positive cells (as observed in my experiments) is supported by data previously published by our group (McEwan et al, 2012) and also by work published by Quintin et al. 2014 (Quintin et al, 2014). Work by Schmidt and colleagues (Schmidt et al, 2010) additionally showed that RAD21 was necessary for oestrogen-mediated G1-S cell cycle transition, but its depletion did not affect ERα protein levels. Hence, it seems likely that cohesin’s influence on oestrogen-dependent gene regulation in MCF7 cells is not via alteration of ERα levels but other alternative mechanisms. These mechanisms may involve modulation of ERα binding to chromatin by recruitment of other co-factors (like mediator complex), chromatin modifiers or by direct interaction with the receptors (discussed in detail in Chapter 4).

In Drosophila, 7.7% of all the EcR-responsive genes were differentially regulated after cleavage of cohesin (Pauli et al, 2010). Cohesin depletion in MCF7 cells appeared to affect 12-22% of all the oestrogen-responsive genes in the first 6 hours of oestrogen stimulation, a reasonably larger percentage of genes than expected by chance alone. Gene expression was both up- and downregulated, consistent with previous findings in Drosophila. The gene responses were fairly rapid. Hence, the chances of these being indirect effects of cohesin removal are minimal. However, as MYC expression is dependent on cohesin, one possibility of indirect gene regulation (although minimal, owing to the shorter oestrogen stimulatory periods) was posed by cohesin’s influence on the MYC oncogene. To address this, expression data from both the 3 and 6 hour time points were additionally screened to identify direct MYC transcriptional targets. Expression changes in five (CYP1B1, CXCR4, KYNU, EFEMP1 and PDCD4) known MYC target genes were identified in my dataset, by comparing against a list of MYC targets identified in MCF7 cells by a previous study (Cappellen et al, 2007). However, interestingly, the five identified MYC target genes are bound by cohesin (RAD21) in oestradiol-stimulated MCF7 cells, which raises the possibility that these genes could be potentially co-regulated by both MYC and cohesin.

RAD21 depletion affected genes that were widely dispersed in the genome with implications in diverse biological processes (section 3.6.3). A number of breast cancer associated genes THBS1, DKK1, MMP9 were repressed following RAD21 knockdown and the repressed state of these genes is correlated to mammary tumours as per disease association analysis by
IPKB (Figure 3.11). Other important breast cancer-associated genes that were negatively affected with cohesin depletion include \textit{CXCR4}, \textit{IL20} and \textit{SMAD2}. Interestingly, the transcription factor \textit{GATA3}, known for its high association with ER expression in Luminal A subtype of breast cancers (Voduc et al, 2008), was also identified to be downregulated following RAD21 depletion and its expression level was correlated to ductal breast carcinoma as per IPA analysis (Figure 3.11). Other noteworthy breast cancer-associated genes that were positively regulated by cohesin depletion include \textit{APOA1}, \textit{SERPINF1}, \textit{IDH1} and \textit{PIK3R3}. \textit{PIK3R3} is particularly important because protein phosphoinositide-3-kinase regulatory subunit 3 encoded by this gene belongs to the PIK3 (phosphatidylinositol 3-kinase) protein family and has been shown to interact with IGF-1 (Insulin-like growth factor 1) and IRS1 (Insulin receptor substrate 1) proteins (Mothe et al, 1997) which are upstream nodes pivotal for activating the PI3K/AKT/mTOR cascade of signalling. \textit{IRS1} was also one of the significantly affected genes identified in the microarray analysis. Canonical pathway analysis confirmed PI3K/mTOR pathway (Table 3.4) as one of the most significantly perturbed pathways with overrepresentation of dysregulated genes. This signalling pathway is frequently activated in breast cancers. Hyperactivation of PI3K/mTOR pathway is implicated in fuelling the tumourigenesis process in ER-positive tumours (Ciruelos Gil, 2014) and leads to the development of endocrine therapy resistance (Cavazzoni et al, 2012; Miller et al, 2011). Of note, the PI3K/mTOR pathway has a deep-rooted interdependence with ER signalling. Therefore, altered expression of genes that constitute crucial nodes of this pathway following RAD21 depletion indicates that cohesin may be important for the PI3K/mTOR pathway to function in ER-positive breast cancer cells (Leiserson et al, 2015).

The microarray gene expression analysis identified differentially expressed genes following RAD21 depletion, and oestrogen stimulation that belong to nodes functional in the pathway crosstalk between ER and PI3K/mTOR pathways are presented in Figure 3.12. Crosstalk between the two pathways occurs at multiple levels. Reciprocal signalling between these pathways is proposed to mediate endocrine resistance, as ER becomes activated by direct phosphorylation by growth factor signalling molecules. Findings from the microarray thus support the idea that cohesin function might be important for tumour cell proliferation and/or development/sustenance of endocrine-resistance in ER-positive breast tumour cells. Lastly, it must be noted that in addition to the significantly affected gene expression changes that likely mediate the functions of cohesin in ER-positive breast tumour cells, it is possible that multiple modest transcript level changes (statistically insignificant in these experiments) may also contribute in eliciting the cohesin-dependent biologic responses.
Figure 3.12

Figure 3.12 Model of ER and growth factor signalling pathway crosstalk in human breast cancer

Reciprocal signalling is observed between ER and growth factor ErbB (erythroblastic leukemia viral oncogene homolog) or EGF (epidermal growth factor) and IGF (Insulin-like growth factor) networks leading to the development of a cross-activating loop. This cross-activation is thought to be the mechanism behind ligand-independent ER signalling resulting in treatment resistance to endocrine therapy. Pathways marked with red stars are identified to be dysregulated upon RAD21 knockdown in oestrogen-stimulated conditions.
CHAPTER 4
How does cohesin affect transcription of ER target genes?

4.1 Introduction

Oestrogen-mediated gene transcription occurs via ER. The canonical pathway is the genomic, ligand-dependent pathway, in which oestrogen binds to ER and initiates transcription of target genes by interaction with specific DNA sequences called the oestrogen response elements (EREs) located in or near the gene promoters. However, a significant proportion of oestrogen-mediated transcriptional regulation happens via alternative pathways (also genomic) where ER tethers to other transcription factors and co-regulatory proteins to indirectly influence gene expression (Marino et al, 2006; Wall et al, 2014a; Zwart et al, 2011). Such non-canonical pathways involve modulation of the chromatin milieu by epigenetic modifications (such as histone modifications) and chromatin remodelling. Several genome-wide ERα mapping studies in the past have highlighted that majority of the ERα binding sites are localised to regions distant (>10^5 bp) from transcribed genes (Carroll et al, 2006; Gao & Dahlman-Wright, 2011; Krum et al, 2008; Lin et al, 2007; Ross-Innes et al, 2010; Welboren et al, 2007). Seventy-eighty percent of ERα binding events occur in intronic and intergenic regions, and only up to 7% is confined to 5 kb region upstream of TSS.

In MCF7 cells, cohesin’s control of ER target genes can occur via one or a combination of the following possible mechanisms:

(i) By altering binding of ERα to its target genes. Altered ERα expression as a contributor, is unlikely because ERα levels remained unaffected upon cohesin depletion in MCF7 cells (section 3.5.7, Figure 3.8).

(ii) By modifying the chromatin environment via recruitment of chromatin remodellers (Kawazu et al, 2011) and histone modifiers (HATS) (Reid et al, 2009) that help ERα to facilitate gene transcription.

(iii) By enabling looping between distal ERα binding sites to promoters of target genes to initiate gene transcription (DeMare et al, 2013).

With the microarray analysis, I was able to identify a signature of cohesin-dependent genes (as outlined in Chapter 3) and establish that cohesin is involved in the regulation of a restricted set of ER target genes in MCF7 cells. The next goal was to select candidate genes for downstream analysis, to validate results obtained from the microarray, and to further
interrogate the mechanism by which cohesin mediates bidirectional regulation of candidate genes responsive to oestrogen.

4.2 Aim

To experimentally validate microarray expression data, and select candidate genes to further explore their regulation by cohesin. To test this, a ChIP-qPCR approach was taken to first determine if cohesin and ERα bind directly to select candidate genes, and second, if ERα binding is altered in response to loss of cohesin.

4.3 Results

4.3.1 Validation of microarray data by qRT-PCR

To validate expression data obtained from the microarray, a total of 14 genes (IL20, SOX4, DKK1, BMPR2, TFF2, PIM2, THBS1, BAG1, CXCL12, IRS2, CCNG2, PDCD4, PAK1, INSR) were chosen from the differentially expressed gene list of the 6 hour time point, as it served as a good intermediate timepoint to validate cohesin-influenced transcriptional events. 11 genes were chosen from the adjusted p-value cut-off list of <0.005 (n=77, 14.3% chosen for validation) and an additional 3 genes were chosen from the adjusted p-value cut-off list of <0.05. The selection of the list of 14 significantly differentially expressed genes (relative to oestradiol-treated control in the microarray) for follow-up analysis was based on their previously published associations to breast cancer aetiology.

Transcript levels were determined relative to reference genes PPIA and GAPDH, which remained unaltered under all experimental conditions (Appendix A). Out of the 14 genes chosen for validation by qRT-PCR across all three biological replicates (same biological replicates were used for the microarray), transcript levels of 11 genes were confirmed experimentally. However, expression level changes for 3 genes (INSR, PDCD4, PAK1) were inconclusive due to high experimental variation. Hence, these three genes were excluded from the final validated analyses. Details of the eleven final validated genes are summarised in Table 4.1.

Under conditions of oestrogen stimulation, cohesin depletion had a rather qualitative effect on the expression of a distinct subset of genes (Figures 4.1 and 4.2) showing some interesting trends. SOX4, a gene that was relatively insensitive to oestrogen, gained oestrogen sensitivity upon RAD21 depletion indicating that the normal function of cohesin is to restrain oestradiol-mediated transcription of this gene. However, DKK1 that was normally
repressed by RAD21 depletion was further repressed in RAD21-depleted conditions. The expression trend was different for another oestrogen-repressed gene *BMPR2*, which was de-repressed following cohesin depletion. Interestingly, for the highly oestrogen-induced gene *IL20*, loss of RAD21 was accompanied by a significant decrease in its expression level. A number of other oestrogen-activated genes behaved in a similar way, namely *THBS1, PIM2, IRS2, CXCL12* and *BAG1*. On the contrary, oestrogen activation of *TFF2* was dramatically enhanced, whereas repression of *CCNG2* was reduced, in response to loss of RAD21.

Upon cohesin depletion, differential transcriptional effects were notable within the first 6 hours of oestrogen addition. While some transcriptional effects were transient and restricted to the first 6 hours (*DKK1, BAG1*, Figures 4.1 and 4.2), most responses persisted for up to 24 hours. Altered regulation was less apparent in vehicle-treated conditions indicative of basal regulation not entirely accounting for differential transcription of the identified genes. Taken together, qRT-PCR analyses validated expression data obtained from the microarray and confirmed that cohesin does not merely facilitate oestrogen responsiveness of target genes, but can also impede ER-dependent transcription in a gene-specific manner.
Figure 4.1 Depletion of cohesin (RAD21) promotes as well inhibits expression of oestradiol-responsive genes

Oestradiol-deprived MCF7 cells were transfected with 5 nM of siRNA targeting RAD21 (or non-targeting control) for 48 hours. Oestradiol (100 nM) was added, and cells were harvested at 3, 6 and 24 hours post-stimulation for RNA isolation. Quantitative RT-PCR analysis of selected genes revealed unique transcriptional profiles following oestradiol stimulation of cohesin (RAD21)-depleted MCF7 cells. (A) RAD21 depletion sensitised SOX4 to stimulation by oestradiol. (B) Depletion of RAD21 prevented oestradiol induction of IL20. (C) RAD21 depletion prevented oestradiol-mediated repression of BMPR2 expression. (D) RAD21 depletion further enhanced repression of DKK1 by oestradiol. Transcript levels for the indicated genes were quantitated by qRT-PCR and normalised to the mean of two reference genes PPIA and GAPDH. Error bars represent the SEM of three biological replicates. Significance was determined using two-way ANOVA.
Figure 4.2 Expression changes of seven additional cohesin-regulated candidate genes identified from the microarray
Depletion of cohesin subunit RAD21 modulated oestrogen responsiveness of TFF2, PIM2, THBS1, BAG1, CXCL12, IRS2 and CCNG2. MCF7 cells were oestrogen starved and transfected with control or RAD21 siRNA as described in the previous experiments. Cells were treated with vehicle or stimulated with 100 nM of 17-β-oestradiol for 24 hours. Transcript levels were determined by qRT-PCR and are represented relative to the mean of two reference genes, PPIA and GAPDH. Error bars represent the SEM from three independent experiments. Significance was determined using two-way ANOVA.
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>AFFIMETRIX_IDs</th>
<th>Description of function</th>
<th>Log2 Fold Change (microarray)</th>
<th>Adjusted P-value (microarray)</th>
<th>Fold Change (qRT-PCR)</th>
<th>Adjusted P-value (qRT-PCR)</th>
<th>Effect of RAD21 loss on oestrogenic response (microarray &amp; qRT-PCR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFF2</td>
<td>11721090_at</td>
<td></td>
<td>-1.90801</td>
<td>-0.90999</td>
<td>0.91521</td>
<td>1.59449</td>
<td>Up</td>
</tr>
<tr>
<td>SOX4</td>
<td>11721810_at</td>
<td></td>
<td>-0.90999</td>
<td>0.91521</td>
<td>1.59449</td>
<td>0.003847</td>
<td>Up</td>
</tr>
<tr>
<td>PIM2</td>
<td>11719561_s_at</td>
<td></td>
<td>0.90999</td>
<td>0.91521</td>
<td>1.59449</td>
<td>0.003847</td>
<td>Up</td>
</tr>
<tr>
<td>TFF2</td>
<td>11723751_at</td>
<td></td>
<td>0.90999</td>
<td>0.91521</td>
<td>1.59449</td>
<td>0.003847</td>
<td>Up</td>
</tr>
<tr>
<td>BAG1</td>
<td>11753334_x_at</td>
<td></td>
<td>-0.77667</td>
<td>0.91521</td>
<td>1.59449</td>
<td>0.003847</td>
<td>Up</td>
</tr>
</tbody>
</table>

Table 4.1 Dysregulated genes identified by cDNA microarray and experimentally validated by qRT-PCR, in cohesin (RAD21)-depleted MCF7 cells exposed to 6 hours of oestradiol.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
<th>Gene</th>
<th>Description</th>
<th>Gene</th>
<th>Description</th>
<th>Gene</th>
<th>Description</th>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCNG2</td>
<td>Cyclin G2</td>
<td>BMPR2</td>
<td>Bone morphogenetic protein receptor Type II</td>
<td>DKK1</td>
<td>Dickkopf 1 homolog</td>
<td>IRS2</td>
<td>Insulin receptor substrate 2</td>
<td>CXCL12</td>
<td>Chemokine (C-X-C) motif ligand 12</td>
</tr>
<tr>
<td>11726624_a_at</td>
<td></td>
<td>11756105_a_at</td>
<td></td>
<td>11718141_at</td>
<td></td>
<td>11718013_at</td>
<td></td>
<td>11753257_a_at</td>
<td></td>
</tr>
<tr>
<td>One of the eight mammalian cyclins, negatively regulated by ER</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Encoded protein has been shown to promote breast cancer microcalcification. Reported to have tumour-suppressive function in mammary epithelia and microenvironment, and its disruption can lead to increased metastases in mammary carcinoma</td>
<td></td>
<td>Protein encoded by this gene is a member of the dickkopf family and is a regulator of the Wnt signalling pathway. Reported to be expressed in tumours from women with a family history of breast cancers</td>
<td></td>
<td>Expressed in breast carcinoma cells and acts as a positive regulator of metastasis in breast cancers</td>
<td></td>
<td>Encoded chemokine is implicated in the metastasis of breast cancers</td>
<td></td>
<td>Encoded cytokine plays pivotal roles in breast tumour progression and its expression in breast cancer tissues is associated with poor clinical outcome</td>
<td></td>
</tr>
<tr>
<td>0.70897</td>
<td>0.80872</td>
<td>-1.057509</td>
<td>-1.185881</td>
<td>-0.97075</td>
<td>-2.607999</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.024979</td>
<td>0.001695</td>
<td>0.004766</td>
<td>0.028756</td>
<td>0.0110567</td>
<td>0.000461</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.9</td>
<td>1.9</td>
<td>0.6</td>
<td>0.3</td>
<td>0.4</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2142</td>
<td>0.0064</td>
<td>0.033</td>
<td>0.0001</td>
<td>0.0003</td>
<td>0.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Up</td>
<td>Up</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Selected genes that were up- or downregulated in response to RAD21 depletion in oestradiol-stimulated MCF7 cells, when compared to oestradiol-stimulated control MCF7 cells. The table represents genes that had evidence of altered expression after adjustment for multiple testing (Benjamini-Hochberg method) as identified by cDNA microarray, followed by experimental validation by qRT-PCR. The ratio of differential expression is presented on a log2 scale for microarray and linear scale for qRT-PCR. Corrected statistical significance (adjusted p-values) and a description of function of the encoded proteins (most relevant to their role in the aetiology of breast cancer) have been included for each gene.

4.3.2 Experimentally validated candidate genes co-bind ERα and RAD21

Although they do not recognise a specific DNA sequence element, cohesin and cohesin loaders were found to mark the promoters of transcriptionally active genes in *Drosophila* (Misulovin et al, 2008), mouse embryonic stem cells (Kagey et al, 2010), and also in the zebrafish (Jenny Rhodes and Michael Meier, Horsfield lab, unpublished data). In MCF7 cells, cohesin binds in isolation and also with ERα at multiple sites, and cohesin binding is highly enriched at ERα bound locations that harbour anchors for inter-chromosomal loops (Schmidt et al, 2010). Many of these binding sites were found to mark the promoters and enhancers of transcriptionally active genes and this preferential binding at active genes is speculated to have a direct influence on their transcription (Dorsett & Merkenschlager, 2013). Hence, to differentiate between direct versus indirect cohesin target genes, and as per the rationale that cohesin-bound genes would undergo the most rapid changes in expression upon cohesin depletion, the relative locations of RAD21 and ERα binding were mapped for the candidate genes. UCSC genome browser [http://genome.ucsc.edu/cgi-bin/hgGateway](http://genome.ucsc.edu/cgi-bin/hgGateway) (2006 NCBI36/hg18 version) and existing global binding data from (Schmidt et al, 2010) was used for this analysis (Table 4.2). All 11 candidate genes had binding sites for ERα, RAD21 and oestrogen-induced RAD21, with 9/11 genes having oestrogen-induced RAD21 binding within 15 kb of their TSS. At *SOX4, PIM2, THBS1, IL20, CXCL12*, and *DKK1*, ERα and RAD21 shared co-localised binding (<1 kb) and were confined within 7 kb of the TSS. In addition, binding sites for ERα and RAD21 were found at a distance from other genes, such as *IRS2*, up to 17.5 kb and *CCNG2*, up to 62 kb from the TSS. Evidence of coincident binding is consistent with the idea that cohesin and ERα can cooperate to regulate transcription of these genes. Cohesin binding has been shown to increase in response to oestrogen addition in MCF7 cells (Schmidt et al, 2010), however, whether cohesin and ERα directly influence each other’s binding was still unclear and warranted further investigation.
Figure 4.3

Schematic representation of genomic binding sites for RAD21, oestradiol-stimulated RAD21 and ERα, previously identified by ChIP-seq (Schmidt et al, 2010) for SOX4, IL20, BMPR2 and DKK1

Binding sites for cohesin subunit RAD21 and ERα were mapped on the UCSC genome browser using the 2006 NCBI36/hg18 version. All the eleven candidate genes demonstrated one or more binding sites for ERα and cohesin (RAD21) (Table 4.2). Schematics of genomic binding sites for four of the eleven candidate genes are shown here.
Table 4.2 Proximity of RAD21 and ERα binding sites to genes that are dysregulated in oestrogen-stimulated RAD21-depleted MCF7 cells

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Ref Seq summary ID</th>
<th>Chromosomal Location (including UTR's)</th>
<th>ERα binding site(s)</th>
<th>Nearest RAD21 binding site(s)</th>
<th>Nearest RAD21 binding site(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Oestrogen Present Absent</td>
<td>Oestrogen Present Absent</td>
</tr>
<tr>
<td>TFF2</td>
<td>NM_005423</td>
<td>Chr 21: 42,639,536 - 42,644,176</td>
<td>-15.31 kb</td>
<td>-15.310 kb</td>
<td>+13.135 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+124 bp</td>
<td>+13.128 kb</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+12.92 kb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOX4</td>
<td>NM_003107</td>
<td>Chr 6: 21,701,951 - 21,706,828</td>
<td>-6.023 kb</td>
<td>-6.444 kb</td>
<td>-4.558 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-4.958 kb</td>
<td>-5.647 kb</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+2.919 kb</td>
<td>-5.012 kb</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-4.482 kb</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+420 bp</td>
<td></td>
</tr>
<tr>
<td>PIM2</td>
<td>NM_006875</td>
<td>Chr X: 48,655,403 - 48,661,228</td>
<td>+6.945 kb</td>
<td>+1.824 kb</td>
<td>+1.836 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+124 bp</td>
<td>+13.128 kb</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+12.92 kb</td>
<td>+420 bp</td>
<td></td>
</tr>
<tr>
<td>THBS1</td>
<td>NM_003246</td>
<td>Chr 15: 37,660,572 - 37,676,960</td>
<td>-1.466 kb</td>
<td>-903 bp</td>
<td>-1.203 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+629 bp</td>
<td>+653 bp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+2.919 kb</td>
<td>+540 bp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+436 bp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAG1</td>
<td>NM_004323</td>
<td>Chr 9: 33,242,470 - 33,254,648</td>
<td>+436 bp</td>
<td>-22.419 kb</td>
<td>-22.456 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+124 bp</td>
<td>+13.128 kb</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+12.92 kb</td>
<td>+420 bp</td>
<td></td>
</tr>
<tr>
<td>IL20</td>
<td>NM_018724</td>
<td>Chr 1: 205,105,777 - 205,109,191</td>
<td>-744 bp</td>
<td>-810 bp</td>
<td>-831 bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+124 bp</td>
<td>+13.128 kb</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+12.92 kb</td>
<td>+420 bp</td>
<td></td>
</tr>
<tr>
<td>CXCL12</td>
<td>NM_000609</td>
<td>Chr 10: 44,185,611 - 44,200,548</td>
<td>-1.461 kb</td>
<td>-1.457 kb</td>
<td>571 bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+124 bp</td>
<td>+13.128 kb</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+12.92 kb</td>
<td>+420 bp</td>
<td></td>
</tr>
<tr>
<td>IRS2</td>
<td>NM_003749</td>
<td>Chr 13: 109,204,185 - 109,236,915</td>
<td>-17.305 kb</td>
<td>-4.810 kb</td>
<td>-4.827 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+1.913 kb</td>
<td>+420 bp</td>
<td></td>
</tr>
<tr>
<td>DKK1</td>
<td>NM_012242</td>
<td>Chr 10: 53,744,047 - 53,747,423</td>
<td>-4.480 kb</td>
<td>-6 kb</td>
<td>-6.033 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-987 bp</td>
<td>-101 bp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+103 bp</td>
<td></td>
</tr>
<tr>
<td>BMPR2</td>
<td>NM_001204</td>
<td>Chr 2: 202,949,295 - 203,140,719</td>
<td>-14.198 kb</td>
<td>-1.354 kb</td>
<td>-468 bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+124 bp</td>
<td>+13.128 kb</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+12.92 kb</td>
<td>+420 bp</td>
<td></td>
</tr>
<tr>
<td>CCNG2</td>
<td>NM_004354</td>
<td>Chr 4: 78,297,381 - 78,310,237</td>
<td>-21.677 kb</td>
<td>-36.252 kb</td>
<td>-61.791 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+124 bp</td>
<td>+13.128 kb</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+12.92 kb</td>
<td>+420 bp</td>
<td></td>
</tr>
</tbody>
</table>

Sites of RAD21 and ERα binding in closest proximity to dysregulated genes in oestrogen-stimulated RAD21-depleted MCF7 cells. The UCSC genome browser (2006 NCBI36/hg18 assembly) and publicly available
genome-wide binding data (ChIP-seq) (Schmidt et al, 2010) was used to map one or more nearest binding sites for RAD21 and ERα, for the qRT-PCR validated candidate genes. Distance from the start site of a gene is used as a reference point for binding locations (up- or downstream of the gene).

4.3.3 **SOX4** and **IL20** as candidate cohesin-regulated genes

After experimentally validating candidate genes from the microarray, I next sought to understand how cohesin interacts with genes to regulate expression in opposite directions. For this, I decided to select two genes, one gene that is activated, and another that is repressed following cohesin depletion and stimulation by oestradiol, which can then represent models to understand cohesin’s control of oestrogenic gene activity.

Oestrogen activation of the **SOX4** gene (normally relatively oestrogen insensitive) was dramatically enhanced by cohesin knockdown, indicating that loss of cohesin sensitised **SOX4** to oestrogenic response (Figure 4.1A). In contrast, expression of **IL20** was robustly reduced upon cohesin depletion (Figure 4.1B). Both these genes are implicated in breast cancer, where in a cohort of 959 breast invasive carcinoma tumours (TCGA Provisional dataset), alterations in **SOX4** gene was found in 6% and that of **IL20** was encountered in 16% of all the cases investigated (Figure 4.4) (Cerami et al, 2012; Gao et al, 2013).

**SOX4** (Sex determining region Y) is a member of the **SOX** family of developmental transcription factors and its overexpression in breast cancers is reported to orchestrate epithelial-mesenchymal transition (EMT) contributing to cancer progression and metastasis. Abnormal overexpression of **SOX4** is found predominantly in aggressive triple-negative breast cancers (TNBCs) but around 19.7% of ER-positive samples also showed higher expression of **SOX4** indicative of a fundamental role for the gene in this malignancy (Zhang et al, 2012). On the other hand, **IL20** (Interlukin 20) is a pro-inflammatory cytokine and its overexpression in invasive breast ductal carcinomas positively correlated to tumour progression and poor clinical outcome (Hsu et al, 2012). In MCF7 cells, following depletion of RAD21, the expression of **IL20** was significantly reduced, which is otherwise robustly activated by oestrogen. **SOX4** and **IL20** were thus chosen as candidate genes for further investigation of cohesin’s bidirectional control of oestrogen-mediated gene responses.
Figure 4.4 Oncoprints of SOX4 and IL20

Data were generated using breast invasive carcinoma data from cBioPortal (TCGA Provisional dataset) (Cerami et al, 2012; Gao et al, 2013). Alterations in SOX4 and IL20 were annotated using the largest available dataset of breast invasive carcinoma comprising 959 tumour samples. Alterations in the SOX4 gene was identified in 6% of all the tumours whereas IL20 was misregulated in a higher proportion, 16% of the tumour samples.

4.3.4 Recruitment of ERα and RAD21 to candidate genes

As described in section 4.3.2, the comprehensive mapping of genome-wide ChIP-seq data to gene loci of interest identified genomic locations of ERα and RAD21 binding (Figure 4.3, Table 4.2). To determine whether cohesin influences ERα binding at ER target genes, I sought to investigate if previously identified ERα binding at SOX4 and IL20 are affected in response to loss of cohesin. For this purpose, a ChIP approach was taken to assess ERα’s interactions with the two candidate genes, in the presence and absence of cohesin.

4.3.4.1 Primers designed to determine RAD21 and ERα binding at SOX4 and IL20

Two candidate genes were chosen to investigate for cohesin (RAD21) and ERα binding, and to also delineate the consequences on binding of ERα, following loss of cohesin. For this, prior knowledge of genome-wide binding sites for ‘RAD21’, ‘oestradiol-induced RAD21’ and ‘ERα’ identified in MCF7 cells (Schmidt et al, 2010) coupled with mapping of these sites to these candidate genes, served as a guide for designing site-specific primers for SOX4 and IL20.
A total of 8 binding sites for RAD21 and ERα were identified for SOX4 located within 6.5 kb upstream of the gene start site (GSS), and one ERα binding site was located within the 3’ UTR of SOX4 (Table 4.2). Site-specific primer pairs (4 pairs) were designed spanning up to 6.5 kb upstream to 4.1 kb downstream of the GSS of SOX4, to amplify ChIP-enriched DNA to determine effects on ERα binding following loss of cohesin (Figure 4.5B). Two of these primer pairs were used to amplify regions common to previously identified ‘oestradiol-induced RAD21’ and ‘ERα’ binding (primer pairs 3 and 4) and the other two primer pairs (1 and 2) were used to amplify individual binding sites of ‘ERα’ and ‘oestradiol-induced RAD21’, respectively. For IL20, three binding sites (one for oestradiol-induced RAD21, RAD21 and ERα each) were identified within 1 kb upstream of GSS (Table 4.2) and a single primer pair spanning the region common to all three binding sites was designed to amplify the genomic region (Figure 4.6B). So, overall 4 primer pairs were designed for SOX4 and one primer pair was designed for IL20 to amplify regions of DNA previously identified to bind RAD21 and ERα by qPCR using ChIP-ed DNA as a template. ERα binding at these previously identified sites was used as a reference point, to investigate if loss of cohesin binding has had an effect on the level of ERα binding. RAD21 binding sites were used to confirm the actual enrichment of RAD21 binding at these sites.

4.3.4.2 Characterisation of primer specificity and efficiency

To accurately determine the enrichment of target proteins (ERα, RAD21 and RNA Pol II) by ChIP followed by qPCR, it was critical to first validate primer specificity and efficiency, to ensure reproducibility and precision of my ChIP measurements. These have been discussed in section 2.5.3.

4.3.5 Cohesion (RAD21) binding restricts oestrogen activation of SOX4

As mentioned previously in section 4.3.3, SOX4 was relatively insensitive to oestrogen in control MCF7 cells. However, upon cohesin depletion, expression of SOX4 was increased significantly (Figure 4.1A). The increase in transcript levels of SOX4 was apparent as early as post 3 hours of oestrogen stimulation and the expression change was robustly persistent until 24 hours (Figure 4.5A). This suggests that wild-type levels of cohesin impede oestrogen activation of SOX4 in MCF7 cells. Since cohesin prevents oestrogen activation of SOX4, it is possible that RAD21 binding at nearby regulatory chromosomal sites prevent ERα binding to hinder oestradiol activation of SOX4.
Quantitative ChIP was performed using an anti-ERα antibody to determine if cohesin modulates ERα binding at the SOX4 gene. ChIP primers were designed to detect ERα binding at sites previously identified by ChIP-seq, relative to positions 1, 3 and 4 as indicated in Figure 4.5B. The primers located at position 3 and 4 were enriched for DNaseI hypersensitive clusters (indicative of accessible chromatin) and enhancer- as well as promoter-associated histone marks (H3K4me1 and H3K4me3) as determined by the ENCODE project (Ernst et al, 2011). In addition to sites 1, 3 and 4, the ChIP amplicon spanning position 2 was designed as this site immediately upstream of the TSS recruits RAD21 upon oestradiol stimulation, and is enriched for the promoter-associated histone mark (H3K4me3) (Ernst et al, 2011).

All ChIP analyses were carried out as described in section 2.5.3.2. A period of 45 minutes of oestrogen treatment was chosen because many groups (Carroll et al, 2005; Carroll et al, 2006; Shang et al, 2000) have reported maximal recruitment of ERα to chromatin (at ER target genes) at the end of this brief treatment period. Chromatin was immunoprecipitated and prepared as described in section 2.5.2 and four biological replicates were generated. Considering SOX4 is not an oestrogen sensitive gene in normal conditions, following stimulation by oestrogen, there was only a minor increase (statistically non-significant) in ERα binding at all three previously identified ERα binding sites (primers 1, 3 and 4) in control-stimulated cells. However, it was interesting to note that in RAD21-depleted cells, binding of ERα was significantly enhanced at sites upstream of its TSS (primers 3 and 4, adjusted p-value <0.05). Although statistically insignificant, an increase in ERα binding was also observed at the site downstream of its TSS (primer 1). The increase in ERα binding at these locations is consistent with oestrogen-induced increase in transcript levels observed for SOX4, in RAD21-depleted cells. The enrichment of ERα binding at these distal sites in RAD21-depleted cells shows that cohesin might normally restrict ERα-mediated transcription of the SOX4 gene.

I next sought to determine if RAD21 depletion affects RNA Pol II recruitment at SOX4. As RNA Pol II transcribes all protein-coding genes (Egloff & Murphy, 2008), this could help in understanding the reason behind enhanced transcription of SOX4 upon RAD21 depletion. An antibody detecting phosphorylated serine-5 of RNA Pol II (RNAPIIser5-P) was used for this purpose. The carboxy terminal domain (CTD) of the largest subunit of RNA Pol II is composed of unique heptapeptide repeats, of which serine5 is a component (Heidemann et
Extensive posttranslational modifications (including phosphorylation) of these repeats orchestrate appropriate temporal transcription of corresponding genes. RNAPIIser5-P is part of the early transcription initiation complex and accumulates at the TSS of genes prior to activation of transcription (Egloff & Murphy, 2008; Heidemann et al, 2013). The antibody raised against serine-5-phosphorylated RNA Pol II detects the early initiation as well as the paused form of RNA Pol II. Following oestrogen addition, I observed a slight decrease in RNAPIIser5-P binding at site 4 (distal regulatory site) in RAD21-depleted cells compared to control. In contrast, RNAPIIser5-P binding at sites 3 and 2 (immediately upstream of the TSS, promoter-proximal region) remained unaltered in RAD21-depleted cells in comparison to control-stimulated cells. In oestrogen-stimulated RAD21-depleted cells, there was no observable enrichment of RNAPIIser5-P binding, despite significant enhancement in expression of this gene, which could perhaps reflect the transition of paused polymerase to productive elongation. However, further observation of decreased enrichment of RNAPIIser5-P at site 1 (in the gene body) potentially signifies that RNA Pol II is paused at promoter-proximal regions (perhaps at additional sites not examined in this study).

Upon oestrogen stimulation of control transfected cells, increased enrichment of RAD21 binding was observed at sites 2 and 4, and also at site 3, albeit to a moderate extent. RAD21 binding was particularly enriched at site 2 (adjusted p-value <0.05) immediately upstream of the TSS. On the contrary, enrichment of RAD21 binding at these two sites, along with the third site (primer 3) was markedly reduced in RAD21-depleted cells. Notably, sites 3 and 4 that lose RAD21 binding demonstrated gain in ERα binding in cells depleted of RAD21. It is possible that RAD21 binding at site 2 establishes a repressive chromatin loop with ERα binding sites at 3 and/or 4 and thus contributes to restriction of ERα-mediated transcriptional activation of SOX4 by preventing communication between ERα-bound distal enhancers and the proximal promoter. However, to prove this, effect of RAD21 depletion on chromatin interactions at the SOX4 gene loci would require additional examination using the ‘chromosome conformation capture (3C)’ technique. Taken together, these results suggest that cohesin impedes oestrogen activation of SOX4 possibly by inhibiting enrichment of ERα binding.
Figure 4.5

Oestrogen-deprived MCF7 cells were transfected with control or RAD21 siRNA for 48 hours, then treated with vehicle or 100 nM oestradiol for 3, 6 and 24 hours for qRT-PCR and 45 minutes for ChIP. (A) RAD21 knockdown in the presence of oestradiol enhanced oestrogen activation of SOX4 transcripts. Transcript levels
were determined using qRT-PCR relative to the mean of two reference genes, *PPIA* and *GAPDH*. Error bars are SEM of three biological replicates. Significance was determined using two-way ANOVA. (B) Schematic of *SOX4* gene showing RAD21, oestradiol-induced RAD21 and ERα binding, mapped on the UCSC genome browser (2006 NCBI36/hg18 version) using previously published ChIP-seq data in MCF7 cells. Numbers represent positions of the ChIP primers used to amplify various sites of the *SOX4* gene locus (C) ChIP analyses of ERα, RNAPIIser5-P and RAD21 binding at the specified locations for the *SOX4* gene. RAD21 depletion significantly enhanced ERα binding at locations (primer 3 and 4) upstream of *SOX4*. RAD21 depletion decreased RAD21 binding at both these sites and also at an additional site (primer 2). Binding is shown relative to input chromatin. Error bars are SEM of four independent biological replicates. Significance was determined using two-way ANOVA.

### 4.3.6 Cohesin (RAD21) binding at the promoter facilitates oestrogen activation of *IL20*

The pro-inflammatory cytokine *IL20* was identified as one of the most positively regulated genes by oestrogen activation of ERα in a different ER-positive cell line T47D (Williams et al, 2008) and its promoter was shown to contain an ERE motif binding ERα in MCF7 cells (Laganiere et al, 2005). Thus, it did not come as a surprise when my microarray data also demonstrated *IL20* to be massively oestrogen-induced (vehicle-treated control compared to oestradiol-stimulated control: log2FC: 2.04127, 3 hour time point; log2FC: 2.2305, 6 hour time point). Oestrogen induction for *IL20* persisted significantly for up to 24 hours in control-stimulated cells (Figure 4.6A, adjusted p-value <0.0001). Interestingly, RAD21-depleted cells failed to upregulate *IL20* in response to oestrogen. Therefore, there must be dependence on cohesin for oestrogen activation of the *IL20* gene. In addition to detecting binding of ERα to the promoter of *IL20*, another genome-wide ChIP analysis in MCF7 cells (Schmidt et al, 2010) identified a coinciding RAD21 binding site at the same genomic location. Since, *IL20* recruits both proteins, and RAD21 depletion blocked oestrogen induction of *IL20*, it seemed likely that RAD21 binding at the promoter could influence recruitment of ERα to this gene.

I examined the binding profiles for RAD21, ERα and RNAPIIser5-P at the promoter of *IL20* by quantitative ChIP (as described in section 2.5.3.2) using antibodies specific to each, in the presence and absence of RAD21 siRNA. Following oestrogen addition, there was a statistically significant increase in ERα binding (Figure 4.6C, adjusted p-value <0.05) in control cells stimulated by oestradiol. Compared with controls, there was marked reduction in binding levels of ERα in oestradiol-stimulated cells depleted of the RAD21 subunit of cohesin. Reduction in ERα binding was not statistically significant but showed a definite
trend of decrease in cells lacking RAD21. Importantly, I observed a very strong oestradiol-induced enrichment of RNAPIIser5-P binding at the promoter, in control-stimulated cells and this enrichment was significantly reduced in RAD21-depleted cells implying that RNA Pol II is no longer recruited to IL20 upon depletion of RAD21. It was interesting to note that the observed binding profile of these proteins was consistent with blockade of IL20 expression upon RAD21 depletion.

No apparent change in the binding levels of the RAD21 protein was observed in between vehicle versus oestradiol-treated states of the control MCF7 cells, implying RAD21 is pre-bound to this location in the absence of oestrogen. Upon siRNA depletion of RAD21, followed by oestrogen-treatment, the enrichment of RAD21 binding was remarkably decreased (Figure 4.6C, adjusted p-value <0.05), matching the corresponding expression pattern of IL20. Together these results suggest that cohesin positively regulates transcription of IL20 and facilitates oestrogen responsiveness of the gene by allowing enrichment of ERα and recruitment of RNAPIIser5-P binding.
Figure 4.6 RAD21 depletion repressed oestradiol responsiveness of *IL20*

Oestrogen-deprived MCF7 cells were transfected with control or RAD21 siRNA for 48 hours, then treated with vehicle or 100 nM oestradiol for 3, 6 and 24 hours for qRT-PCR and 45 minutes for ChIP. (A) RAD21 knockdown in the presence of oestradiol reduced transcript levels of *IL20*. Transcript levels were determined using qRT-PCR relative to the mean of two reference genes, *PPIA* and *GAPDH*. Error bars are SEM of three biological replicates. Significance was determined using two-way ANOVA. (B) Schematic of *IL20* gene showing RAD21, oestradiol-induced RAD21 and ERα binding, mapped on the UCSC genome browser (2006 NCBI36/hg18 assembly) using previously published ChIP-seq data in MCF7 cells. The number represents position of the ChIP primer used to amplify the genomic site. (C) ChIP analyses of ERα, RNAPIIser5-P, and RAD21 binding at the promoter of *IL20*. RAD21 knockdown decreased ERα and RNA Pol II binding at the *IL20* promoter. RNAPIIser5-P and RAD21 binding were significantly reduced with depletion of RAD21. Binding is shown relative to input chromatin. Error bars are SEM of four independent biological replicates. Significance was determined using two-way ANOVA.
4.4 Discussion

Results described in Chapters 3 and 4 have shown that cohesin regulates expression of a wide variety of genes in oestradiol-stimulated MCF7 cells. Transcript levels of select candidate genes from the microarray were experimentally verified using qRT-PCR, and reflected transcriptional responses identified in the microarray. These results suggest that cohesin influences transcriptional responsiveness to oestrogen for a distinct subset of ER target genes.

The genes chosen for experimental validation were found to recruit both RAD21 and ERα, consistent with the idea that genes with binding sites for cohesin have a greater propensity to be directly regulated by cohesin. From the microarray data, it emerged that cohesin exhibits bidirectional regulation of oestrogen-responsive genes but how these transcriptional programs might be operating was enigmatic. I interrogated this transcriptional regulatory role of cohesin in greater detail for two genes, \textit{SOX4} and \textit{IL20} that are regulated in opposite direction by cohesin depletion. ChIP analyses were carried out for RAD21, ERα and RNAPII\text{ser5-P} binding in the presence and absence of RAD21 siRNA. RAD21 depletion enriched ERα binding to the regulatory sites of \textit{SOX4} and greatly enhanced oestrogen-mediated expression of the gene. On the contrary, in the absence of RAD21, ERα and RNAPII\text{ser5-P} binding to the promoter of \textit{IL20} was weakened and corresponded with downregulation of the gene. Together these results indicate that RAD21 depletion results in opposite effects at different gene loci via modulation of other regulatory proteins that govern transcription of the genes. These observations suggest that cohesin has context-specific effects on the expression of individual genes and serve to better understand cohesin’s dual role as a facilitator or obstructer of transcriptional response to oestrogen.

In an attempt to summarise these findings, a hypothetical model is presented in Figure 4.7 that might explain the unique differential responses observed for the two genes. With regards to cohesin’s role in organisation of functional chromatin domains together with the insulator factor CTCF (Dowen et al, 2014; Le Dily et al, 2014; Zuin et al, 2014), it is possible that oestrogen-responsive genes reside in discrete regulatory domains. Upon cohesin depletion, it seems likely that integrity of such domains is disturbed, such that some cohesin-associated local chromatin interactions are disrupted while some others are gained in a gene- and context-dependent manner, affecting the transcriptome as a consequence. For example, it may be possible that \textit{SOX4} gained sensitivity to oestrogen following cohesin depletion.
potentially by the elimination of transcription-hindering loops previously held by cohesin, as a result of a disrupted chromatin boundary. This in turn, may have triggered enrichment of ERα binding at the distal enhancers and favoured interactions between ERα-bound enhancers and promoter of SOX4 leading to upregulation of this gene in response to oestrogen. In contrast, IL20 lost its sensitivity to oestrogen following loss of cohesin by possible weakening of ERα binding together with marked reduction in transcription initiating polymerase (RNAPIIser5-P) at the promoter. It may also be possible that cohesin aids in the recruitment of ERα at the IL20 gene, since cohesin is pre-bound to IL20 (upstream of TSS) before MCF7 cells are stimulated with oestradiol. Cohesin depletion might have also triggered modifications of the epigenetic signature of IL20, for example by alteration of histone modification patterns in a manner that impedes expression of the gene. Furthermore, it also remains possible that cohesin depletion resulted in the overall loss of transcription-favouring enhancer-promoter loop(s) at IL20, as oestradiol-induced RAD21 binds at additional sites upstream of the IL20 gene (Schmidt et al, 2010). Such sites were marked with the enhancer-associated histone mark H3K4me1 (Ernst et al, 2011). In summary, my data suggests that cohesin promotes transcription of certain genes while inhibits transcription of some others by modulating the integrity of local chromatin interactions.

However, to substantiate the observed findings at both these gene loci, future studies could attempt to examine a broader range of chromatin marks, on a background of wild-type/depleted-cohesin (RAD21). Investigation of such chromatin marks may include the enhancer-associated histone marks H3K4me1/2 at the putative enhancers of SOX4, along with the presence of H3K27ac, which could reveal the active state of the enhancers. To determine if a chromatin boundary existed upstream of the SOX4 gene, the repressive H3K9me3 and the active histone mark H3K9ac could be profiled across the SOX4 locus. The promoter of IL20 could be similarly investigated for the combined presence of the promoter-associated histone marks H3K4me2/3 and H3K27ac. The histone modifications associated with transcriptional repression, such as H3K27me3, H3K4me3 (both representative of polycomb silencing pathway) and H3K9me3 could be additionally investigated in the presence/absence of cohesin. Furthermore, elongating RNA Pol II (RNA Pol II phosphorylated at serine 2), which represents the productive elongation phase of transcription, could be examined at the gene bodies of both SOX4 and IL20. This would help to correlate better with the enrichment patterns observed for RNAPIIser5-P, and then to be
able to explicitly conclude whether or not promoter proximal pausing was released and productive elongation initiated.

**Figure 4.7**

**Figure 4.7 Hypothetical model of how cohesin (RAD21) might influence oestrogenic response of SOX4 and IL20**

The chosen candidate genes SOX4 and IL20 are presented (shown linked here only for convenience) as part of a hypothetical model to illustrate the potential mechanisms by which RAD21 regulates expression of specific ER target genes. Oestradiol stimulation of cells (plus ER representation) causes binding of ERα to its target sites (putative enhancers of SOX4 and promoter of IL20) and promotes recruitment of cohesin to additional sites (upstream of the TSS of SOX4 and IL20, and also at the promoter of IL20). ERα binding has no effect on the transcription of SOX4 (which is normally oestrogen insensitive). However, overlapping binding of ERα and RAD21 at the promoter of IL20, together with RAD21 binding at distal sites upstream of IL20, accelerates gene transcription by potential stabilisation of ERα binding by cohesin and also by the formation of potential transcription-favouring enhancer-promoter loop(s) at IL20. Upon RAD21 depletion, local chromatin organisation is altered, wherein some interactions are weakened (IL20) while new interactions are gained due to increased flexibility of local chromatin (SOX4). RAD21 depletion weakens binding of ERα to IL20’s promoter (represented by a smaller icon of ER) and also potentially results in the loss of the transcription-favouring enhancer-promoter loop(s) at IL20. However, ERα binding is strengthened to SOX4’s enhancers thereby allowing distal enhancers to interact with the proximal promoter (represented by a bigger icon of ER) due to the possible loss of transcription-hindering/repressive loop(s) previously held by cohesin. Thus allowing ectopic gain of oestrogen sensitivity for SOX4. Figure is partially adapted from Antony et al (2014).
CHAPTER 5
Targeting cohesin with an inhibitor that disrupts the acetylation cycle of cohesin subunit SMC3

5.1 Introduction
After having identified cohesin-dependent genes in the oestrogen response pathway, the next goal was to assess efficacy of the cohesin inhibitor (PCI-34051), and determine if this inhibitor can be used to target the transcriptional role of cohesin. To understand how this inhibitor works, it is imperative to revisit the cell division function of cohesin.

Cohesin contributes to maintaining integrity of the genome by ensuring faithful segregation of chromosomes, which is meticulously mediated by a group of accessory proteins/factors that coordinate the accurate loading and unloading of cohesin from chromatin, throughout the entire cell division process (Nasmyth & Haering, 2009). In higher eukaryotes including humans, cohesin is loaded onto chromosomes by the cohesin loading complex NIPBL-MAU2 in telophase immediately after sister chromatids separate in anaphase of the preceding cell cycle. However, establishment of cohesion between duplicated sister chromatids occur only during S phase, when a pair of lysine residues (K105, 106 in vertebrates) of the SMC3 subunit’s nucleotide-binding domain (NBD) is acetylated by the acetyl transferase ESCO1/2 (Beckouet et al, 2010). This acetylation step is proposed to lock the exit gate for cohesin and prevent the premature release of cohesin from chromatin. Cohesin (~85-90% of mitotic cohesin) from the chromosomal arms gets removed via the prophase pathway. However, the fraction of cohesin (~10-15%) in the pericentromeric region stays protected and remains chromatin-bound until the start of anaphase (Feeney et al, 2010; Waizenegger et al, 2000). At the onset of the metaphase-anaphase transition, separase cleavage of RAD21 cleaves the pericentromeric cohesin and allows eventual disjunction of the sister chromatids to opposite spindle poles (Remeseiro et al, 2013), following which, the class I HDAC8 reverses acetylation of SMC3 (Beckouet et al, 2010; Borges et al, 2010; Deardorff et al, 2012; Xiong et al, 2010). This deacetylation step is crucial for dissolution of the cohesive complex and recycling of ‘used’ cohesin in order for it to be reloaded onto chromatin, for the continuation of the subsequent cell cycles (Beckouet et al, 2010; Remeseiro & Losada, 2013; Xiong et al, 2010). Failure of SMC3 deacetylation is also evidenced to hamper cohesin-mediated gene transcription (Deardorff et al, 2012).
Reassociation of cohesin to chromatin occurs by three distinct modes. One cohesin pool (one-third of nuclear cohesin) binds stably to chromatin during DNA replication in the S phase. However, a second cohesin pool (~30-40% of total nuclear cohesin) binds in a highly dynamic fashion all throughout interphase, which is proposed to be important for functions such as gene transcription and regulation of chromatin structure (Gause et al, 2010; Gerlich et al, 2006). In addition, a third cohesin pool was recently recognised to interact with chromatin in an extremely transient mode with chromatin-interaction time ranging in seconds (Ladurner et al, 2014).

The rationale of using the HDAC8 inhibitor PCI-34051 in MCF7 cells was to block deacetylation of acetylated-SMC3 to force accumulation of ac-SMC3. This was anticipated to alter the stoichiometry and recycling of cohesin interrupting its association to chromatin including its association to the regulatory sites that control expression of oestrogen-responsive genes. I hypothesised that this would disrupt cohesin’s transcriptional regulatory role prior to perturbing its cell division functions for the successive cell cycles.
Figure 5.1

**Cohesin binding to chromatin in non-mitotic cells**

- Telophase: Cohesin associates to chromatin in telophase "Cohesin loading"
  - NIPBL-MAU2
- Anaphase: Sister chromatid separation and dissociation of cohesin from mitotic chromatin "Anaphase pathway of cohesin removal"
  - HDAC8
- Interphase (G1 - S - G2):
  - Dynamic binding of cohesin to chromatin "Dynamically bound cohesin"
  - ESCO1/2
- Metaphase: Centromeric cohesin removal by separate cleavage of RAD21 subunit of cohesin
- Prophase: Cohesin removal from chromosomal arms "Prophase pathway of cohesin removal" "Stably bound cohesin"
- G2 phase: Cohesin bound stably to chromatin

**Cohesin's association to chromatin for sister chromatid cohesion**

- Deacetylation of SMC3 by HDAC8
- Removal of ac. Release of RAD21

---

---
Figure 5.1 Functions mediated by cohesin in mitotic and non-mitotic cells

Refreshed cohesin containing non-acetylated-SMC3 is loaded onto chromatin in telophase in the presence of the NIPBL-MAU2 heterodimer. This association to chromatin is highly dynamic marked with relatively short residence time, 25 minutes on an average (Gerlich et al, 2006). Cohesin’s chromatin association in non-cycling cells is thought to be important for the topological organisation of the genome including transcriptional regulation throughout interphase (Gibcus & Dekker, 2013; Sofueva et al, 2013). To prevent locking the ring prior to the start of DNA replication in S phase, additional factors like WAPL (Wings-apart like) and PDS5 (Precocious dissociation of sisters 5) attach to chromatin-bound cohesin in the G1 phase. Cohesion establishment takes place by acetylating two lysine residues of SMC3, catalysed by the cohesin acetyl transferase(s) ESCO1/ESCO2, and also by binding of an additional factor called Soronin to PDS5. Acetylation of SMC3 is thought to mediate stable binding of cohesin to chromatin with the proportion increasing gradually with the progression of S phase. At the onset of mitosis, cohesin bound to chromosomal arms is removed via the prophase pathway (Buheitel & Stemmann, 2013; Nakajima et al, 2007). The rest of the cohesin bound to mitotic chromosomes is removed by separase cleavage of RAD21, thereby allowing separation of the sister chromatids. Once removed from chromosomes (via the prophase or anaphase pathways), HDAC8 deacetylates SMC3 to help release RAD21 and Soronin (Deardorff et al, 2012), thereby refreshing cohesin used in the previous cell cycle.

5.2 Aim

To block deacetylation of cohesin subunit SMC3 using an HDAC8 inhibitor (HDAC8i) PCI-34051 and investigate if this can recapitulate transcriptional dysregulation observed with RAD21 siRNA-mediated depletion of cohesin.

5.3 Results

5.3.1 HDAC8 inhibitor PCI-34051 blocks deacetylation of SMC3 subunit of cohesin in MCF7 cells

In an extensive screen of all known human deacetylases (HDACs), HDAC8 was identified to be the vertebrate deacetylase for cohesin subunit SMC3 (Deardorff et al, 2012). Interestingly, siRNA-mediated loss of HDAC8 or treatment with the HDAC8-specific inhibitor PCI-34051 resulted in increased levels of ac-SMC3 in synchronised as well as asynchronous cultures of HeLa cells without notable alterations in the progression of cell cycle of these cells (Deardorff et al, 2012). Using these observations as a starting point, I assessed the ability of the inhibitor to block deacetylation of SMC3 in my model system MCF7 cells, to be certain that the inhibitor is functioning as reported by the previous study. For this, asynchronous MCF7 cells were seeded in full media and allowed overnight attachment, following which cells were treated with varying concentrations of the inhibitor (10 µM, 25 µM and 50 µM) for up to 72 hours. The 25 µM dose was sufficient to block
deacetylation of SMC3 in HeLa cells. However, I chose to include two more doses (one lower and one higher than 25µM) in addition to the 25 µM dose, to further characterise the dosage-dependent effects of the inhibitor, if any. Levels of total- and ac-SMC3 were assessed by SDS-PAGE followed by immunoblotting with antibodies specific to each. A mouse monoclonal anti-acetyl-SMC3 antibody that only detects acetylated form of the human SMC3 peptide (sequence corresponding to the amino acid residues 97 to 115 of human SMC3), was kindly provided by Dr. Katsumi Shirahige (Nishiyama et al, 2010). After treatment for the indicated time periods (24, 48 and 72 hours), levels of ac-SMC3 were assessed by normalisation to the levels of total-SMC3, as total-SMC3 levels were found to remain unchanged for the entire course of the progression of cell cycle in HeLa cells (Deardorff et al, 2012). My results indicated that there was no detectable ac-SMC3 by 24 hours of PCI-34051 treatment (Appendix I). However, by 48 hours, there seemed to be a dose-dependent increase in the levels of ac-SMC3 with the highest level of accumulation observed in cells treated with the 50 µM dose (Figure 5.2A). After 72 hours of treatment with PCI-34051, accumulation of ac-SMC3 persisted for the 50 µM dose compared to control, however, the 10 and 25 µM doses were inconclusive (Appendix I). Several attempts were made to replicate results obtained with immunoblotting. However, due to reasons unknown, the anti-acetyl-SMC3 antibody failed to detect any protein bands in further immunoblot attempts. Hence, keeping in mind the limited availability of this antibody, immunofluorescence detection of ac-SMC3 was considered over immunoblotting because the volumes needed for immunofluorescence detection was lower than that needed for immunoblots.

To better capture the cell cycle effects mediated by PCI-34051, MCF7 cells were synchronised to G1-S phase by double thymidine block and released in the presence of different doses of the inhibitor. This time a further higher dose of 100 µM was also included in addition to the other three. Synchronised cells were fixed with 4% paraformaldehyde and immunostained for DNA (Hoechst-H33342), total- and ac-SMC3 and confocal microscopy was used to visualise the level/location of the two versions of the SMC3 protein. Confocal images indicated that as expected, the two proteins were located in the nucleus, with the level of ac-SMC3 becoming robustly apparent with increasing concentration of the inhibitor. The level of ac-SMC3 was low in cells treated with the 10 µM dose and of the highest intensity for cells treated with 100 µM of the inhibitor. Images from the 6 hour treatment period are presented in Figure 5.2B and additional images from the 12 and 24 hour treatment period.
times are included in Appendix J. Taken together, these results confirmed that PCI-34051 works as expected, by resulting in the accumulation of ac-SMC3.

**Figure 5.2**

(A) Asynchronous MCF7 cells grown in full media were incubated with different concentrations of PCI-34051 for 48 hours. Total cell lysates were prepared, and levels of total- and ac-SMC3 were analysed by SDS-PAGE and immunoblotting. The bar plot indicates quantification of protein levels of ac-SMC3 normalised to total-SMC3 (n=1). Immunoblots from the 24 and 72 hours treatment periods are available in Appendix I. (B) MCF7 cells were synchronised to G1-S phase by double thymidine block and released in the presence of different concentrations of the inhibitor for 6 hours. Cells were fixed with 4% paraformaldehyde, permeabilised, and immunostained for DNA (Hoechst-H33342, blue, 1 µg/ml), total-SMC3 (green, 1:200 dilution) and ac-SMC3
Cells were then analysed by Nikon C2 confocal microscope; images were taken at 20x magnification (z-series) using the NIS elements software, z-stack and processed using FIJI. The data indicate that deacetylation of SMC3 subunit of cohesin is hindered in response to the inhibitor and has led to the over-accumulation of ac-SMC3 in a dose-dependent manner, at the time points tested. Shown are representative immunofluorescence images from the 6 hour time point. Images at later time points (12 and 24 hours post-release) are available in Appendix J.

### 5.3.2 PCI-34051 treatment did not affect expression of oestrogen-responsive genes

After confirming that the inhibitor blocked deacetylation of SMC3 in both synchronous and asynchronous cultures of MCF7 cells grown in full media, my next goal was to assess whether the inhibitor blocked SMC3 deacetylation in MCF7 cells grown in the same oestrogen-starved conditions used previously for the microarray and ChIP experiments in Chapters 3 and 4. The rationale was to compare the transcriptional signature of cohesin-dependent oestrogen-responsive genes that occurred upon RAD21 siRNA-mediated depletion of cohesin, with any transcriptional changes observed with blocking SMC3 deacetylation.

MCF7 cells were grown in oestrogen-starved conditions according to the same protocol described in section 3.4. For gene expression analysis, I sought to use the lower doses of the inhibitor, 10 and 25 µM only, as the higher doses could have been potentially cytotoxic for the period of study and hence could have off-target effects on transcription. Oestrogen-starved MCF7 cells were grown for a day, followed by treatment with the inhibitor for 48 hours in serum-starved conditions to elicit accumulation of ac-SMC3 (Figure 5.3B). Cells were then stimulated with 100 nM of oestradiol for 3, 6 and 24 hours prior to harvesting cells for RNA. A schematic representation of the experiment design is presented in Figure 5.3A.

Immunofluorescence results indicated that cells treated with the inhibitor had accumulated ac-SMC3 in comparison to vehicle-treated control cells prior to oestradiol stimulation (Figure 5.3B). It was important to confirm this, as when growing cells in serum-starved (steroid-hormone free) conditions, there is always an increased risk of slowing growth of the cells to an extent of making them quiescent or resulting in their complete growth arrest (Hamelers et al, 2003). More so in this case as the inhibitor itself was potentially capable of arresting the cell cycle.
Next, serum-starved MCF7 cells treated with PCI-34051 for 48 hours were stimulated with 100 nM of oestrogen and processed for transcriptional analyses. Three independent biological replicates were generated, and accumulation of ac-SMC3 was confirmed by immunofluorescence analyses. As described in Chapter 4 (Figures 4.1 and 4.2), a subset of oestrogen-responsive genes were dysregulated upon siRNA depletion of RAD21. Following PCI-34051 treatment, the transcript levels of ten genes that had altered transcription upon RAD21 depletion were analysed (Figure 5.4). For ease of comparison, the transcriptional profiles from RAD21 knockdown experiments presented earlier (Figures 4.2 and 4.3) have been included together with expression profiles from PCI-34051 treatment in Figure 5.4. Transcriptional profiles of genes using the 25 µM dose of PCI-34051 is presented in Figure 5.4, whereas that of the 10 µM dose is included in Appendix K. With both the doses, the overall patterns of gene expression changes were relatively similar.

Unexpectedly, most oestrogen-responsive genes that were downregulated by RAD21 siRNA such as MYC, IL20, THBS1, IRS2, CXCL12, TFF2, PIM2 retained sensitivity to oestradiol in the presence of PCI-34051. On the other hand, expression of other genes such as SOX4, CCNG2, BMPR2 were unaffected (results were statistically insignificant) by PCI-34051 at 3 and 6 hours post-stimulation by oestradiol (Figure 5.4). Notably, baseline expression of CCNG2, an inhibitor of cell cycle progression (Ahmed et al, 2012) was significantly upregulated in response to treatment with PCI-34051, perhaps reflecting the initiation of cell cycle blockade by the inhibitor. However, leaving aside CCNG2, basal expression of most genes remained unaffected by PCI-34051, reflecting that block of SMC3 deacetylation did not affect gene expression in the absence of oestrogen as well. Thus, both basal and oestrogen-responsive expression of the selected cohesin-dependent genes remained unaffected upon treatment with PCI-34051.

Since RAD21 siRNA treatment and blockade of SMC3 deacetylation would be predicted to hinder cohesin function similarly, the finding that these two strategies for cohesin disruption had different consequences for gene expression was surprising.
Figure 5.3 Accumulation of ac-SMC3 in serum-starved MCF7 cells in response to treatment with PCI-34051

(A) Schematic of experiment design used for immunofluorescence and transcriptional analyses of PCI-34051-treated MCF7 cells grown in steroid hormone depleted medium. (B) Representative confocal immunofluorescence images from immunostainings of acetylated- and total-SMC3 in PCI-34051-treated MCF7 cells. MCF7 cells were starved of oestrogen for 24 hours, followed by treatment with 10 or 25 µM of PCI-34051 for 48 hours. Cells were then fixed with 4% paraformaldehyde, and permeabilised, followed by immunostaining for DNA (Hoechst H33342, 1 µg/ml, blue), total-SMC3 (1:200 dilution, green) and ac-SMC3 (1:100 dilution, red). Cells were then analysed by Nikon C2 confocal microscope; images were taken at 20x magnification (z-series) using the NIS elements software, z-stacked and processed using FIJI. The results are consistent with work published by the Deardorff group (Deardorff et al, 2012). Levels of total-SMC3 remained relatively stable in control and PCI-34051-treated cells, however, there is accumulation of ac-SMC3 upon treatment with PCI-34051.
Figure 5.4
Figure 5.4 PCI-34051 treatment did not affect expression of oestrogen-responsive genes despite accumulation of ac-SMC3

Oestradiol-deprived MCF7 cells were transfected with 5 nM of siRNA targeting RAD21 (or non-targeting control) or 25 µM of PCI-34051 blocking deacetylation of SMC3, for 48 hours. Oestradiol (100nM) was added and cells were harvested at 3 and 6 hours post-stimulation, for RNA isolation. Transcript levels for the indicated genes were quantitated by qRT-PCR and normalised to the mean of two reference genes PPIA and GAPDH. Transcript levels of reference genes for PCI-34051-treated cells are included in Appendix K. Error bars represent the SEM of three independent biological replicates for each treatment type (siRNA transfection or treatment with PCI-34051, separately). Significance was determined using two-way ANOVA. (A.i and B.i). Transcript levels of cohesin-dependent genes upon treatment with RAD21 siRNA or (A.ii and B.ii), after PCI-34051 treatment and stimulated with 3 (A.i and ii) and 6 (B.i and ii) hours of oestrogen, respectively. The data indicate that upon RAD21 depletion, oestrogen responsiveness of the indicated genes is affected significantly; however, responsiveness to oestrogen is retained in cells treated with PCI-34051 despite an accumulation of ac-SMC3.

5.3.3 PCI-34051 delayed cell cycle progression of MCF7 cells without altering ERα levels

The transcription and cell cycle functions carried out by the cohesin protein complex were shown to be separable in many studies (Gard et al, 2009; Rollins et al, 2004). Since PCI-34051 did not appear to influence oestrogen-dependent gene transcription, I next sought to determine if cell cycle progression was affected. Oestrogen-starved MCF7 cells treated with PCI-34051 were stained with propidium iodide and subjected to flow cytometry following 6 or 24 hours of stimulation by oestrogen. Results from cell cycle analysis using the 25 µM dose is presented in Figure 5.5 and that of the 10 µM dose is included in Appendix L. Cell cycle analysis revealed that PCI-34051-induced block of SMC3 deacetylation resulted in delay of transition of cells from G₁ to S phase (Figure 5.5A); this delay was also observed in RAD21-depleted cells (Figure 3.6, Chapter 3). After 24 hours of oestrogen stimulation, cells from control and PCI-34051 treatments had progressed into the S phase with PCI-34051-treated cells showing a lesser extent of transition (Figure 5.5Aii). This suggests that PCI-34051-treated cells become cell cycle delayed but are still viable and can still respond to the growth stimulus of oestrogen. This further suggested that although the inhibitor caused a generalised delay in cell cycle progression, which was rightly expected, however it did not arrest cells in any particular phase. This observation indicates that hyperacetylation of SMC3 did not impair transition of cells from G₁ to S phase.
To further determine if the differences in transcriptional responses for the two variables (RAD21 depletion and ac-SMC3 accumulation) are caused by differential effects on the level of ERα, I evaluated the expression level of the hormone receptor in cells with accumulated ac-SMC3. The rationale was to investigate if maintenance of oestrogenic transcriptional response following treatment with PCI-34051 is caused due to increased transcription of the ESRI gene, by any chance. As described in Chapter 3 (Figure 3.8), in oestrogen-stimulated conditions, RAD21-depletion did not alter levels of the ERα transcript or its protein. I investigated the effect of PCI-34051 on the level of ERα mRNA and protein in oestrogen-starved MCF7 cells. RNA from the 6 and 24 hours time points (n=3) was assessed for the level of ERα transcripts by qRT-PCR by normalising transcript levels to the mean of PPIA and GAPDH. Treatment with the inhibitor did not alter the basal or oestrogen-induced expression levels of ERα mRNA at either of the time points analysed (Figure 5.5Bi). Protein lysates from the 6 and 24 hours time points were immunoblotted with anti-ERα and γ-tubulin antibodies. Upon normalising ERα protein levels to that of γ-tubulin, consistent with ERα transcript levels, no statistically significant up/downregulation of the ERα protein was observed in cells treated with 10 µM (Appendix L) or 25 µM (Figure 5.5Bii) of PCI-34051. Therefore, altered ERα levels cannot account for the observed oestrogen-responsive transcriptional disparity between RAD21 depletion and inhibition of SMC3 deacetylation.

Taken together, these results revealed that abrogating cohesin function by targeting different subunits (via siRNA depletion or chemical inhibition) could have different consequences for transcriptional regulation. RAD21 depletion impaired oestrogen responsiveness of a specific subset of genes, but when SMC3 function was blocked chemically, the same genes did not lose their responsiveness to oestrogen.
Figure 5.5 Assessment of cell cycle progression and level of ERα in oestrogen-starved MCF7 cells treated with 25 µM PCI-34051

Oestrogen-deprived MCF7 cells were treated with 25 µM of PCI-34051 for 48 hours, and then treated with vehicle or 100 nM of 17-β-oestradiol for 6 and 24 hours, respectively. (A) Flow cytometry analyses of cell cycle distributions. Following incubation with PCI-34051 and oestradiol stimulation, cells were fixed and
stained with propidium iodide and cell cycle analyses were carried out on a Beckman Coulter Galios Flow
Cytometer and processed using the FlowJo software (version 9.7). Histograms represent the intensity of
fluorescence on the X-axis and cell number on the Y-axis for (i) 6 hours and (ii) 24 hours time points. Bar
graphs show the percentage of cells quantified in the G0/G1, S and G2/M phases of the cell cycle in control and
PCI-34051-treated conditions for vehicle/oestradiol stimulation, separately. Error bars represent SEM from two
independent biological replicates. Significance was determined by two-way ANOVA. (B) Assessment of
transcript and protein levels of ERα in response to treatment with PCI-34051. (i) Transcript levels of ERα in
MCF7 cells. RNA levels were determined using qRT-PCR and are shown relative to the mean of two reference
genes, PPIA and GAPDH. Error bars represent the SEM from 3 independent experiments. Significance was
determined using two-way ANOVA. (ii) Representative immunoblots showing the level of ERα protein after
incubation with PCI-34051 for the indicated time points. The level of ERα protein was analysed by SDS-PAGE
and immunoblotting. The bar graph depicts quantification of total ERα protein levels, normalised to γ-tubulin.
Error bars represent the SEM from three independent experiments. PCI-34051 delayed cell cycle progression
but did not alter ERα mRNA or protein levels.

5.3.4. PCI-34051 delayed cell cycle progression of MCF7 cells in a
concentration-dependent manner

Having found that a PCI-34051-mediated block of SMC3 deacetylation did not affect
oestrogen-responsive transcription of selected cohesin-dependent genes, I next sought to
characterise the cell cycle modifying properties exerted by PCI-34051.

The flow cytometry data from oestrogen-starved MCF7 cells indicated that PCI-34051
treatment delayed cell cycle progression of cells grown in hormone-deprived media.
Although delayed, when stimulated by oestrogen, PCI-34051-treated cells were able to
progress from G1 to S phase (Figure 5.5A). This indicates that DNA synthesis can proceed,
but whether these cells were able to complete mitosis or impaired cohesin function causes
cell death, was a possibility I thought was worth testing in cells grown in full supplemented
media. Starved-media was replaced by full supplemented media to eliminate confounding
possibilities arising from reduced basal cellular activity due to growing cells in serum-
starved medium. As mentioned before, the SMC3 acetylation cycle is a tightly cell cycle
regulated event. Therefore, instead of using exponentially growing cells, I decided to
synchronise MCF7 cells grown in full supplemented media to induce these to cycle in a
homogeneous manner. PCI-34051 blocks HDAC8 from deacetylating SMC3 following sister
chromatid separation in anaphase. Hence, I synchronised the cells at the G1-S phase border
using a double thymidine block (in a phase prior to the actual phase to be investigated as
recommended for cell synchronisation experiments) (Harper, 2005). This was achieved by
inhibiting DNA synthesis by using two sequential rounds of exposure to thymidine, separated by an interval of recovery time (specific for the doubling time of the cell line used, 9 hours of recovery time for MCF7 cells (Sutherland et al, 1983) in this case) from the thymidine block. Details of the experiment are included in Figure 5.6A.

To determine concentration-dependent effects of PCI-34051 on the cell cycle progression of MCF7 cells, apart from the lower doses of 10 and 25 µM (used for expression studies), two further higher doses of 50 and 100 µM were included. After release of second thymidine block and subsequent treatment with PCI-34051, cells were monitored for up to 36 hours with cells harvested for cell cycle analysis at every 2 hours interval for the first 12 hours, thereafter every 12 hours until up to 36 hours. Cells were harvested for up to 36 hours as cells harvested beyond this time would have lost synchrony as the synchronisation procedure is recommended for one cycle passage only. Cell synchronisation experiments revealed that PCI-34051 delayed cell cycle progression of MCF7 cells in a dose-dependent manner (Figure 5.6B). Transition to successive cell cycle phases was delayed at all doses tested, with effects for the lower doses (10, 25 µM) showing up at reasonably late exposure times (8 and 6 hours of treatment for 10 µM, 25 µM doses, respectively) while that of the higher doses (50, 100 µM) were distinctively apparent relatively early (as early as 2 hours of treatment). The higher doses also appeared to exert cytotoxic effects as evidenced by cells in sub-G₁ phase. This synchronised cell cycle study supported my previous observation (Figure 5.5A) that exposure to varying doses and treatment periods of PCI-34051 slowed growth rate of MCF7 cells by curtailing successive progression of cells throughout the entire course of the cell cycle, in a concentration-dependent manner.
Figure 5.6

A. Cells seeded, allowed overnight attachment Thymidine Block release Thymidine Block release & PCI34051 treatment Harvest cells for Flow cytometry

18 hours 9 hours 17 hours 2 4 6 8 10 12 24 35 hours

B. Hours post release 0 hours 2 hours 4 hours 6 hours 8 hours 10 hours 12 hours 24 hours 36 hours

Cell count

- PCI 100 µM
- PCI 50 µM
- PCI 25 µM
- PCI 10 µM
- Control
Figure 5.6 Characterisation of cell cycle kinetics of PCI-34051-treated synchronised MCF7 cells

Asynchronous MCF7 cells were synchronised to G1-S phase by double thymidine block. Thymidine was removed, and cells were released in the presence of different concentrations of PCI-34051 for the indicated time points. Cells were fixed in ethanol and stained with propidium iodide. Cells were gated to exclude aggregates and doublets, and identical gates were applied to the cell populations of every time point relative to the vehicle-treated control of that time point. Cell cycle distributions were assessed by flow cytometry on a Beckman Coulter Gallios Flow Cytometer and processed using the FlowJo software (version 9.7). (A) Schematic of experiment design used for synchronising cells to G1-S phase by double thymidine block. (B) Histograms of flow cytometry analyses showing fluorescence intensity on the X-axis and cell counts on the Y-axis. DNA content analysis of vehicle-treated control cells and PCI-34051-treated cells is presented from 0, 2, 4, 6, 8, 10, 12, 24 and 36 hours, post treatment. Lines are shown for ease of comparison of gated cell populations between G1 and S/G2-M phases of control and PCI-34051-treated cells. These data indicate that PCI-34051 delayed cell cycle progression in a concentration-dependent manner. Cell cycle delay was apparent earlier at higher concentrations of PCI-34051.

5.3.5 PCI-34051 suppressed proliferation of MCF7 cells

After characterising cell cycle progression in response to PCI-34051, I next determined the effects of this inhibitor on MCF7 cell proliferation and survival. Since HDAC8 promotes tumour cell proliferation (Vannini et al, 2004; Wu et al, 2013) it is possible that inhibition of HDAC8 could be deleterious for breast cancer cell survival. MCF7 cells treated with different doses (10, 25, 50, 100 µM) of the inhibitor were monitored for cell growth in real time, using the Essen IncuCyte automated digital live imaging system (described earlier in section 2.8.2.2). Using this system, cell proliferation was quantified in the presence of varying concentrations of PCI-34051 relative to vehicle-treated control cells, at regular intervals until seven days. At the end of the treatment period, integrated growth curves were generated using the confluence software (Figure 5.7Ai). Phase-contrast images of cellular morphology were also taken at regular intervals using proprietary algorithms of the software (Figure 5.7Aii). As shown in Figure 5.7Ai, treatment with PCI-34051 was significantly anti-proliferative for MCF7 cells for all doses tested, except for the lowest dose of 10 µM which did not appear as inhibitory (until day 5) when compared with the higher doses. The graph shows temporal and dose-dependent anti-proliferative effects of the inhibitor over a treatment window of 7 days. Also shown in Figure 5.7Aii are the phase-contrast morphological images of PCI-34051-treated cells from treatment day 5 and 7 showing that prolonged treatment with PCI-34051 is lethal to MCF7 cells.
In addition to Incucyte-based growth monitoring, effects of PCI-34051 on cell growth were also measured by the colourimetric, metabolic assay MTT. Results from the MTT assay confirmed dose-dependent reduction in MCF7 cell proliferation at all times tested (Figure 5.7B). Two-way ANOVA indicated a significant treatment × time interaction (p-value <0.0001), together with significant time (p-value <0.0012) and treatment (p-value <0.0001) effects. Treatment with lower doses of PCI-34051 (10 and 25 µM) for 24 hours caused no detectable inhibition of cell growth. However, with increasing time, even the lower doses started to exert growth-hindering properties. After day 3, all concentrations of PCI-34051 caused significant growth suppression of MCF7 cells; however, effects with higher doses were pronounced from early on (end of day 1). Sidak’s multiple comparisons test indicated that treatment with PCI-34051 for 1 day caused significant growth suppression with 50 µM (p-value <0.0009) and 100 µM (p-value <0.0001). From 3rd day onwards, maximal growth inhibition (p-value <0.0001) was encountered with doses 25, 50 and 100 µM. However, the 10 µM dose was relatively less growth suppressive (range of p-value <0.0057 to p-value <0.0215). Taken together, these results indicated that PCI-34051 treatment suppressed MCF7 cell proliferation in a concentration-dependent manner.
Figure 5.7

MCF7 cells were grown in full media and seeded in quadruplet on 96 well plates at a cell seeding density of 9900 cells for treatment periods of days 1 and 3, 6250 cells for day 5, and 4000 cells for day 7 treatments. Cells were allowed to attach overnight and then treated with varying concentrations of PCI-34051. For the 5-day treatment, culture medium containing the inhibitor was replaced on day 3, and for the day-7 treatment culture medium containing the inhibitor was replaced on days 3 and 5. (A) PCI-34051 is cytotoxic to MCF7 cells at all doses tested. The growth of cells was periodically monitored using the Essen IncuCyte FLR digital imaging system. Data points collected every 10 hours using monolayer confluence are presented in the graph (i) Percentage of cell confluence is plotted on the Y-axis with treatment times on the X-axis and show that the inhibitor reduced growth of MCF7 cells in a dose-dependent manner. Error bars are SEM of two independent experiments. (ii) Representative morphological phase-contrast images from the IncuCyte showing cells treated with different doses of PCI-34051, at the end of treatment periods of days 5 and 7. This indicated prolonged...
treatment with PCI-34051 was lethal to MCF7 cells. (B) MTT assessment of cell proliferation indicated significant growth suppression of MCF7 cells in response to PCI-34051. MTT assay at the end of each treatment period indicated significant decreases in cell proliferation at all times tested. Two-way ANOVA indicated a significant treatment × time interaction (p-value <0.0001). Sidak’s multiple comparisons test revealed significant differences between control and treated cells. Error bars represent the SEM of two independent experiments.

5.3.6 PCI-34051 induced cell death in MCF7 cells

Having shown that PCI-34051 suppresses the growth of MCF7 cells, I next investigated cytotoxic effects of this inhibitor. For this, I made use of a fluorescent dye YOYO-1, which is a dimeric cyanine nucleic acid stain that only permeates cells that have lost plasma membrane integrity. YOYO-1 binds to dsDNA of dying cells, allowing for the kinetic evaluation of cytotoxicity when monitored in real time using IncuCyte FLR. For this assay, cells were seeded in triplicate on 6 well plates and allowed to attach to the culture vessel overnight. On the following day, along with varying concentrations of PCI-34051, YOYO-1 was added to the growth medium at a final dilution of 1:10,000 and cytotoxic effects of the inhibitor was monitored using IncuCyte’s ‘fluorescent and phase-contrast scans’. The IncuCyte FLR object counting algorithm was used to quantify cytotoxicity, which takes into consideration parameters such as object count/mm² and object confluence. Consistent with the profiles of proliferation inhibition induced by PCI-34051 (see section 5.3.5), a gradual increase in fluorescence confluence was observed with increased dose and treatment time (Figure 5.8Aii). Cell non-viability was greater with higher doses and with increased treatment times (Figure 5.8Aiii). To control for cell proliferation and to calculate the cytotoxic index, triton-X 100 was added at a concentration of 0.0625% at the end of the assay period. Triton X-100 was added to permeabilise all viable cells and to label all DNA-containing objects with the YOYO-1 reagent. However, this triggered onset of a high fluorescent background owing to the culture medium (DMEM) containing high levels of riboflavin. The idea was to normalise the number of YOYO-1 stained objects to the total number of DNA-containing objects as a means for end point normalisation. However, this could not be accomplished due to excessive rise in background fluorescence, which interfered with computation of accurate downstream readouts. Regardless, the YOYO-1 assay confirmed that PCI-34051 causes cell death in a dose-dependent manner.

To confirm results obtained from the YOYO-1 measure of membrane integrity, cell death triggered by PCI-34051 was also assessed by one of the oldest but also extensively used methods, trypan blue exclusion count test. Dead cells that have lost membrane integrity stain
blue whereas live cells with intact membranes do not take up the dye. Cells exposed to different doses of PCI-34051 were stained with 0.4% solution of trypan blue dye, and the percentage of stained cells was counted microscopically on a Neubauer chamber. Percentage of viable cells was calculated as per the following:

\[
% \text{ viable cells (or } % \text{ survival)} = \left( \frac{\text{Number of unstained cells}}{\text{total number of cells}} \right) \times 100
\]

Consistent with cell growth analysis by IncuCyte FLR, proliferation suppression assessment by the MTT metabolic assay, and cell viability estimation by membrane integrity analysis using YOYO-1 dye, results from the Trypan Blue exclusion test (Figure 5.8B) indicated that increasing concentrations of PCI-34051 were accompanied by increased cell death.

Gene expression analyses (Figure 5.4) were made following treatment with 10 µM and 25 µM of PCI-34051 for 51 hours (3 hour time point; 48 hours of PCI-34051 treatment + 3 hours of stimulation with oestradiol) and 54 hours (6 hour time point; 48 hours of PCI-34051 treatment + 6 hours of stimulation with oestradiol). The possibility that cell death affected downstream gene expression profiles of cells treated with 10 µM of PCI-34051 is low, as the 10 µM dose appears non-cytotoxic for the time period assessed. However, some degree of cell death (~20-30%) is observed with 25 µM of PCI-34051. Hence, there remains a possibility that changes in gene expression are caused by cytotoxicity induced by this dose of PCI-34051.
Figure 5.8 PCI-34051 induced cell death in MCF7 cells

MCF7 cells were seeded at a density of 300,000 in triplicate on 6 well plates and allowed to attach overnight. (A) Assessment of cell viability using YOYO-1 dye. Cells were treated with varying concentrations of the inhibitor and cell viability was estimated by the fluorescent dye YOYO-1, which only permeates non-viable cells, fluorescently stains their nuclear DNA, and allows for the kinetic evaluation of non-viability. Phase contrast and fluorescent images were taken every 3 hours using the IncuCyte FLR (Essen, V1.5). Using the IncuCyte FLR object counting algorithm, non-viability was quantified by plotting fluorescence confluence versus time (ii) and to estimate the degree of non-viability fluorescence confluence/cell confluence was plotted over time (iii). Representative fluorescent phase contrast images from the IncuCyte, from day 3 and day 4 of treatment are presented in Figure 5.8Ai. Error bars represent the SEM of two independent experiments. (B) Trypan Blue exclusion assay for assessment of cell death in response to treatment with PCI-34051. After treatment with varying concentrations of the inhibitor for the indicated time periods, cells were stained with trypan blue and the number of stained (dead) and unstained cells (live) were assessed microscopically. % viable
cells (% survival) was calculated by dividing the total number of unstained by the total number of cells (stained + unstained) and represent the mean of three experiments. Error bars are SEM. These data indicate that PCI-34051 caused dose-dependent death of MCF7 cells over time.

5.3.7 Investigating the mechanism of cell death induced by PCI-34051 in MCF7 cells

After assessing the anti-proliferative and cytotoxic effects of PCI-34051, I next sought to investigate the mechanism by which PCI-34051 triggers cell death. The inhibitor was detected to induce caspase-mediated apoptosis in T-cell lymphomas, however for doses less than 20 µM and up to 3 days of treatment with PCI-34051, apoptosis could not be detected in solid tumour cell lines (including breast tumour lines) as measured by Annexin-V based flow cytometric analyses (Balasubramanian et al, 2008). However, exposure of phosphatidylserine on the external surface of the plasma member of an apoptotic cell (which is measured by Annexin-V staining) is a late apoptotic event, the timing of which can differ depending on the events that take place prior to it. The biochemical progression of apoptosis is dependent on a number of additional factors including nature and concentration of the death-inducing agent. To determine the kinetics of apoptotic cell death, I monitored MCF7 cells treated with varying concentrations of PCI-34051 at multiple time points and by a few different methods.

Members of the BCL2 family of genes such as BAX and PUMA, have been shown to exert pro-apoptotic activity (Chittenden et al, 1995; Dudgeon et al, 2010; Nakano & Vousden, 2001) resulting in increased mitochondrial membrane potential and release of Cytochrome C, which is counteracted by overexpression of the anti-apoptotic gene BCL2 (Tsujimoto, 1998; Yip & Reed, 2008). Interestingly, in human breast cancer cells, BCL2 and BAX genes are constitutively expressed to tightly regulate the apoptotic process (Adams & Cory, 1998; Kumar et al, 2000). Transcript levels of the three BCL2 family of apoptosis-regulating genes were evaluated for up to 72 hours following treatment with different doses of PCI-34051. Due to the high level of cytotoxicity exerted by the prolonged use of 100 µM of PCI-34051, this dose was excluded from the expression analysis, as it was difficult to collect RNA from dying cells. Transcript analysis showed that expression levels of BAX and PUMA were unaltered (statistically non-significant for all the doses) and that of BCL2 in PCI-34051-treated samples was downregulated in comparison to vehicle-treated control cells (Figure 5.9Ai). Although BCL2 expression was suppressed in PCI-34051-treated MCF7 cells, absence of upregulation of the pro-apoptotic genes BAX and PUMA did not indicate that
mitochondrial-mediated apoptosis was occurring in PCI-34051-treated MCF7 cells. The absence of a correlation between the expression levels of the pro- (BAX, PUMA) and anti-apoptotic (BCL2) genes perhaps suggest that regulation of the BCL2 gene in breast cancer cells is complex and probably influenced by several factors. However, I am aware that measuring their absolute levels may not always be contextually meaningful; instead assessing their activation states (like phosphorylation or conformational change) may be more appropriate.

I next assessed cleavage of PARP (poly ADP-ribose polymerase), a biochemical event preceding DNA fragmentation, which is the endpoint characteristic of the apoptosis pathway. For this, I analysed the early time point samples (treated for 6 and 12 hours only) as such an event is expected to occur sooner in the pathway. When uncleaved, a 113 kDa band appears in non-apoptotic cells and when cleaved, an 89 kDa band shows up in apoptotic Jurkat cells (T-cell leukaemia cell line). The Odyssey infrared detection system failed to detect the anti-PARP antibody. Hence, I used chemiluminescence-based immuno-detection for the anti-PARP immunoblot. The positive control (Jurkat cells) produced an 89 kDa band in the anti-PARP Western blot (Appendix M), however neither 113 nor 89 kDa bands appeared for control and PCI-34051-treated MCF7 cells (treated from 6 to 12 hours of PCI-34051). Rather an unexpected band of size ~60 kDa appeared in addition to the 48 kDa band of the loading control γ-tubulin for the untreated and treated cells (Appendix M). This made analysis of the state of PARP protein (cleaved or uncleaved) in PCI-34051-treated MCF7 cells inconclusive.

I moved on to next evaluate if there was internucleosomal DNA fragmentation (which is followed by PARP cleavage) in response to treatment with PCI-34051. For this, genomic DNA (gDNA) was extracted from MCF7 cells treated with different doses of PCI-34051 (from 24, 48 and 72 hours) and was evaluated for cleavage of gDNA by separation on agarose gels. If apoptotic, the samples are predicted to demonstrate a distinct ladder pattern characteristic of oligonucleosomal fragments of the size range of ~200 bp. DNA gels of the PCI-34051-treated samples did not demonstrate a DNA ladder, although samples from treatment day 4 showed smears of DNA (Figure 5.9Aii), which could be representative of necrosis. However, in the absence of a positive control for necrotic cells, this could not be confirmed.
As a further qualitative measure of cell death, I next assessed physical changes to nuclear morphology by microscopic examination of the PCI-34051-treated cells stained with the Hoechst (H33342) dye. As mentioned previously, Hoechst is a bisbenzimide DNA-binding dye that fluoresces blue when in contact with DNA. Nuclear condensation and fragmentation are also representative of cells undergoing apoptosis. Confocal microscopy showed that PCI-34051-treated cells did not demonstrate signs of fragmented nuclei (Figure 5.9Aiii). Taken together, these results indicate that PCI-34051-dependent cell death in MCF7 cells seems to be mediated by mechanisms other than apoptosis.

Next, the morphology of PCI-34051-treated cells was examined by phase contrast microscopy. Membrane blebbing, cell and nuclear shrinkage, formation of apoptotic bodies are representative of an apoptotic phenotype (Saraste & Pulkki, 2000). Although the morphology of the PCI-34051-treated cells appeared different to untreated controls (Figure 5.9Bi), this was not suggestive of an apoptotic phenotype.

As PCI-34051 interferes with cell cycle progression (Figure 5.6), I sought to determine its effect on cellular microtubule organisation. Microtubules are principal building block of mitotic spindles which are essential for cell division (Musch, 2004). To test the hypothesis that cell death in PCI-34051-treated MCF7 cells is accompanied by defects in microtubule assembly, cells were stained for α-tubulin and examined by confocal microscopy. Synchronised (G1-S) MCF7 cells were treated with varying doses of PCI-34051 and stained with an anti-α-tubulin antibody. Confocal images of MCF7 cells treated for 24 hours with 50 µM of PCI-34051 are presented in Figure 5.9B. Confocal immunofluorescence images indicated that after treatment with 24 hours of PCI-34051, the cytoskeleton of the PCI-34051-treated cells did not appear to be perturbed (Figure 5.9Bi; Bii). The data indicated that PCI-34051 treatment of concentrations up to 50 µM for a period of 24 hours did not affect mitotic microtubule assembly in G1-S synchronised MCF7 cells.
Figure 5.9 Evaluating the mechanism of cell death induced by PCI-34051 in MCF7 cells

(A) MCF7 cells grown in full media were incubated with different concentrations of PCI-34051 for up to 72 hours. (i) Transcript levels of BAX, PUMA and BCL2 in response to treatment with different doses of PCI-34051. Total RNA was extracted at the indicated time points, and transcript levels of the indicated genes were quantified by qRT-PCR and normalised to the mean of two reference genes PPIA and GAPDH. Data shown are SEM (n=2). (ii) Agarose gel electrophoresis analysis of DNA fragmentation. gDNA was extracted after 24 hours (lanes 1: DNA ladder, 2: Control + vehicle, 3: PCI-34051 10 µM, 4: PCI-34051 25 µM and 5: PCI-34051 50 µM).
50 µM); 48 hours (Lanes 6: PCI-34051 50 µM, 7: Control + vehicle, 8: PCI-34051 10 µM, 9: PCI-34051 25 µM) and 72 hours (Lanes 10: Control + vehicle, 11: PCI-34051 10 µM, 12: PCI-34051 25 µM and 13: PCI-34051 50 µM) of treatment with varying concentrations of PCI-34051. 1 µg of gDNA was separated on 1.8% agarose gels and electrophoresed to assess for DNA fragmentation. (iii) Confocal microscopy of nuclear morphology. G1-S synchronised MCF7 cells were treated with 50 µM of PCI-34051 for 24 hours, then washed in PBS, fixed, permeabilised and stained with Hoechst (H33342 to visualise DNA, 1 µg/ml). Cells were then analysed by confocal microscopy. (B) (i) Morphological analysis of PCI-34051-treated MCF7 cells. MCF7 cells were treated with varying concentrations of PCI-34051 for 3 days and visualised on an IX71 Olympus inverted microscope at 20x magnification. Images were taken using the Olympus DP71 camera and DC software and digitally zoomed and processed using Image J. (ii) Confocal immunofluorescence analysis of tubulin organisation of PCI-34051-treated MCF7 cells. G1-S synchronised MCF7 cells were treated with 50 µM of PCI-34051 for 24 hours, then washed in PBS, fixed, permeabilised, and stained with Hoechst (H33342 to visualise DNA, 1 µg/ml, blue) and α-tubulin (1:500 dilution, red) and detected with ‘Alexa fluor 488’ (goat polyclonal anti-mouse secondary antibody, 1:2000 dilution). Cells were analysed by Nikon C2 confocal microscope; images were taken at 60x magnification (z-series) using the NIS elements software, digitally zoomed (ii), z-stacked and processed using FIJI.

5.4 Discussion

The identification of HDAC8 as the human SMC3 deacetylase and the subsequent characterisation of the HDAC8 inhibitor PCI-34051 in HeLa cells (Deardorff et al, 2012), offered the opportunity to investigate whether blocking SMC3 deacetylation has similar consequences to knocking down RAD21 in MCF7 cells. Although, the aftereffects of a perturbed SMC3 acetylation cycle are incompletely understood, prior studies have indicated that this acetylation switch is crucial for dissolving used cohesin (cohesive) and recycling it for its optimal functionality in the subsequent cell cycles (Beckouet et al, 2010; Xiong et al, 2010). In this regard, work by Deardorff and colleagues provided the first interesting evidence that siRNA-mediated loss of HDAC8 in HeLa cells resulted in a unique pattern of altered transcription that was similar to CdLS cell lines harbouring mutations in either NIPBL or HDAC8. Remarkably, the HDAC8-specific inhibitor PCI-34051 reproduced results obtained using a HDAC8 siRNA. From the viewpoint of substrate selectivity, other than the Deardorff study, confirmation of PCI-34051 inhibiting HDAC8’s SMC3 deacetylase activity came from two recent studies, where SMC3 acetylation was identified to be a key target, reinstating PCI-34051’s target specificity for cohesin subunit SMC3 (Olson et al, 2014; Scholz et al, 2015).

In MCF7 cells, treatment with PCI-34051 resulted in the accumulation of ac-SMC3 (Figures 5.2, 5.3), which is consistent with PCI-34051 blocking HDAC8 activity in these cells.
However, inhibition of SMC3 deacetylation did not affect oestrogen responsiveness of selected cohesin-dependent genes. Transcription of oestrogen-responsive genes that were altered upon siRNA depletion of RAD21 remained normal in cells with accumulated ac-SMC3 (Figures 5.3, Appendix K). This led to the conclusion that a chemical block of cohesin recycling is perhaps not comparable with siRNA depletion of RAD21, in terms of effects on oestrogen-dependent transcription. Moreover, a recent study has identified HDAC8 to be the deacetylase for a few other proteins in MCF7 cells, in addition to ac-SMC3 (Olson et al, 2014). Hence, there remains a possibility that the transcriptional responses observed in PCI-34051-treated MCF7 cells could be the result of targeting these additional substrates of HDAC8, outside of the cohesin system. The other HDAC8 substrates identified in MCF7 cells include the transcription factor RAI1, epigenetic regulators ARID1A (member of the SWI/SNF family) and MLL2 (member of the trithorax family), transcriptional co-activator NCOA3, alongside CENPF and ZRANB2 proteins. Reflecting on the primarily nuclear nature of the identified substrates, it is plausible that some of these proteins might be, at least in part, driving the transcriptional events observed in PCI-34051-treated MCF7 cells. Taken together, these results argue for caution in assuming equivalence between RNAi-mediated perturbation versus other means of impairing cohesin (chemical block of cohesin recycling in this case). Therefore, due to the interference of confounding variables, different modes of disrupting cohesin function may not necessarily recapitulate all functions of the cohesin complex; an interpretation also reported by a previous study (Laugsch et al, 2013).

Loss of HDAC8 in HeLa cells was associated with 17% reduction in chromatin-associated cohesin peaks, which might explain the characteristic changes in the expression of classifier genes associated with the clinical manifestation of CdLS. Whether inhibition of HDAC8 in MCF7 cells also affects the distribution of cohesin binding to the regulatory regions of its target genes is not known as yet, and is definitely worth attempting in future studies. With RAD21 depletion, there was loss of cohesin (RAD21) binding at cohesin-dependent oestrogen-responsive genes (as explored for SOX4 and IL20 in Chapter 4, Figures 4.5 and 4.6), so it would be interesting to test binding of RAD21 in cells with accumulated ac-SMC3. If RAD21 binding persists at the regulatory regions of cohesin-dependent genes, even when cells have accumulated ac-SMC3 that might explain maintenance of oestrogen responsiveness of these genes.
In synchronised cultures of HeLa cells, cell cycle progression remained unaltered for up to 12.5 hours of treatment with 25 µM of PCI-34051. However, in MCF7 cells, PCI-34051 caused a dose-dependent cell cycle delay of cells grown in oestrogen-starved (Figure 5.5A) as well as in full growth medium (Figure 5.6). Delay of cell cycle was most conspicuous at higher doses and manifested with a time delay at lower doses of the inhibitor. Unlike observations made in HeLa cells (Deardorff et al, 2012), increasing concentrations of PCI-34051 delayed progression of MCF7 cells through different phases of the cell cycle.

PCI-34051 was significantly anti-proliferative in a dose-dependent manner consistent with HDAC8’s role in facilitating cell survival in tumour cells. PCI-34051 suppressed cell growth (Figure 5.7Ai, 5.7B) and induced cell death (Figures 5.7) in a concentration-dependent manner. Consistent with previous reports (Balasubramanian et al, 2008) and with preliminary qualitative analyses, PCI-34051-treated cells did not demonstrate features indicative of apoptosis at doses up to 50 µM (Figure 5.9A). However, considering biochemical processes of dying cells can be incredibly complex, a thorough analysis of intermediary events in the timeline of an apoptotic pathway would be essential before confirming this. The mechanism of cell death may also involve crosstalk between several cell death pathways. Hence, to be sure of the specific mechanism of cell death induced by PCI-34051 in MCF7 cells, additional qualitative and quantitative assays would be necessary for addressing other modes of cell death as well. Alternatively, the Olson study identified that PCI-34051 treatment of MCF7 cells correlated with an overexpression of the cyclin-dependent kinase inhibitor p21. Furthermore, biological pathway analysis revealed a direct connection of the three HDAC8 substrates RAI1, ARID1A and MLL2 with p21. Hence, it could be possible that the acetylation statuses of these three substrates might influence the expression level of this important cell cycle regulating protein. Since p21 can induce growth arrest alongside cellular senescence, it is reasonable to speculate that some of the growth inhibitory effects seen in my experiments could be partly mediated by p21.

In order to circumvent some of the HDAC8i PCI-34051 associated limitations observed in my study (unintended off-targeting of other HDAC8 substrates, leaving aside ac-SMC3), a SMC3 knockdown approach could have been taken, which could have served as a control to then compare its transcriptional output with that of RAD21 knockdown. However, one limitation to such an approach would have been that more than 30% knockdown of SMC subunits have been shown to induce gross effects on cell proliferation and viability in
mammalian cells (Bauerschmidt et al, 2010; Kim et al, 2002b). However, upto ~75% of RAD21 knockdown did not lead to such deleterious effects in MCF7 cells. As cohesin is known to exert biphasic effects owing to cohesin dosage (Schaaf et al, 2009), comparing the transcriptome obtained on a background of less than 30% reduction in SMC3 with that of 75% reduction in RAD21 would have been technically inappropriate. Hence, such an approach was not considered as part of this study.

Taken together, my results indicate that targeting cohesin by blocking deacetylation of SMC3 delayed cell cycle progression but did not affect cohesin (RAD21)-mediated regulation of oestrogen-responsive genes. The inhibitor PCI-34051 also did not alter levels of ERα transcript and protein. However, PCI-34051 decreased cell proliferation and led to cell death without showing features characteristic of apoptosis. Further characterisation of PCI-34051, including genome-wide analyses in additional cell lines would be crucial before it can be considered as a therapeutic option for the treatment of oestrogen fuelled breast cancers with overexpression of the RAD21 transcript/protein.
CHAPTER 6
Discussion

With the advent of genome-wide transcriptome analysis techniques, we are beginning to understand that cohesin’s tissue-specific regulatory functions are well separated from its role fundamental to maintaining genome integrity. Several studies in the last decade have indicated that disruption of cohesin function can create a conduit to tumourigenesis and cell death (Losada, 2014). Cohesin dysregulation may involve overexpression, underexpression and mutation of cohesin subunits and cohesin-interacting proteins, either alone or collectively (Dolnik et al, 2012; Kim et al, 2012; Oikawa et al, 2004; Solomon et al, 2011).

One of the hallmarks of cancer (including those that harbour cohesin alterations) is a perturbed transcriptome, yet what continues to remain enigmatic is whether transcriptional dysregulation generated by cohesin malfunction underpins initiation or sustenance of tumourigenesis. It is possible that dysregulation of downstream genes resulting from cohesin malfunction underlies development and progression of cancers that harbour alterations in cohesin (Losada, 2014; Rhodes et al, 2011; Xu et al, 2011a; Xu et al, 2011b).

With the identification of a SNP in intron 1 of the RAD21 gene as being significantly associated with breast cancer risk (Kammerer et al, 2004) and subsequent studies correlating RAD21 overexpression with treatment resistance and delayed prognosis (Xu et al, 2011b; Yan et al, 2012) in breast cancer, it seemed appealing to speculate if RAD21 contributes to this malignancy. Since, cohesin transcriptionally regulates ecdysone response in Drosophila salivary glands (Pauli et al, 2010) and RAD21 co-localises with ERα at multiple genomic sites in breast cancer cells (Schmidt et al, 2010), it seemed plausible that cohesin’s role in steroid hormonal control is conserved. Furthermore, the MYC oncogene requires cohesin for its oestrogenic transcriptional activation in breast cancer cells (McEwan et al, 2012). These findings formed the basis of my PhD research, which aimed to investigate cohesin-dependent gene regulation pertaining to oestrogen response in the breast cancer model system MCF7. I investigated if RAD21 depletion had functional consequences for the expression of other breast cancer associated genes, and/or, on the molecular pathways indispensable for breast cancer cell survival. Results from my research confirmed that cohesin transcriptionally regulates a subset of oestrogen-responsive genes. The first phase of my PhD involved taking an unbiased genome-wide approach to identify transcriptional signatures on a background of depleted cohesin (siRNA knockdown of RAD21). I then undertook a candidate gene approach and interrogated the potential mechanism of cohesin-
mediated transcriptional alteration, for two of the candidate genes, *SOX4* and *IL20*. This helped to further understand how cohesin might promote or impede transcription of oestrogen-sensitive genes. I next investigated the efficacy of a cohesin inhibitor PCI-34051 (HDAC8i), in targeting the transcriptional function of cohesin, pertaining to its role in oestrogen signalling. My results revealed that despite treatment with PCI-34051, oestrogen-sensitive genes that were affected by RAD21 knockdown remained normally responsive to oestrogen. This highlighted that different modes of cohesin disruption may not always bring about similar transcriptional outcomes (as witnessed in this case). Following this, experiments were carried out to characterise the cell cycle-centric functions and assess the oncostatic potency of PCI-34051.

6.1 Cohesin depletion leads to bidirectional transcriptional response to oestrogen in MCF7 cells

The effect of cohesin depletion on the expression of oestrogen-sensitive genes was analysed in MCF7 cells using a cDNA microarray. In order to map direct regulatory effects on oestrogen responsiveness, two early time points, 3 and 6 hours post oestrogen stimulation were chosen. Oestrogen-mediated transcription occurs after a time lag of 2-3 hours post ligand addition (Farach-Carson & Davis, 2003; Marino et al, 2005; Reid et al, 2009). Thus, it was important to pick genes that demonstrate an early change in transcription in response to oestrogen because with longer time points other secondary regulatory events are possible. Microarray analysis in MCF7 cells uncovered dependence on cohesin for normal response to oestrogen for around 12-22% of all oestrogen-responsive genes following 3 to 6 hours of oestrogen stimulation. By comparison, 7.7% of ecdysone-responsive genes were affected in cohesin-depleted *Drosophila* salivary glands (Pauli et al, 2010). Furthermore, in support of direct regulation by cohesin (Misulovin et al, 2008), all validated genes displayed cohesin binding sites within ~20 kb of their TSS, suggesting cohesin has a potential direct role in the regulation of ER-dependent transcription.

Consistent with cohesin-dependent ecdysone-responsive gene regulation in *Drosophila* (Pauli et al, 2010), cohesin depletion in MCF7 cells resulted in bidirectional regulation of oestrogen-responsive genes. Out of the 153 differentially expressed genes at the 3 hour time point, 60 genes were up-, 93 genes were downregulated. However, out of the 77 differentially expressed genes of the 6 hour time point, 32 genes were up- and 45 genes were downregulated. The expression signature recapitulated the pattern of cohesin-dependent
transcriptional regulation observed for ecdysone-responsive genes where out of 33 differentially expressed genes, 18 were up and 15 were down. This drew attention to cohesin’s role as an activator as well as a repressor of ER-mediated transcription. It remains possible that the observed complex effects on transcription resulted from an incomplete knockdown of the RAD21 protein (~75% across all experiments). However, since modest changes in cohesin activity are sufficient to alter gene expression (Borck et al, 2006; Dorsett, 2011; Kawauchi et al, 2009) without affecting chromatid cohesion, I contend that the level of knockdown was appropriate for this transcriptional analysis. Moreover, RAD21-depleted cells remained viable and only exhibited minor delays in cell cycle progression during the entire course of this study. Results from these analyses suggest that the regulatory role of cohesin in steroid hormonal response could be evolutionarily conserved. Cohesin depletion in MCF7 cells led to dichotomous changes in gene expression. This observation is consistent with the conclusion that despite sharing co-incident binding with ERα, cohesin is not merely a facilitator but may also act as a repressor of ER-dependent transcription, conditional on gene and context.

6.2 Cohesin (RAD21) depletion alters transcription of oestrogen-sensitive genes that actively contribute to breast cancer

As described in section 3.6.2, the identified signature of cohesin-dependent oestrogen-responsive genes is widely dispersed in the genome and implicated in a variety of biological processes (Table 3.3). Molecular pathway analysis identified participation of many of the significantly dysregulated genes in key signalling pathways that are known to drive cellular growth, proliferation and survival in breast cancer. These include pathways such as β1 integrin cell surface interactions, ErbB1 downstream signalling, Class I PI3K signalling events, mTOR signalling (Table 3.4) (Baselga, 2011; Howe & Brown, 2011; Hynes & Boulay, 2006; Yao et al, 2007). The PI3K/mTOR signalling axis is notably relevant to ER-positive breast tumours. Because PI3K/mTOR pathway is frequently activated in breast cancer, and when hyperactivated, it accelerates tumourigenesis and contributes to the development of endocrine resistance in these tumours (Ciruelos Gil, 2014). Interestingly, in a recent network analysis across a pan-cancer cohort, statistically significant frequency of co-occurrence was identified between cohesin mutations and the PI3K signalling sub-network, raising the possibility that cohesin interacts with this signalling pathway in cancer (Leiserson et al, 2015). Furthermore, stimulation of the mTOR pathway has been shown to overcome growth deficiency of cells devoid of cohesin (Xu et al, 2013).
Also noteworthy was the finding that many of the dysregulated genes (such as *IRS1, IGF1, PIK3R3*) belonged to upstream nodes (such as ErbB, IGF, PI3K, mTOR) that are considered crucial for activating the cascade of breast cancer signalling networks in mammary tumours (Figure 3.12). As discussed in section 1.3.5.2, oestrogen-fuelled proliferation of mammary cells is an output of coordinated signalling between ER and growth factor signalling networks (Azuma & Inoue, 2012; Farach-Carson & Davis, 2003; Marino et al, 2006). This reciprocal signalling signifies a high level of inherent interdependence and synchrony between the genomic and non-genomic actions carried out by ER. Interestingly, some of the identified cohesin-dependent oestrogen-responsive genes were found to constitute crucial nodes of this network, suggesting that cohesin participates in breast cancer signalling, at multiple levels within the network (Figure 3.12, section 3.7). Results from the microarray analysis thus suggest that cohesin perturbation perhaps not only affected the genomic (transcriptional) but also some of the non-genomic (growth factor-induced) actions mediated by ER.

Endocrine resistance is one of the biggest clinical challenges towards the effective treatment of ER-positive breast tumours. Out of the several mechanisms proposed, reciprocal signalling between ER and growth factor signalling cascades (PI3K/mTOR pathway being one) is thought to be the mechanism behind emergence of endocrine resistance in breast cancer (Osborne & Schiff, 2011). Due to increased growth factor signalling, ER-positive tumours gradually adapt to ligand-independent (non-oestrogenic) stimulation and start responding to growth factor-mediated cues for proliferation and survival. PI3K/mTOR pathway has been implicated in the development of endocrine resistance in breast tumours (Barone et al, 2009; Cavazzoni et al, 2012; Miller et al, 2011). Such studies have also shown that ER levels were inversely correlated with PI3K pathway activation. Cohesin depletion affected expression of upstream genes of the PI3K/mTOR network. Results from this study thus raise the question whether cohesin function can be therapeutically targeted to downregulate PI3K pathway activation. In this regard, cohesin’s potential as a therapeutic target in delaying or reversing any degree of endocrine resistance in ER-positive tumours with RAD21 overexpression warrants further investigation.
6.3 Cohesin modulates chromatin binding of proteins important for oestrogen-responsive gene transcription

As discussed in section 4.1, and reported by our group in a previous study in 2012, RAD21 depletion impeded oestrogen activation of the MYC gene, which was accompanied by loss of ERα binding from key regulatory sites in MCF7 cells (McEwan et al, 2012). In contrast, our recent paper reported that RAD21 depletion greatly enhanced oestrogen activation of genes of the TFF locus (TFF1, TFF2, TMPRSS3 except TFF3) and was accompanied by increased binding of ERα to regulatory sites in MCF7 cells (Antony et al, 2014). Both MYC and TFF locus genes are normally positively regulated by oestrogen, however, RAD21 depletion had opposite effects at both these gene loci.

To understand how cohesin depletion can have divergent effects on oestrogen-responsive gene expression, I investigated one of the potential mechanisms by which this control might be implemented (section 4.1). In MCF7 cells, ERα levels (transcript and protein) remained unaltered with RAD21 depletion (Figure 3.8). However, a paper published in 2012 showed SMC3-mediated knockdown of cohesin downregulated levels of ERα transcript and protein (Prenzel et al, 2012). Despite this, my results reinforce findings from previous studies showing cohesin depletion does not affect hormone receptor levels in vertebrates (Liu et al, 2009; McEwan et al, 2012; Monnich et al, 2011; Quintin et al, 2014). Thus, downregulation of oestrogen response for cohesin-dependent genes in MCF7 cells does not involve decreasing ESR1 expression or reducing levels of the ERα protein.

An alternative mechanism could involve altering binding of ERα to cognate regulatory sites of the target genes (as evidenced for MYC and TFF gene loci). Using genome-wide binding data of ERα and RAD21 in MCF7 cells (Table 4.2), I found that all experimentally validated candidates genes (including SOX4 and IL20) harboured one or more binding sites for both ERα and cohesin (RAD21). To determine the functional significance of cohesin binding at these genes, I investigated the behaviour of other factors that govern transcription of oestrogen-sensitive genes. For this, binding of ERα and RNAPIIser5-P (this antibody detects the early initiation form of the polymerase) proteins at regulatory sites were assessed for two model genes, SOX4 and IL20. Both these genes served as model candidates for delineating cohesin’s divergent control of oestrogenic responses. SOX4 is relatively oestrogen insensitive, however, it gained significant responsiveness to oestrogen following RAD21 depletion. In contrast, IL20 is highly oestrogen-induced a gene (Williams et al, 2008),
however, RAD21 depletion abrogated oestrogenic response of this gene. Depletion of RAD21 was accompanied by significant increase in ERα binding (sites 3, 4 Figure 4.5) to upstream putative enhancers for the SOX4 gene. In contrast, recruitment of ERα and RNA Pol II to the promoter of IL20 in response to oestrogen stimulation was reduced upon RAD21 depletion (Figure 4.6). Although decrease in ERα binding was statistically non-significant, reduced RNA Pol II binding (statistically significant) was consistent with downregulation of IL20 in the absence of cohesin. Therefore, taken collectively, ERα and RNA Pol II binding at SOX4 and IL20 reflected the transcriptional changes of both these genes in response to depletion of RAD21.

Taken together, cohesin binding at regulatory regions of SOX4 restricts oestradiol activation of SOX4 by possibly impeding enrichment of ERα binding to putative distal enhancers of SOX4. On the other hand, cohesin binding at the promoter of IL20 facilitates oestradiol activation of IL20 by allowing strengthening of ERα and facilitating recruitment of RNA Pol II. These results showcase that cohesin binding can have gene-specific influences via divergent effects on the chromatin binding levels of ERα and RNA Pol II, resulting in appropriate directionality of oestrogenic responses.

6.4 Other potential mechanisms for cohesin-dependent influences on ER-associated transcription

With regards to ER-associated transcription, it has emerged in the recent years that formation of long-range chromosomal loops between spatially distant EREs is characteristic of many oestrogen-responsive genes (Carroll et al, 2005; Carroll et al, 2006; Fullwood et al, 2009; Klinge, 2001; Pan et al, 2008; Theodorou et al, 2013). ER-anchored chromatin loops have functional implications for ER-dependent transcription (Fullwood et al, 2009; Pan et al, 2008; Theodorou et al, 2013). Also, cohesin co-binds with Mediator to coordinate long-range enhancer-promoter communication in embryonic stem cells (Kagey et al, 2010). In another study, cohesin was found to localise to genomic regions that harboured anchors for ER-associated chromatin interactions (Fullwood et al, 2009). Results from these studies raise the possibility that cohesin’s role in the formation of chromosomal loops (between spatially distant regulatory elements and proximal promoters) could contribute to its control of the oestrogenic response. Dysregulated expression of cohesin-dependent genes could thus be accompanied by, or even caused by, changed chromatin architecture in the neighbourhood of these genes upon cohesin depletion. In support of this idea, cohesin depletion alters
chromatin interactions in MCF7 cells (Li et al, 2013; Quintin et al, 2014). For example, for the MYC gene, work by our group showed that RAD21 depletion reduced chromatin interactions between the distal enhancer (67 kb upstream) and the MYC promoters, commensurate with oestradiol repression of MYC transcript levels, following RAD21 depletion (Antony et al, 2014). In contrast, we observed increased chromatin interactions between the TMPRSS3 ERE (enhancer) and the TFF1 promoter with depletion of RAD21, which was accompanied by an increase in transcript levels of TFF genes (Antony et al, 2014). However, contradictory results were reported by Quintin group, who found diminished frequency of interaction between oestrogen receptor binding sites and target promoters within the TFF gene cluster upon RAD21 depletion. Intriguingly, another paper also reported observing a global decrease in chromatin interactions and subsequent decrease in ER-mediated transcription on a background of depleted RAD21 (Li et al, 2013). Inconsistency with some of the data published by our group with the above mentioned studies could be partly explained by biphasic effects exerted by cohesin. In such a situation, transcriptional response to the same gene can differ with changes in cohesin doses owing to the putative formation of alternative chromatin arrangements (Schaaf et al, 2009). The amount of cohesin binding to chromosome is also thought to be important for the formation of such chromatin structures, which in turn regulates transcription of cohesin-bound genes (Gause et al, 2010). Another possible speculation is that RAD21-depleted cells from the Quintin paper were completely quiescent, whereas RAD21-depleted cells in our experiment were still viable and actively cycling (albeit with a slight delay in cell cycle, Figure 3.6). Hence, chromatin architecture observed in the Quintin study could have also been affected by growth/cell cycle arrest. Overall, chromosomal conformation at MYC and TFF gene loci reflected transcription of these genes and could be explained by cohesin depletion either enhancing (as for the TFF locus) or repressing (like for the MYC locus) chromatin looping.

In addition, there is evidence that cohesin is involved in modifying the chromatin environ to augment or impede transcription of a gene. Cohesin has been shown to influence activity of chromatin modifiers such as the PcG epigenetic silencing proteins in Drosophila, where cohesin was found to physically and also functionally interact with the PcG proteins (Schaaf et al, 2013b; Strubbe et al, 2011). In yeast, the chromatin remodelling complex RSC was shown to be critical for establishment of cohesin binding (Baetz et al, 2004). Moreover, in humans, cohesin was shown to interact with the chromatin remodelling complex ISWI-SNF2h, which was essential for cohesin to bind to certain specific DNA sequences (Hakimi
et al, 2002). These studies suggest that cohesin modulates the activity of chromatin modifiers. Furthermore, in a recent study, the histone demethylase JMJD2B was identified as a cofactor for ERα in breast cancer cells (Kawazu et al, 2011). Taken together, it seems possible that cohesin also interacts with complexes recruited by ERα and thus helps to modify chromatin in response to stimulation by oestrogen.

A hypothetical model for how cohesin might regulate oestrogen-responsive genes is presented using SOX4 and IL20 as examples (discussed in section 4.4, Figure 4.7). Cohesin and CTCF are involved in compartmentalising the genome (Dowen et al, 2014; Sofueva et al, 2013), it is therefore likely that these two proteins also form the functional chromatin domains that oestrogen-responsive genes reside in (Hu et al, 2008; Le Dily et al, 2014). There is a strong possibility that the integrity of such functional domains is compromised upon loss of cohesin. As a result, cohesin-mediated interactions within such chromosomal compartments could be affected by RAD21 depletion, such that some local chromatin interactions are lost while others are gained. For example, for SOX4 new interaction(s) between the ER-bound putative enhancers and the proximal promoter could be gained as a consequence of chromatin boundary disruption upon depletion of RAD21 (Figure 4.7) thus conferring oestrogen sensitivity. However, for IL20, loss of RAD21 may have destabilised/displaced ERα binding at the promoter, and also potentially abrogated the transcription-favouring enhancer-promoter loop(s). This could be accompanied by chromatin compaction by possible recruitment of repressive chromatin modifiers and loss of transcriptionally engaged RNA Pol II (Schaaf et al, 2013a) (Figure 4.6) and thus loss of oestrogen sensitivity. To support the possibility of chromatin compaction at IL20 following loss of cohesin, there is evidence of alterations in histone modification patterns with cohesin depletion in Drosophila (Schaaf et al, 2013b) and also in the zebrafish (Rhodes et al, 2010).

6.5 Chemical disruption of SMC3 deacetylation did not mimic oestrogenic transcriptional changes caused by RAD21 depletion

It has been long considered in the field of cohesin biology that since cohesin subunits together form a ring-like structure, disruption of any one of its subunits violates stoichiometric integrity of the cohesin complex and compromises cohesin function as a whole. Accordingly, I expected disruption of RAD21 (by siRNA depletion) or SMC3 (by using a chemical inhibitor which blocks deacetylation of acetylated-SMC3) subunits to cause
very similar if not identical functional consequences on the corresponding oestrogen-responsive transcriptome (section 5.1).

As discussed in section 5.1, most of the cohesin released from chromatin during mitosis is recycled by the SMC3 deacetylation pathway, catalysed by the SMC3 deacetylase HDAC8 (Figure 5.1). HDAC8 removes acetyl groups from SMC3 and refreshes the protein complex for reloading in telophase, for its functions in the regulation of gene expression in interphase and mitosis (Borges et al, 2010; Deardorff et al, 2012; Xiong et al, 2010). In HeLa cells, loss of HDAC8 activity and accumulation of ac-SMC3 did not interfere with cell cycle progression; however it resulted in a pattern of transcriptional alterations unique to CdLS (Liu et al, 2009) cases that harbour mutations in HDAC8 or NIPBL (Deardorff et al, 2012). Transcriptional dysregulation was accompanied by 17% loss in total-cohesin (tested using an anti-RAD21 antibody) and 16% loss in ac-SMC3 (tested using an anti-ac-SMC3 antibody) genome-wide binding, as identified by ChIP-seq analysis in HDAC8 depleted HeLa cells.

The SMC3 acetylation cycle is important for cohesion establishment (Beckouet et al, 2010); however, the consequences of accumulated ac-SMC3 post mitosis are not very well understood. Work by the Deardorff group provided the first evidence that inhibition or depletion of HDAC8 led to the accumulation of ac-SMC3 and resulted in the reduction of chromatin-bound cohesin leading to global transcriptional dysregulation. Interestingly, the HDAC8-specific inhibitor PCI-34051 reproduced results obtained using HDAC8 siRNA (Deardorff et al, 2012). Thus, this study uncovered that blockade of HDAC8 activity (either using a siRNA or an inhibitor) affected transcriptional regulation in HeLa cells without significantly disrupting the cell division process.

Based on these previous studies, the HDAC8 inhibitor PCI-34051 seemed to have the potential to be a ‘cohesin-inhibitor’, which aligned with the rationale of my project. Therefore, I used PCI-34051 with the goal to impair the transcriptional role of cohesin in MCF7 cells. Treatment of MCF7 cells with PCI-34051 resulted in the accumulation of ac-SMC3 in a dose-dependent manner, as expected (Figures 5.2, 5.3, Appendix J). However, unexpectedly, the oestrogen response of cohesin-dependent genes remained unaffected in response to PCI-34051 treatment (Figure 5.4). Loss of RAD21 was accompanied by loss of cohesin (RAD21) binding from the regulatory sites of cohesin-bound genes, as observed for SOX4 and IL20 (Figures 4.5, 4.6), which consequently affected gene expression of such
genes. However, it is not known as yet, whether PCI-34051-treated cells that accumulate ac-SMC3 lose cohesin binding from the regulatory sites of cohesin-dependent genes in MCF7 cells. If PCI-34051-treated cells retain their cohesin binding despite accumulating ac-SMC3, then such an observation would back up maintenance of oestrogenic responses of these genes. As per the existing notion, cohesin’s chromatin re-association efficiency should be relatively reduced (if not eliminated) when cohesin retains accumulated ac-SMC3 (Deardorff et al, 2012). However, this has not been tested as yet, and warrants further investigation, which could potentially be one of the future directions of this study. This is important, as it would help to delineate the mechanism behind retained oestrogenic response in MCF7 cells having accumulated ac-SMC3.

Speculating on the above possibility, it is known that at any given time, cells contain two separate pools of cohesin, one that is ‘chromatin-bound’ and the other serving more like a cohesin reservoir for the cell, the pool of ‘soluble cohesin’. The chromatin-bound pool of cohesin is further divided into three sub-pools. One pool (~30-40% of total nuclear cohesin) binds to chromatin in a highly dynamic manner (dynamic-binding pool) with a mean residence time of 25 minutes or so (Gerlich et al, 2006). This dynamic-binding pool is proposed to contribute to regulation of gene expression and chromatin organisation during interphase (Chan et al, 2012; Tedeschi et al, 2013). The second pool (~30% of total nuclear cohesin) binds in a highly stable manner, ‘stable-binding pool’. This pool requires acetylation of SMC3 by ESCO1/2 for stable binding to chromatin during and after S phase through to sister chromatid separation in anaphase. The stable-binding pool has a longer residence time of ~6 hours (Gerlich et al, 2006), which is proposed to mediate SCC (Nishiyama et al, 2010; Schmitz et al, 2007). A recent study (Ladurner et al, 2014) reported presence of a third sub-pool of chromatin-bound cohesin, which interacts with chromatin in a highly transient fashion (transient-binding pool) in the range of seconds. This pool corresponds to the transient cohesin-chromatin interactions previously observed in the Drosophila model system (Gause et al, 2010).

Upon completion of one round of the cell cycle, it is plausible that blockade of SMC3 deacetylation by PCI-34051 would only affect recycling of the stable-binding pool of mitotic cohesin (~30% of total cohesin) used in the previous cell cycle. It seems likely that failure of deacetylation of ac-SMC3 under PCI-34051-treated conditions, compromised functionality of this pool of cohesin (as evidenced by high levels of ac-SMC3 in PCI-34051-treated MCF7
cells, Figures 5.2 and 5.3). However, it remains possible that alongside ‘the pool of soluble cohesin’, the ‘dynamic-binding pool’ and the ‘transient-binding pool’ of chromatin-bound cohesin would remain unaffected, as cohesin of these pools did not require acetylation of their SMC3 subunits. Hence, the cumulative amount of available nuclear cohesin in these pools might be sufficient to drive transcription of cohesin-dependent genes during interphase of the successive cell cycle. Moreover, the cohesin reservoir in MCF7 cells is relatively high compared to HeLa cells because MCF7 cells already have plentiful of cohesin owing to overexpression of cohesin (RAD21) (Atienza et al, 2005). Thus, it may be possible that in HeLa cells, in one cell cycle, the proportion of non-recycled cohesin might be limiting to the total-cohesin content. Hence, absence of sufficient cohesin levels reduced cohesin binding to chromatin and disrupted the transcription regulatory programs driven by cohesin. On the contrary, for transcription to be affected by accumulation of ac-SMC3 in MCF7 cells, these cells would probably need to undergo multiple cell cycles to an extent that the functional cohesin reservoir becomes limiting due to non-recycling of used cohesin harbouring accumulated ac-SMC3. However, since PCI-34051-treated cells become cell cycle delayed and succumb to cell death over time, growing these cells for multiple cell cycles would be difficult. A schematic representation of this hypothesis is presented in Figure 6.1.

Together with the above hypothesis, it is also possible that transcriptional differences between the two approaches are affected by PCI-34051’s interference with cell cycle progression (Figure 5.6) and suppression of cell proliferation of MCF7 cells (Figure 5.7Ai and ii, Figure 5.8). Significant upregulation of basal levels of the cell cycle inhibitor CCNG2 was particularly indicative of this cell cycle blockade. This further aligns with the dose-dependent increase in the level of the cyclin-dependent kinase inhibitor p21, observed in response to treatment of MCF7 cells with PCI-34051, as per the Olson study (Olson et al, 2014). Hence, the resulting gene expression responses in PCI-34051-treated MCF7 cells are likely to be affected by stress arising from cell cycle blockade as well as cell death. Additionally, as discussed in section 5.4, the confounding off-target effects generated as a result of unintended inhibition of deacetylation of the newly discovered HDAC8 substrates will also have to be taken into consideration with regards to interpreting the observed transcriptional responses. These substrates may include the histone variants H2A, H2B, H3 (Buggy et al, 2000; Van den Wyngaert et al, 2000), H4 and the non-histone target tumour suppressor p53 (Dose et al, 2011). However, treatment with up to 25 µM of PCI-34051 was not observed to induce acetylation of histones or tubulin in Jurkat cells (Balasubramanian et
al, 2008) or acetylation of H4 in HeLa cells (Deardorff et al, 2012). Evidence from another study confirmed that histone acetylation remains unaffected upon knockdown of HDAC8 (Oehme et al, 2009). Furthermore, a recent acetylome profiling study carried out in PCI-34051-treated MCF7 cells confirmed that histones were not identified as HDAC8 substrates as part of the acetylome coverage (Olson et al, 2014). However, the same study identified six other HDAC8 substrates, namely RAI1, ARID1A, MLL2, NCOA3, CENPF and ZRANB2 that have roles in epigenetic regulation, chromatin remodelling as well as RNA splicing. Hence, these proteins are well capable of modulating gene expression in a number of different ways. Although, siRNA disruption of HDAC8 in a series of human cancer cell lines (colorectal, pancreatic, lung carcinoma) have been shown to decrease the expression of both wild-type and mutant p53, suggesting that HDAC8 transcriptionally regulates the expression of p53 (Yan et al, 2013). However, p53 was not identified as a HDAC8 substrate in the acetylome profiling experiments carried out by the Olson study. Therefore, direct modification of p53 by HDAC8 in MCF7 cells is not confirmed as yet. However, as per biological pathway analysis, the HDAC8 substrate SMC3 and the tumour suppressor p53 were found to be directly connected to HDAC8. Therefore, it is plausible that inhibition of HDAC8 by PCI-34051 in MCF7 cells could affect transcription of wild-type p53 via SMC3 (MCF7 cells have wild-type p53). This in turn could also contribute, to the observed transcriptional disparity (as p53 is able to regulate gene expression) (Woods & Vousden, 2001). Taken together, it is suspected that transcriptional responses observed in PCI-34051-treated MCF7 cells are unlikely to be simple and could be the complex outcome of one or more of the above-mentioned possibilities. For clarity of comparison, a table summarising the differences observed in response to the two variables, RAD21 siRNA and PCI-34051 treatment in MCF7 cells is presented in Table 6.1.

In addition to the above perspectives, these unresolved findings also highlight our limited understanding of ‘how the cohesin complex stoichiometry affects transcriptional regulation’. This is an area of cohesin research that continues to remain underinvestigated, confronting many scientists of this field. Mild cohesin deficits (in the core subunits or its regulators) have been shown to drastically affect gene expression and development, leading to striking pathologies (Borck et al, 2006; Kawauchi et al, 2009). However, severe reduction in cohesin levels do not have a detrimental impact on sister-chromatid cohesion and DNA damage repair (cohesion-based functions) (reviewed in Dorsett (2011); Heidinger-Pauli et al (2010); Schaaf et al (2009). In order to figure out how cohesin affects different functions in the cell,
it is fundamentally important to better our existing understanding of the stoichiometry of the cohesin complex. In this regard, two groups have shown that there exists a 1:1:1:1 stoichiometry between the core cohesin subunits SMC1: SMC3: RAD21:STAG1/2 (Ding et al, 2011; Holzmann et al, 2011). However, as most of the investigations analysing the transcriptional functions of cohesin have relied on the depletion of one of its subunits using RNAi, we are still far from understanding how knockdown of one of the cohesin subunits might affect the stoichiometric integrity/localisation/functions of the rest of the cohesin partners. In this regard, intriguing results were published by Laugsch and colleagues, showing highly distinct phenotypes corresponding to RNAi-mediated depletion of either one of the heterodimeric partners, SMC1 or SMC3 (Laugsch et al, 2013). In their study, knockdown of SMC1 led to the cytoplasmic mislocalisation of SMC3 (demonstrating poor chromatin association) alongwith degradation of RAD21, whereas knockdown of SMC3 caused degradation of both SMC1 as well as RAD21. These results showed a highly distinct SMC protein-specific phenotype that drew attention to the importance of a 1:1 balance between the heterodimeric partners. Results from this study thus challenge the existing notion of deriving inferences on specific biological effects based on the removal of one of the cohesin subunits. Future studies may thus warrant the stoichiometric reduction of the individual cohesin subunits in parallel, followed by comparison of the respective phenotypes, for drawing biologically meaningful conclusions in the context of cohesin-affected human disorders and pathologies.
Figure 6.1

- Plasma Membrane
- Nucleus
- Estradiol
- Estrogen Receptor
  - ER
  - ER

**Interphase**
- Non-cohesive Cohesin
- Non-cohesive pool of Interphase cohesin unaffected by PCI-34051 treatment
- One pool (30%) of Cohesin becomes cohesive for SSC

**Cell Cycle**
- Separase
- Metaphase

**Stages**
- Anaphase Pathway
- Cohesin removal
- Blockade of HDAC3 deacetylation
- PCI-34051

**Gene Expression**
- Cohesin-dependent estrogen-responsive gene expression unaffected
Figure 6.1 Model of how oestrogenic transcriptional responses of cohesin-dependent genes in MCF7 cells might be maintained despite accumulating ac-SMC3 upon PCI-34051 treatment

Upon completion of one round of the cell cycle, treatment with PCI-34051 would inhibit deacetylation of now dissociated (via prophase and anaphase pathway of cohesin removal) stably-bound pool of mitotic cohesin (30% of total cohesin) resulting in accumulation of non-recycled cohesin. However, dynamic- and transient-binding pools of interphase cohesin, together with the reservoir of soluble cohesin might be enough to carry on with the transcriptional functions of cohesin for the successive (2 or 3) cell cycles. In such a situation, upon oestrogenic stimulation functional cohesin would continue to regulate expression of oestrogen-responsive cohesin-dependent genes.

Table 6.1 Comparison of the consequences of siRNA depletion of RAD21 or blockade of SMC3 deacetylation by PCI-34051 in MCF7 cells

<table>
<thead>
<tr>
<th></th>
<th>siRNA depletion of RAD21</th>
<th>Blockade of SMC3 deacetylation by PCI-34051</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oestrogenic response of cohesin-dependent genes</td>
<td>Lost</td>
<td>Retained</td>
</tr>
<tr>
<td>Morphology</td>
<td>Normal relative to control cells</td>
<td>Abnormal</td>
</tr>
<tr>
<td>Cell cycle progression</td>
<td>Minor delay</td>
<td>Major dose-dependent delay</td>
</tr>
<tr>
<td>Cell proliferation</td>
<td>Unaffected</td>
<td>Suppressed</td>
</tr>
<tr>
<td>Cell death and cytotoxicity</td>
<td>Not present</td>
<td>Relatively high</td>
</tr>
<tr>
<td>ERα transcript and protein levels</td>
<td>Unaffected</td>
<td>Unaffected</td>
</tr>
<tr>
<td>Genomic binding of RAD21 to candidate genes SOX4 and IL20</td>
<td>Decreased</td>
<td>Not known as yet</td>
</tr>
<tr>
<td>Genomic binding of ERα and RNA Polymerase II to candidate genes SOX4 and IL20</td>
<td>Altered</td>
<td>Not known as yet</td>
</tr>
</tbody>
</table>

6.6 PCI-34051 delayed progression of cell cycle, suppressed proliferation, and induced cell death in MCF7 cells

After uncovering that PCI-34051 did not induce the same gene expression changes as RAD21 depletion in MCF7 cells, functional assays were performed to better characterise the activity of this inhibitor. Abrogation of HDAC8 in human cell lines of different origin (HeLa-cervical, A549-lung, HCT116-colorectal carcinoma cell lines) was found to significantly inhibit cell proliferation highlighting the importance of HDAC8 in tumour cell proliferation (Vannini et al, 2004). Another study in cultured neuroblastoma cell lines showed consistent growth-hindering properties (Oehme et al, 2009) in addition to inducing differentiation and cell cycle arrest in response to HDAC8 knockdown/treatment with a
selective inhibitor. These findings suggested that HDAC8 may have an important role in the physiology of solid tumours, and inhibition of its function compromises cell survival. To characterise if HDAC8 has similar roles in breast cancer, the effect of HDAC8 inhibition (by using the chemical inhibitor PCI-34051) on cell cycle progression, growth, proliferation and viability was assessed in MCF7 cells. As discussed in sections 5.3.4, 5.3.5 and 5.3.6, PCI-34051 delayed progression of the cell cycle, reduced the rate of cell proliferation and induced cell death. However, it is not known whether cell death induced by PCI-34051 in MCF7 cells is a consequence of inhibition of HDAC8 activity or caused by cytotoxicity triggered by the chemical inhibitor. One speculation for the significant loss of cell viability could be that over time, on one hand cell proliferation is hampered due to scarcity of functional cohesin and on the other, limited cohesin availability leads to incompetent DNA damage repair. Compromised DSB repair then subsequently increases the cellular sensitivity of MCF7 cells to cytotoxicity induced by PCI-34051. Alternatively, it could also be possible that the oncostatic effects seen in MCF7 cells, is atleast in part, mediated by p21 overexpression through the PCI-34051-mediated blockade of deacetylation of the newly discovered substrates of HDAC8 (ARID1A, RAI1 and MLL2) (discussed in section 5.4). Although the definitive mechanism of cell death induced by PCI-34051 remains enigmatic; however, my data is consistent with data published by Balasubramanian and group in 2008, showing that apoptotic hallmarks could not be detected in cells treated with PCI-34051 (discussed in section 5.3.7, Figure 5.9).

In summary, this study indicates that depletion of RAD21 affected transcriptional response to oestrogen for a considerable (12-22%) proportion of oestrogen-responsive genes, some of which are enriched in signalling pathways associated with breast cancer cell proliferation and survival (PI3K/mTOR, ErbB1). Considering overexpression of RAD21 is associated with poor prognosis and treatment resistance, results from this study support the idea that RAD21 through its regulation of a subset of ER target genes may contribute, at least in part, to the poor prognostic profile of ER-positive breast cancer. Cohesin expression may thus be functionally important for tumour progression in this class of breast cancer. Future studies are therefore warranted to further explore whether cohesin function can be therapeutically curtailed to antagonise progression of ER-positive breast tumours.

While investigating the cohesin-targeting efficacy of the HDAC8 selective inhibitor PCI-34051, it emerged that the inhibitor could not target the transcription-specific role of cohesin
pertaining to oestrogen signalling in MCF7 cells. However, it led to significant concentration-dependent cell death. Although inhibition of proliferation and rise in cell death in PCI-34051-treated MCF7 cells seems exciting, the underlying mechanism is not very well understood. Therefore, keeping in mind the possibility of generation of potential pleiotropic effects in response to treatment with PCI-34051, consideration of this inhibitor in the context of breast cancers with RAD21 overexpression should be approached with caution. A thorough understanding of the molecular mechanisms underlying PCI-34051-induced cell death (p21 upregulation being one) would thus be pivotal, before considering it as an anti-cancer agent for this class of breast cancer.

6.7 Future Directions

Results from my studies have provided evidence that cohesin is an important contributor to the transcriptional regulation of oestrogen-dependent genes in breast cancer. A subset of oestrogen-responsive genes affected by RAD21 depletion participates in breast cancer signalling networks. Using a candidate gene approach, I showed that RAD21 binding at regulatory sites of SOX4 and IL20 is significantly reduced upon RAD21 depletion, and loss of RAD21 impaired binding of ERα and RNA Polymerase II at some of these sites as well. One of the future directions of this study could be to perform ChIP-seq in RAD21-depleted cells, and analyse genome-wide changes in the pattern and level of RAD21 and ERα binding by comparing profiles to previously identified localisation sites of these proteins (Schmidt et al, 2010). RAD21 binding would be naturally expected to reduce in RAD21-depleted cells. However, evaluating the precise proportion/percentage of decrease in RAD21 binding from previously identified sites would help to determine cohesin’s impact on modulation of ERα binding levels. Binding profiles of RNA Polymerase II and histone modifications (representative of active and repressive chromatin, as discussed in section 4.4) could also be additionally mapped to delineate changes in chromatin environ conditional on cohesin loss. Existing gene expression data from the microarray could then be meaningfully correlated to the pattern of protein binding and chromatin environment, adding a lot more clarity to cohesin’s precise role (direct versus indirect) in controlling transcriptional machinery of oestrogen-responsive genes in the MCF7 model system.

For exploring the relationship between cohesin and enhancer-promoter interactions, pertinent to cohesin’s role in chromatin configuration for influencing transcription, 3C could be performed for select candidate gene loci (e.g SOX4, see the model presented in Figure
4.7) or Hi-C could be attempted to examine genome-wide chromatin contacts. Hi-C is a genome-wide derivative of 3C and allows for unbiased, high-resolution mapping of physical chromatin interactions on a genome-wide scale (Belton et al, 2012). The model presented in section 4.4 (see Figure 4.7) hypothesises that functional chromatin domains of cohesin-dependent oestrogen-responsive genes are disrupted upon cohesin depletion, and the Hi-C technique could test this hypothesis. The Hi-C technique could also help to simultaneously identify other interacting loci, including those gene loci that we do not currently have a prior knowledge of. This would be valuable for adding further evidence to our current understanding of cohesin-mediated interactions, which appears to be strictly gene locus- and context-specific rather being generalised across the genome. Current data suggest that although cohesin can mediate long-range chromosomal interactions, this may not be needed for all cohesin-regulated genes (Antony et al, 2014; Lavagnolli et al, 2015). It would be reassuring to see additional evidence for this.

It would be worthwhile to map localisation sites for RAD21, ac-SMC3 and ERα in MCF7 cells having accumulated ac-SMC3 upon PCI-34051 treatment. This could be done by performing ChIP followed by qPCR, for instance, at the SOX4 and IL20 gene loci. Comparing the binding patterns to those obtained on a background of depleted RAD21 (existing data, Figures 4.5, 4.6) would help to understand why genes continue to respond to oestrogen in PCI-34051-treated MCF7 cells. Another interesting experiment would be to consider performing metaphase spreads for analysing chromosome cohesion in PCI-34051-treated cells. This would determine whether PCI-34051 affects chromosome cohesion as opposed to transcription in MCF7 cells, and would partly explain the disparity in transcriptional response observed with the two approaches.

Furthermore, in order to comprehensively bypass some of the inherent limitations associated with RNAi or HDAC8i PCI-34051-mediated abrogation of cohesin, the modern, unbiased, sequence-specific genome-engineering tool such as the RNA-guided DNA targeting system, the CRISPR-Cas9 system could be used (Mali et al, 2013). For a subunit-specific cohesin inactivation, the CRISPR-Cas9 gene knockout system could be a good alternate strategy to RNAi/chemical inhibitor based cohesin loss-of-function analysis. Building on the current body of work, such an approach could serve as a useful control for analysing RAD21 versus SMC3 gene knockout effects. This technology could also be used to interrogate the functional relevance of chromosomal loop anchors that are usually held by cohesin, together
with CTCF and mediator proteins (Baranello et al, 2014; Seitan et al, 2013; Zuin et al, 2014). As an example, the significance of cohesin (RAD21) binding at the speculated activating enhancer-promoter loop at the IL20 promoter and the repressive loop upstream of the SOX4 gene (as discussed in section 4.4, Figure 4.7) could be investigated by deleting the cohesin (RAD21) binding sites at these two gene loci using the CRISPR system. This principle could also be extended to interrogate the genome-wide implications of cohesin binding (together with CTCF) in chromosomal boundary function delineation in MCF7 cells, as has been investigated for CTCF in a previous study (Dowen et al, 2014).

In terms of treatment of ER-positive breast cancers using PCI-34051 (either as a monotherapeutic or as part of combinatorial approaches), complete understanding of the underlying molecular mechanism of its activity is crucial. This would warrant thorough functional characterisation (of its on and off-target effects) in additional ER-positive model systems.

Finally, although findings from this study have been greatly insightful, repeating this work in additional ER-positive breast cancer cell lines/breast tumour tissues could be a good future direction, so as to expand the generality of these results for paving a way towards cohesin-based therapeutic interventions.
References


Beatson G (1896) ON THE TREATMENT OF INOPERABLE CASES OF CARCINOMA OF THE MAMMA: SUGGESTIONS FOR A NEW METHOD OF TREATMENT, WITH ILLUSTRATIVE CASES.1. The Lancet **148**: 104-107


Cirillo LA, Zaret KS (1999) An early developmental transcription factor complex that is more stable on nucleosome core particles than on free DNA. Mol Cell 4: 961-969


mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. *Nature* **489**: 313-317


Hamblers IH, Van Schaik RF, Sussenbach JS, Steenbergh PH (2003) 17beta-Estradiol responsiveness of MCF-7 laboratory strains is dependent on an autocrine signal activating the IGF type I receptor. Cancer Cell Int 3: 10


receptor in breast cancer proliferation and mammary gland development. *PLoS One* **6**: e0017830


Kim ST, Xu B, Kastan MB (2002b) Involvement of the cohesin protein, Smc1, in Atm-dependent and independent responses to DNA damage. *Genes Dev* **16**: 560-570


chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev* **28**: 2151-2162


Cohesin is required for expression of the oestrogen receptor-alpha (ESR1) gene. *Epigenetics Chromatin* **5**: 1756-8935

Dynamic oestrogen receptor interactomes control oestrogen-responsive trefoil Factor (TFF) locus cell-specific activities. *Mol Cell Biol* **34**: 2418-2436

GreB 1 is a critical regulator of hormone dependent breast cancer growth. *Breast Cancer Res Treat* **92**: 141-149

ERs associate with and regulate the production of caveolin: implications for signaling and cellular actions. *Mol Endocrinol* **16**: 100-115

Cell membrane and nuclear oestrogen receptors (ERs) originate from a single transcript: studies of ERAlpha and ERGBeta expressed in Chinese hamster ovary cells. *Mol Endocrinol* **13**: 307-319

Proximal events in signaling by plasma membrane oestrogen receptors. *J Biol Chem* **278**: 2701-2712

Marking time: the dynamic role of chromatin and covalent modification in transcription. *Int J Biochem Cell Biol* **41**: 155-163


Molecular profiling: moving away from tumor philately. *Sci Transl Med* **2**: 3001329

Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres. *Embo J* **31**: 2076-2089

Cohesin in development and disease. *Development* **140**: 3715-3718


Stehelin D, Varmus HE, Bishop JM, Vogt PK (1976) DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. Nature 260: 170-173


Welboren WJ, Sweep FC, Span PN, Stunnenberg HG (2009a) Genomic actions of oestrogen receptor alpha: what are the targets and how are they regulated? Endocr Relat Cancer 16: 1073-1089


Appendix A

Appendix A. Transcript levels of reference genes PPIA and GAPDH in MCF7 cells following transfection with Control or RAD21 siRNA

Oestrogen-deprived MCF7 cells were transfected with 5 nM of non-targeting control or RAD21 siRNA for 48 hours, then treated with vehicle or 100 nM of 17-β-oestradiol for 3, 6 and 24 hours. Transcript levels were quantitated by qRT-PCR. Depletion of RAD21 did not alter expression levels of the reference genes used to quantitate differential transcript levels. Error bars represent the SEM of three to five independent experiments.
Appendix B

Appendix B. Electropherograms from Bioanalyzer profiling of samples from the 6 hours time point

RNAs extracted from vehicle/oestrogen-treated, control/RAD21-depleted MCF7 cells were assessed for quality and integrity by Bioanalyzer profiling. RNAs were size-separated by voltage and detected by laser-induced fluorescence. The RIN algorithm calculated integrity by taking into account, the total electrophoretic separation of the RNA samples. High-quality intact RNAs were represented by distinct peaks of their corresponding 18s and 28s ribosomal components and also by their <1.8 28s/18s ratios (see Table 3.1).
Appendix C

A.

3 HOUR TRANSCRIPTS

RAD21 siRNA.E
Biological Replicates

Control siRNA.E
Biological Replicates

Adjusted p value <0.005
(Transcripts: 1 to 50)
B.

3 HOUR TRANSCRIPTS

Adjusted p value < 0.005

(Transcripts: 51 to 109)
D.

**3 HOUR TRANSCRIPTS**

![Heatmap Image](image)

*Adjusted p value <0.005 (Transcripts: 151 to 218)*

**RAD21 siRNA.E**

- Biological Replicates: 1, 2, 3

**Control siRNA.E**

- Biological Replicates: 1, 2, 3
Appendix C. Heatmaps from microarray analysis showing differentially expressed transcripts of control versus RAD21-depleted MCF7 cells treated with 3 or 6 hours of 100 nM of 17-β-oestradiol

(A-D) Heatmaps of genes differentially expressed at 3 the hour time point post oestradiol treatment with adjusted p-value <0.005 (A) Transcripts 1 to 50; (B) Transcripts 51-100; (C) Transcripts 101-150; (D) Transcripts 151-218. (E) Heatmap of genes differentially expressed at the 6 hour time point post oestradiol treatment with adjusted p-value <0.001.
Appendix D. Differentially expressed genes at the 3 hour time point

I. Genes activated by oestradiol but repressed by RAD21 knockdown

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCB10</td>
<td>0.006722116</td>
<td>-1.248704</td>
</tr>
<tr>
<td>ABCC1</td>
<td>0.135429</td>
<td>-1.036484</td>
</tr>
<tr>
<td>ADRA2A</td>
<td>1.553493</td>
<td>-1.963922</td>
</tr>
<tr>
<td>AFAP1L2</td>
<td>0.496375</td>
<td>-1.213463</td>
</tr>
<tr>
<td>ANKRD6</td>
<td>0.4241579</td>
<td>-0.5855179</td>
</tr>
<tr>
<td>CYP24A1</td>
<td>0.735844</td>
<td>-0.7929069</td>
</tr>
<tr>
<td>DBF4</td>
<td>0.1877524</td>
<td>-0.9860047</td>
</tr>
<tr>
<td>DENND5B</td>
<td>0.8032585</td>
<td>-1.405617</td>
</tr>
<tr>
<td>FEM1A</td>
<td>0.05833824</td>
<td>-0.9557717</td>
</tr>
<tr>
<td>FGFBP2</td>
<td>1.664813</td>
<td>1.641386</td>
</tr>
<tr>
<td>FMR1</td>
<td>1.395121</td>
<td>1.226514</td>
</tr>
<tr>
<td>FYCO1</td>
<td>0.01866587</td>
<td>-0.7012802</td>
</tr>
<tr>
<td>GAB2</td>
<td>0.9894096</td>
<td>-0.790672</td>
</tr>
<tr>
<td>IL20</td>
<td>1.810717</td>
<td>-2.226669</td>
</tr>
<tr>
<td>KCNK6</td>
<td>1.880904</td>
<td>0.8671912</td>
</tr>
<tr>
<td>LYAR</td>
<td>0.3198165</td>
<td>-0.8652976</td>
</tr>
<tr>
<td>MMP9</td>
<td>0.07124668</td>
<td>-1.397175</td>
</tr>
<tr>
<td>MYC</td>
<td>0.9089035</td>
<td>-0.8484117</td>
</tr>
<tr>
<td>NRCAM</td>
<td>0.07232265</td>
<td>-1.009462</td>
</tr>
<tr>
<td>NRXN3</td>
<td>0.1186084</td>
<td>-1.332161</td>
</tr>
<tr>
<td>PDSS1</td>
<td>0.1644583</td>
<td>-0.788353</td>
</tr>
<tr>
<td>PRTG</td>
<td>0.230315</td>
<td>-0.6640307</td>
</tr>
<tr>
<td>RAD21</td>
<td>0.06503017</td>
<td>-2.706914</td>
</tr>
<tr>
<td>RASSF8</td>
<td>0.004443065</td>
<td>-0.7323298</td>
</tr>
<tr>
<td>RBM24</td>
<td>0.959047</td>
<td>-1.694153</td>
</tr>
<tr>
<td>RNF152</td>
<td>0.03756132</td>
<td>-1.14579</td>
</tr>
<tr>
<td>RINGT</td>
<td>0.05140399</td>
<td>-0.7387052</td>
</tr>
<tr>
<td>SCLY</td>
<td>0.1629167</td>
<td>-1.106398</td>
</tr>
<tr>
<td>SELT</td>
<td>0.1937063</td>
<td>-0.741176</td>
</tr>
<tr>
<td>SERBP1</td>
<td>0.06172825</td>
<td>-0.9622356</td>
</tr>
<tr>
<td>SLC22A5</td>
<td>0.8754084</td>
<td>0.8102008</td>
</tr>
<tr>
<td>SLC25A30</td>
<td>0.01399165</td>
<td>-1.086615</td>
</tr>
<tr>
<td>SMN1</td>
<td>1.883652</td>
<td>1.369825</td>
</tr>
<tr>
<td>SPCS3</td>
<td>0.1756354</td>
<td>-0.795414</td>
</tr>
<tr>
<td>SPG20</td>
<td>0.001297441</td>
<td>-0.886185</td>
</tr>
<tr>
<td>STC1</td>
<td>0.04595645</td>
<td>-1.674343</td>
</tr>
<tr>
<td>SUSD4</td>
<td>0.2590951</td>
<td>-1.268916</td>
</tr>
<tr>
<td>TAF5L</td>
<td>0.264946</td>
<td>-1.26309</td>
</tr>
<tr>
<td>THBS1</td>
<td>1.402654</td>
<td>-2.271033</td>
</tr>
<tr>
<td>TMEM64</td>
<td>0.1718111</td>
<td>-1.026575</td>
</tr>
<tr>
<td>TRIM8</td>
<td>0.11107253</td>
<td>-0.893768</td>
</tr>
<tr>
<td>VASN</td>
<td>1.309232</td>
<td>1.271569</td>
</tr>
</tbody>
</table>
### II. Genes repressed by oestradiol but de-repressed by RAD21 knockdown

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C11orf75</td>
<td>-0.10798</td>
<td>0.937538</td>
</tr>
<tr>
<td>CLCN4</td>
<td>-0.05041746</td>
<td>1.092938</td>
</tr>
<tr>
<td>CSRP2</td>
<td>-0.2583986</td>
<td>1.014068</td>
</tr>
<tr>
<td>CYP1B1</td>
<td>-0.1188178</td>
<td>1.143478</td>
</tr>
<tr>
<td>GRAMD1C</td>
<td>-0.1835065</td>
<td>1.079258</td>
</tr>
<tr>
<td>IDH1</td>
<td>-0.1283513</td>
<td>0.8537367</td>
</tr>
<tr>
<td>IGSF5</td>
<td>-0.3559875</td>
<td>1.228061</td>
</tr>
<tr>
<td>MAP4K4</td>
<td>-0.1566738</td>
<td>0.952193</td>
</tr>
<tr>
<td>METTL7A</td>
<td>-0.1913541</td>
<td>1.051026</td>
</tr>
<tr>
<td>MR1</td>
<td>-0.1818423</td>
<td>1.134833</td>
</tr>
<tr>
<td>MSRB3</td>
<td>-0.5023226</td>
<td>1.334027</td>
</tr>
<tr>
<td>PAM</td>
<td>-0.05890297</td>
<td>0.9080108</td>
</tr>
<tr>
<td>PCDHB14</td>
<td>-0.2150384</td>
<td>1.419692</td>
</tr>
<tr>
<td>PCDHB16</td>
<td>-0.1912991</td>
<td>1.10964</td>
</tr>
<tr>
<td>PCDH8B3</td>
<td>-0.2956034</td>
<td>0.9246768</td>
</tr>
<tr>
<td>PIK3R3</td>
<td>-0.8280526</td>
<td>0.9089381</td>
</tr>
<tr>
<td>RGNEF</td>
<td>-0.04352762</td>
<td>0.8289012</td>
</tr>
<tr>
<td>RPL31</td>
<td>-0.1115596</td>
<td>0.9480966</td>
</tr>
<tr>
<td>SWAP70</td>
<td>-1.024725</td>
<td>1.020991</td>
</tr>
<tr>
<td>UBE2E3</td>
<td>-0.1428595</td>
<td>0.6767747</td>
</tr>
<tr>
<td>UTRN</td>
<td>-0.216511</td>
<td>1.334515</td>
</tr>
</tbody>
</table>
III. Genes activated by oestradiol and further upregulated by RAD21 knockdown

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>APOA1</td>
<td>0.4706476</td>
<td>1.518329</td>
</tr>
<tr>
<td>B4GALT1</td>
<td>0.7676265</td>
<td>0.8425373</td>
</tr>
<tr>
<td>LINC00341</td>
<td>1.105054</td>
<td>1.821377</td>
</tr>
<tr>
<td>C9orf95</td>
<td>0.6126757</td>
<td>0.8396569</td>
</tr>
<tr>
<td>CACNA1I</td>
<td>0.2995278</td>
<td>0.8829806</td>
</tr>
<tr>
<td>CASP7</td>
<td>0.0577794</td>
<td>1.108593</td>
</tr>
<tr>
<td>CBFA2T3</td>
<td>0.4487255</td>
<td>0.7230302</td>
</tr>
<tr>
<td>CELSR2</td>
<td>0.7059313</td>
<td>1.354819</td>
</tr>
<tr>
<td>CHST8</td>
<td>1.051885</td>
<td>1.111126</td>
</tr>
<tr>
<td>DOK7</td>
<td>0.2276471</td>
<td>0.8366411</td>
</tr>
<tr>
<td>FAM63A</td>
<td>0.2984808</td>
<td>0.7135791</td>
</tr>
<tr>
<td>FHDC1</td>
<td>0.2208408</td>
<td>0.9631203</td>
</tr>
<tr>
<td>FOSL2</td>
<td>0.5315235</td>
<td>0.9848575</td>
</tr>
<tr>
<td>FREM2</td>
<td>0.116067</td>
<td>0.9361321</td>
</tr>
<tr>
<td>GATSL3</td>
<td>0.398409</td>
<td>1.3469</td>
</tr>
<tr>
<td>GDPD1</td>
<td>0.04869935</td>
<td>1.024916</td>
</tr>
<tr>
<td>IL17RB</td>
<td>0.141736</td>
<td>1.027365</td>
</tr>
<tr>
<td>KIF21A</td>
<td>0.2755946</td>
<td>0.759118</td>
</tr>
<tr>
<td>KLF9</td>
<td>0.1471309</td>
<td>0.9002208</td>
</tr>
<tr>
<td>LRG1</td>
<td>0.1342415</td>
<td>0.8208717</td>
</tr>
<tr>
<td>MAML3D</td>
<td>0.8017642</td>
<td>0.8280011</td>
</tr>
<tr>
<td>OSTF1</td>
<td>0.3161927</td>
<td>0.7435348</td>
</tr>
<tr>
<td>PCDHB13</td>
<td>0.008845581</td>
<td>0.9911355</td>
</tr>
<tr>
<td>PDZK1</td>
<td>0.5465602</td>
<td>1.822205</td>
</tr>
<tr>
<td>PSKH1</td>
<td>0.1354299</td>
<td>1.031479</td>
</tr>
<tr>
<td>SERPINB9</td>
<td>1.30266</td>
<td>1.921402</td>
</tr>
<tr>
<td>SERPINF1</td>
<td>0.04484977</td>
<td>1.286993</td>
</tr>
<tr>
<td>SIDT2</td>
<td>0.3023343</td>
<td>0.6769137</td>
</tr>
<tr>
<td>SLC6A6</td>
<td>0.2744938</td>
<td>0.8132419</td>
</tr>
<tr>
<td>SNX24</td>
<td>0.5403931</td>
<td>0.9750667</td>
</tr>
<tr>
<td>SOX4</td>
<td>0.6272081</td>
<td>0.7783175</td>
</tr>
<tr>
<td>SPATS2L</td>
<td>0.5414724</td>
<td>0.9854646</td>
</tr>
<tr>
<td>SYNPO2L</td>
<td>0.796826</td>
<td>1.25521</td>
</tr>
<tr>
<td>SYTL5</td>
<td>0.7105045</td>
<td>1.397373</td>
</tr>
<tr>
<td>TLN2</td>
<td>0.2289171</td>
<td>1.021306</td>
</tr>
<tr>
<td>TMPRSS3</td>
<td>0.8834247</td>
<td>1.114383</td>
</tr>
<tr>
<td>VWF</td>
<td>0.768808</td>
<td>1.540244</td>
</tr>
<tr>
<td>ZFP36L1</td>
<td>0.07645114</td>
<td>0.6364939</td>
</tr>
<tr>
<td>ZNF185</td>
<td>0.3639288</td>
<td>1.177212</td>
</tr>
</tbody>
</table>
IV. Genes repressed by oestradiol and further repressed by RAD21 knockdown

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AKAP8</td>
<td>-0.300051</td>
<td>-0.8186504</td>
</tr>
<tr>
<td>AMI02</td>
<td>-0.8676876</td>
<td>-0.963489</td>
</tr>
<tr>
<td>ATP6V1C1</td>
<td>-0.05796468</td>
<td>-0.7863551</td>
</tr>
<tr>
<td>ATP8A1</td>
<td>-0.1367624</td>
<td>-0.8400338</td>
</tr>
<tr>
<td>BAGALT6</td>
<td>-0.08816636</td>
<td>-0.8904561</td>
</tr>
<tr>
<td>C4orf32</td>
<td>-0.1419186</td>
<td>-0.8438435</td>
</tr>
<tr>
<td>CAPZA1</td>
<td>-0.000192968</td>
<td>-0.8700503</td>
</tr>
<tr>
<td>CCKG1</td>
<td>-0.1531419</td>
<td>-1.588532</td>
</tr>
<tr>
<td>CDKN3</td>
<td>-0.09025574</td>
<td>-0.8060759</td>
</tr>
<tr>
<td>CENPQ</td>
<td>-0.1649475</td>
<td>-0.8895469</td>
</tr>
<tr>
<td>CLDN1</td>
<td>-0.1733701</td>
<td>-0.8501864</td>
</tr>
<tr>
<td>CCP110</td>
<td>-0.01296913</td>
<td>-0.7395859</td>
</tr>
<tr>
<td>CXC4</td>
<td>-0.1148074</td>
<td>-1.654622</td>
</tr>
<tr>
<td>DKK1</td>
<td>-0.815863</td>
<td>-1.472408</td>
</tr>
<tr>
<td>EFEMP1</td>
<td>-0.04762954</td>
<td>-1.628294</td>
</tr>
<tr>
<td>FAM171B</td>
<td>-0.409179</td>
<td>-1.385027</td>
</tr>
<tr>
<td>FAM20C</td>
<td>-0.2907603</td>
<td>-1.265996</td>
</tr>
<tr>
<td>FRMD6</td>
<td>-0.7552282</td>
<td>-0.8016427</td>
</tr>
<tr>
<td>GATA5</td>
<td>-0.7785428</td>
<td>-0.6902632</td>
</tr>
<tr>
<td>HSP90AA1</td>
<td>-0.04174932</td>
<td>-1.06709</td>
</tr>
<tr>
<td>IRS1</td>
<td>-0.2345978</td>
<td>-1.216147</td>
</tr>
<tr>
<td>IRX3</td>
<td>-0.1599901</td>
<td>-0.7116723</td>
</tr>
<tr>
<td>ITPRIPL2</td>
<td>-0.8314561</td>
<td>-0.975207</td>
</tr>
<tr>
<td>KITLG</td>
<td>-0.6330535</td>
<td>-0.839995</td>
</tr>
<tr>
<td>KNUV</td>
<td>-0.1986026</td>
<td>-0.7603519</td>
</tr>
<tr>
<td>LXN</td>
<td>-0.06880144</td>
<td>-1.560793</td>
</tr>
<tr>
<td>MMD</td>
<td>-0.66699</td>
<td>-1.295338</td>
</tr>
<tr>
<td>MTX3</td>
<td>-0.04957882</td>
<td>-0.8871936</td>
</tr>
<tr>
<td>NUDT4</td>
<td>-0.1219187</td>
<td>-1.701655</td>
</tr>
<tr>
<td>NUDT4P1</td>
<td>-0.153465</td>
<td>-1.883846</td>
</tr>
<tr>
<td>NUMBL</td>
<td>-0.1834098</td>
<td>-1.083011</td>
</tr>
<tr>
<td>PDCCD4</td>
<td>-0.04109919</td>
<td>-1.272385</td>
</tr>
<tr>
<td>PHF10</td>
<td>-0.04965022</td>
<td>-1.457634</td>
</tr>
<tr>
<td>PVRL2</td>
<td>-0.06722011</td>
<td>-1.354319</td>
</tr>
<tr>
<td>RPRM</td>
<td>-0.3256156</td>
<td>-1.696365</td>
</tr>
<tr>
<td>RPS6KA5</td>
<td>-0.1661258</td>
<td>-1.035104</td>
</tr>
<tr>
<td>RTN4</td>
<td>-0.004714775</td>
<td>-1.207813</td>
</tr>
<tr>
<td>RUNDCC3B</td>
<td>-0.1910299</td>
<td>-1.199729</td>
</tr>
<tr>
<td>S1PR3</td>
<td>-1.096446</td>
<td>-1.800269</td>
</tr>
<tr>
<td>SGCG</td>
<td>-0.2129794</td>
<td>-1.811148</td>
</tr>
<tr>
<td>SGMS1</td>
<td>-0.03947649</td>
<td>-1.344764</td>
</tr>
<tr>
<td>SMAD2</td>
<td>-0.05747803</td>
<td>-0.628095</td>
</tr>
</tbody>
</table>
Appendix D. Categorised lists of significantly (adjusted p-value <0.005) dysregulated genes of the 3 hour time point

The tables represent categories of microarray identified genes significantly (adjusted p-value cut-off <0.005) dysregulated upon RAD21 depletion in oestradiol-stimulated (RAD21.E) conditions in comparison to control-stimulated (Control.E) cells at the 3 hour time point. Expression changes are presented as fold change on a log2 scale. Benjamini-Hochberg method was used to correct the p-values for multiple testing. Column three represents log2FC of significantly dysregulated genes (adjusted p-value <0.005) obtained by comparing Control.E versus RAD21.E cells. Column two represents log2FC of the same genes obtained by comparing vehicle-treated control (Control.V) versus oestradiol-stimulated control (Control.E) cells.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SNRPA1</td>
<td>-0.005195817</td>
<td>-1.640003</td>
</tr>
<tr>
<td>ST8SIA4</td>
<td>-1.032166</td>
<td>-0.99553</td>
</tr>
<tr>
<td>SUMF1</td>
<td>-0.01811672</td>
<td>-1.144648</td>
</tr>
<tr>
<td>SYNPO2</td>
<td>-0.2017283</td>
<td>-2.044823</td>
</tr>
<tr>
<td>TPST2</td>
<td>-0.07544708</td>
<td>-0.798426</td>
</tr>
<tr>
<td>SYTL2</td>
<td>-0.4322197</td>
<td>-0.8269582</td>
</tr>
<tr>
<td>TACC1</td>
<td>-0.2991319</td>
<td>-1.290091</td>
</tr>
<tr>
<td>TRMT6</td>
<td>-0.06416272</td>
<td>-1.216575</td>
</tr>
<tr>
<td>TXNRD1</td>
<td>-0.1910572</td>
<td>-1.564066</td>
</tr>
</tbody>
</table>
Appendix E. Differentially expressed genes at the 6 hour time point

I. Genes activated by oestradiol but repressed by RAD21 depletion

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ACACB</td>
<td>0.04528536</td>
<td>-1.346346</td>
</tr>
<tr>
<td>ADORA2B</td>
<td>0.461313</td>
<td>-0.9767842</td>
</tr>
<tr>
<td>ADRA2A</td>
<td>1.705066</td>
<td>-1.684131</td>
</tr>
<tr>
<td>BAG1</td>
<td>0.4127149</td>
<td>-0.7766677</td>
</tr>
<tr>
<td>C14orf182</td>
<td>3.139943</td>
<td>1.707877</td>
</tr>
<tr>
<td>CELSR2</td>
<td>1.193262</td>
<td>1.087758</td>
</tr>
<tr>
<td>FEM1A</td>
<td>0.1561891</td>
<td>-0.8943885</td>
</tr>
<tr>
<td>FGFBP2</td>
<td>2.777325</td>
<td>1.524184</td>
</tr>
<tr>
<td>FMN1</td>
<td>1.654229</td>
<td>1.270497</td>
</tr>
<tr>
<td>IL20</td>
<td>2.230501</td>
<td>-2.607999</td>
</tr>
<tr>
<td>INSR</td>
<td>0.6064472</td>
<td>-0.9916648</td>
</tr>
<tr>
<td>KCNK6</td>
<td>1.549113</td>
<td>0.8465848</td>
</tr>
<tr>
<td>LRRC3</td>
<td>0.1687778</td>
<td>-1.051362</td>
</tr>
<tr>
<td>MED8</td>
<td>0.07120226</td>
<td>-0.9067429</td>
</tr>
<tr>
<td>NAT8L</td>
<td>0.7845537</td>
<td>-1.14758</td>
</tr>
<tr>
<td>NRXN3</td>
<td>0.04523663</td>
<td>-1.561702</td>
</tr>
<tr>
<td>MIR4680</td>
<td>0.1773167</td>
<td>-1.662629</td>
</tr>
<tr>
<td>PIM2</td>
<td>0.4566174</td>
<td>-0.9099945</td>
</tr>
<tr>
<td>RASSF8</td>
<td>0.2937465</td>
<td>-0.9264572</td>
</tr>
<tr>
<td>RBM24</td>
<td>1.190349</td>
<td>-1.238233</td>
</tr>
<tr>
<td>SNRPA1</td>
<td>0.1864441</td>
<td>-1.879401</td>
</tr>
<tr>
<td>SYTL5</td>
<td>1.799412</td>
<td>1.111749</td>
</tr>
<tr>
<td>THBS1</td>
<td>1.769204</td>
<td>-1.908007</td>
</tr>
<tr>
<td>TMEM64</td>
<td>0.2572414</td>
<td>-1.212098</td>
</tr>
<tr>
<td>TRMT6</td>
<td>0.2684299</td>
<td>-1.471756</td>
</tr>
</tbody>
</table>
II. Genes repressed by oestradiol but de-repressed by RAD21 knockdown

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BMPR2</td>
<td>-0.2912326</td>
<td>0.8087214</td>
</tr>
<tr>
<td>CYP1B1</td>
<td>-0.5736935</td>
<td>1.083446</td>
</tr>
<tr>
<td>DPY19L1</td>
<td>-0.1316039</td>
<td>0.8167661</td>
</tr>
<tr>
<td>GPR87</td>
<td>-0.4859768</td>
<td>1.891309</td>
</tr>
<tr>
<td>IDH1</td>
<td>-0.2945893</td>
<td>0.893639</td>
</tr>
<tr>
<td>MAP4K4</td>
<td>-0.1974471</td>
<td>0.9592062</td>
</tr>
<tr>
<td>METTL7A</td>
<td>-0.5905114</td>
<td>1.045458</td>
</tr>
<tr>
<td>PAK1</td>
<td>-0.02843225</td>
<td>1.255965</td>
</tr>
<tr>
<td>PCDHB10</td>
<td>-0.3597716</td>
<td>1.338604</td>
</tr>
<tr>
<td>PCDHB14</td>
<td>-0.3481348</td>
<td>1.514998</td>
</tr>
<tr>
<td>PCDHB16</td>
<td>-0.2873094</td>
<td>1.199895</td>
</tr>
<tr>
<td>PTPN21</td>
<td>-0.03773918</td>
<td>1.161039</td>
</tr>
<tr>
<td>RPL31</td>
<td>-0.3089176</td>
<td>1.229425</td>
</tr>
<tr>
<td>UTRN</td>
<td>-0.2282075</td>
<td>1.343746</td>
</tr>
<tr>
<td>ZNF117</td>
<td>-0.4644906</td>
<td>0.9852949</td>
</tr>
</tbody>
</table>

III. Genes activated by oestradiol and further upregulated by RAD21 knockdown

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CHST8</td>
<td>1.234353</td>
<td>1.460441</td>
</tr>
<tr>
<td>FREM2</td>
<td>0.1164459</td>
<td>1.006511</td>
</tr>
<tr>
<td>IL17RB</td>
<td>0.47702</td>
<td>1.110841</td>
</tr>
<tr>
<td>KIF21A</td>
<td>0.2890981</td>
<td>0.804269</td>
</tr>
<tr>
<td>LINC00341</td>
<td>0.616909</td>
<td>1.966462</td>
</tr>
<tr>
<td>GATSL3</td>
<td>0.1944399</td>
<td>1.66265</td>
</tr>
<tr>
<td>PLEKHH1</td>
<td>0.4355047</td>
<td>0.9210815</td>
</tr>
<tr>
<td>RNF144A</td>
<td>0.1362978</td>
<td>0.9217994</td>
</tr>
<tr>
<td>SAMD4A</td>
<td>0.4597087</td>
<td>1.486741</td>
</tr>
<tr>
<td>SERPINB9</td>
<td>0.5094082</td>
<td>2.349131</td>
</tr>
<tr>
<td>SPATS2L</td>
<td>1.026446</td>
<td>1.02149</td>
</tr>
<tr>
<td>SPRY2</td>
<td>0.1449207</td>
<td>0.8167387</td>
</tr>
<tr>
<td>SNX24</td>
<td>0.973464</td>
<td>1.003592</td>
</tr>
<tr>
<td>SOX4</td>
<td>0.7549291</td>
<td>0.9152076</td>
</tr>
<tr>
<td>TFF2</td>
<td>0.7950682</td>
<td>1.59449</td>
</tr>
<tr>
<td>VWF</td>
<td>1.022039</td>
<td>1.964601</td>
</tr>
<tr>
<td>ZNF185</td>
<td>0.5128296</td>
<td>1.252261</td>
</tr>
</tbody>
</table>
IV. Genes repressed by oestradiol and further repressed by RAD21 knockdown

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CCNG1</td>
<td>-0.4005639</td>
<td>-1.489069</td>
</tr>
<tr>
<td>CENPQ</td>
<td>-0.02298448</td>
<td>-0.8490369</td>
</tr>
<tr>
<td>DKK1</td>
<td>-0.8373635</td>
<td>-1.057509</td>
</tr>
<tr>
<td>EFEMP1</td>
<td>-0.6237884</td>
<td>-1.393859</td>
</tr>
<tr>
<td>HSP90AA1</td>
<td>-0.01877955</td>
<td>-1.134274</td>
</tr>
<tr>
<td>ITPR1L2</td>
<td>-0.9348347</td>
<td>-1.09125</td>
</tr>
<tr>
<td>LXN</td>
<td>-0.1865958</td>
<td>-1.406428</td>
</tr>
<tr>
<td>NUDT4P1</td>
<td>-0.07466145</td>
<td>-1.812237</td>
</tr>
<tr>
<td>NUDT4</td>
<td>-0.1669589</td>
<td>-1.346543</td>
</tr>
<tr>
<td>TMEM150C</td>
<td>-1.325283</td>
<td>-1.374861</td>
</tr>
<tr>
<td>NRCAM</td>
<td>-0.2394238</td>
<td>-1.036307</td>
</tr>
<tr>
<td>PCBP2</td>
<td>-0.177293</td>
<td>-0.7302281</td>
</tr>
<tr>
<td>PDCD4</td>
<td>-0.04838488</td>
<td>-1.506582</td>
</tr>
<tr>
<td>PVRL2</td>
<td>-0.3559753</td>
<td>-0.9325231</td>
</tr>
<tr>
<td>RAD21</td>
<td>-0.2533168</td>
<td>-2.692121</td>
</tr>
<tr>
<td>RTN4</td>
<td>-0.000505089</td>
<td>-1.140429</td>
</tr>
<tr>
<td>SGCG</td>
<td>-0.3994526</td>
<td>-1.64163</td>
</tr>
<tr>
<td>SUMF1</td>
<td>-0.07599363</td>
<td>-1.619865</td>
</tr>
<tr>
<td>SYNPO2</td>
<td>-0.6631919</td>
<td>-1.667412</td>
</tr>
<tr>
<td>TXNRD1</td>
<td>-0.214324</td>
<td>-1.4543</td>
</tr>
</tbody>
</table>

Appendix E. Categorised lists of significantly (adjusted p-value <0.005) dysregulated genes of the 6 hour time point

The tables represent categories of microarray identified genes significantly (adjusted p-value cut-off <0.005) dysregulated upon RAD21 depletion in oestradiol-stimulated (RAD21.E) conditions in comparison to control-stimulated (Control.E) cells at the 6 hour time point. Expression changes are presented as fold change on a log2 scale. Benjamini-Hochberg method was used to correct the p-values for multiple testing. Column three represents log2FC of significantly dysregulated genes (adjusted p-value cut-off <0.005) obtained by comparing Control.E versus RAD21.E cells. Column two represents log2FC of the same genes obtained by comparing vehicle-treated control (Control.V) versus oestradiol-stimulated control (Control.E) cells.
Appendix F

Functional annotation of dysregulated genes in oestrogen-stimulated RAD21-depleted MCF7 cells at the 6 hour time point

Upregulated genes:

<table>
<thead>
<tr>
<th>DAVID functional annotation term</th>
<th>Genes</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell-cell adhesion</td>
<td>FREM2, CELSR2, PCDHB10, PCDHB14, PCDHB16</td>
<td>2.4E-3</td>
</tr>
<tr>
<td>Extracellular structure organisation</td>
<td>PCDHB10, PCDHB14, PCDHB16, UTRN</td>
<td>4.1E-3</td>
</tr>
<tr>
<td>Developmental process</td>
<td>FREM2, SOX4, BMPR2, CELSR2, CHST8, IDH1, PCDHB10, PCDHB14, PCDHB16, SPRY2, UTRN, VWF</td>
<td>4.3E-2</td>
</tr>
</tbody>
</table>

Downregulated genes:

<table>
<thead>
<tr>
<th>DAVID’s functional annotation term</th>
<th>Genes</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regulation of protein kinase activity</td>
<td>ADRA2A, CCNG1, ADORA2B, INSR, IL20, PDCD4, THBS1</td>
<td>4.4E-4</td>
</tr>
<tr>
<td>Apoptosis</td>
<td>BAG1, PIM2, PDCD4, RTN4, THBS1</td>
<td>5.2E-3</td>
</tr>
<tr>
<td>Activation of MAPK activity</td>
<td>ADRA2A, ADORA2B, INSR, THBS1</td>
<td>5.0E-4</td>
</tr>
<tr>
<td>Positive regulation of leukocyte activation</td>
<td>ADORA2B, PVRL2, THBS1</td>
<td>1.7E-2</td>
</tr>
<tr>
<td>Positive regulation of phosphorylation</td>
<td>INSR, IL20, THBS1</td>
<td>1.5E-2</td>
</tr>
<tr>
<td>Regulation of cell migration</td>
<td>INSR, RTN4, THBS1</td>
<td>4.1E-2</td>
</tr>
</tbody>
</table>

Classification of dysregulated genes based on their molecular and biological functions. Functional clustering tool of NIH-DAVID (version 6.7) was used to identify enrichment of functionally related group of genes (p-value <0.005) in order to understand the biological and molecular processes regulated by RAD21 in oestrogen-stimulated MCF7 cells at the 6 hour time point.
## Appendix G

### I. IPA functional network association analysis of differentially regulated genes at the 3 hour time point

<table>
<thead>
<tr>
<th>Functional Network</th>
<th>Focus genes</th>
<th>Score</th>
<th>Focus genes in network</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Visual system development and function, Tissue development, Cardiovascular system development and function</strong></td>
<td>CAPZA1, CBF42T3, CHST8, C5SRP2, CYPS21, EFEMP1, FAM20C, KYNU, MAP4K4, RPRM, S1T2, SLCD646, SMAD2, SLPG20, ST8SIA4, STC1, SPRNO2, TACCC1, TACCC2, VASN, ZFP36L1</td>
<td>20</td>
<td>CAPZA1, CBF42T3, CHST8, C5SRP2, CYPS21, EFEMP1, FAM20C, KYNU, MAP4K4, RPRM, S1T2, SLCD646, SMAD2, SLPG20, ST8SIA4, STC1, SPRNO2, TACCC1, TACCC2, VASN, ZFP36L1</td>
</tr>
<tr>
<td><strong>Lipid metabolism, Small molecule biochemistry, Organismal injury and abnormalities</strong></td>
<td>ABCB10, APOA1, B4GALT1, B4GALT6, CYP24A1, DK1, GAB2, IL17RB, KITLG, MMD, MMP9, NRCAM, RF56K, SIRPFO1, SGCG, SGMS1, SLCD2A5</td>
<td>17</td>
<td>ABCB10, APOA1, B4GALT1, B4GALT6, CYP24A1, DK1, GAB2, IL17RB, KITLG, MMD, MMP9, NRCAM, RF56K, SIRPFO1, SGCG, SGMS1, SLCD2A5</td>
</tr>
<tr>
<td><strong>Cell-to-cell signaling and interaction, Tissue development, Cardiovascular system development and function</strong></td>
<td>AMIGO2, CASC7, DCG4, DFR4, FOSL2, GATA1, IL20, IRS1, MB1, NUMBL, PCDHB14, PCDSC4, SCLY, SERPINE8, SOX4, TRIM8</td>
<td>16</td>
<td>AMIGO2, CASC7, DCG4, DFR4, FOSL2, GATA1, IL20, IRS1, MB1, NUMBL, PCDHB14, PCDSC4, SCLY, SERPINE8, SOX4, TRIM8</td>
</tr>
<tr>
<td><strong>Cancer, Hematological disease, Cellular movement</strong></td>
<td>C4orf32, DENND5B, FAM63A, GDPD1, GRAMD1C, METTL7A, MX3, PCDHB14, PDS11, PRAT, RBM24, SLC25A30, SNX24, TMPRSS3, TRMT6, ZNF185</td>
<td>16</td>
<td>C4orf32, DENND5B, FAM63A, GDPD1, GRAMD1C, METTL7A, MX3, PCDHB14, PDS11, PRAT, RBM24, SLC25A30, SNX24, TMPRSS3, TRMT6, ZNF185</td>
</tr>
<tr>
<td><strong>Cell-to-cell signaling and interaction, Cellular compromise, Tissue development</strong></td>
<td>ABC11, ARHGEF28, CCNG1, CLDN1, KL9, LRG1, OSTF1, PIKR31, PVRL2, SPAT70, THBS1, TXNRD1, UTRN, VWF</td>
<td>14</td>
<td>ABC11, ARHGEF28, CCNG1, CLDN1, KL9, LRG1, OSTF1, PIKR31, PVRL2, SPAT70, THBS1, TXNRD1, UTRN, VWF</td>
</tr>
<tr>
<td><strong>Cellular development, Developmental disorder, Hereditary disorder</strong></td>
<td>ATP6V1C1, CDKN3, CCL34, FHDC1, FRMD6, ITPR1LP2, MAML3, MRB13, PCDHB14, RASSF8, SERPBP1, SVRPA1, SUSD4, UBE2E3</td>
<td>14</td>
<td>ATP6V1C1, CDKN3, CCL34, FHDC1, FRMD6, ITPR1LP2, MAML3, MRB13, PCDHB14, RASSF8, SERPBP1, SVRPA1, SUSD4, UBE2E3</td>
</tr>
<tr>
<td><strong>Lipid metabolism, Small molecule biochemistry, Cellular compromise</strong></td>
<td>CACNA1H, CENPQ, FAM171B, FGFBP2, LYAB, NMRK1, NUMBL, RN1SF32, RPL31, RNT4, SMTNL2, SYNPO2L, SYTL5, TSTP2</td>
<td>14</td>
<td>CACNA1H, CENPQ, FAM171B, FGFBP2, LYAB, NMRK1, NUMBL, RN1SF32, RPL31, RNT4, SMTNL2, SYNPO2L, SYTL5, TSTP2</td>
</tr>
<tr>
<td><strong>Developmental disorder, Hereditary disorder, Metabolic disease</strong></td>
<td>ANKKRD6, ATP8A1, CCPI10, FICO1, IGS5F, KIF21A, NRNX3, RUND38, SELT, SPCS3, SUMF1, SYITL, TNL2</td>
<td>13</td>
<td>ANKKRD6, ATP8A1, CCPI10, FICO1, IGS5F, KIF21A, NRNX3, RUND38, SELT, SPCS3, SUMF1, SYITL, TNL2</td>
</tr>
<tr>
<td><strong>Cancer, Hematological disease, Immunological disease</strong></td>
<td>ADRAR2A, AFA41L2, AKAP8, CELSR2, FEM1A, IDH1, IRX3, MYC, PHF10, S1PR3, SPATS2L</td>
<td>11</td>
<td>ADRAR2A, AFA41L2, AKAP8, CELSR2, FEM1A, IDH1, IRX3, MYC, PHF10, S1PR3, SPATS2L</td>
</tr>
<tr>
<td><strong>Connective tissue disorders, Hereditary disorder, Metabolic disease</strong></td>
<td>DOK7, HSP90AA1, LNX, NUDT4, PAM, PSK11, RAD21, RMRG7, SMCO4, TAF5L, TME64</td>
<td>11</td>
<td>DOK7, HSP90AA1, LNX, NUDT4, PAM, PSK11, RAD21, RMRG7, SMCO4, TAF5L, TME64</td>
</tr>
</tbody>
</table>
II. IPA functional network association analysis of differentially regulated genes at the 6 hour time point

<table>
<thead>
<tr>
<th>Functional Network</th>
<th>Focus genes</th>
<th>Score</th>
<th>Focus genes in network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digestive system development and function, Hepatic system development and function, Organ morphology</td>
<td>21</td>
<td>48</td>
<td>BAG1, BMPR2, CCNG1, CHST8, DKK1, FEM1A, IDH1, IL17RB, INSR, MAP4K4, PCBP2, PCDH14, PDCD4, PTPN21, PVRL2, SRY2, TFF2, THBS1, TXNRD1, UTRN, VWF</td>
</tr>
<tr>
<td>Neurological disease, Psychological disorders, skeletal and muscular disorders</td>
<td>16</td>
<td>33</td>
<td>ADORA2B, ADRA2A, CELSR2, CYP1B1, GPR87, HSP90AA1, IL20, NRCAM, PAK1, PIM2, RAD21, RPL31, RTN4, SERPINB9, SOX4, SPATS2L</td>
</tr>
<tr>
<td>Developmental disorder, Hereditary disorder, Ophthalmic disease</td>
<td>15</td>
<td>31</td>
<td>ACACB, CENPQ, EFEMP1, FMN1, FREM2, KCNN6, KIF21A, METTL7A, PCDHB14, RASSF8, SNRPA1, SNX24, SUMF1, SYNPO2, TRMT6</td>
</tr>
<tr>
<td>Cardiovascular system development and function, Tissue development, Tissue morphology</td>
<td>13</td>
<td>26</td>
<td>DPY9L1, FGFBP2, IFTRIP1L2, LRRC3, NAT8L, PCDHB10, PCDHB14, RBM24, SYTL5, TMEM64, TMEM150C, ZNF117, ZNF185</td>
</tr>
<tr>
<td>Hereditary disorder, Skeletal and Muscular disorders, Cancer</td>
<td>7</td>
<td>12</td>
<td>LNX, MED8, NRXN3, PIM2, RNF144A, SAMD4A, SGCG</td>
</tr>
</tbody>
</table>
Appendix H

Molecular pathway enrichment analysis of dysregulated genes in oestrogen-stimulated RAD21-depleted MCF7 cells at the 6 hour time point

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Number of genes</th>
<th>Total number of reference genes in the pathway</th>
<th>Ratio of enrichment</th>
<th>Raw p-value</th>
<th>Adjusted p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>β1 integrin cell surface interactions</td>
<td>13</td>
<td>1351</td>
<td>5.68</td>
<td>3.98E-07</td>
<td>2.08E-06</td>
</tr>
<tr>
<td>ErbB1 downstream signalling</td>
<td>12</td>
<td>1288</td>
<td>5.50</td>
<td>1.64E-06</td>
<td>2.08E-06</td>
</tr>
<tr>
<td>EGFR-dependent endothelin signaling events</td>
<td>12</td>
<td>1289</td>
<td>5.50</td>
<td>1.65E-06</td>
<td>2.08E-06</td>
</tr>
<tr>
<td>Class I P13K signalling events</td>
<td>12</td>
<td>1288</td>
<td>5.50</td>
<td>1.64E-06</td>
<td>2.08E-06</td>
</tr>
<tr>
<td>mTOR signalling pathway</td>
<td>12</td>
<td>1288</td>
<td>5.50</td>
<td>1.64E-06</td>
<td>2.08E-06</td>
</tr>
<tr>
<td>Insulin pathway</td>
<td>12</td>
<td>1288</td>
<td>5.50</td>
<td>1.64E-06</td>
<td>2.08E-06</td>
</tr>
<tr>
<td>IGF1 pathway</td>
<td>12</td>
<td>1291</td>
<td>5.49</td>
<td>1.68E-06</td>
<td>2.08E-06</td>
</tr>
<tr>
<td>IL5-mediated signalling events</td>
<td>15</td>
<td>1292</td>
<td>5.49</td>
<td>1.69E-06</td>
<td>2.08E-06</td>
</tr>
<tr>
<td>PDGF receptor signalling network</td>
<td>12</td>
<td>1293</td>
<td>5.48</td>
<td>1.71E-06</td>
<td>2.08E-06</td>
</tr>
<tr>
<td>IFN-gamma pathway</td>
<td>12</td>
<td>1296</td>
<td>5.47</td>
<td>1.75E-06</td>
<td>2.08E-06</td>
</tr>
<tr>
<td>Plasma membrane oestrogen receptor signalling</td>
<td>12</td>
<td>1301</td>
<td>5.45</td>
<td>1.82E-06</td>
<td>2.08E-06</td>
</tr>
<tr>
<td>ErbB receptor signalling network</td>
<td>12</td>
<td>1312</td>
<td>5.40</td>
<td>1.98E-06</td>
<td>2.08E-06</td>
</tr>
<tr>
<td>TRAIL signalling pathway</td>
<td>12</td>
<td>1328</td>
<td>5.34</td>
<td>2.25E-06</td>
<td>2.31E-06</td>
</tr>
</tbody>
</table>

Signalling cascades in which cohesin-regulated oestrogen-responsive genes are over-represented. WebGestalt was used to identify the enriched pathways of dysregulated genes at the 6 hour time point. Statistically significant pathway enrichment was determined using a hypergeometric test with a significance cut-off of p-value <0.001. Benjamini-Hochberg method was used for multiple test corrections.
Appendix I

Appendix I. Assessment of levels of total- and acetylated-SMC3 in MCF7 cells treated with HDAC8 inhibitor PCI-34051 by immunoblotting

Asynchronous MCF7 cells grown in full media were incubated with different concentrations of PCI-34051 for 24 and 72 hours. Total cell lysates were prepared and levels of total- and acetylated-SMC3 were analysed by SDS-PAGE and immunoblotting by using antibodies specific to each. Acetylated-SMC3 was non-detectable for the 24 hour time point. Acetylated-SMC3 was detected for the 50 µM dose of the 72 hour time point, however 10 and 25 µM doses were inconclusive.
Appendix J

A. Control

PCI 10 μM

PCI 25 μM

PCI 50 μM

PCI 100 μM

B. Control

PCI 10 μM

PCI 25 μM

PCI 50 μM

PCI 100 μM
Appendix J. Immunofluorescence analyses of total- and acetylated-SMC3 in G₁-S synchronised MCF7 cells treated with HDAC8 inhibitor PCI-34051

MCF7 cells were synchronised to G₁-S phase by double thymidine block and released in the presence of different concentrations of the inhibitor for 12 (A) and 24 hours (B). Cells were fixed with 4% paraformaldehyde and immunostained for DNA (Hoechst H33342), total-SMC3 (green) and acetylated-SMC3 (red). Confocal fluorescence images of immunostained MCF7 cells are shown. These data indicate that deacetylation of SMC3 subunit of cohesin is hindered in response to treatment with PCI-34051 and leads to the accumulation of acetylated-SMC3 in a dose dependent manner.
Appendix K

I. Transcript levels of reference genes *PPIA* and *GAPDH* in MCF7 cells treated with 10 µM of PCI-34051
II. Transcript levels of ten RAD21-dependent oestrogen-responsive genes in MCF7 cells treated with 10 µM of PCI-34051 and stimulated for 3 and 6 hours with oestradiol

![Graph showing transcript levels of genes at 3 hours](image)

![Graph showing transcript levels of genes at 6 hours](image)

Appendix K. PCI-34051 did not affect response to oestradiol, despite accumulation of ac-SMC3 after treatment with 10 µM of PCI-34051

Oestradiol-deprived MCF7 cells were treated with 10 µM of PCI-34051 blocking deacetylation of SMC3 for 48 hours, prior to harvesting RNA, post treatment with 100 nM of oestradiol for 3 (II.A.i) and 6 hours (II.A.ii). Transcript levels of indicated genes were quantitated by qRT-PCR and normalised to the mean of two reference genes PPIA and GAPDH (I) that remained unaltered upon treatment with PCI-34051. Error bars represent the SEM of three independent experiments. Significance was determined using two-way ANOVA.
Appendix L

A. i. 6 hours time point:

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>PCI 10 µM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estradiol</td>
<td><img src="image1" alt="Flow cytometry" /></td>
<td><img src="image2" alt="Flow cytometry" /></td>
</tr>
<tr>
<td>Vehicle</td>
<td><img src="image3" alt="Flow cytometry" /></td>
<td><img src="image4" alt="Flow cytometry" /></td>
</tr>
</tbody>
</table>

A. ii. 24 hours time point:

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>PCI 10 µM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estradiol</td>
<td><img src="image5" alt="Flow cytometry" /></td>
<td><img src="image6" alt="Flow cytometry" /></td>
</tr>
<tr>
<td>Vehicle</td>
<td><img src="image7" alt="Flow cytometry" /></td>
<td><img src="image8" alt="Flow cytometry" /></td>
</tr>
</tbody>
</table>

B. i. ERα transcript levels

<table>
<thead>
<tr>
<th>Time (hours)</th>
<th>6</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative m-RNA</td>
<td><img src="image9" alt="Graph" /></td>
<td><img src="image10" alt="Graph" /></td>
</tr>
</tbody>
</table>

B. ii. 6 hours 24 hours

<table>
<thead>
<tr>
<th>Condition</th>
<th>CV</th>
<th>CE</th>
<th>P10</th>
<th>P10E</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERα</td>
<td><img src="image11" alt="Bands" /></td>
<td><img src="image12" alt="Bands" /></td>
<td><img src="image13" alt="Bands" /></td>
<td><img src="image14" alt="Bands" /></td>
</tr>
<tr>
<td>γ-Tubulin</td>
<td><img src="image15" alt="Bands" /></td>
<td><img src="image16" alt="Bands" /></td>
<td><img src="image17" alt="Bands" /></td>
<td><img src="image18" alt="Bands" /></td>
</tr>
</tbody>
</table>

Normalised signal over time

Appendix L. Assessment of cell cycle progression and level of oestrogen receptor α, in oestrogen-starved MCF7 cells treated with 10 µM of PCI-34051

Oestrogen-deprived MCF7 cells were treated with 10 µM of PCI-34051 for 48 hours, followed by treatment with vehicle or 100 nM of 17-β-oestradiol for 6 and 24 hours, respectively. (A) Flow cytometry analyses of cell cycle distributions. Following incubation with PCI-34051 and oestradiol stimulation, cells were fixed and stained with propidium iodide and cell cycle analyses were carried out on a Beckman Coulter Galios Flow Cytometer and processed using the FlowJo software (version 9.7). Histograms represent the intensity of fluorescence on the X-axis and cell number on the Y-axis for (i) 6 hour and (ii) 24 hour time points. Bar graphs show the percentage of cells quantified in the G0/G1, S and G2/M phases of the cell cycle in control and PCI-34051-treated conditions for vehicle/oestradiol stimulation, separately. Error bars represent the SEM from two
independent biological replicates. Significance was determined by two-way ANOVA. (B) Assessment of transcript and protein levels of ERα in response to treatment with PCI-34051. (i) Transcript levels of ERα in MCF7 cells. RNA levels were determined using qRT-PCR and are shown relative to the mean of two reference genes, PPIA and GAPDH. Error bars represent the SEM from three independent experiments. Significance was determined using two-way ANOVA. (ii) Representative immunoblots showing level of ERα protein after incubation with PCI-34051 for the indicated time points. Level of ERα protein was analysed by SDS-PAGE and immunoblotting. The bar graph depicts quantification of total ERα protein levels, normalised to γ-tubulin. Error bars represent the SEM from three independent experiments. PCI-34051 delayed cell cycle progression but did not alter ERα mRNA or protein levels.
Appendix M

Asynchronous MCF7 cells grown in full supplemented medium were synchronised to G1-S phase by double thymidine block. Thymidine was removed and cells were released in the presence of increasing concentrations of PCI-34051 for 6 and 12 hours, respectively. Total cell lysates were prepared and 30 µg of proteins were loaded onto 10-12 % pre-cast (Invitrogen, USA) SDS-polyacrylamide gel. Following SDS-PAGE, proteins were transferred to a PVDF membrane (Amersham, GE Healthcare Life Sciences, NZ) and immunoblotted with anti-PARP (1:2000) and γ-Tubulin (1:5000) antibodies. Developed bands were detected using the WesternBreeze™ chemiluminescent system (Invitrogen, USA). The positive control (apoptotic Jurkat cells) produced an 89 kDa band representative of cleaved-PARP, however neither 113 (representative of intact PARP) nor 89 kDa bands appeared for the PCI-34051-treated samples. Rather, an unexpected band of size ~60 kDa was visualised for control as well as PCI-34051-treated MCF7 cells which made analysis of the state of PARP protein (cleaved or uncleaved) inconclusive for PCI-34051-treated MCF7 cells.
Appendix N
Publication from this work
Cohesin modulates transcription of estrogen-responsive genes

Jisha Antony a,1, Tanushree Dasgupta a,1, Jenny M. Rhodes a, Miranda V. McEwan a, Cristin G. Print b, Justin M. O’Sullivan a, Julia A. Horsfield a,⁎

a Department of Pathology, Dunedin School of Medicine, The University of Otago, Dunedin, 9016, New Zealand
b Department of Molecular Medicine and Pathology, School of Medical Sciences, The University of Auckland, Auckland, 1142, New Zealand

Abstract

The cohesin complex has essential roles in cell division, DNA damage repair and gene transcription. The transcriptional function of cohesin is thought to derive from its ability to connect distant regulatory elements with gene promoters. Genome-wide binding of cohesin in breast cancer cells frequently coincides with estrogen receptor alpha (ER), leading to the hypothesis that cohesin facilitates estrogen-dependent gene transcription. We found that cohesin modulates the expression of only a subset of genes in the ER transcription program, either activating or repressing transcription depending on the gene target. Estrogen-responsive genes most significantly influenced by cohesin were enriched in pathways associated with breast cancer progression such as PI3K and ErbB1. In MCF7 breast cancer cells, cohesin depletion enhanced transcription of TFF2 and TFF1, and was associated with increased ER binding and increased interaction between TFF1 and its distal enhancer situated within TMR553. In contrast, cohesin depletion reduced c-MYC mRNA and was accompanied by reduced interaction between a distal enhancer of c-MYC and its promoters. Our data indicates that cohesin is not a universal facilitator of ER-induced transcription and can even restrict enhancer-promoter communication. We propose that cohesin modulates transcription of estrogen-dependent genes to achieve appropriate directionality and amplitude of expression.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

The control of gene transcription operates at multiple levels, including the recruitment of transcription factors, transcriptional co-factors and RNA polymerase. Transcription is also accompanied by changes in chromatin conformation and is informed by epigenetic modifications of chromatin and spatial organization of the genome [1,2].

Genome-scale chromosome capture assays have shown that mammalian chromosomes are partitioned into discrete topologically associated domains (TADs) [3,4]. TADs can range in size from several hundred kilobases to megabases; they represent compartments within which genes interact frequently, and thus can be regulated by a common set of factors [3-5]. TADs could therefore represent the framework for interactions between non-protein-coding gene regulatory elements and gene promoters, thereby influencing gene transcription [3-5]. Compartimentalization of genes into TADs could contribute to transcriptional control by restricting the number and types of regulatory elements that can interact with genes [3-6].

The cohesin complex is best characterized for its role in mediating sister chromatid cohesion during mitosis, however it also has important roles in DNA damage repair, ribosome biogenesis and gene regulation [7-9]. Cohesin consists of four core subunits, SMC1A, SMC3, RAD21 and STAG1/2/Stromalin [7-9], indicating that intact cohesin function is important for normal human development. The gene regulatory function of cohesin is independent of its role in cell cycle [12-15] and can be remarkably tissue-specific [16–19], indicating that it has an important but as yet poorly understood role in cell type-specific transcription.

Although the exact mechanism by which cohesin regulates gene expression is unclear, it has been shown to facilitate long-range interactions between DNA elements, including interactions between...
enhancers and gene promoters [19–28]. Recently, cohesin was shown to be important for chromatin interactions within TADs [26,27], and to contribute to the formation of insulated neighborhoods [29], thereby potentiating its contribution to transcription via global genome organization.

Cohesin genes are frequently altered in cancer [30,31], although strikingly, not always in the same way. In leukemia anomalies are inevitably responsible for the development of cancer. It frequently mutated, whereas in ovarian and breast cancer, these genes are more commonly upregulated or amplified (particularly RAD21 near B3q/4) [31,36,37]. Although cohesin gene alterations could lead to aneuploidy and genome instability [38], it is not clear that these anomalies are inevitably responsible for the development of cancer. It may also be possible that dysregulation of downstream genes resulting from cohesin alteration underlies the development of cancers that harbor cohesin alterations [30,31,38].

Estrogen receptor alpha (ER)-associated transcription is the output of a ligand-dependent signaling pathway that drives gene expression and tumor growth in majority of breast cancers [39]. ER-mediated gene transcription involves many layers of regulation, including the binding of ER to estrogen response elements (EREs), and recruitment of co-factors such as FOXA1 and GATA3 [39,40]. In recent years, the formation of chromatin loops that juxtapose regulatory elements and promoters has emerged as an essential feature of ER-dependent transcription [40–42]. Many estrogen-responsive genes interact with spatially distant EREs via long-range interactions [40–45].

 Genome-wide binding analysis in estrogen-responsive MCF7 breast cancer cells revealed estrogen-induced cohesin binding of chromatin frequently coincides with sites bound by ER [25]. This concordance in cohesin and ER binding was particularly striking at estrogen-regulated genes, and also at regions that are involved in ER-associated chromatin interactions [25,41]. These findings raise the possibility that cohesin is involved in regulating ER-dependent transcription.

Previously, we demonstrated that cohesin is required for the estrogen-induced transcription of the oncogene c-MYC (subsequently referred to as MYC) in MCF7 cells [46]. Cohesin binds at the MYC gene promoters and at enhancers upstream of MYC in an estrogen-dependent manner [46]. Moreover, depletion of cohesin hampered ER binding at the MYC gene enhancers and promoters [46]. Myc is also regulated by cohesin in zebrafish and Drosophila [47] and is downregulated in lymphoblastoid cells from CD5 patients [48]. Despite the clear relationship between cohesin and ER in the regulation of MYC gene expression, it remains unknown whether there is a general transcriptional dependence for cohesin at estrogen-responsive genes.

Here we aimed to identify other ER-regulated breast cancer genes that are dependent upon cohesin for transcriptional activity. We found that cohesin only influences the estrogen response of a subset of genes and is not a universal facilitator of ER associated transcription. Cohesin depletion in MCF7 cells elicited a range of responses, including up-regulation of estrogen responsive genes and gain in estrogen sensitivity. Within the TFF1 gene cluster, cohesin depletion, enhanced expression of TFF1, TFF2 and TFF3 genes, was accompanied by increased ER recruitment, and increased interaction between estrogen-induced cohesin binding sites at the TFF3 enhancer and the TFF1 promoter. By contrast, decreased MYC transcription upon cohesin depletion was associated with diminished enhancer–promoter interaction at the MYC locus. Together our results suggest that cohesin can both positively and negatively influence ER-mediated transcription in a gene-dependent manner.

2. Material and methods

2.1. Cell culture and siRNA transfections

MCF7 cells (ATCC) were cultured in Dulbecco’s Modified Eagle Medium (DMEM, Life Technologies or Sigma) supplemented with 10% fetal bovine serum (FBS) in a 37 °C humidified incubator at 5% CO2. T47D cells (ATCC) were cultured in RPMI-1640 media containing insulin (5 μg/ml) and 10% FBS in a 37 °C incubator at 10% CO2. For hormone depletion, cells were cultured for 3 days in phenol-free media supplemented with 10% charcoal dextran (Sigma)-treated FBS. Hormone-depleted cells were treated with 17β-estradiol (Sigma) at a final concentration of 100 nM for various time periods. Estradiol was dissolved in absolute ethanol, and as controls, hormone depleted cells were also treated with the same volume of ethanol (vehicle) for the same time periods. To deplete cohesin in MCF7 cells, we used the ON-TARGET plus siRNA DHA-J-006832-06 (GE Dharmacon) that targets the cohesin subunit, RAD21. The non-targeting control siRNA D-001810-01 (GE Dharmacon) was used as a negative control. RAD21 and control siRNAs were used at final concentrations of 5 nM for microarray experiments and 10 nM for chromatin immunoprecipitation (ChIP) and chromosomal conformation capture (3C) experiments. Cells were reverse-transfected using Lipofectamine RNAiMAX (Life technologies) following 24 h of hormone depletion. After transfection, cells were cultured in hormone-depleted conditions for a further 48 h prior to estradiol stimulation.

2.2. RNA isolation, cDNA synthesis and gene expression analysis

Total RNA was isolated using the Machery Nagel Nucleospin RNA Isolation kit according to manufacturer’s guidelines (Machery Nagel). 0.5 μg of total RNA was reverse-transcribed into first strand cDNA using the Superscript III first strand synthesis system (Life Technologies). Quantitative PCR (qPCR) was performed in technical duplicates with either Platinum SYBR Green qPCR SuperMix-UDG with Rox (Life Technologies) or TaqKaRa SYBR Premix Ex TaqTM (Clontech) on an ABI 7300 (Applied Biosystems) or on the LightCycler 480 (Roche Diagnostics). Each primer was first validated for efficiency with an external standard curve generated by serial dilutions of the cDNA. Gene expression analyses were carried out on qPhase Plus (Biogazelle) and were normalized relative to the mean of reference genes glyceraldehyde-3-phosphate dehydrogenase (GAPDH), cyclophilin and ribosomal protein L13a (RPL13A). Primer details are included in Supplementary Table S3.

2.3. Antibodies and immunodetection

Cells were lysed in RIPA extraction buffer containing protease inhibitors and proteins were quantified using the Bicinchoninic Acid assay kit (Pierce). For immunoblotting, 30–60 μg of protein lysates was resolved on a 10% SDS–PAGE gel and transferred onto a nitrocellulose membrane. The Odyssey infrared detection (LI-COR Biosciences) system was used for protein detection and quantitation. Antibodies used were rabbit anti-RAD21 antibody (1:1000 dilution, A592, Abcam) and mouse anti-γ-Tubulin antibody (1:5000 dilution, T5226, Sigma) and were detected with IRDye 680-labeled goat-anti-rabbit IgG, or IRDye 800-labeled goat-anti-mouse IgG (LI-COR Biosciences) at 1:15,000 dilution. Detected bands were quantified using the Image Studio Lite software (LI-COR Biosciences).

2.4. Microarray and transcript profiling

MCF7 cells growing in hormone-depleted conditions for 24 h were transfected with RAD21 or control siRNA and cultured in hormone-depleted conditions for a further 48 h. Cells were subsequently treated with 100 nM 17β-estradiol or vehicle for 3 and 6 h. Cells were harvest ed for RNA and protein at both time points. Knockdown of RAD21 mRNA and protein were verified by qPCR and immunoblot respectively. RNA from three independent biological replicates for each time point was analyzed for quality and integrity by using the RNA 6000 Nano kit (Agilent Technology) run on a 2100 Bioanalyzer system (Agilent Technology) as per the manufacturer’s instructions. Microarray expression
analyses were carried out at New Zealand Genomics Limited (NZGL). Briefly, RNAs from 3 independent experiments were labeled and hybridized to Affymetrix PrimeView 3’IVT Human Gene Expression Microarrays according to the manufacturer’s guidelines and were analyzed for global changes in gene expression between vehicle/estradiol treated control and RAD21 knockdown cells. The analysis was performed using the statistical software R and Bioconductor packages. Normalization of the arrays was performed using the RMA algorithm. Limma was used to calculate potentially differentially expressed transcripts between the treatment groups. A linear model was used to estimate the fold change, and an empirical Bayes method was used to moderate the standard deviations of the estimated fold changes. The Benjamini-Hochberg method was used to correct the p-values for multiple testing. The Affymetrix probe IDs were annotated using the Bioconductor package hgs219.db, 2.7.1. Gene functions were annotated with Gene Ontology (GO) and pathway analysis used WebGestalt (WEB based Gene SeT Analysis Toolkit). Data have been deposited at GEO under identification GSE59808.

2.5 Chromatin immunoprecipitation

ChIP was performed and analyzed by qPCR as previously described [46]. Briefly, MCF7 cells were hormone depleted and transfected with RAD21 or control siRNA as described above. Cells were then stimulated with 100 nM estradiol or vehicle for 45 min, fixed with 1% formaldehyde and quenched with 125 mM glycine. Fixed cells were lysed in SDS lysis buffer containing protease inhibitor prior to sonication. Equal amounts of diluted chromatin were pre-cleared with Protein A Agarose/Salmon Sperm DNA (Millipore). Immunoprecipitations were performed over-night at 4 °C with anti-RAD21 (Ab992, Abcam), ER (HC-20, Santa Cruz) RNA polymerase II serine 5 phosphorylation (ab5131, Abcam) antibodies. Immunocomplexes were immobilized onto Protein A Agarose/Salmon Sperm DNA, washed, eluted, reverse-crosslinked at 65 °C overnight, and purified. For qPCR analyses, 1 μl of pre-cleared or immunoprecipitated chromatin was used for each reaction. Binding at each site was expressed relative to the pre-cleared input chromatin after subtraction from no antibody control. Data from each replicate were subjected to mean centering and results are represented as an average of three biological replicates. Primer sequences are included in Table S3.

2.6 Chromosome conformation capture assay

MCF7 cells were hormone depleted and transfected with RAD21 or control siRNA as described above. Cells were then stimulated with 100 nM estradiol or vehicle for 3 h. Cells were fixed with 2% formaldehyde for 10 min at room temperature, quenched with 125 mM glycine, scraped off the plate and pelleted. Nuclei were harvested by lysing the cell pellets in ice-cold lysis buffer (10 mM Tris pH 8.0, 10 mM NaCl, 0.25% NP-40) containing protease inhibitors. Nuclei were then resuspended in 1:2 of Dpn II restriction enzyme buffer (New England Biolabs) and 0.3% SDS and incubated for 1 h at 37 °C while shaking. Triton X-100 was then added to the mix at a final concentration of 1.8% and the reaction was left at 37 °C shaking for another hour. Chromatin was digested with 800 U of Dpn II overnight at 37 °C while shaking. Dpn II was inactivated by adding SDS at a final concentration of 1.3% and incubating at 65 °C for 20 min. Nuclei were diluted into volume of 7 ml containing 1.4 T4 DNA ligase buffer (Life Technologies) and Triton X-100 at final concentration of 1% and incubated at 37 °C for 1 h. Ligations were carried out with 200 U of T4 ligase (Life Technologies) for 4 h at 16 °C followed by 30 min at room temperature. Samples were protease K treated and reverse-crosslinked overnight at 65 °C. The next day, samples were treated with RNase A at 37 °C for 30 min. DNA was purified and ethanol precipitated. qPCR was performed with 3C primers via Taqman assay using the KAPA PROBE FAST Universal 2× mix (KAPA Biosystems) on the Light Cycler 480 (Roche Diagnostics). The probe and 3C primer sequences are provided in Tables S4 and S5. Interaction frequency at each site was determined from a standard curve generated by serially diluted (10−1 to 10−5) of donor cell type (estrogen) or vehicle, and ven- chor and interacting fragment, analyzed by Taqman assay using the KAPA PROBE FAST Universal 2× mix (KAPA Biosystems). Values were then normalized for total input using primers that do not span the Dpn II restriction site. Data from each replicate were subjected to mean centering and results are represented as an average of three biological replicates.

2.7 Statistical analysis

The software GraphPad Prism (GraphPad software, San Diego, CA, USA) was used to graph and analyze all data. qPCR, ChIP and 3C results were analyzed using two-way ANOVA with Sidák’s or Tukey multiple comparisons test, set at a threshold of p ≤ 0.05. The cell cycle data was analyzed by two-way ANOVA with a significance level set at p ≤ 0.05. RAD21 immunoblots were analyzed by ratio-paired t-tests, significance level set at p ≤ 0.05.

3. Results

3.1. RAD21 deletion alters the estrogen-responsive nature of a subset of genes that regulate growth pathways in MCF7 cells

We depleted the RAD21 subunit of cohesin in MCF7 cells using siRNA to investigate cohesin’s role in the regulation of estrogen-responsive genes [46]. Estrogen-deprived, RAD21-depleted MCF7 cells were stimulated with 17β-estradiol (estrogen) or vehicle, and were harvested at 3, 6 and 24 h post-stimulation for analysis. Depletion of RAD21 was confirmed for all samples by reverse-transcriptase quantitative PCR (RT-qPCR) (Fig. S1A) and immunoblot (Figs. S1B and S7B), which indicated a decrease by 70–75% of RAD21 protein at the time of estrogen stimulation. Consistent with our previous work, we observed that total ER protein levels were not altered by RAD21 depletion [46]. Cell cycle analysis, performed 24 h post-estrogen stimulation in RAD21-depleted MCF7 cells indicated that RAD21 depletion delayed transition of MCF7 cells from the G0/G1 to S/M phase in both the presence and absence of estrogen. RAD21-depleted cells showed no evidence of a sub-G0 population (Fig. 1A) and exhibited no significant changes in overall growth (data not shown) or morphology compared with controls (Fig. 1B) within 72 h after siRNA treatment.

Global changes in gene expression were analyzed using Affymetrix PrimeView 3’IVT Human Gene Expression Microarrays in RAD21-depleted estradiol/vehicle stimulated MCF7 cells. At 3 and 6 h post-estradiol stimulation, several genes were identified as differentially regulated in comparisons of estrogen-stimulated control versus estrogen-stimulated RAD21-depleted cells (Figs. 2 and S2). More tran- scripts were differentially regulated at 3 h (218 probes representing 153 genes, p < 0.005) than at 6 h (110 probes representing 77 genes, p < 0.005) (Fig. S2A), indicating statistically significant differential regu- lation of genes that respond transcriptionally within the first few hours of estrogen stimulation. Seventy-four probes (representing 53 genes) were similarly differentially regulated at both 3 and 6 h post-estrogen stimulation, indicating consistency of the early transcriptional response to estrogen stimulation in RAD21-depleted cells. The fraction of differentially regulated genes in RAD21-depleted cells was 22% of the total number of estradiol-responsive transcripts at 3 h post- stimulation (n = 72, p < 0.005), and 12% at 6 h post-stimulation (n = 169, p < 0.005). This indicates that not all estrogen-induced genes de- pend on cohesin for their estrogenic response.

We undertook functional annotation clustering using NIH- DAVID to identify biological processes involving RAD21-regulated estrogen-responsive genes (Table S1). Statistically significant over- representation of protein networks was identified using web-based
gene set analysis toolkit (WebGestalt) (Table 1). This analysis identified the ErbB1 downstream signaling network and the PI3K/mTOR signaling pathways as two major molecular pathways affected by cohesin depletion. Pertinent to our results, stimulation of the mTOR pathway can overcome growth deficiency of cohesin-depleted cells [49].

3.2. RAD21 depletion both represses and enhances ER-mediated transcription of cohesin-bound genes

Estrogen-responsive transcripts were both up- and downregulated upon RAD21 depletion in MCF7 cells (Figs. 2B and S2). We used RT-qPCR to confirm dysregulated expression of 11 selected genes that
had been previously implicated in breast cancer [51,53–60] (Table S2). Transcript levels were determined relative to those of the reference genes GAPDH and CyclinH1, which remained unaltered under all experimental conditions (Fig. S3). All 11 differentially expressed genes in the microarray were validated by RT-qPCR (Table S2; Figs. 2C, S4 and S5). Genes with transcriptional activation or repression influenced by cohesin appear to fall into at least three different regulatory categories (Fig. 2C). In one category, the transcriptional activity was blocked transcriptional activation by ER; for example, RAD21 depletion significantly increased estrogen stimulation of S0X4, and led to de-repression of BMPR2. In a second category, cohesin appears to be required for ER’s transcriptional activity; for example, RAD21 depletion prevented estrogen-mediated activation of IL20. In a third category, cohesin appears to restrict transcriptional repression mediated by ER, for example, RAD21 depletion prevented BMPR2 enhanced estrogen stimulation of BAG1. While some differential transcriptional effects persisted to 24 h post-stimulation with estradiol, others were transient and restricted to within the first 6 h post-stimulation (e.g., DKK1, Fig. S4).

Differential regulation verified by RT-qPCR of 7 other genes (THBS1, PIN2, IRS2, OCX1L2, BAG1, CCNG2 and TF2F) showed that RAD21 depletion reduced estrogen stimulation of THBS1, PIN2, IRS2, OCX1L2 and BAG1, enhanced estrogen activation of TF2F, and reduced repression of CCNG2 (Fig. S5). Cohesin-altered regulation of genes was much more pronounced in estrogen-induced than non-induced conditions. Overall, our results show that cohesin does not merely facilitate ER-dependent transcription, but can function in a gene-dependent manner to modulate ER-mediated transcriptional output.

We examined the relative location of RAD21 and ER binding of the cohesin-regulated genes in MCF7 cells using the UCSC genome browser (2006 NCBI36/g18 version) and existing global binding data [25] (Table S1). All 11 co-regulated genes have binding sites for ER, RAD21 and estrogen-induced RAD21, with 9/11 having estrogen-induced RAD21 binding within 15 kb of the transcription start site (TSS). At S0X4, THBS1, IRS2, OCX1L2 and DKK1, ER and RAD21 were co-localized (<1 kb) and within 5 kb of the TSS. In addition, binding sites for ER and RAD21 were found at a distance from other genes, such as CCNG2, up to 67 kb from the TSS. Evidence of coincident binding is consistent with the idea that cohesin and ER can cooperate to regulate transcription of these genes.

3.3. Cohesin binding restricts ER-mediated transcription within the TFF cluster in MCF7 cells

The TFF gene cluster is a well-characterized ER-regulated region [40,42,43] and was chosen to interrogate the role of cohesin in estrogen-dependent transcription in greater detail. Moreover, TF2F and TMPRSS3 were identified in the microarray analyses as significant cohesin-regulated estrogen-responsive genes. We used RT-qPCR to examine the expression of TFF1, TF2, TF3 and TMPRSS3 in MCF7 cells following RAD21 depletion (Fig. S7) and estrogen stimulation. Estradiol treatment enhanced the expression of all four genes to varying degrees (Fig. 3A). Transcript levels for TMPRSS3, TF1 and TF2F increased at both 3 and 6 h post-estrogen treatment, while TF3 transcript levels demonstrated a much smaller increase and only at 6 h.

Compared with controls, RAD21 depletion in estradiol-stimulated MCF7 cells significantly upregulated the transcription of TMPRSS3 by 2.3-fold and TF2F by 3.4-fold (3 and 6 h post-stimulation; p < 0.0001 for both), and of TF1 by 1.4 fold (3 h post-stimulation; p = 0.005). No significant alteration was seen for TF3F expression upon RAD21 depletion.

Genome-wide binding data in MCF7 cells indicate that RAD21 and ER localize both independently and together at several sites within the TFF cluster (Fig. 3B) [25]. We performed quantitative chromatin immunoprecipitation (ChIP) using an anti-ER antibody to determine if cohesin modulates ER binding within the TFF cluster. ChIP primers were used to detect binding at the promoters of TFF1 (primer 4), TF2F (primer 3) and TF3F (primer 1) and at two other sites; one encompassing theERE within TMPRSS3 gene body (primer 5), and the other incorporating theERE downstream ofTF2F (primer 2) (Fig. 3B). The EREs detected by primers 2 and 5 were previously shown to be important for demarcating the boundary of extranuclear responsiveness for the TFF genes [61]. The ERE within TMPRSS3 was chosen for analysis because it is enriched for enhancer histone marks [40] and is known to function as a distal enhancer forTF1 [42].

All ChIP analyses were carried out in MCF7 cells stimulated with estrogen or vehicle for 45 min as described previously [25,46]. Following estrogen stimulation, ER binding increased, and was strongest at the TF1F promoter (primer 4) and at the TMPRSS3 ERE (primer 5) (Fig. 3C). In estrogen-stimulated RAD21-depleted cells, ER binding was significantly enhanced at the TF1F and TF2F promoters (p < 0.005) and at the TMPRSS3 ERE (p < 0.01) compared to controls. The increase in ER binding at these loci is consistent with the increase in transcript levels we observed following RAD21 depletion (Fig. 3A). No significant changes were detected for ER binding at the TF3F promoter, consistent with the finding that its expression is not enhanced by RAD21 depletion (Fig. 3A).

We next sought to determine if RAD21 depletion affects RNA polymerase II (RNAPII) binding at the TFF locus in MCF7 cells. ChIP was carried out using antibody raised against serine-5-phosphorylated RNAPII (SSP), which detects the early initiation form of the polymerase, and also with antibody raised against serine-2-phosphorylated RNAPII (SP2), which detects the productive elongation form of RNAPII. Estrogen treatment increased levels of both SSP and SP2 RNAPII at all sites examined (Fig. 3B), consistent with previous reports [62]. In estrogen-stimulated RAD21-depleted cells, RNAPII-SSP binding was reduced at the promoter ofTF1F and at the ERE downstream of TF2F (both p < 0.05), compared with controls. In contrast, RNAPII-SP2 binding increased significantly at the promoters of TF1F and at the ERE within the TMPRSS3 gene body (both p < 0.05) while no significant changes were observed at the ERE downstream of TF2F. The early initiation SSP RNAPII complex accumulates at the promoter proximal region during polymerase pausing [63]. Cohesin was previously found to influence its transition to the elongating form [64,65]. The decrease in the early initiation complex at the promoters and the associated increase in productive initiation complex of RNAPII indicate that RAD21 depletion results in release from pausing at theTF1F promoter, reflecting the observed enhanced transcription of these loci.

We observed enrichment of chromatin-bound RAD21 upon estrogen treatment at sites previously found to recruit RAD21 in response to estradiol treatment [25] (Fig. S8). Upon siRNA depletion of RAD21 (Fig. S7), enrichment of RAD21 markedly decreased at promoters 2, 4, and 5 (Fig. S8B), all of which correspond to estrogen-induced cohesin binding sites identified previously (Fig. S8A). Notably, sites at TF1F and TMPRSS3 (primers 4 and 5) that lose RAD21 binding exhibited enhanced ER binding (p < 0.05) upon RAD21 depletion (Fig. 3C). This indicates that RAD21 is not required for ER to bind these locations, and furthermore, that RAD21 loss may even facilitate ER binding. Together these results suggest that cohesin impedes estrogen responsiveness of the TFF cluster, possibly by restricting ER binding.

3.4. Cohesin influences chromatin interactions within the TFF cluster in MCF7 cells

It is possible that dysregulated expression of the TFF gene cluster and downregulation of MYC upon cohesin depletion are accompanied by changes in chromatin interactions in the vicinity of these genes. Altered chromatin interactions could reflect, or even account for, the observed changes in transcription. Using chromosome conformation capture (3C), we examined the impact ofRAD21 depletion on selected chromatin interactions within the TFF gene cluster and at MYC.

Chromatin interactions at the TFF gene cluster in the presence and absence of estradiol have been well characterized previously [40,61]. The TMPRSS3 ERE functions as a distal enhancer for TFF1 and also
interacts with other sites within the TFF cluster [40–42]. We performed 3C analyses to determine the frequency of interactions with the TMPRSS3 ERE in MCF7 cells that were transfected with either control or RAD21 siRNA, and stimulated with either estrogen or vehicle for 3 h.

Robust estradiol-induced interactions occurred between the TMPRSS3 enhancer and three sites near and within TFF1, located between 8.9 and 12.5 kb from the TMPRSS3 ERE (Fig. 4A, all \( p < 0.01 \)). The nearest interacting site is separated from the TMPRSS3 ERE by 8.9 kb and 25 Dpn II restriction sites (Table S6), indicating that...
interactions are unlikely to simply be due to TMPS3 proximity. These results are consistent with previous data showing that chromatin interactions between the TMPS3 ERE and the TFF1 promoter correspond with transcriptional activation in MCF7 cells [42].

Most estradiol-induced chromatin interactions between TMPS3 and TFF1 were unaffected by RAD21 depletion (Fig. 4A, B). However, one interaction between TMPS3 and the TFF1 promoter (10 kb downstream) was significantly enhanced by RAD21 depletion in the presence of estrogen (p = 0.0001). This location within the TFF1 promoter corresponds to the TSS of TFF1 and incorporates previously identified estrogen-induced RAD21 and ER binding sites (Fig. 3B). Significantly, RAD21 siRNA caused loss of RAD21 binding (Fig. 5A), increased ER binding and enhanced transition of RNAPII to the elongation phase (Fig. 3C), indicating that enhanced chromatin looping is likely to be associated with transcription activation by ER. This result also implies that RAD21 is dispensable for the formation of chromatin loops anchored by ER, at least within the TFF complex. Interactions between TMPS3 and TFF2 were unaffected by RAD21 depletion.

Together our results suggest that RAD21 depletion is associated with altered local chromatin interactions within the TFF region that are consistent with increased transcriptional activation by ER. Surprisingly, RAD21 appears not to be essential for bolstering ER-anchored loop formation.

3.5. Cohesin depletion disrupts enhancer–promoter communication at the MYC locus in MCF7 cells

We previously reported that cohesin positively regulates Myc expression [31,46] and that RAD21 depletion blocked activation of MYC by estradiol in MCF7 cells [31,46]. An ERE 67 kb upstream of the MYC promoters (P1 and P2) is involved in estrogen-mediated activation of MYC [66], and we previously observed that RAD21 depletion decreased ER binding at this –67 kb ERE [46]. Here we used 3C to determine if RAD21 depletion alters a long-range interaction between this –67 kb ERE and the MYC promoters (P1 and P2) in MCF7 cells (Fig. 5).

Robust estradiol-induced interactions were detected between the –67 kb ERE and sites within MYC gene (Fig. 5A, p < 0.005). In estradiol-stimulated, RAD21-depleted cells, there was a statistically significant reduction in interactions between the –67 kb ERE and regions located at +66.8 (i.e., P2 region) and +67.2 kb (p = 0.0001 and p < 0.001 respectively). Since the –67 kb ERE is necessary for estrogen-mediated activation of MYC [66], loss of interaction between this ERE and MYC is consistent with MYC’s decreased transcription in RAD21-depleted cells.

4. Discussion

4.1. Cohesin depletion modulates transcription of a subset of estrogen responsive genes in MCF7 cells

Coincidence between cohesin and ER binding in estrogen-stimulated MCF7 cells [25] is consistent with the hypothesis that cohesin has a genome-wide role in facilitating ER-dependent transcriptional regulation, which is supported by studies demonstrating a partial requirement for cohesin in estrogen-dependent transcription [67,68].

Our study analyzed the qualitative influence of cohesin on estrogen-dependent genes, generating some unexpected findings. Cohesin depletion only affected around 12–22% of estrogen-responsive genes in MCF7 cells during the first 6 h post estrogen stimulation. Baseline expression of a subset of genes (e.g., CCNG2, an inhibitor of cell cycle-progression [69]; Fig. S6) was altered by RAD21 depletion in the absence of estrogen, perhaps reflecting inhibition of entry into the cell cycle. Under conditions of estrogen stimulation, we found that cohesin depletion had a rather qualitative effect on the expression of a modest numbers of genes (Figs. 2, S2, S4 and S5). Most validated differentially regulated genes had cohesin binding sites within ~20 kb of gene transcription start sites (Table 2), indicating potential for direct regulation of these genes by cohesin. We expected to observe an overall decrease in expression of genes that normally respond to estrogen, and some genes did behave in this way (e.g., IL20, Fig. 2C). However, we also found that other genes that were relatively insensitive to estrogen induction (e.g., SOX4) gained sensitivity upon RAD21 depletion.
Distance from the start site of gene is used as a reference point for binding locations (up- or downstream of the gene).

Sites of RAD21 and ER binding in closest proximity to dysregulated genes in estrogen-stimulated, RAD21-depleted MCF7 cells. The UCSC genome browser [Mar. 2006 (NCBI/hg18)] (e.g., while others that were normally repressed by estrogen signaling [31]).

Proximity of RAD21 and ER binding sites to genes that are dysregulated in estrogen stimulated, RAD21-depleted MCF7 cells.

Table 2

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>UCSC ID</th>
<th>Ref Seq summary ID</th>
<th>Chromosomal Location (including UTR)</th>
<th>ER binding site (s)</th>
<th>Neared RAD21 Binding site (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFF2</td>
<td>u0002zw1</td>
<td>NM_005423</td>
<td>Chr 21: 42,630,536–42,644,176</td>
<td>−15,310 kb</td>
<td>−15,310 kb +13,135 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+124 bp</td>
<td>+13,128 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+12,500 kb</td>
<td></td>
</tr>
<tr>
<td>S0X4</td>
<td>u003inl1</td>
<td>NM_003107</td>
<td>Chr 6: 21,701,951–21,796,828</td>
<td>−6,025 kb</td>
<td>−5,012 kb −4,558 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>−4,958 kb</td>
<td>−4,462 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+2,919 kb</td>
<td>+420 bp</td>
</tr>
<tr>
<td>PIM2</td>
<td>u004dls1</td>
<td>NM_006875</td>
<td>Chr X: 48,655,403–48,661,123</td>
<td>+6,945 kb</td>
<td>+1,824 kb +1,836 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFB55</td>
<td>u001ahb1</td>
<td>NM_003246</td>
<td>Chr 15: 37,668,572–37,676,960</td>
<td>−1,466 kb</td>
<td>−903 bp −1,201 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+629 bp</td>
<td>+653 bp +540 bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+2,481 bp</td>
<td>−22,419 bp −22,456 kb</td>
</tr>
<tr>
<td>IL20</td>
<td>u0010er1</td>
<td>NM_018724</td>
<td>Chr 1: 205,105,777–205,109,191</td>
<td>−744 bp</td>
<td>−810 bp −831 bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CKX1L2</td>
<td>u0001b1</td>
<td>NM_000890</td>
<td>Chr 10: 44,185,611–44,200,548</td>
<td>−1,461 bp</td>
<td>−1,457 bp 571 bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IRS2</td>
<td>u001popv1</td>
<td>NM_003749</td>
<td>Chr 13: 109,204,185–109,236,015</td>
<td>−17,365 bp</td>
<td>−4,810 kb +8,111 kb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>−6,031 kb −4,827 kb</td>
</tr>
<tr>
<td>DKK1</td>
<td>u001jjr1</td>
<td>NM_012242</td>
<td>Chr 10: 53,744,047–53,747,423</td>
<td>−4,480 bp</td>
<td>−987 bp −101 bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+103 bp</td>
<td>−5,496 bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMPR2</td>
<td>u002zuf2</td>
<td>NM_001204</td>
<td>Chr 2: 202,849,295–201,140,719</td>
<td>−14,198 kb</td>
<td>−1,354 kb −468 bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENG2</td>
<td>u003hjq2</td>
<td>NM_004554</td>
<td>Chr 4: 78,297,381–78,310,237</td>
<td>−21,677 kb</td>
<td>−36,252 kb −61,791 kb</td>
</tr>
</tbody>
</table>

Sites of RAD21 and ER binding in closest proximity to dysregulated genes in estrogen-stimulated, RAD21-depleted MCF7 cells. The UCSC genome browser [Mar. 2006 (NCBI/hg18)] (e.g., while others that were normally repressed by estrogen signaling [31]).

Sites of RAD21 and ER binding in closest proximity to dysregulated genes in estrogen-stimulated, RAD21-depleted MCF7 cells. The UCSC genome browser [Mar. 2006 (NCBI/hg18)] (e.g., while others that were normally repressed by estrogen signaling [31]).

Pros and cons for breast cancer progression [70].

Expression in T47D, broadly –52,71,72 [70].

Transcription of TFF1, TFF2, and BMPR2 in T47D cells is associated with increased expression of ERα, suggesting that this effect may be mediated by estrogen. The increase in ERα expression in T47D cells is consistent with previous observations [67,68].

In a previous study, we reported that RAD21 depletion prevented estrogen induction of the MYC locus, did not affect ER levels, and was accompanied by loss of ER binding to key regulatory sites [46]. Here we showed that RAD21 depletion enhanced estrogen-induced transcription of TFF1, TFF2, and BMPR2 in MCF7 cells. Like MYC, these genes are positively regulated by ER, therefore collectively, our results show that RAD21 depletion has divergent effects at loci that are normally transcriptionally induced by estrogen induction. The finding that cohesin can both positively and negatively influence ER-mediated gene transcription was unexpected, therefore we sought to better understand the mechanisms involved.

We found that the behavior of factors that govern transcription of the TFF gene cluster reflected the effects of RAD21 depletion, augmenting TFF1, TFF2, and BMPR2 transcription in MCF7 cells. Strikingly, RAD21 depletion from binding sites within the TFF gene cluster was accompanied by enhanced ER recruitment to regulatory sites within the cluster (including those normally co-bound by RAD21), and by release of RNPAP1 from pausing at TFF1 (Fig. 3B, C).

If we identify an estrogen-induced enhancer-promoter connection between BMPR2 and TFF1 that occurred significantly more frequently when RAD21 was depleted in MCF7 cells (Fig. 4, p < 0.0001). The BMPR2 and TFF1 sites that exhibited increased interaction normally recruit estrogen-induced RAD21 binding, which was lost upon its depletion by siRNA (Fig. 5B). Therefore, since enhancer-promoter interactions are thought to promote transcription [74,75], the increased interaction frequency we observed is consistent with upregulated TFF
transcription, enhanced ER recruitment and release of RNAPII pausing upon RAD21 depletion.

RAD21 depletion in MCF7 cells enhanced transcription of TFF2 much more than it did TFF1 (Fig. 3A), despite the increased chromatin interaction being between TMPRSS3 and TFF1. It is possible that other chromatin interactions not examined in our study contributed to elevated TFF2 transcription upon RAD21 depletion. In support of this idea, it has been reported that ER-anchored loops exist between TFF1 and TFF2 [76]. Moreover, TFF1 is highly enriched for enhancer-associated histone marks [40], and previous studies in mouse gastric cells show

**Fig. 3.** RAD21 depletion enhances estradiol responsiveness within the TFF cluster in MCF7 cells. Estrogen-deprived MCF7 cells were transfected with control or Rad21 siRNA for 48 h, then treated with vehicle or 100 nM estradiol for 3 and 6 h for RT-qPCR and 45 min for ChIP. (A) RAD21 knockdown in the presence of estradiol enhanced levels of transcripts within the TFF cluster. Transcript levels were determined using RT-qPCR relative to the mean of three reference genes (Cyclophilin, GAPDH and RPL13A; Fig. S4). Bar graphs represent the mean ± SEM from 6–8 independent experiments. Significance was determined using two-way ANOVA. (B) Schematic of the TFF cluster showing RAD21, estradiol-induced RAD21 and ER binding, mapped on the UCSC genome browser (2006 NCBI36/hg18 version) using previously published ChIP-Seq data in MCF7 cells [25]. Numbers represent the positions of the ChIP primers used to amplify various sites within the TFF cluster. (C) ChIP analysis of ER, RNAPII-S5-P and RNAPII-S2-P binding at specific regions within the TFF locus. Binding is shown relative to input chromatin. Bar graphs represent mean ± SEM from three independent experiments. Significance was determined using two-way ANOVA.
that TFF1 can activate TFF2 expression [77]. In addition, we observed enhanced recruitment of ER to the promoter of TFF2 upon RAD21 depletion, although this site is not co-bound by RAD21. It is therefore possible that altered chromatin structure across the TFF domain resulting from RAD21 depletion facilitated increased ER binding and enhanced transcription of TFF2 in MCF7 cells.

We had previously determined that RAD21 binds to the MYC locus gene promoters and a cognate enhancer 67 kb upstream of the MYC gene [46,66]. In contrast to our observations at the TFF cluster, RAD21 depletion decreased the frequency of interactions between the −67 kb enhancer for MYC and the MYC gene. This decrease in interaction strength is consistent with previous data that showed loss of ER binding to both the −67 kb enhancer and MYC, commensurate with downregulated MYC expression following RAD21 depletion [46]. However, since Myc expression is RAD21-dependent in a wide range of cell types and species [47], it is not clear whether the enhancer–promoter loop is causative or merely associated with downregulated MYC expression.

4.3. Model for the influence of cohesin on transcription of estrogen-responsive genes

We observed a wide diversity in altered transcription of estrogen-dependent genes upon cohesin depletion in MCF7 cells (Figs. 2, 3A, S4 and S5). We propose a hypothetical model that might explain the differential response of ER regulated genes to RAD21 depletion (Fig. 6).

Based on previous studies, it is likely that estrogen-responsive genes reside in functional chromatin domains that are organized by cohesin and CTCF [29,78,79]. It is therefore possible that domain integrity is compromised by cohesin depletion. RAD21 depletion may result in the loss of some intra-domain chromatin interactions, and as a consequence, provide opportunities for new interactions to form, and make available additional binding sites for ER. In the case of SOX4, a gene that gains sensitivity to estrogen upon RAD21 depletion, a new interaction with an ER-bound enhancer might be gained as a result of chromatin boundary disruption. In the case of the TFF gene cluster, overall chromatin domain decompaction as a result of RAD21 depletion may
strengthen some ER-associated loops and increase ER binding, thereby enhancing transcription. For \textit{MYC}, reduced enhancer–promoter communication associated with downregulation of transcription results from RAD21 depletion (Fig. 6). These effects can be compounded by the fact that cohesin interacts with chromatin modifiers that alter transcription and its depletion can alter histone modification patterns \cite{47, 80}.

An important further consideration is that cohesin is also likely to exert dose-dependent effects on transcription. In human syndromes resulting from cohesin mutations, small changes in the amount of functional cohesin can lead to developmental pathology \cite{11}, and in cancers, heterozygous mutations in cohesin genes appear to contribute to disease progression \cite{30, 33}. Furthermore, cohesin exerts biphasic effects on gene transcription in Drosophila \cite{81} thus; the same gene can increase or decrease its transcriptional response according to cohesin dose. It is thought that such biphasic effects might arise from the formation of alternative chromatin structures. Biphasic effects might explain in part why some of our data is inconsistent with reports showing that cohesin depletion globally decreased chromatin interactions and ER-mediated transcription in breast cancer cell lines \cite{67, 68}.

In summary, our results indicate that cohesin’s role in estrogen-stimulated transcription is unlikely to be simple. The complex effects on gene transcription that result from the interplay between cohesin and ER provide cells with the opportunity to respond differentially to estrogenic signals, in a gene- and context-dependent manner.

Overall, our data are consistent with the following conclusions: 1) that enhancer–promoter communication can be either augmented or abrogated upon depletion of cohesin; 2) that cohesin is not essential to anchor ER-mediated regulatory loops; and 3) that the normal function of cohesin might be to restrict formation of selected ER-anchored loops in a context-dependent manner.

Acknowledgments

We are grateful to Anita Dunbier, Mike Eccles and Antony Braithwaite for helpful discussions. Thanks to Nicholas Fleming and
Imogen Roth for critical reading of the manuscript. Research was supported by grants from the Breast Cancer Research Trust (JAH), The Cancer Society of NZ (JMO/JAH), The Health Research Council of NZ (JMO/JAH), and the Genesis Oncology Trust (JAH). JAH and JMO acknowledge support from the Cancer Society of NZ (JMO/JAH), The Health Research Council of NZ (JMO/JAH), and the Genesis Oncology Trust (JAH). JAH and JMO are investigators of the Maurice Wilkins Centre for Biodiscovery and Gravida Centre for Reproduction, Growth and Development.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbagrm.2014.12.011.

References


[40] V. Thumposiri, R. Stark, S. Menon, J.S. Carroll, CATS acts upstream of FOXL1 in modulating ERα binding by shaping enhancer accessibility, Genome Res. 23 (2013) 12–22.


