Visualizing the Structural Basis of Genome Silencing

by

Eden Margaret Fussner

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Biochemistry
University of Toronto

© Copyright by Eden Margaret Fussner 2012
Eukaryotic genomes must be folded and compacted to fit within the restricted volume of the nucleus. This folding, and the subsequent organization of the genome, reflects both the transcription profile of the cell and of the specific cell type. A dispersed, mesh-like chromatin configuration, for example, is characteristic of a pluripotent stem cell. Here we show that the acquisition of the pluripotent state during somatic cell reprogramming is coincident with the disruption of compact heterochromatin domains. Using Electron Spectroscopic Imaging (ESI), I made the surprising observation that the heterochromatin domains of the induced pluripotent and of the parental somatic cell contained 10 nm chromatin fibres. Since ESI generates projection images, the precise three-dimensional organization of all chromatin fibres within these domains could not be elucidated. To circumvent this limitation, I developed an electron microscopy technique that combines ESI with tomography. Using this approach, I found that both heterochromatin domains and the surrounding euchromatin of murine pluripotent cells, fibroblasts, and somatic tissues are in fact organized entirely as 10 nm chromatin fibres. This challenges the current paradigm that most, if not all, of the genome exists as 30 nm and higher-order chromatin fibre assemblies. Rather than transitions between 10 nm and 30 nm fibres, I propose that the organization and thus the regulation of the genome is achieved by the bending and folding of 10 nm chromatin fibres into discrete domains in a cell type-specific manner.
Preface

There once was a chromatin fibre
Endowed with a 30 nm diameter
Zigzagged or coiled
Had Scientists foiled
But its presence could never be doubted

So David an electron microscopist
With James of iPS fame
Set out to determine
The stubborn solution
A paper of greatest acclaim!

And so two students did toil
To find the 30 nm coil
They searched and they measured
Every cranny, each nook
But the project was mission impossible

Then David, the electron microscopist
Developed a startling hypothesis
This fibre it seemed
Despite lack of genes,
Had shrunk at least three full sizes

Then along came dear Mike from Max Planck
Who tilted and montaged….
Calculated, transformed and reconstructed
Every chromatin fibre within the tiniest cell
Now in exquisite 3-D detail!

Beyond all reasonable doubt
All the fibres about
Had diameters of simply 10 nm
The fibres it seems, had been feeling a bit stout
And had fasted since the late 70’s

So with this decree
Eden and Ugi earned their degrees
But now my dear reader, please do not despair
For the rest of this thesis
Is not in iambic pentameter
Acknowledgments

My deepest thanks to my adviser, mentor, supervisor, David Bazett-Jones, could fill volumes greater than the data within this thesis… David, I am most grateful for your patience, your kindness, your wisdom, your brilliant perspective, and your boundless enthusiasm for science – and mostly for sharing these wonderful attributes with me. When I felt discouraged and insecure, a visit to your office was the only remedy required. I can honestly say that you brought out the absolute best in me, and through your guidance, I am the scientist before you today. I feel humbled, as more than these words can express, so I offer the most sincere respect and thanks for all you have done.

To my advisory committee members, Drs. Craig Smibert and James Ellis, I thank you for your consistent efforts. You challenged me by setting high expectations, and with your help, I achieved those goals. Craig your words of wisdom from my undergraduate days continue to shape the way in which I communicate and divulge my scientific arguments. James, you are much more than just an outstanding committee member. In many ways I consider you my other advisor (aka the other master). I’ve enjoyed immensely our frequent and frank discussions. And I especially admire your zingy abstracts….

Many thanks go to all my labmates, both past and present. Team Bazett-Jones is an inspired, fun and engaged group and we really do finely vet one another’s data, ideas and arguments both freely and fully. Ren Li, my EM sensei; Lindsy Rapkin, random-thoughts-both related and unrelated to all things, scientific springboard; Reagan Ching, portable pubmed; David Anchel, instant smile; Kashif Ahmed, embryo maestro; Liron Even-Faitelson, new blood; and Vahideh Hassan-Zadeh and Calvin Tang, team ES chromatin!

To my closest collaborators Ugljesa Djuric and Dr. Mike Strauss, as with David I could go on and on here… Ugi, you are an incredible and brilliant scientist. I truly admire your ability to self evaluate and find the holes in your own (our own) work and in finding effective and creative ways to address them. You are the best thing that happened to me in graduate school (after having the dumb luck to end up in David’s lab that is) and I hope we get the opportunity to work together for many years to come. I dream that one day you, Mike and I will have neighboring labs! Mike, I am still in awe that you spent your honeymoon at the microscope with me. I will
never forget that first chromocentre reconstruction that was acquired and processed at 10PM, and the pure joy at seeing those chromatin fibres in three dimensions. I am also so very pleased that your lovely wife Leanne still likes me.

To my family, huge, fun, loud, brilliant, and creative – your love and support bolstered me to such great heights. Without your love, I cannot imagine having the strength, courage or perspicacity to have accomplished what I have. Specific thanks to my parents: Tim, Pamela, Susan and Wayne; and my brothers: Andrew, Dave and Pete. Dad, who would have imagined that clipping an article about naturopathy and bringing it with you to the health food store would have resulted in a lifetime obsession with academic science. There was a study published while I was in graduate school that correlated paternal support with the success of women in science – we have just added another supporting data point to their curve. Mum, science is draining and emotionally demanding, and every time I felt down you were there to shoulder my sorrow. Often, the mere sound of your voice made everything seem okay. Susan, my other mom, and Papa Wayne – I am so truly blessed – most kids only have two parents, but I have four, and your love and support was more than additive – it was synergy. To my brother, my “real brother” Andrew, it was always the two of us – you were the one constant growing up in my life. I can only imagine that it was annoying having a know-it-all goodie-goodie for a sister and so I especially appreciate all your love and support – it means the world to me. To my stepbrothers, Dave and Pete, thanks for the lifetime of torture; I mean love (smiley face). And to my grandparents, aunts and uncles, both genetic and pseudo – thank you thank you thank you.

Friends are the family you choose, and this family is pure awesomeness – I am so very rich indeed. Costin Antonescu, Allison Guy and Andrei Burcescu… you guys have been there through thick and thin and are the best friends anyone could have. Lindsy Rapkin, I feel I have known you all my life – when things were tough and spirits down – you were always there for me – and hilarity in fact ensued – vive la project hilarious! Jonathan Steels, founding member of the triumvirate – may we reign forever – or at least enjoy several more years of notoriety. With you, Dean Steels, and Prof. Antonescu, bravely blazing the trail ahead – I took solace in the idea that I might really be able to do this Ph.D thing! I love you guys. Nick Shah, perpetual cure for all things crustafarian, Shah 2009. Wesley Errington, there are not too many ideas in this thesis that I didn’t discuss with you in great depth before they ever saw the light of day. Thank you for your insights, patience, support and friendship. Mathew Estey and Jess DiCiccio, beers of
celebration, beers of disappointment, beers of student council frustration, beers of friendship – mmmm beer. Janice Lewis, you believed in me when I was dropping out of high school – you’ve heard it all – and for your eternal friendship I cannot thank you enough. Ugljesa Djuric, much, much more than an awesome scientist and collaborator, I am honored to call you a friend. And last, but by no means least, Justin Dupas; my graduate school bookend, political foe and longtime friend – you now hold the esteemed title of “world expert in all things Eden-un-togetherness”. Who would have guessed how this thesis business would inspire such pure lunacy… El amor enorme, apoyo, abrazos mágicos y besos, hizo un mundo de diferencia. Usted está un gran hombre y yo muchas gracias - camino encima y más allá. To all my friends in and through graduate school your thoughts, smiles, coffee breaks, and time spent; lightened my spirits, inspired and revived me – thank you all.
# Table of Contents

Preface................................................................................................................................. iii
Acknowledgments.................................................................................................................. iv
Table of Contents................................................................................................................... vii
List of Tables .......................................................................................................................... x
List of Figures ......................................................................................................................... xi
Abbreviations......................................................................................................................... xiii

Chapter 1 ................................................................................................................................ 1
  1 Introduction ....................................................................................................................... 1
    1.1 The hierarchy of DNA organization ................................................................. 1
      1.1.1 30 nm chromatin fibre models ............................................................... 2
      1.1.2 Methods to elucidate chromatin fibre structures....................................... 5
      1.1.3 Chromosome territories and nuclear organization..................................... 10
      1.1.4 Epigenetics and chromatin structure....................................................... 12
      1.1.5 Higher order chromatin fibre assemblies and gene transcription ............ 14
    1.2 Chromatin reorganization in cellular differentiation ........................................... 15
      1.2.1 Chromatin organization varies between cell types .................................... 15
      1.2.2 The unique nuclear organization of pluripotent cells .................................. 18
      1.2.3 Epigenetic changes associated with the acquisition of induced pluripotency ..... 20
    1.3 Summary and Hypothesis ....................................................................................... 21

Chapter 2 ................................................................................................................................ 22
  2 Materials and Methods ................................................................................................. 22
    2.1 Cell culture and tissue preparation ..................................................................... 22
    2.2 Immunofluorescence microscopy and immunostaining .................................... 22
    2.3 Fixation conditions ............................................................................................... 23
2.4 Correlative LM/ESI microscopy ................................................................. 23
2.5 Electron tomography .................................................................................. 24
2.6 Image analysis ............................................................................................... 25
  2.6.1 Optical methods to measure chromatin density ........................................ 25
  2.6.2 DAPI line scans ...................................................................................... 25
  2.6.3 Phosphorus density analyses ................................................................. 26
  2.6.4 Chromatin fibre measurements and Fourier Analysis ............................ 26
  2.6.5 Detection Limits of ESI ........................................................................ 26
2.7 Flow cytometry .............................................................................................. 27
2.8 Quantitative RT-PCR analysis ..................................................................... 27
2.9 Bisulfite sequencing ...................................................................................... 27
2.10 Chromatin Immunoprecipitation (ChIP) ..................................................... 28
2.11 Microarray analysis ..................................................................................... 29
2.12 Western blot analysis .................................................................................. 29
Chapter 3 ........................................................................................................... 30
  3 Constitutive heterochromatin reorganization during somatic cell reprogramming .... 30
  3.1 Introduction ................................................................................................ 31
  3.2 Results ......................................................................................................... 33
    3.2.1 Chromocentre compartmentalization differences in partial and full iPS cell lines ........................................................................................................ 33
    3.2.2 Chromocentres in full iPS cells are composed of 10 nm chromatin fibres........ 39
    3.2.3 Heterochromatin reorganization during 2i conversion of partial iPS cells ...... 47
    3.2.4 Heterochromatin reorganizes in Nanog high pluripotent cells .................. 50
  3.3 Discussion .................................................................................................... 53
Chapter 4 ........................................................................................................... 55
  4 “Open” and “closed” domains in the mouse genome are configured exclusively as 10 nm chromatin fibres .................................................................................. 55
4.1 Introduction........................................................................................................................................... 55
4.2 Results and Discussion ............................................................................................................................... 57
  4.2.1 30 nm chromatin fibre assemblies of starfish sperm nuclei....................................................... 57
  4.2.2 MEF chromatin within both constitutive heterochromatin domains and surrounding chromatin are exclusively 10 nm..................................................... 59
  4.2.3 Mouse tissue cells are comprised exclusively of 10 nm chromatin fibres........... 65
  4.2.4 Chromatin compaction versus condensation .............................................................................. 71
4.3 Concluding remarks ................................................................................................................................. 73
Chapter 5...................................................................................................................................................... 74
5 Perspectives and Future Directions ............................................................................................................ 74
  5.1 Significant findings ................................................................................................................................. 74
  5.2 Future directions ..................................................................................................................................... 75
    5.2.1 What are the factors that maintain chromatin in the characteristic open state in pluripotent cells? ...................................................................................... 75
    5.2.2 Is the pluripotent genome organized in an open configuration to provide genomic plasticity? ........................................................................................................ 78
    5.2.3 What is the relationship between epigenetic modifications and chromatin structure? ......................................................................................................................... 80
References....................................................................................................................................................... 82
List of Tables

Table 1. Primers used for qRT-PCR and Bisulfite sequencing ................................................. 28

Table 2. Calculation of chromatin content based on ESI phosphorus signal of individual nucleosome particles .......................................................................................................................... 45

Table 3. Packing and volume calculations of the 2 nm, 10 nm and 30 nm chromatin fibre....... 72
List of Figures

Figure 1. 30 nm chromatin fibre folding models. ................................................................. 3

Figure 2. Interphase chromatin described by a single polymer type ................................. 8

Figure 3. Diversity in chromatin organization observed in mouse nuclei. ...................... 17

Figure 4. Characterization of induced pluripotent stem cells. ........................................... 35

Figure 5. Microarray and pluripotent gene expression analysis of partial and full iPS cells. .... 37

Figure 6. Constitutive heterochromatin chromocentre organization differs between partial and full iPS cells. ................................................................. 38

Figure 7. Chromocentres in J1 ES, EOS3F-29, -28, -24 and parental MEF cells. .............. 39

Figure 8. Correlative LM/ESI procedure used to identify chromocentres in all cell types. .... 41

Figure 9. Chromocentres in full iPS cells are composed of 10 nm heterochromatin fibres. .... 42

Figure 10. Constitutive heterochromatin organization differs between partial and full Yamanaka iPS cells........................................................................................................ 44

Figure 11. J1 ES cells are comprised entirely of 10 nm chromatin fibres both within constitutive heterochromatin regions and in the surrounding euchromatin. ................................. 47

Figure 12. Sorting and characterization of the 2i-treated EOS3F-24 converted iPS cell line. .... 48

Figure 13. Characterization of 2i conversion of partial EOS3F-24 iPS cells into full iPS cells. 49

Figure 14. Heterochromatin reorganizes during 2i-mediated conversion of partial iPS cells..... 50

Figure 15. Nanog expression level impacts chromatin organization in J1 ES cells. .......... 52

Figure 16. Heterochromatin reorganizes in Nanog-high pluripotent cells in vivo in the early mouse blastocyst. .......................................................................................... 53
Figure 17. 30 nm chromatin fibres of starfish sperm nuclei are detected by ESI and ESI-tomography ................................................................. 58

Figure 18. ESI-tomographic analysis detects both 30 and 10 nm chromatin fibres in situ. ....... 60

Figure 19. 10 nm chromatin fibres populate both the surrounding chromatin and the constitutive heterochromatin domains in MEFs................................................................. 62

Figure 20. High-resolution images “beads-on-a-string” 10 nm chromatin fibres of a mouse embryonic fibroblast chromocentre. ................................................................. 63

Figure 21. Higher magnification images of chromatin fibres in a mouse embryonic fibroblast chromocentre showing both side views and end-on views of nucleosomes in situ......................... 64

Figure 22. Structures observed by cryo-fixation are consistent with chemical fixation.............. 65

Figure 23. Diversity of chromatin organization is evident in mouse tissues assessed by ESI. ... 66

Figure 24. Chromatin fibres in mouse tissue lymphocyte nuclei are comprised exclusively of 10 nm chromatin fibres................................................................. 67

Figure 25. Compact chromatin domains in mouse tissues are organized in 10 nm chromatin fibre assemblies ..................................................................................... 68

Figure 26. Chromatin fibre organization in both open and closed chromatin domains of liver nuclei in mouse tissue are comprised of 10 nm chromatin fibres................................. 69

Figure 27. Nanog expression in Epi-Stem cells correlates with a disruption in chromatin....... 77
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEK-GSK inhibitor</td>
<td>2i</td>
</tr>
<tr>
<td>Chromatin conformation capture assay</td>
<td>3C</td>
</tr>
<tr>
<td>Adenine- and thymine-rich</td>
<td>AT-rich</td>
</tr>
<tr>
<td>Chromatin immunoprecipitation</td>
<td>ChIP</td>
</tr>
<tr>
<td>Conventional electron microscopy</td>
<td>CTEM</td>
</tr>
<tr>
<td>4’,6-Diamidino-2-phenylindole</td>
<td>DAPI</td>
</tr>
<tr>
<td>Deoxyribonucleic acid</td>
<td>DNA</td>
</tr>
<tr>
<td>Electron microscopy</td>
<td>EM</td>
</tr>
<tr>
<td>ETn-Oct4-Sox2 vector</td>
<td>EOS</td>
</tr>
<tr>
<td>Embryonic stem</td>
<td>ES</td>
</tr>
<tr>
<td>Electron spectroscopic imaging</td>
<td>ESI</td>
</tr>
<tr>
<td>Fetal bovine serum</td>
<td>FBS</td>
</tr>
<tr>
<td>Fluorescence in situ hybridization</td>
<td>FISH</td>
</tr>
<tr>
<td>Glutaraldehyde</td>
<td>GA</td>
</tr>
<tr>
<td>Green fluorescence protein</td>
<td>GFP</td>
</tr>
<tr>
<td>Glycogen synthase kinase-3</td>
<td>GSK3</td>
</tr>
<tr>
<td>Tri-methylated histone 3 on lysine residue 36</td>
<td>H3K36me3</td>
</tr>
<tr>
<td>Term</td>
<td>Abbreviation</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Tri-methylated histone 3 on lysine residue 9</td>
<td>H3K9me3</td>
</tr>
<tr>
<td>Tri-methylated histone 4 on lysine residue 20</td>
<td>H4K20me3</td>
</tr>
<tr>
<td>Heterochromatin protein 1</td>
<td>HP1</td>
</tr>
<tr>
<td>Heat shock protein 70</td>
<td>Hsp70</td>
</tr>
<tr>
<td>Induced pluripotent stem</td>
<td>iPS</td>
</tr>
<tr>
<td>Kilobase</td>
<td>kb</td>
</tr>
<tr>
<td>Leukemia inhibitory factor</td>
<td>LIF</td>
</tr>
<tr>
<td>Light microscopy</td>
<td>LM</td>
</tr>
<tr>
<td>Methyl CpG binding protein 2</td>
<td>MeCP2</td>
</tr>
<tr>
<td>Mouse embryonic fibroblasts</td>
<td>MEFs</td>
</tr>
<tr>
<td>Mitogen activated protein/Erk kinase</td>
<td>MEK</td>
</tr>
<tr>
<td>Millimeter</td>
<td>mm</td>
</tr>
<tr>
<td>Nanometer</td>
<td>nm</td>
</tr>
<tr>
<td>Phosphate buffer saline</td>
<td>PBS</td>
</tr>
<tr>
<td>Paraformaldehyde</td>
<td>PFA</td>
</tr>
<tr>
<td>Quantitative polymerase chain reaction</td>
<td>qPCR</td>
</tr>
<tr>
<td>Quatitative reverse transcriptase polymerase chain reaction</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>Ribonucleic acid-protein</td>
<td>RNP</td>
</tr>
<tr>
<td>Transmission electron microscopy</td>
<td>TEM</td>
</tr>
</tbody>
</table>
Chapter 1

1 Introduction

1.1 The hierarchy of DNA organization

The groundbreaking discovery of the structure of the DNA double helix (Franklin & Gosling, 1953; Watson & Crick, 1953) prompted the question of how a 2-meter length of negatively charged polymer is packaged within the nucleus, having an average diameter of about 10 microns. Roger Kornberg biochemically isolated the chromatin subunit, consisting of a unique histone-DNA complex of discrete molecular weight and stoichiometry, which became known as the nucleosome (Kornberg, 1974). Electron micrographs from Don and Ada Olins and also from Chris Woodcock presented at the 1973 meeting of the American Society for Cell Biology clearly demonstrated that these same chromatin subunits corresponded to the beads, which they called nu bodies, of the “beads-on-a-string” morphology of the 10 nm chromatin fibre (Olins & Olins, 1974; Woodcock et al, 1976). X-ray diffraction, neutron scattering and energy loss electron microscopy (EM) experiments of these 10 nm fibres revealed that the DNA wrapped 1.75 turns around the outside of the nucleosome core particles (Bazett-Jones & Ottensmeyer, 1981; Hjelm et al, 1977; Luger et al, 1997; Richmond et al, 1984). The 10 nm chromatin fibre accomplished two functions: i) the positive charge on the histones of the core particle neutralized the negative charge on the DNA backbone, and, ii) the wrapping of the DNA around the histone core provided the first level of DNA compaction required to fit the genome into the confines of the nucleus. Although packaging DNA into a 10 nm fibre contributes to a 6-fold level of compaction, significantly more folding was needed to account for the condensation observed in mitotic chromosomes. It was not until Aaron Klug proposed that chromatin fibres formed higher-order structures with a regular periodicity that a mechanism for both DNA compaction and gene silencing based on chromatin structure really came to the forefront of chromatin biology (Finch & Klug, 1976; Levy & Noll, 1981). He proposed that the vast majority of the genome would be composed of 30 nm fibres, formed by coiling a 10 nm fibre into a solenoid, he and others suggested it could be further compacted into higher-order or chromonema structures of 200-300 nm in mitotic chromosomes (Rattner & Lin, 1985; Widom & Klug, 1985) or discrete
chromatin fibre assemblies ranging from 60-130 nm in interphase nuclei (Belmont & Bruce, 1994).

1.1.1 30 nm chromatin fibre models

Several models have been proposed to describe the molecular details of the 30 nm fibre, including the one-start solenoid (Finch & Klug, 1976; Widom & Klug, 1985), two-start helix zigzag (Horowitz et al, 1994; Woodcock et al, 1984), the cross-linker (Williams & Langmore, 1991) and supranucleosome (Zentgraf & Franke, 1984) models. The solenoid model involves coiling of the 10 nm fibre around a central axis of symmetry with nucleosomes packed face-to-face and 6 or 7 nucleosomes per helical turn. The two-start zigzag model predicts a zigzagging of two nucleosomes, which in turn, coil into a helical conformation. The cross-linker model is similar to the two-start model except that the linker DNA crisscrosses back and forth across the helical axis, with adjacent nucleosomes positioned across the axis from each other. The supranucleosome model involves clumps of nucleosomes separated by linker sequences. The last two models did not gain widespread acceptance. The cross-linker model was dismissed owing to the folding complexity and precision in linker length and nucleosome spacing that would be required to maintain these structures in vivo. Opponents to the aggregated supranucleosome model argued that these structures were simply partially unfolded 30 nm solenoid fibres (Felsenfeld & McGhee, 1986). A summary of these models are shown in Figure 1. Early evidence for each of these models was largely based on EM of buffer-extracted chromatin fibres and microccocal nuclease digested material. Current research efforts continue to explore the prevalence and structural details of both the solenoid and two-start helix zigzag models.
Figure 1. 30 nm chromatin fibre folding models.

A Cryo-EM micrograph of reconstituted nucleosome arrays supports the solenoid model of 30 nm chromatin folding, adapted from (Robinson et al, 2006). B conventional EM-tomography of 30 nm chromatin fibres in situ from starfish sperm chromatin, adapted from (Horowitz et al, 1994). C Shadow-casting EM of “Beads-on-a-string” 10 nm chromatin fibres, adapted from (Olins & Olins, 1974). D-G models of 30 nm chromatin fibres shown along horizontal (top) and
vertical axis (bottom) adapted from (Dorigo et al, 2004). D solenoid model, E two-start (zigzag) model, F cross linker model, and G supranucleosome model.

In addition to the high-resolution in situ EM studies (Horowitz et al, 1994), support of the two-start zigzag model is found in the crystal packing of the high-resolution structure of the tetranucleosome demonstrated that nucleosomes could be organized in a zigzag pattern (Schalch et al, 2005). This crystal structure is consistent with other biochemical studies including disulfide cross-linking studies of nucleosome arrays (Dorigo et al, 2004) and EM-assisted nucleosome interaction capture experiments combined with Monte Carlo simulations of nucleosome chain folding (Grigoryev et al, 2009). Together, these approaches support a two-start model of the 30 nm chromatin fibre. The rise and pitch of nucleosome in the asymmetric unit of the crystal structure (Schalch et al, 2005), however, predicts a 24 nm rather than a 30 nm fibre diameter structure (reviewed in Tremethick, 2007). Potentially indicating a crystal-packing artifact, rather than a true reflection of the in vivo organization of chromatin fibres.

Recent support for the more classical solenoid model comes from a molecular tweezers experiment (Kruithof et al, 2009). In this experiment, single force spectroscopy was used to measure the force required to convert a 30 nm chromatin fibre consisting of 25 nucleosome repeat arrays into a fully extended 10 nm fibre. This molecular tweezers experiment provides sub-pico-Newton force resolution, enabling precise mathematical modeling of the transitions between chromatin fibre configurations. The results indicate that the 30 nm fibre conforms to a regular helical structure, like a coiled spring, which obeys Hooke’s Law (which states that the extension of a spring is proportional to the applied force) (reviewed in Zlatanova et al, 2000). These molecular tweezers experiments are consistent with high-resolution EM studies of in vitro reconstituted nucleosome arrays (Robinson et al, 2006). However, although in vitro evidence supports both the two-start and the solenoid models, it is unclear how the regular spacing of the 30 nm solenoid chromatin fibre could accommodate the irregular nucleosome spacing that is present in situ (Kornberg, 1981). The two-start zigzag model can accommodate a variety of linker length sequences without steric hindrance between nucleosomes and thus convincingly mimics the irregularity of nucleosome spacing known to exist in situ (Horowitz et al, 1994).
Despite these many observations in cell-free *in vitro* systems that have suggested the existence of 30 nm fibres, evidence for such organization in cells remains scarce.

The first *in situ* evidence of the 30 nm fibre was shown in starfish sperm by both conventional EM studies (Horowitz et al, 1994) and by Electron Spectroscopic Imaging (ESI) (Bazett-Jones, 1992). These studies demonstrated that the 10 nm fibre was folded and twisted into a fibre that was approximately 30 nm in width. EM studies using extracted *Xenopus laevis* mitotic chromosomes demonstrated that higher-order folding of 30 nm chromatin is required to establish condensed mitotic chromosomes (Konig et al, 2007). By contrast, cryo-EM studies of mitotic chromosomes demonstrated that 30 nm fibres are absent in the fully condensed state of chromatin (Eltsov et al, 2008; Maeshima et al, 2010). The principal advantage of analyzing chromatin structure by cryo-EM is that the samples are fully hydrated and do not require chemical fixation with cross-linking agents; however, direct imaging can be difficult to interpret due to the low contrast intrinsic to biological specimens. A contrast transfer function algorithm was applied to the unstained cryo-bright field images of the mitotic chromosomes were used to derive accurate measurements of particle size and spacing. An 11 nm peak in the power spectrum supported an 11 nm diameter particle, but the absence of a peak at 30 nm indicated that no higher-order chromatin fibres are present within the mitotic chromosome. Because they found no evidence of higher-order chromatin fibres in the mitotic chromosome, either the 30 nm fibre does not exist in the interphase mammalian nucleus or the higher-order chromatin fibres are disrupted prior to the onset of mitosis, or the 30 nm fibres undergo a “polymer melt” transition (Eltsov et al, 2008).

### 1.1.2 Methods to elucidate chromatin fibre structures

Several biochemical-based approaches have been employed to elucidate the nature of chromatin fibre organization in mammalian nuclei. The vast majority of chromatin structural studies have utilized buffer and extracted samples or reconstituted nucleosome arrays using recombinant histone proteins and an engineered DNA template, mainly due to the technical limitations of studying chromatin structures *in situ*. In addition to direct contrast-enhanced EM studies of isolated chromatin described above, chromatin has been studied extensively by microccocal nuclease digest experiments, whereby higher-order chromatin structures are more refractory to
digest than the open or active 10 nm euchromatin structures (Griffith, 1975). In addition to these enzymatic assays, ultracentrifugation experiments have also provided a powerful in vitro cell-free system for elucidating higher-order chromatin structure, where 10 nm, 30 nm and higher-order oligomers can be easily resolved (Gilbert et al, 2004; Zheng et al, 2005). Studying chromatin assemblies at high-resolution in situ has presented a greater challenge, but some recent technical advances have provided some insights into native chromatin organizations.

A study utilizing the chromosome conformation capture (3C) technique reported that a sub-chromosomal domain within the yeast genome did not exist as a compact 30 nm fibre, but instead as an extended fibre (Dekker, 2008). The linear mass density obtained with this method, approximately 2 nucleosomes per 11 nm, was significantly less than that derived by DNA fluorescence in situ hybridization (FISH) experiments, also in yeast (Bystricky et al, 2004). The lower mass density, obtained by the higher resolution 3C method, is not consistent with a 30 nm chromatin fibre. More recently, Dekker and colleagues applied the Hi-C technique, an adaptation of the 3C method that allows for the unbiased identification of long-range chromatin interactions, to a human lymphoblastoid cell line (Lieberman-Aiden et al, 2009). The 3C technique (Dekker et al, 2002) and its extensions, 4C (Simonis et al, 2006) and 5C (Dostie et al, 2006), suffer from the same technical limitations because it is only possible to determine the long-range chromatin interactions from a locus of interest. Hi-C, however, avoids this limitation by digesting cross-linked DNA with a restriction enzyme to produce 5’ overhangs, which are filled with biotinylated nucleotides. These blunt ends are ligated together and isolated using streptavidin beads. The interacting fragments are then identified using massive parallel sequencing methods, providing enormous coverage as well as high resolution. With the Hi-C lymphoblastoid library, it was demonstrated that chromatin could be modeled as a polymer behaving as a fractal globule (Lieberman-Aiden et al, 2009). The interpretation is that chromatin exists as a series of globules, which cluster together forming yet larger globules, until the final stage of clumping represents a complete chromosome territory. This study supports a model of uniform chromatin structure based on a single fibre type, which can give rise to both open and closed chromatin compartments. Whether in a closed or open context, no change in the underlying chromatin configuration is detected. This model is consistent with the observations from both cryo-EM of mitotic chromosomes (Eltsov et al, 2008) and ESI of compact chromatin domains consisting of densely-packed 10 nm fibres (Ahmed et al, 2010) (Figure 2).
Figure 2. Interphase chromatin described by a single polymer type

Behaviour of chromatin within an interphase mammalian nucleus can be described by a compartmentalized model; made up entirely of 10 nm chromatin fibres. A Hi-C data can be modeled as a fractal globule, which is based on a single polymer type. In addition to chromosome territories, local domains of both open and closed chromatin are observed. Data in A adapted with permission from the AAAS from (Lieberman-Aiden et al, 2009). B Electron spectroscopic image of a region of a MEF (i) Chromatin is represented in yellow and protein and RNA-protein structures in blue, based on nitrogen and phosphorus content. Most of the chromatin is observed as short strands of 10 nm fibres that are represented in the thin (70 nm) section. The fibres are illustrated in the cartoon (ii) at higher magnification of the indicated region. This fibre cartoon is also superimposed (blue) on the chromatin map (white). (iii) Open and closed chromatin domains are observed in the nuclear interior, and closed domains also accumulate along the nuclear envelope (NE). Scale bar is 0.5 µm in (i) and 0.25 µm in (ii,iii). Entire figure is adapted from (Fussner et al, 2011a).

The giant loop model which describes the mega-base folding and organization of chromatin fibres in the interphase nucleus proposed by Trask and colleagues (Yokota et al, 1995) could accommodate chromatin organized as either 10 or 30 nm fibres. This model is not inconsistent with the fractal globule cluster model of open and closed compartments, which would comprise these giant constrained loops. However, several studies relying heavily on FISH, micrococcal nuclease digest methodologies, or cell-free in vitro biophysical assays, have assumed that the 30 nm fibre is the lowest order of chromatin folding in vivo (Belmont et al, 1999; Bystricky et al, 2004; Gilbert et al, 2004; Strukov et al, 2003). For example, studies in yeast using FISH have modeled the data to a flexible 30 nm-like polymer with a persistence length of approximately 200 nm and a mass density of approximately 150 bp/nm (Bystricky et al, 2004; Hagerman, 1988). These data, however, could be fitted to multiple models. I would argue that these light microscopy-based FISH studies, in contrast to 3C molecular biology approaches, lack sufficient resolution to rule out a highly bent and kinked 10 nm fibre versus a less bent or kinked 30 nm fibre.
To directly visualize chromatin fibres in interphase nuclei, higher-contrast imaging technologies that do not rely on heavy metal contrast agents are needed. In spite of the power of conventional transmission electron microscopy (CTEM), this technology is limited, as it unfortunately relies on heavy atom contrast agents, such as uranium acetate, since biological materials scatters electrons poorly. The contrast agents, however, create problems interpreting images because they do not bind all biochemical components with equal or linear affinity (Bazett-Jones et al, 2008). Thus, some structures appear denser and others less dense in a manner that does not reflect the actual biological mass density. Of particular relevance to nuclear studies, dispersed chromatin of either 10 or 30 nm in diameter, RNP-complexes and nuclear bodies are poorly contrasted in situ, whereas compact chromatin along the nuclear envelope or on the periphery of the nucleolus are well contrasted, although potentially exaggerated in their apparent chromatin density. Due to the low contrast inherent in biological samples and the requirement for contrast enhancing agents in CTEM it has not been possible to visualize 10 or even 30 nm chromatin fibres in situ owing to these technical limitations. Hence, the organization of DNA beyond the 10 nm fibre in vivo remains an important and unanswered question.

To address this longstanding question and overcome the limitations of CTEM it is possible to utilize ESI, which is a high-contrast technique that is not reliant on heavy metal contrast agents. This technique, based on electron energy loss spectroscopy, takes advantage of the fact that some incident electrons will ionize elements of the specimen. When this occurs, the incident electron will lose a characteristic amount of energy, dependent on the specific element. Hence with an imaging electron spectrometer, element-specific maps can be generated (Bazett-Jones et al, 1999; Bazett-Jones et al, 2008; Bazett-Jones & Ottensmeyer, 1981). This produces a dark field image of high contrast. Moreover, because of the high phosphorus content of DNA, a phosphorus map delineates chromatin from other nuclear structures. These phosphorus images can be used to visualize nucleosomes and even linker DNA of 10 nm chromatin fibres in situ (Ahmed et al, 2009; Efroni et al, 2008; Hiratani et al, 2010). For a complete description of ESI and the advantages and limitations of this technology see Chapter 2. Using ESI the global chromatin organization is readily detected and in many cell types. Although only 10% of the genome is thought to be euchromatin, the majority of the genome is not necessarily compacted into higher-order chromatin fibres, as demonstrated in a MEF, where much of the chromatin is highly dispersed in a chromatin mesh configuration, comprising 10 nm fibres (Figure 2).
Moreover, even in condensed chromatin domains typically defined as condensed heterochromatin, such as domains along the nuclear envelope or at the periphery of the nucleolus, 10 nm chromatin fibres are prevalent (Ahmed et al, 2010). Whether 30 nm fibres or higher levels of packing are also present in such domains would require the development of three-dimensional energy loss imaging techniques. If, however, 10 nm fibres are well represented in compact heterochromatin domains, the question arises as to whether 30 nm fibres are required at all for chromatin compaction in such domains.

1.1.3 **Chromosome territories and nuclear organization**

It remains unclear whether chromatin packing in interphase nuclei involves the obligate formation of 30 nm chromatin fibres. Other mechanisms of folding chromatin into meso-scale (mega-base chromatin clusters on the order of DNA replication domains) and large-scale (individual chromosomes) chromatin fibre structures could provide the compartmentalization and segregation required for functional organization of the nucleus. The large-scale organization of the genome includes the clustering of individual chromosomes into discrete structures during interphase known as chromosome territories. Carl Rabl postulated this clustering of individual chromosomes into distinct structures during interphase in 1885 (reviewed in Meaburn & Misteli, 2007). He speculated that this organization would be established in anaphase and would result in a linear orientation of specific chromosomes with a distal arrangement of telomeres and centromeres. This chromosome territory model is known as the Rabl conformation, and is evident in both plants and Drosophila (reviewed in Cremer & Cremer, 2001). This model of chromosome organization has since been elegantly confirmed by Cremer and colleagues in multiple studies employing high resolution imaging techniques and in a seminal experiment using a laser micro-irradiation approach (Cremer et al, 1983; Cremer & Cremer, 2001; Cremer et al, 2006). In higher eukaryotes, however, chromosome territories are positioned with a more globular and radial organization rather than in the classic extended Rabl configuration (Habermann et al, 2001). These territories have conserved radial positioning patterns, demonstrated by the more central location of the gene-rich chromosome 19, whereas the gene-poor chromosome 18 associates with the nuclear envelope (Croft et al, 1999) in mouse, chicken and human nuclei (Tanabe et al, 2002). The evolutionary conservation of these positional
preferences implies a conserved mechanism for establishing chromosome territories. A classic example of relative chromosome territory positioning and association is the clustering of chromosomes encoding the rRNA repeats. These chromosome territories associate during interphase through their nucleolus organizing regions (Henderson et al, 1972). The clustering of these specific chromosomes forms the basis for the establishment of the largest sub-nuclear structure, the nucleolus, the site of ribosome biogenesis. These studies indicate that chromatin is spatially segregated on a large-scale to achieve functional control of gene expression.

Evidence for both radial organization and pair-wise associations of chromosome territories alone is not sufficient to exclusively support a discrete model of chromosome territory organization. Robust support was derived from the coupling of 3D SKY FISH experiments, providing simultaneous multi-colour visualization of every chromosome territory, with advancements in image analysis (Bolzer et al, 2005). Fluorescence intensities in these experiments now served as a proxy for local chromatin density, which on the whole, was found to be significantly greater at the centroid of the chromosome territory.

There are two possible models to account for how gene expression may be subjected to unique regulation within distinct large-scale nuclear compartments. The first, termed the discrete chromosome territory is the predominant and widely accepted model for describing the behaviour of individual chromosomes in interphase (Cremer & Cremer, 2001). In this model repetitive or silenced densely-packed heterochromatin is confined to the restricted space of the territory interior. Conversely, gene-rich and active euchromatin regions would be located at the territory periphery, proximal to the inter-territory channels spatially facilitating access to regulatory nuclear sub-compartments. Gene-rich regions in some contexts are relegated to the periphery of chromosome territories, notably in the active versus inactive X chromosomes (Clemson et al, 2006). The behaviour of the Hox gene loci provides further evidence that chromosome territories may have an inaccessible core and an active periphery. The Hox gene loci loop sequentially out of their chromosome territory just prior to transcription activation, and this looping is thought to be crucial in establishing appropriate temporal expression of these critical developmentally regulated genes (Chambeyron & Bickmore, 2004). Together these data support the hypothesis that chromosome territory organization confers some degree of constraint on gene regulation. However, several examples of highly transcribed genes located within the interior of a chromosome territory have also been described, indicating that strict gene-
chromosome territory spatial relationships are not an absolute requirement for appropriate gene expression (Mahy et al, 2002).

The alternative model for chromosome territory organization is known as the *intermingling fibre model* (Branco & Pombo, 2006). This model does not require clear well-defined boundaries between chromosome territories, and does not impose strict chromatin-free channels filled with regulatory protein-bodies. Instead it allows for significant intermingling of chromatin fibres between territories, especially at the periphery of neighboring territory boundaries. In support of this model Pombo and colleagues utilized a combination of CTEM and FISH experiments and demonstrated, at least in human fibroblasts and peripheral blood cells, that significant intermingling of chromatin fibres did occur between adjacent chromosome territories and did so with regular frequency (Branco & Pombo, 2006). The implications of these data are that discrete channels enriched in regulatory factors would not separate individual chromosome territories and thus regulatory protein factors would need to be able to access both the peripheral and interior regions of chromosome territories. Consistent with an intermingling model of chromatin, at least at the periphery of territories, are studies involving fluorescently tagged proteins, which have demonstrated that the mobility of biological molecules within the nucleus are not in fact significantly constrained or excluded from specific sub-nuclear domains (Darzacq et al, 2007). The possibility remains that a combination of both models may represent chromosome territories in a cell type- or cell cycle- dependent manner.

1.1.4  Epigenetics and chromatin structure

Regardless of the absolute model of chromosome territory organization a distinct structural relationship exists between silenced and transcribed regions of the genome. Heterochromatin was initially defined by CTEM studies that revealed the compact chromatin domains present in many somatic cell types (Hyde, 1965). Large blocks of densely contrasted chromatin was observed at the periphery of the nucleus, later shown to be associated with the nuclear lamina, around the nucleolus and in specific cell types as blocks of chromatin within the interior of the nucleus (reviewed in Salina et al, 2001). Heterochromatin is characterized as always silenced, or constitutive heterochromatin, and facultative, or conditionally silenced.
We are now aware that an extensive network of post-translational modification of both histone proteins and DNA, collectively termed epigenetics, regulates gene expression. Epigenetic modifications can impact gene expression by altering the underlying chromatin structure through modulations of either 1) histone-DNA interactions or 2) nucleosome-nucleosome contacts. As a result, the original cytological definitions of heterochromatin are now known to be manifestations of epigenetic regulation (Jenuwein & Allis, 2001).

Epigenetic modifications dictate cell- and tissue-specific gene expression. Known epigenetic modifications of amino acids residues on histone tails are diverse and include: methylation, phosphorylation, acetylation, and ubiquination. The protruding tails on histone core proteins are extensively modified in vivo and multiple modifications of single nucleosomes define a complex histone “code” regulating both chromatin structure directly and by recruitment of histone-binding complexes (Jenuwein & Allis, 2001; Strahl & Allis, 2000). DNA modification occurs through cytosine methylation and more recently described hydroxymethylation (Tahiliani et al, 2009).

Histone and DNA modifications are associated with gene activation and silencing. A critical regulator of genome silencing is achieved through the modification of a key amino acid residue on the H3 core histone protein tail, lysine 9. A series of complexes are responsible for modifying this specific residue, Setdb1 (Loyola et al, 2009), G9a (Tachibana et al, 2002) and Suv39h1/2 (Peters et al, 2001), which ultimately gives rise to H3K9me3 modifications. H3K9me3-defined constitutive heterochromatin is widely thought to silence these modified genomic regions through the formation of 30 nm and higher-order chromatin structures. In addition to direct structural modulation imparted by this modification, H3K9me3 serves as a platform to enhance recruitment of proteins such as HP1 that are affiliated with genomic silencing (Fischle et al, 2005). Facultative heterochromatin is classically defined by H3K27me3 modified histone proteins, and is generated by the polycomb enzyme complex containing EZH2 (Cao et al, 2002), this modification is also thought to function through structural conversion of 10 nm chromatin fibres into 30 nm and higher-order chromatin structures. This was demonstrated in elegant shadow-casting EM experiments, Polycomb-complex binding alone in vitro directly impacts chromatin structure and dramatically enhances the formation of higher-order chromatin fibre assembles (Francis et al, 2004). A counterpart to polycomb-complex, trithorax complexes are chromatin regulators, which promote gene activation. Trithorax group
complexes acetylate chromatin (Jenuwein & Allis, 2001), increasing electrostatic repulsions between histone core proteins and associated DNA, casing local opening in the chromatin at specific nucleosome sites (Gould, 1997). Trithorax-group complexes are also principally responsible for writing the H3K4me modification, also classically associated with gene activation. In addition, this family of proteins utilizes ATP-dependent mechanisms to directly impact nucleosome spacing (Tamkun et al, 1992), potentially exposing regulatory elements to the transcription factor machinery. Together these modifications function to open up chromatin and allow access of the transcription machinery to the underlying DNA sequence. Various histone modifications can also be combined on individual nucleosomes to expand the available specificity and complexity of gene regulation. These histone modifications are modifiable, by enzymes that can erase even cytosine methylation marks (Kangaspeska et al, 2008; Reik, 2007). Thus, posttranslational modifications enable a dynamic and reversible genome regulatory system.

Integral to this concept of epigenetics is that these various post-translational modifications, either individually or in assemblages of multiple modifications and affiliated binding-partners, modulates chromatin structure directly, thereby influencing the gene transcriptional potential of specific genomic loci. However, it is currently not known how epigenetic modifications impacts chromatin organization into structures on a scale larger than the 10 nm fibre in situ.

1.1.5 Higher order chromatin fibre assemblies and gene transcription

Transitions between 10 and 30 nm chromatin fibres have been assumed to be required for gene regulation functioning through changes in accessibility of the DNA template to regulate factors and polymerase machinery. Support for a less accessible template for silenced chromatin is well supported by nuclease digestion experiments, where sensitivity of specific loci correlates with transcription activation, such as observed with the Hsp70 gene (Levy & Noll, 1981). Moreover, the compaction of artificial loci, such as tandem arrays observed by fluorescence microscopy with GFP-tagged reports, reveals changes in compaction of the arrays as a function of transcription or in their conversion to transcriptionally poised states (Hu et al, 2009; Muller et al, 2004; Muller et al, 2001). The measured size of such arrays has been modeled to a 30 nm and higher-order chromatin configurations. Similarly, distances obtained from transcription-
dependent decondensation of native chromatin fibres observed by FISH experiments have also been reconciled with an organization based on 30 nm chromatin fibres (Hu et al, 2009; Muller et al, 2004). Whereas these distances can be modeled by a 30 nm fibre, frequent and extensive bending and folding could produce the same degree of effective shortening. This would be accomplished by trans-acting factors that contribute to extensive bending, looping and cross-linking of 10 nm fibres. Likewise, regulatory protein factors or possibly structural RNA-based complexes could also function to minimize the space between 10 nm fibres. Indeed, factors that clearly modulate 10 to 30 nm transitions observed in vitro, such as methyl CpG binding protein 2 (MeCP2) (Nikitina et al, 2007) and heterochromatin protein 1 (HP1) (Daujat et al, 2005) chromatin binding factors, the presence and modification of linker histones (McBryant et al, 2010), and histone tail modifications of core histones (Kan et al, 2009), could all affect the degree of compaction of 10 nm fibres without a transition to the 30 nm fibre.

1.2 Chromatin reorganization in cellular differentiation

1.2.1 Chromatin organization varies between cell types

During embryogenesis, a series of epigenetic modifications occur in conjunction with the progression of a totipotent two-cell embryo to a pluripotent eight-cell state. Lineage commitment and cell fate decisions are established at the pre-implantation blastocyst stage. Cells isolated from these various stages provide a powerful in vitro cell culture model system for studying both pluripotency and differentiation. Embryonic stem (ES) cells derived from the inner cell mass of the 3.5 day embryo, for example, can be stably maintained in culture in the presence of leukemia inhibitory factor, and have the ability to self renew and differentiate into all the three germ layers. An advantage of working with ES cells is the circumvention of the technical challenges and limiting material associated with working with whole embryos. Since ES cells have the potential to become every cell type in the developing fetus, experimentation with this system may reveal crucial insights into fundamental biological processes associated with differentiation and the pluripotent state.

Epigenetic modifications are assumed to impact chromatin structure and organization. It is, therefore, not surprising, that tissues with unique transcriptional profiles and epigenetic
signatures also differ in their global chromatin organization. This has been convincingly demonstrated by both CTEM and ESI of various cell and tissue types. Four examples of global chromatin organization in mouse nuclei are shown in Figure 3. Although each nucleus contains the same genetic material, the nuclear configurations differ dramatically between cell types. In ES cells, for example, the chromatin is nearly uniformly dispersed with few or no regions of chromatin compaction (Figure 3A). The nucleoplasmic space, with the exception of the nucleolus, is filled with dispersed chromatin fibres (Efroni et al, 2008; Fussner et al, 2010; Hiratani et al, 2010). In contrast, in MEFs, chromatin is densely packed into constitutive heterochromatin domains or chromocentres, rich in major satellite repeat sequences. These domains are highly contrasted over the nucleoplasmic background, consisting of very dispersed 10 nm chromatin. The chromatin configuration of kidney epithelium nuclei contains both open dispersed and closed compact domains, with chromatin compacting at the nuclear envelope and at the periphery of the nucleolus. With a few chromatin fibres throughout the nucleoplasmic space between these compact domains. The lymphocyte represents an extreme opposite state relative to the chromatin configuration of an ES cell. These nuclei have very large, very densely packed chromatin in blocks at both the nuclear envelope and within the nucleoplasmic space. Hypothetically, if only a small fraction of the genome were to be modulated in response to tissue-specific gene transcription, than as a consequence only a small fraction of the genome would be structurally distinct. These striking differences make it tempting to speculate that chromatin configurations may play an integral role in defining specific cell-type states. Of note is the absence of compact chromatin domains in ES cells, in fact they are seemingly not required and uniform dispersed chromatin may be a hallmark of pluripotency.
Figure 3. Diversity in chromatin organization observed in mouse nuclei.

When analyzed using electron spectroscopic imaging (ESI) techniques chromatin structures range from an open 10 nm mesh-like configuration observed in embryonic stem (ES) cells A to the discrete domain organization of mouse embryonic fibroblasts (MEFs) B. Kidney nuclei have compact chromatin domains at their nuclear envelope C, which are more extensive in mouse lymphocytes D. Chromatin in all images is pseudo-coloured yellow and protein- or RNA-based regions cyan. Scale bar represents 0.5 µm in all panels. Figure is adapted from (Fussner et al, 2010).
1.2.2 The unique nuclear organization of pluripotent cells

Pluripotent cells have the unique and remarkable ability to both self-renew and give rise to any cell-type in the adult mammalian organism. Epigenetic mechanisms play pivotal roles to establish the pluripotent state, and in dictating the finely regulated differentiation program that occurs during embryogenesis. Large-scale epigenetic events accompany embryogenesis at specific developmental stages, such as the rapid global modifications in DNA methylation status (Monk et al, 1987), changes in histone variant dynamics (Chang et al, 2005), and deposition of linker proteins (Clarke et al, 1992). Understanding these epigenetic changes are essential to our understanding of developmental processes. These studies rely on indirect measurements of chromatin structure and of equal importance is direct visualization of large-scale chromatin organization in pluripotent cells to garner an understanding of structural features that define this developmental state.

To this end, Ahmed and colleagues examined the changes in the global chromatin configuration during lineage commitment in the early mouse embryo. Major changes from the one-cell to the E5.5 post-implantation stage were observed (Ahmed et al, 2010). At the eight-cell stage, for example, the chromatin landscape is characterized by highly-dispersed chromatin fibres and little or no indication of the need to form chromatin domains. Epiblast nuclei of the E3.5 embryo also display a dispersed 10 nm chromatin fibre distribution. It is not surprising that the chromatin distribution in nuclei of the E3.5 epiblast is indistinguishable from that of ES cells, having been derived from this embryonic stage and cell type. This is in contrast to neighbouring cells of the primitive endoderm or trophectoderm, which display blocks of compact chromatin along the nuclear envelope and in the nucleoplasmic volume. The dispersed chromatin phenotype in the epiblast requires expression of pluripotency factors since epiblast cells in embryos that do not express Oct4 display a chromatin configuration similar to the lineage committed trophoblast nuclei. Though the Oct4 null cells will eventually express trophoblast lineage specific factors, the change in chromatin occurs before they begin to express the differentiated cell-type transcription factors. Thus, changes in the global chromatin architecture precede exiting the pluripotent state and lineage commitment.
This unique “open” chromatin organization prevalent in the pluripotent epiblast cells of the mouse embryo is also reflected in the in vitro cell culture model system of the embryonic stem cell (Ahmed et al, 2010; Efroni et al, 2008; Hiratani et al, 2010). Consistent with this “open” chromatin state is the hyperdynamic association of proteins, such as HP1 and H1 linker proteins with chromatin (Meshorer et al, 2006) and an increase in the globally permissive transcriptional environment (Efroni et al, 2008) in ES cells. A more comprehensive understanding of this signature epigenetic landscape of pluripotent stem cells is emerging. Regulatory factors required in maintenance of heterochromatin in somatic cells, such as Suv39h (Peters et al, 2001), the enzyme responsible for generating H3K9me3 modified histones or even the DNA methyltransferase enzymes (Okano & Li, 2002) are not essential in ES cells, in fact these null mutants are in most respects indistinguishable from wild type ES cells. These mutations, consistent with their role in maintaining heterochromatin, however, do cause severe defects in differentiation and in exiting the pluripotent state. What maintains ES cell chromatin in this unique state is even more uncertain. However, a few potential key factors have been identified to date, notably Chd1, a chromatin remodeling protein known to regulate and associate with euchromatin. Chd1-knockdown results in an increase in prominence of heterochromatin domains, specifically chromocentres, suggesting that this remodeling protein is integral in maintaining the unique ES-cell chromatin state (Gaspar-Maia et al, 2009).

By maintaining the ES cell genome in an open chromatin configuration this may provide the genomic plasticity required to rapidly reorganize chromatin into cell-type specific configurations upon differentiation cues. Support for this hypothesis is derived from the presence of ES-cell specific bivalent domains. These domains are found at lineage-specific gene promoters and are marked with both the repressive H3K27me3 and with the active H3K4me3 histone modifications (Bernstein et al, 2006). This bivalency functions to effectively silence these lineage-specific genes, but allows for rapid activation of specific genes during development. Further evidence of an ES cell specific epigenetic landscape is evident in female mouse ES cells, which maintain two active X chromosomes and exhibit globally hypo-methylated DNA (Maherali et al, 2007).

In concert, these findings support the hypothesis that pluripotent cells have a unique epigenetic landscape that is less reliant on heterochromatin domains for viability then their somatic cell counterparts. However, a systematic study examining the precise chromatin structures associate with heterochromatin in pluripotent ES cells has yet to be undertaken.
1.2.3 Epigenetic changes associated with the acquisition of induced pluripotency

*In vitro* culture model systems for studying pluripotency are no longer relegated exclusively to ES cells derived from the early blastocyst. Pluripotent cells can also be generated by reprogramming somatic cells through the exogenous expression of four pluripotency-associated transcription factors, commonly referred to as the “Yamanaka” factors: Oct4, c-Myc, Klf4 and Sox2 (Takahashi & Yamanaka, 2006), transcription factors that are key regulators of the pluripotent state. This seminal discovery has provided a remarkable tool both to generate patient-specific pluripotent cells and to further our understanding of the fundamental biological processes that regulate the pluripotent state (reviewed in Plath & Lowry, 2011; Wu & Hochedlinger, 2011).

The generation of these induced pluripotent stem (iPS) cells requires large-scale global epigenetic changes to accompany reprogramming in order to reach the final fully reprogrammed iPS cell state (Maherali et al, 2007). The exogenous factors used to generate the iPS cells must be silenced, and the endogenous pluripotent transcription factor network is concomitantly reactivated (reviewed in Djuric & Ellis, 2010). Lineage-specific genes expressed in the parental cells are silenced and bivalent domains at these promoters are established. Female mouse iPS cells also undergo chromosome X reactivation and re-establish ES cell-like DNA methylation patterns (Sridharan et al, 2009). The addition of inhibitors that affect the regulation of heterochromatin marks often aid in the efficiency and establishment of iPS cells during reprogramming (Huangfu et al, 2008; Mikkelsen et al, 2008; Shi et al, 2008). Together, these studies have firmly established that induction of the pluripotent state requires epigenetic changes. However, for an iPS cell to fully recapitulate an ES cell, these changes would be expected to accompany dramatic chromatin rearrangements during the fibroblast to iPS transition (Figure 3). Whether these significant epigenetic modifications are sufficient to revert the iPS chromatin architecture to the characteristic “open” configuration observed in ES cells is a research area that remains to be explored.
1.3 Summary and Hypothesis

Conventional models of genome organization predict that the genome is largely organized into 30 nm and higher-order chromatin fiber-assemblies, which are refractory to transcription factors thereby silencing the vast majority of the genome. This model extends to pluripotent stem cells, despite a lack of evidence for compact chromatin domains characteristic of differentiated cell-types. My research seeks to address whether epigenetic regulation at the meso-scale level of genome organization plays a role in cellular reprogramming and to determine the precise structure of chromatin fibre assemblies involved in genome silencing in situ. I hypothesize that the pluripotent state is defined by a unique chromatin organization and that acquisition of the pluripotent state during somatic cell reprogramming will involve a transition to this disrupted chromatin architecture, which lacks discrete open and closed domains. In addition, I hypothesize that the mouse genome is organized exclusively as 10 nm chromatin fibres, and 10 nm fibres comprise both open and closed domains in somatic cells, without the requirement for 30 nm chromatin fibre or higher-order chromatin fibre assemblies.
2 Materials and Methods

2.1 Cell culture and tissue preparation

J1 ES cells and iPS cell clones were grown on feeders (mitomycin C treated MEFs day 15.5) in DMEM supplemented with 10% FBS, L-glutamine, penicillin-streptomycin antibiotics, sodium pyruvate, beta-mercaptoethanol, non-essential amino acids and recombinant LIF. EOS-generated iPS cell lines were maintained under 1 μg/mL puromycin selection in regular ES cell DMEM. Cells were passaged by trypsinization every 48 hours. 2i conversion was performed using MEK inhibitor (PD0325901, 0.5 μM) and GSK3 inhibitor (CHIR99021, 3 μM) (StemGent), represented as [1x] concentration in the text, in LIF supplemented ES media in the absence of puromycin selection.

Blastocysts were collected from uteri of ICR outbreed mice at embryonic day 3.75.

Primary mouse embryonic fibroblasts (MEFs) were isolated from day 15.5 embryos and cultured in standard DMEM and 10% FBS. Immortalized MEFs were cultured under similar culture conditions.

Patiria miniata testis and mouse tissues were isolated by dissection. Mouse tissues were cut into approximately 1mm³ prior to fixation, testis were cut after fixation.

2.2 Immunofluorescence microscopy and immunostaining

Cells were prepared for immunolabeling as described (Ahmed et al, 2009), except for blastocysts which were washed in PBS-Tween and permeabilized with 0.2% Triton-X 100. Primary antibodies used were: rabbit anti-H3K9me3, rabbit anti-H4K20me3 (both generous gifts from Dr. Prim Singh), human anti-CREST (Immunovision), rabbit anti-Nanog (Cosmobio), Cdx2 (BioGenex), Gata6 (R&D) and SSEA1 (Developmental Studies Hybridoma Bank). Secondary antibodies used were: donkey anti-rabbit or mouse Cy2, Cy3 and Cy5 (Jackson Laboratories) and PE-Cy5 (eBiosciences), Alexa488 and Alexa633 (Molecular Probes). Cells used for DAPI line
scan analyses were mounted in 1 mg/mL paraphenlenediamine and 1 µg/mL 4',6-Diamidino-2-phenylindole (DAPI) or soaked in 35 ng/mL DAPI. Images were collected on a Leica Microsystems DMRA2 microscope with a Hamamatsu ORCA-ER camera or on an Olympus IX81 inverted microscope with a Cascade II CCD (Photometrics) camera with or without the spinning disc. OpenLab 3.5.1 (Improvision) or InVivo (Media Cybernetix) was used to collect images. Fluorescence images of embryos were processed with Image-Pro Plus 6.2 3D-blind deconvolution software. All other images were processed with Photoshop 7.0 (Adobe) or ImageJ.

2.3 Fixation conditions

Preparing biological samples can have significant consequences on preserving chromatin fibre structures. In situ chromatin fibre structure measurements vary significantly depending on the sample preparation strategy (Belmont & Bruce, 1994; Dehghani et al, 2005; Horowitz et al, 1994; Langmore & Paulson, 1983). To ensure to the greatest possible degree that our sample preparations were not causal agents in disrupting higher-order chromatin fibre assemblies we employed two additional preparation strategies to our standard 2% PFA for 10 minutes at room temperature. Samples were also fixed in 2-2.5% glutaraldehyde (GA) overnight at 4°C in HEPES or PBS. These samples were then washed in PBS prior to dehydration and embedding in Quetol (EMS) or LRWhite (EMS). Tissues were all prepared in this manner, as GA fixation is widely accepted as the most rigorous and most biologically relevant fixation for preserving chromatin structure in biological specimens. Unfortunately GA fixation is not compatible with immunolabeling, so this fixation strategy cannot be employed for correlative LM/ESI experiments. Immortalized MEFs in addition to the fixation described above, were also fixed with GA as per the tissues. Additionally, these cells were also cryo-fixed these cells in liquid ethane, and then used freeze-substitution methods to dehydrate and infiltrate samples with resin.

2.4 Correlative LM/ESI microscopy

Sample preparation and detailed ESI procedure is described fully in (Ahmed et al, 2009). Cells were fixed with 2% paraformaldehyde for 10 minutes at room temperature, washed in PBS, permeabilized with 0.5% Triton-X 100 for 5 minutes, washed in PBS and blocked in 5% donkey serum for 20 minutes. Cells were labeled with primary antibodies overnight at 4 °C. Cy3 secondary antibodies were used to fluorescently label primary antibodies, as Cy3 is especially
amenable to correlative LM/ESI as the fluorescence is conserved throughout the embedding process and the fluorophore is easily visualized in the ultra thin 70 nm EM sections after embedding. Following immunolabeling with H3K9me3 or H4K20me3 cells were post fixed in 1% GA, dehydrated and embedded in Quetol (Electron Microscopy Sciences) or LRWhite (EMS) according to manufactures procedures. Samples were sectioned by an ultramicrotome (Leica) into 70 nm sections onto finder grids. Heterochromatin domains were identified after samples were prepared for EM using a Leica epifluorescence microscope with a 63X objective lens. Samples were carbon coated with 3 nm carbon films to improve stability.

Energy filtered electron micrographs were taken on a transmission electron microscope (Tecnai 20, FEI) operated at 200 kV and images were collected using a GATAN post column imaging filter at 120 and 155, and 385 and 415 eV to generate the phosphorus and nitrogen images, respectively. Pre- and post-edge images were recorded on a CCD camera. Digital micrograph software was used to collect images. Nitrogen images were subtracted from phosphorus images, so that the net nitrogen signal in chromatin structures were normalized to zero. These phosphorus subtracted nitrogen images are pseudo-coloured blue. The phosphorus images are pseudo-coloured yellow and overlaid onto the phosphorus subtracted nitrogen image. Images were processed with Digital micrograph and Photoshop 7.0 (Adobe).

2.5 Electron tomography

10 nm gold particles were used on section as fiducial markers for tomographic reconstructions, ensuring that 15-30 gold particles were well represented in every image field. Phosphorus ESI jump ratio tilt series of H3K9me3-enriched regions and surrounding chromatin of sectioned J1 ES cells, or of GA fixed tissue samples, were acquired at 120 eV and 155 eV using SerialEM (Mastronarde, 2005) with 2° increments over a tilt range of +/-60°. The images in the series were aligned and processed into jump ratios using a combination of IMOD (Kremer et al, 1996), SPARX (Hohn et al, 2007) and ImageJ. 3-dimensional phosphorus maps were reconstructed using the IMOD implementation of SIRT from the aligned jump ratio tilt series, previously demonstrated to maintain the quantitative information of ESI (Aronova et al, 2007). Rendering of the volumes was done using UCSF Chimera (Pettersen et al, 2004).
2.6 Image analysis

2.6.1 Optical methods to measure chromatin density

Constitutive heterochromatin was identified throughout these studies on the basis of the biochemical enrichment of H3K9me3, a classic modification associated with constitutive heterochromatin domains and enriched in chromocentres in MEF cells. In fact, H3K9me3 is a common feature of chromocentres in all analyzed cell types. However, H3K9me3 immunofluorescence microscopy could be problematic for making conclusions about differences in chromatin density between differentiated and pluripotent cell types. The potential problem exists if primary antibody access in compact versus dispersed chromatin differs by even a small factor. Moreover, we do not know whether chromatin modifications (such as H3K9me3) in constitutive heterochromatin are absolutely conserved upon reprogramming. It is possible that changes in the distribution of histone modifications between repetitive and non-repetitive chromatin sequences accompany the acquisition of the pluripotent state. Reorganization of the H3K9me3 mark would preclude it from being used to make conclusions about changes in chromatin density. On the other hand, even though DAPI binds AT-rich DNA sequences preferentially, its signal intensity is a better reflection of chromatin density than immunofluorescence signals of specific histone modifications. Consequently where image analysis was preformed on light microscopy images it was done on the DAPI signal. ESI, however, is superior to both immunofluorescence microscopy of histone modifications and DAPI imaging for measuring chromatin fibre density. Not only does ESI provide much higher spatial resolution, but it also measures chromatin directly, without the caveats introduced by stains or contrast agents.

2.6.2 DAPI line scans

DAPI line scan analyses were performed using ImageJ on optical sections where DAPI foci were at optimal focal planes. Pixel width depended on the microscope and camera used to acquire the images, but was selected to ensure approximately half the diameter of the average chromocentre analyzed was integrated into the signal. Histograms were generated through the foci and the background (outside the nucleus), which was subtracted from the nucleoplasmic and chromocentre signals. The ratio of the normalized signal intensity from within the chromocentre domain was calculated relative to the signal from the normalized nucleoplasmic DAPI signal.
Variations in these data were calculated as a ratio of chromocentre peak height to nucleoplasmic signal.

2.6.3 Phosphorus density analyses

Integrative phosphorus analyses of ESI images were preformed on unprocessed phosphorus jump ratio maps in ImageJ. The integrated average phosphorus intensities within H3K9me3-enriched regions were compared to the integrated average phosphorus intensity outside the H3K9me3-enriched region within the same image field. Phosphorus intensities were background subtracted using a small region within the field devoid of nucleic acid or phosphorus signal.

2.6.4 Chromatin fibre measurements and Fourier Analysis

Chromatin fibre dimensions were measured on tomographic slabs averaged from 4 tomographic slices. Measurements were made using Digital micrograph or ImageJ, pixels were converted to nm using a spacing grid measurement at individual magnifications.

To measure fibre densities and spacings, line scans were generated using Digital micrograph. Fourier analyses was carried out by applying an edge detection filter to central slabs of the tomogram and summing the power spectra of overlapping sub-areas of the most compact chromatin regions in each tomogram. The spectra were rotationally averaged using SPARX to yield a 1-dimensional trace, which were normalized by integral area and averaged to give a single representative profile of reciprocal distances. RNA-containing structures were avoided in these analyses.

2.6.5 Detection Limits of ESI

To calculate the total content of chromatin within a physical section and ultimately within a mouse diploid cell we used the phosphorus content of individual nucleosome particles in 70 nm sections of MEF nuclei to calculate the total phosphorus content of chromatin in the image field of the section. From the volume of the field (70 nm x 2.662 µm x 2.662 µm), we estimated the total chromatin content of a MEF nucleus, assuming a volume of $1 \times 10^3 \mu m^3$. We used the net P images that were obtained from ratio maps of post-edge/pre-edge of the $L_{II,III}$ P edge. Phosphorus signals of RNPs surrounding chromatin fibres were measured. These structures typically have a lower phosphorus density compared to chromatin. This value was subtracted
from the net P image, leaving primarily the phosphorus signal contributed by only the chromatin. Nucleosomes were identified by their morphology (“beads-on-a-string” separated by putative linker DNA) and size (circular or elliptical particles with a long axis of 10-13 nm). The particles used for integrated phosphorus signal measurements had to be well separated from neighbouring or adjacent particles. Integrated phosphorus densities were obtained with Digital Micrograph (Gatan Corporation). Total phosphorus content of the entire image of the nucleus in the field was obtained using the same software. If the nucleus did not fill the field, or if a region of the nucleolus or interchromatin granule cluster comprised part of the field, these regions were not included in the measurements, and the volume contributed by the remainder of the nucleoplasm was measured. The chromatin content of the field of the 70 nm section was calculated from the average integrated phosphorus signals of 10 nucleosomes. The total chromatin content of a nucleus was then estimated by multiplying by the ratio of total nuclear volume (estimated to be 1 x 10³ µm³) relative to the volume of the field of the physical section (70 nm x 2.662 µm x 2.662 µm).

### 2.7 Flow cytometry

Flow cytometry was performed as previously described (Hotta et al, 2009). Briefly, trypsinized cells were suspended in PBS with 5% FBS and were analyzed by FACScan using CellQuest software (Becton Dickinson) and data was analyzed with FlowJo (Tree Star).

### 2.8 Quantitative RT-PCR analysis

1 µg of RNA, collected using TRIZOL (Invitrogen), was reverse transcribed using SSII RT kit with random hexamer primers (Invitrogen) using manufacturer’s instructions. 50 ng of cDNA was used for qRT-PCR using SYBR green master mix (ABI) in triplicate. The primer sequences are shown in Table 2.

### 2.9 Bisulfite sequencing

Bisulfite sequencing was performed on DNA isolated from iPS lines at passage 10. DNA was bisulfite converted and purified using EZ DNA methylation Gold Kit (Zymo Research). 50 ng of converted DNA was subjected to PCR (See Table 2 for primer sequences). PCR reactions were subcloned into TOPO-TA vectors (Invitrogen) and sequenced.
Table 1. Primers used for qRT-PCR and Bisulfite sequencing

<table>
<thead>
<tr>
<th>Primer sequence</th>
<th>Target</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGA AGA AAT GCT GAA GTT GGA GAC</td>
<td>Rex-1 forward</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>AGT CCC CAT CCC CTT CAA TAG C</td>
<td>Rex-1 reverse</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>TCTTTCCACCCGCCCCGGGCTC</td>
<td>Oct4 forward</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>TGCAGCCGGACATGGGAGATCC</td>
<td>Oct4 reverse</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>CAGGAGTTGGAGGTTAGCTC</td>
<td>Nanog forward</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>CGGTTTTGACATGTAAGTC</td>
<td>Nanog reverse</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>GTGAAAGCTTGGAGGTGGC</td>
<td>Dnmt3b forward</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>GAGCAATGGAGGCAAGTCTCTG</td>
<td>Dnmt3b reverse</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>CTT TGT TGT CGG TGC TGA AA</td>
<td>Dppa3 forward</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>GGA AGA AAT GCT GAA GTT GGA GAC</td>
<td>Dppa3 reverse</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>CATGAGAAGCAAGTACTGGGCAAG</td>
<td>Sox2 forward</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>CCAACGATACACCTCGATGG</td>
<td>Sox2 reverse</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>GCTTGACAACTCTGGGTGCA</td>
<td>c-Myc forward</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>CCTTTTTTCGGAGACTAATAA</td>
<td>c-Myc reverse</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>CCAGCTTCTGCTGTTACACATGA</td>
<td>pMX-Klf4 forward</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>CTTTTATATACGTTGCCACC</td>
<td>pMX-Klf4 reverse</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>TCTCCATGCACTCAGACTG</td>
<td>pMX-Oct4 forward</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>CTTTTATATACGTTGCCACC</td>
<td>pMX-Oct4 reverse</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>CTGCAGGAGATGACAGC</td>
<td>pMX-Sox2 forward</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>CTCTTATATACGTTGCCACC</td>
<td>pMX-Sox2 reverse</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>GATTTTGTAGGGGGATTATGTAATTTGTAAATTT</td>
<td>Nanog forward</td>
<td>promoter bisulfite sequencing</td>
</tr>
<tr>
<td>ACCAAAAAACCCACACTATACATAATA</td>
<td>Nanog reverse</td>
<td>promoter bisulfite sequencing</td>
</tr>
<tr>
<td>GTATGGTTTTTTTGTGGTTTGGTTTTT</td>
<td>Oct4 forward</td>
<td>promoter bisulfite sequencing</td>
</tr>
<tr>
<td>CCACCCTCTAACCCTAAACCTCTAACC</td>
<td>Oct4 reverse</td>
<td>promoter bisulfite sequencing</td>
</tr>
<tr>
<td>TTATTTTGTAGGGTTTGGTTAGGTAGTTTAA</td>
<td>LTR forward</td>
<td>promoter bisulfite sequencing</td>
</tr>
<tr>
<td>ATATCCCTAACCACAATATCTCCC</td>
<td>LTR reverse</td>
<td>promoter bisulfite sequencing</td>
</tr>
</tbody>
</table>

2.10 Chromatin Immunoprecipitation (ChIP)

Chromatin immunoprecipitation was performed using the Upstate chromatin immunoprecipitation kit. Briefly, 1.0x10^6 cells were crosslinked in 1% PFA, cells were washed, scraped into a conical tube and sonicated for 25 minutes in the Bioruptor™ XL at the high setting with cyclings of 30 seconds ON, 30 seconds OFF. ChIPs were performed using EZ-ChIP kit (Millipore) using manufacturer's instructions with 4 µg of Anti-acetyl histone H3 (Upstate), DNA was purified using phenol/chloroform extraction and DNA was resuspended in 50 µL of H₂O. 1 µL of DNA was subjected to qPCR using SYBR green.
2.11 Microarray analysis

Total RNA isolated from different clones was labeled and hybridized to Affymetrix GeneChIP Mouse Gene 1.0 ST gene expression microarray using standard conditions. Intensity values from biological triplicates were processed by variance stabilization and normalization (Huber et al, 2002) and summarized by RMA (Irizarry et al, 2003), iPS expression profiles were contrasted both to MEFs and J1 ES cell profiles. Differentially expressed transcript in each contrast with an adjusted p-value less than 0.05 were determined using limma (Wettenhall & Smyth, 2004). The raw and processed data with additional information is available at the GEO database website under accession number GSE21595.

2.12 Western blot analysis

Protein was isolated from different clones with RIPA lysis buffer and run on an 8% SDS-PAGE gel. Proteins were transferred to a nitrocellulose membrane, blocked with 5% blocking buffer and probed with anti-Nanog (Cosmobio) and β-actin (Sigma) antibody at 1:500 and 1:2000 dilution, respectively. Secondary HRP-conjugated antibodies were used with SuperSignal West Pico detection reagents (Thermo Scientific).
Chapter 3

3 Constitutive heterochromatin reorganization during somatic cell reprogramming

Based largely on “Constitutive heterochromatin reorganization during somatic cell reprogramming” Fussner E*, Djuric U*, Strauss M, Hotta A, Perez-Iratxeta C, Lanner F, Dilworth FJ, Ellis J# and Bazett-Jones DP# EMBO J. 2011 May 4;30 (9):1778-89 (Fussner et al, 2011b). EF performed the EM and high resolution IF experiments and DAPI line scan analyses, collected and analyzed the tomography data. *These authors contributed equally to this work #Co-corresponding authors

Induced pluripotent stem (iPS) cell reprogramming is a gradual epigenetic process that reactivates the pluripotent transcriptional network by erasing and establishing repressive epigenetic marks. In contrast to loci-specific epigenetic changes, heterochromatin domains undergo epigenetic resetting during the reprogramming process but the effect on the heterochromatin ultrastructure is not known. Here, we characterize the physical structure of heterochromatin domains in full and partial mouse iPS cells by correlative Electron Spectroscopic Imaging. In somatic and partial iPS cells, constitutive heterochromatin marked by H3K9me3 is highly compartmentalized in chromocentre structures of densely packed chromatin fibres. In contrast, chromocentre boundaries are poorly defined in pluripotent embryonic stem and full iPS cells, and are characterized by unusually dispersed 10 nm heterochromatin fibres in high Nanog-expressing cells, including pluripotent cells of the mouse blastocyst prior to differentiation. This heterochromatin reorganization accompanies retroviral silencing during conversion of partial iPS cells by Mek/Gsk3 2i inhibitor treatment. Thus, constitutive heterochromatin is compacted in partial iPS cells but reorganizes into dispersed 10 nm chromatin fibres as the fully reprogrammed iPS cell state is acquired.
3.1 Introduction

The cascade of events in somatic cell reprogramming to a pluripotent state involves large-scale epigenetic remodeling to establish repressive epigenetic marks on tissue-specific genes, and to erase these marks on key members of the pluripotent network (Barrero et al, 2010; Maherali et al, 2007; Takahashi & Yamanaka, 2006; Wernig et al, 2007). Reprogramming proceeds through at least three stages: intermediate, partial and full-induced pluripotent stem (iPS) cell states (Jaenisch & Young, 2008; Yamanaka, 2009). Partial mouse iPS cells attain some aspect of pluripotency including ES cell-like colony morphology, teratoma-forming ability, and partial activation of pluripotency genes accompanied by downregulation of differentiation-specific genes (Mikkelsen et al, 2008; Okita et al, 2007; Sridharan et al, 2009). Partial iPS cells can be converted into full iPS cells using cell signaling or epigenetic inhibitors (Mikkelsen et al, 2008; Silva et al, 2008). As they complete reprogramming, full mouse iPS cells acquire epigenetic marks of pluripotency including X chromosome reactivation and genome-wide establishment of ES cell-like H3K27 and H3K4 trimethylation patterns (Maherali et al, 2007).

ES cells epigenetically inactivate exogenous retroviruses, similarly full iPS cells robustly activate endogenous pluripotency genes during the reprogramming process while silencing the retroviral reprogramming factors (Okita et al, 2007; Silva et al, 2008; Stadtfeld et al, 2008). In contrast, partial iPS cells maintain retroviral gene expression indicating that an ES cell-specific transcriptional network is required for successful silencing of retroviruses in pluripotent cells. Clearly, expression of Oct4, Sox2 and Klf4 in primary MEFs is not sufficient for retrovirus silencing, nor is their persistent expression in partial iPS cells, suggesting that additional endogenous pluripotency factors are required for retroviral silencing in full iPS cells. For example, Nanog, which controls the pluripotency ground state (Silva et al, 2009), is reactivated in full mouse iPS cells (Okita et al, 2007), and is capable of silencing gene expression (Liang et al, 2008). Epigenetic mechanisms normally employed in establishing heterochromatin have prominent functional roles in silencing retrovirus vectors in ES cells. Viral DNA sequences are recognized by a ZFP809/TRIM28 complex that establishes H3K9me2 marks bound by HP1 (Wolf & Goff, 2007; Wolf & Goff, 2009). In addition, H3 is deacetylated (Lorincz et al, 2001; Pannell et al, 2000) and H3K9me3 is deposited by Eset/Setdb1 (Matsui et al, 2010). Additional epigenetic involvement includes recruitment of Swi/Snf components (Golding et al, 2010) and participation of de novo DNA methylases (Cherry et al, 2000; Dodge et al, 2002). These
enzymes also reorganize chromatin structure on a global scale. As retroviral silencing occurs at late time points in the reprogramming process, it can be a useful marker to identify molecular changes that take place in the fully reprogrammed iPS cell state.

ES cells are known to have unique heterochromatin domain organization with hyperdynamic binding of histone and associated heterochromatin structure proteins (Meshorer et al, 2006). In contrast to loci-specific epigenetic changes compatible with altered gene expression, changes to the physical structure of heterochromatin domain organization during reprogramming remain unexplored. Mouse cells are unusual in that pericentric constitutive heterochromatin, comprised of major and minor satellite repeat sequences, cluster into structures known as chromocentres (Guenatri et al, 2004; Joseph et al, 1989; Wong & Rattner, 1988). These chromocentres are easily identified in mouse nuclei by their DAPI-rich staining, and are specifically marked by H3K9me3 and H4K20me3 (Peters et al, 2001). This clustering makes mouse heterochromatin an attractive model system for studying chromatin domain organization. Human cells, on the other hand, contain repetitive sequences that are distributed more evenly across the genome and in most contexts do not cluster to the same degree as in mouse cells. Chromocentre organization has typically been investigated by measuring the amount of clustering, or changes in the number of observable chromocentre foci, within differentiating nuclei (Brero et al, 2005; Meyer-Ficca et al, 1998; Tessadori et al, 2007). Notably, the number of these H3K9me3 foci increases when Chd1 is knocked down in ES cells and pluripotency is simultaneously lost (Gaspar-Maia et al, 2009). Thus, like retrovirus silencing, heterochromatin organization in ES cells also correlates with the pluripotent state. However, the timing of its reorganization during iPS cell reprogramming and the specific structures of the heterochromatin fibres remain to be identified.

As an alternative to using molecular biology approaches or visible light imaging, chromatin organization has been studied using conventional transmission electron microscopy (CTEM). Somatic cell nuclei imaged by CTEM reveals condensed “closed” chromatin domains along the nuclear envelope and at the nucleolus periphery (Belmont et al, 1989; Kireev et al, 2008). These silenced compartments, including chromocentres, are widely accepted to be comprised of 30 nm and higher-order chromatin fibre assemblies (Rego et al, 2008). Currently, evidence for 30 nm fibres in vivo is largely restricted to non-mammalian cell types (Maeshima et al, 2010; Tremethick, 2007; van Holde & Zlatanova, 1995). When visualized by ESI (Ahmed et al, 2009), the only technique that provides high contrast of unstained chromatin at high-molecular
resolution (Bazett-Jones & Ottensmeyer, 1981; Dehghani et al, 2005), the predominant chromatin configuration in ES cells is a mesh of “open” dispersed chromatin fibres, and displays a paucity of the blocks of condensed “closed” chromatin observed in somatic cells (Efroni et al, 2008). Upon differentiation of ES cells to neural progenitor cells, however, some of the dispersed chromatin becomes organized into compact heterochromatin domains, particularly along the nuclear envelope (Hiratani et al, 2010). In addition to cell culture models, we have shown that pluripotent pre-implantation embryos also have globally decondensed chromatin, most strikingly after the 8-cell stage (Ahmed et al, 2010). These observations led us to ask whether reprogramming is accompanied by the loss of compact heterochromatin domains at chromocentres to the more dispersed open structures that is typical of pluripotent ES cells.

The significant differences in global nuclear architecture between somatic and ES cells, implies that large-scale reorganization events may be involved in iPS cell reprogramming. Our primary goal was to identify the timing and structural features of heterochromatin reorganization during reprogramming. We compare H3K9me3 marked constitutive heterochromatin organization in full and partial iPS cells to that of the parental MEFs and the J1 ES cell line. We demonstrate that chromocentres in full iPS cells with high Nanog levels are characterized by dispersed open domains comprised entirely of 10 nm chromatin fibres, like those of ES and high Nanog expressing cells of the inner cell mass (ICM) of the mouse blastocyst prior to differentiation. Failure to disrupt the heterochromatin domains seen in MEFs by retaining tightly packed closed chromatin domains is a characteristic of partial iPS cells. Conversion of partial iPS cells with the mitogen activated protein/Erk kinase and glycogen synthase kinase-3 (Mek/Gsk3) inhibitor (2i) cocktail (Silva et al, 2008) shows that heterochromatin reorganization is a novel signature of the fully reprogrammed state that accompanies retrovirus silencing.

3.2 Results

3.2.1 Chromocentre compartmentalization differences in partial and full iPS cell lines

To study the relationship between global chromatin organization and cellular reprogramming events we compared MEFs, iPS and J1 ES cells. We previously reported the derivation of three
mouse iPS cell lines, EOS3F-24, EOS3F-28 and EOS3F-29, using pMX-based retroviral delivery of three “Yamanaka” factors, \textit{Klf4}, \textit{Sox2} and \textit{Oct4}. All three iPS cell lines were shown to maintain expression of the EOS-EGFP pluripotency reporter vector and have a pluripotent capacity to differentiate \textit{in vitro} and form teratoma \textit{in vivo} \cite{Hotta et al, 2009}. Interestingly, only the EOS3F-24 cell line maintained its teratoma-forming capacity after embryoid body-mediated differentiation, indicative of the inefficient differentiation capacity of partial iPS cells \cite{Jaenisch & Young, 2008; Okita et al, 2007}. In contrast, differentiation of both EOS3F-28 and -29 is accompanied by EOS-EGFP inactivation and subsequent loss of their teratoma-forming capacity \cite{Hotta et al, 2009}.

A reliable molecular marker for full reprogramming is retrovirus silencing, and partially reprogrammed cell lines fail to silence the retroviral transgenes. Indeed, EOS3F-24 iPS cells maintain high level of expression of all three endogenous reprogramming factor transgenes compared to the retrovirus silencing observed in EOS3F-28 and -29 cells (Figure 4A). Accordingly, bisulfite sequencing reveals that the LTR promoters of the reprogramming vectors are hypomethylated in EOS3F-24 iPS cells but hypermethylated in EOS3F-29 and EOS3F-28 (Figure 4B). Furthermore, the endogenous \textit{Oct4} and \textit{Nanog} promoters remain hypermethylated at CpG sites in EOS3F-24 but not in EOS3F-28 and -29 iPS cells (Figure 4B). Chromatin immunoprecipitation (ChIP) analysis with an anti-H3Ac antibody reveals that \textit{Nanog} and \textit{Oct4} promoters are hypoacetylated in EOS3F-24 compared to J1 ES and EOS3F-29 iPS cells while the viral transgenes are hyperacetylated compared to the EOS3F-29 iPS cells (Figure 4C). These data demonstrate that, in contrast to EOS3F-28 and -29, the epigenetic state of the viral transgenes and endogenous pluripotency genes have not been reprogrammed in EOS3F-24 iPS cells.
Figure 4. Characterization of induced pluripotent stem cells.

A qRT-PCR was performed on RNA from infected MEFs (MEF3F) and uninfected controls, EOS3F-29, -28 and -24 iPS cell lines using primers specific for pMX-Klf4, pMX-Oct4 and pMX-Sox2. B Bisulfite sequencing of CpG sites of endogenous Oct4 and Nanog promoters and pMX LTRs. Open and closed circles indicate unmethylated and methylated CpG sites, respectively. C ChIP analysis of histone H3 acetylation of the Oct4 and Nanog promoters and the pMX vector of EOS3F-29 and EOS3F-24 cells. J1 cells are included as a control for ChIP
analyses of endogenous promoters. Error bars represent +/- standard error of the mean of triplicate reactions.

To directly examine expression of endogenous pluripotency genes in the iPS cell lines in comparison to reference pluripotent J1 cells and MEFs global gene expression analyses with microarray chips was performed. These results show that the EOS3F-29 expression profile is similar to J1 cells (354 genes differentially expressed). In contrast, EOS3F-24 cells are distinct from both J1 ES cells (4987 differentially expressed genes) and MEFs (5621 differentially expressed genes) (Figure 5A). The distance relationships are further illustrated in the hierarchical clustering (Figure 5B) of the arrays. Importantly, pluripotency-related genes are expressed at significantly reduced levels compared to J1 ES cells only in EOS3F-24 iPS cells as validated by qRT-PCR (Figure 5C). Nanog is a strong marker of full reprogramming in the mouse system and is less easily detected in EOS3F-24 as analyzed by immuno-cytochemistry and western blot analysis compared to EOS3F-28 and -29 cells (Figure 5D). Furthermore, EOS3F-24 is lacking any detectable levels of the SSEA1 cell surface marker measured by flow cytometry. Collectively, these data show that EOS3F-28 and EOS3F-29 have the molecular marks of full iPS cells whereas EOS3F-24 has molecular features of partial iPS cells.
Figure 5. Microarray and pluripotent gene expression analysis of partial and full iPS cells.

A Microarray data summary of differentially expressed genes in iPS cell lines shows that EOS3F-29, -28 and 2i-treated and not EOS3F-24 iPS cells have expression profiles that are very close to J1 cells. Table indicates the number of genes over- or under-represented in different iPS cells compared to J1 ES cells and MEF controls. B Principal component analysis was performed to compare the microarray expression profiles of genes in MEFs, iPS (EOS3F-24, EOS4F-24 +2i, EOS3F-28, and EOS3F-29), and J1 ES cells. The left panel represents a scatter plot of the microarray expression profiles on the planes spanned by the first and second principal components. The right panel represents a dendogram of the hierarchiacal clustering of expression profiles. C Pluripotency-related gene expression in iPS cells compared to J1 ES cells as measured by qRT-PCR analysis. Error bars represent +/- standard error of the mean of triplicate reactions. D Nanog expression measured by immunofluorescence microscopy and western blot analysis of EOS3F-29, -28 and -24 iPS cell lines. Scale bar, 20 µm.

To determine whether the pluripotent versus differentiated state was reflected in global chromatin organization we compared parental MEFs, partial and full iPS cells and J1 ES cells. These cells all have physically confined heterochromatin regions enriched in H3K9me3 as seen by fluorescence microscopy (Figure 6A). These domains were identified as chromocentres in all analyzed cell lines on the basis of H4K20me3 enrichment, DAPI density and proximity to centromeres, visualized with CREST antisera (Figure 7). Analysis of optical z-stack images and two-channel line scans confirm that DAPI enrichment is always associated with H3K9me3 enrichment in these cells (Figure 6B). These data are consistent with previously published results indicating that both pluripotent and differentiated mouse cells have compartmentalized heterochromatin, which form chromocentres (Ahmed et al, 2009; Gaspar-Maia et al, 2009; Martin et al, 2006). Although H3K9me3 enrichment is useful for identifying chromocentres in all analyzed cell types, it cannot be used to measure chromatin fibre density. On the other hand, though DAPI binds AT-rich DNA sequences preferentially, its signal intensity is a more accurate measure of chromatin compaction than detection of specific histone modifications. Indeed, close inspection of DAPI stained nuclei revealed significant differences between differentiated and pluripotent cell types. A description of the advantages and caveats of these optical methods to
measure chromatin density is described in Chapter 2. DAPI line scan analysis demonstrate that MEF and partial iPS cell chromocentres appear as bright foci while those of J1 ES and full iPS cells contain lower DAPI signal relative to nucleoplasmic background (Figure 6B). These line scan analyses are supported by whole nucleus variance analyses which show a 9-fold increase in the signal variance between MEF feeder and J1 ES cells within the same image field. Thus, based on DAPI distribution chromocentres are poorly compartmentalized in ES and full iPS cells.

Figure 6. Constitutive heterochromatin chromocentre organization differs between partial and full iPS cells.

A H3K9me3 immunofluorescence and DAPI counterstain in iPS, J1 ES and MEF cells. White arrows through DAPI images indicate intensity line scan plot and direction. DAPI (red) and H3K9me3 (blue) line scans are shown below image panels. Scale bar is 5 µm. B Box plot analysis of 100 chromocentres from at least 30 different cells and 3 independent trials of the background to chromocentre DAPI ratio, asterisks indicate significance to $P < 0.001$. These DAPI line scan analyses were entirely consistent with a variance analysis. The average variance of feeder cells (MEF) was measured to be 48,477 +/- 19115 compared to the average variance of
J1 ES cells, which was 5,964 +/- 4005. Sixteen cells of each cell type were measured to calculate the variance.

**Figure 7. Chromocentres in J1 ES, EOS3F-29, -28, -24 and parental MEF cells.**

DAPI-enriched regions in all cell types are co-incident with H4K20me3 enrichment and are proximal to kinetochores as defined by CREST antisera binding, an antibody that binds to several components of the kinetochore complex. Scale bar is 5 µm in all image panels.

### 3.2.2 Chromocentres in full iPS cells are composed of 10 nm chromatin fibres

DAPI is not a good indicator of chromatin fibre density *per se* because it is also affected by AT-richness and nucleosome-repeat length. Therefore, to interpret differences in heterochromatin organization observed by DAPI counterstain we employed ESI. Indirect labeling of H3K9me3 and H4K20me3 (Figure 8) was used to demark the constitutive heterochromatin domains for imaging the underlying chromatin fibre organization by correlative light microscopy (LM)/ESI (Figure 8 and 9A). (Dellaire et al, 2004). Integrative phosphorus density analysis demonstrates
that the chromocentres of ES and full iPS cells are difficult to delineate from the surrounding chromatin (Figure 9B). Therefore, the presence of H3K9me3 is compatible with dispersed chromatin packing density. In contrast, chromocentres in partial iPS cells and MEFs were densely packed, displaying a significantly higher fibre density than the surrounding chromatin.
Figure 8. Correlative LM/ESI procedure used to identify chromocentres in all cell types.

H3K9me3 fluorescence from a physical section of a MEF was imaged by epifluorescence microscopy (top left). Mass image of same physical section (top right) was acquired on the Tecnai20 with the GIF at 155 eV. Overlaid images allow for accurate mapping of H3K9me3 regions of interest in MEFs (middle) with the corresponding high-resolution phosphorus (yellow) and phosphorus-subtracted nitrogen (blue) ESI micrograph of the identified H3K9me3 enriched region to the right. An additional example of correlative LM/ESI in iPS cell (bottom) demonstrating that the disrupted regions are also easily identified using this technique. Scale bar is 0.5 µm.
Figure 9. Chromocentres in full iPS cells are composed of 10 nm heterochromatin fibres.

Correlative LM/ESI with indirectly labeled H3K9me3 delineates chromocentres. **A** White-circles delineate H3K9me3 enrichment. Quantitative phosphorus and nitrogen ratio images were segmented to show chromatin in yellow and protein-based structures in blue. White arrows indicate length and direction of phosphorus line scan analysis shown below image panels. Scale bar, 0.5 µm. **B** Box plot analysis of the phosphorus density shows the distribution of chromatin
compaction within these cells, 50 chromocentres were measured from at least 30 different nuclei from 3 independent trials, asterisks indicate significant differences in fibre density, \( P < 0.001 \). C 10 nm fibres within constitutive heterochromatin in a pluripotent cell and a differentiated cell, chromatin fibres are shown in yellow with zoomed panels to the right. Zoomed regions are shown in non-enhanced image in white boxes. White arrows in the zoomed images indicate 10 nm fibres. Scale bar, 30 nm. D Bisulfite sequencing of CpG sites in major satellite repeats, open and closed circles represent unmethylated and methylated CpG sites, respectively. E Perspective image of Chimera generated model of tomographic reconstruction of H3K9me3-enriched region and surrounding chromatin of a J1 ES cell pseudo-coloured by electron density. Connected nucleosomes and intervening linker sequence of these 10 nm fibres clearly visualized in the zoomed panel to the right. Representative Fourier transform analysis showing an average chromatin fibre size at 10.84 nm.

To investigate whether this disruption of constitutive heterochromatin was a general feature of iPS cell reprogramming we analyzed previously characterized full (Ng-20D17) and partial (Fb-20A10) iPS cell lines generated independently (Okita et al, 2007). When analyzed by DAPI line scan analysis and integrated ESI-phosphorus analysis, the results from these cell lines were in full agreement with those of the EOS-derived iPS cell lines (Figure 10). The fully reprogrammed iPS cell line had disrupted heterochromatin regions whereas the partially reprogrammed cell line had well-delineated closed heterochromatin domains with open dispersed surrounding regions.
Figure 10. Constitutive heterochromatin organization differs between partial and full Yamanaka iPS cells.

A H3K9me3 immunofluorescence and DAPI counterstain in partial iPS cell line Fb-20A10 and fully reprogrammed Ng-20D17 iPS cells. Scale bar, 5 µm. Box plot analysis of at least 30 different cells, asterisks indicate significance to P < 0.0001. B Correlative LM/ESI of indirectly labeled H3K9me3 (white outline) delineates chromocenters. Quantitative phosphorus and nitrogen ratio images were segmented to show chromatin in yellow and protein-based structures in blue.

At higher resolution, we observed an abundance of dispersed chromatin within chromocentre domains of ES and full EOS3F-28 iPS cells that is difficult and sometimes impossible to
distinguish from the surrounding chromatin (Figure 9C). Surprisingly, we also observed a prevalence of 10 nm fibres within the compact heterochromatin domains of MEFs, where only 30 nm or higher-level fibre organization might be expected. Although unexpected, this result is consistent with cryo-EM studies of metaphase chromatin in situ, where only 10 nm chromatin fibres were detected using phase contrast algorithms (Eltsov et al, 2008). These observations can be quantitatively confirmed since the phosphorus signal is a direct measure of chromatin density (Bazett-Jones et al, 1999; Bazett-Jones & Ottensmeyer, 1981). In addition, all of the chromatin in a physical section can be detected by ESI (Bazett-Jones & Hendzel, 1999). We further demonstrate this by using the integrated phosphorus intensities of individual nucleosomes to predict total chromatin content of the image field and ultimately of an entire nucleus. These measurements predict a chromatin content of a MEF to be between approximately 2.4x10^7 and 6.8x10^7 nucleosomes (Table 3). Based on one nucleosome per 200 bp (~50 bp linker DNA), the diploid mouse genome containing approximately 5.4x10^9 bp (Genome.gov) would consist of approximately 2.7 x10^7 nucleosomes. This analysis correctly estimates the real chromatin content of a diploid mouse cell (Table 3). Previous low angle x-ray scattering (Langmore & Paulson, 1983) and EM with conventional heavy atom contrast agents of isolated chromatin experiments have indicated that the genome is comprised entirely of 30 nm and higher-order chromatin fibres (Gilbert et al, 2004; Sinclair et al, 2010). Our observation of 10 nm chromatin fibres in compact H3K9me3-marked heterochromatin structures challenges current models in which chromatin condensation is based on a transition between 10 and 30 nm chromatin fibres. Bisulfite sequencing shows that major satellite DNA repeats located in chromocentres are still highly methylated in the full iPS cells (Figure 9D), indicating that DNA methylation is also compatible with dispersed chromocentre formation.

Table 2. Calculation of chromatin content based on ESI phosphorus signal of individual nucleosome particles.
To confirm that these dispersed fibres were in fact 10 nm we combined ESI with tomography to generate a three-dimensional representation of the phosphorus content in ES cells (Figure 9E and Figure 11). This analysis clearly shows the 10 nm chromatin fibre comprised of both nucleosomes and intervening linker sequence. Both the H3K9me3-defined heterochromatin regions and the surrounding euchromatin were comprised exclusively of 10 nm fibres. Although we detect no 30 nm fibres in these dispersed chromatin regions, 30 nm fibres can be detected by ESI. For example, 30 nm chromatin fibres have been imaged by ESI in starfish sperm (Bazett-Jones, 1992). Fourier analysis of five tomograms containing both H3K9me3-enriched regions and surrounding chromatin demonstrates an average particle size of 10.8 nm with a complete absence of higher-order chromatin fibres of greater diameters (Figure 9E). The 3D analysis clearly demonstrates that 10 nm chromatin fibres exclusively populate both the heterochromatin regions and the surrounding chromatin domains in pluripotent J1 ES cells.

<table>
<thead>
<tr>
<th>Section</th>
<th>Average integrated P signal of 10 nucleosomes</th>
<th>Integrated P signal in the image field</th>
<th>Predicted P content of nucleus based on section volume:nuclear volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31.7 ± 5.3</td>
<td>1.49 X 104</td>
<td>3.0 X 107</td>
</tr>
<tr>
<td>2</td>
<td>30.9 ± 8.9</td>
<td>1.18 X 104</td>
<td>2.4 X 107</td>
</tr>
<tr>
<td>3</td>
<td>29.4 ± 3.0</td>
<td>1.41 X 104</td>
<td>2.8 X 107</td>
</tr>
<tr>
<td>4</td>
<td>31.1 ± 7.7</td>
<td>3.38 X 104</td>
<td>6.8 X 107</td>
</tr>
<tr>
<td>5</td>
<td>29.4 ± 4.7</td>
<td>1.41 X 104</td>
<td>2.8 X 107</td>
</tr>
<tr>
<td>6</td>
<td>44.0 ± 10.9</td>
<td>2.4 X 104</td>
<td>4.8 X 107</td>
</tr>
</tbody>
</table>
Figure 11. J1 ES cells are comprised entirely of 10 nm chromatin fibres both within constitutive heterochromatin regions and in the surrounding euchromatin.

Left H3K9me3-enriched region circled on a zero-tilt net phosphorus image. White circle indicates H3K9me3-enriched region. Middle Tomographic reconstruction slice of the same field shown on left, chromatin fibres are white. Right Zoomed region used to generate chromatin fibre images shown in Figure 9E. Scale bar is 0.5 µm in the first two panels.

Taken together these data indicate that dispersion of blocks of constitutive heterochromatin correlates with complete reprogramming, whereas partial iPS cells maintain compact constitutive heterochromatin domains resembling those in MEFs. This is the first demonstration of 10 nm chromatin fibres within constitutive heterochromatin domains.

3.2.3 Heterochromatin reorganization during 2i conversion of partial iPS cells

To address whether constitutive heterochromatin reorganization is indeed a feature of the acquisition of the fully reprogrammed state, we induced complete reprogramming of EOS3F-24 iPS cell line using a MEK/GSK 2i inhibitor cocktail (Silva et al, 2008). Preliminary readouts of successful conversion after one-week chemical treatments were evaluated by SSEA1 and EOS-EGFP activation in EOS3F-24 iPS cells. Three concentrations of the 2i cocktail resulted in a dosage-dependent emergence of a high EOS-EGFP expressing EOS3F-24 subpopulation of cells that was also SSEA1 positive (Figure 12). To investigate whether molecular changes at the DNA/RNA level were taking place in the newly derived SSEA1+ EOS3F-24 cells, we sorted for the high EGFP expressing cells and cultured them for an additional week in the presence of 2i. Following 2i treatment, the converted EOS3F-24 cells robustly activated pluripotency-related loci (Figure 13A) and silenced the reprogramming retroviral transgenes (Figure 13B). Accordingly, Nanog protein is abundant in the 2i-converted EOS3F-24 colonies (Figure 13C). Bisulfite sequencing demonstrates appropriate changes in DNA methylation at the LTRs and promoter regions of Oct4 and Nanog following the 2i treatment (Figure 13D). These expression patterns were reproduced in independently converted 2i-treated EOS3F-24 cells, and microarray
analysis on these samples strongly indicates that 2i treated cells cluster with ES cells and are as similar to them (368 differentially expressed genes) as the fully reprogrammed EOS3F-29 iPS cells (Figure 5A and B).

Figure 12. Sorting and characterization of the 2i-treated EOS3F-24 converted iPS cell line.

A Levels of EOS-EGFP expression in response to three concentrations of 2i inhibitor cocktail. The bar graph indicates the percent of EOS3F-24 iPS cells expressing high levels of EOS-EGFP, as gated in the left panel. B SSEA1 expression in 2i-treated EOS3F-24 before and 7 days after sorting for high EGFP expressing populations. SSEA1-positive, high EGFP expressing cells were sorted and cultured in the presence of 2i inhibitor cocktail. High EGFP and SSEA1 expression was confirmed after 7 days.
Figure 13. Characterization of 2i conversion of partial EOS3F-24 iPS cells into full iPS cells.

A qRT-PCR of upregulated endogenous pluripotency-associated genes during 2i conversion of EOS3F-24 iPS cells. B qRT-PCR of pMX-Klf4, pMX-Oct4 and pMX-Sox2 transgenes. C Upregulation of Nanog after 2i conversion of EOS3F-24 iPS cells demonstrated by indirect immunofluorescence microscopy, Nanog shown in red and DAPI inset in blue. Scale bar, 40 µm. D Bisulfite sequencing of CpG sites of endogenous and transgene promoters in 2i-treated EOS3F-24 cells.

To determine whether conversion of partial iPS cells was accompanied by global heterochromatin changes, we performed LM/ESI. H3K9me3-enriched regions were significantly more disrupted in the 2i treated cells but not in DMSO treated controls (Figure 14A). These trends were consistent with our LM observations, where the chromocentre DNA density relative to nucleoplasmic background, as measured by DAPI intensity, was significantly different
between the control and 2i treated cells (Figure 14B). Thus, DAPI line scan analysis supports the LM/ESI data and is a simple method to identify dispersed heterochromatin in converted and fully reprogrammed iPS cells. 2i treatment caused dispersion of densely packed heterochromatin fibres during the conversion of partial iPS cells. We speculate that this transition of heterochromatin domain reorganization occurs at a late stage in reprogramming as only the cells which have silenced the transgenes, a previously described late-stage event in iPS reprogramming, have disrupted heterochromatin.

**Figure 14. Heterochromatin reorganizes during 2i-mediated conversion of partial iPS cells**

A Correlative LM/ESI analysis of constitutive heterochromatin in DMSO control and 2i treated cells, H3K9me3 enriched regions outlined in white. B DAPI line scan analysis of 2i versus DMSO control cells indicates a significant reduction in chromocentre partitioning, asterisk indicates significance of $P < 0.001$. Scale bar is 0.5 µm.

### 3.2.4 Heterochromatin reorganizes in Nanog high pluripotent cells

Disruption of constitutive heterochromatin in the 2i-treated EOS3F-24 cells was accompanied by reactivation of endogenous Nanog expression. To determine whether Nanog levels correlate with heterochromatin reorganization, we analyzed ES cell colonies exhibiting Nanog variegation (Chambers et al, 2007) (Figure 15A). To assess chromatin compaction, we measured DAPI
intensity of chromocentres relative to nucleoplasmic intensity in relatively high and low Nanog expressing J1 ES cell nuclei. We observed a correlation of Nanog expression with heterochromatin compaction, where high Nanog expressing cells exhibited the most disrupted and least compartmentalized chromocentres (Figure 15B). The lowest Nanog expressing ES cells had significantly more compacted chromocentres, suggesting that heterochromatin reorganization initiates when Nanog is downregulated as, or even before, ES cells begin to differentiate. It should be noted, however, that Nanog-low ES cells still have more disrupted chromocentres than differentiated cells. Retrovirus silencing also occurs in the presence of Nanog in 2i-converted iPS cells but not Nanog-negative partial iPS cells. This is consistent with other reports showing that retroviral reprogramming factors are specifically silenced in Nanog-GFP positive iPS cell colonies (Nakagawa et al, 2008). These data show that high Nanog expression in pluripotent cells in vitro correlates with heterochromatin reorganization. Furthermore, withdrawal of LIF from J1 ES cells for 72 hours results in rapid reappearance of fully compact chromocentres detected by LM/ESI (Figure 15C). These results indicate that dispersed heterochromatin in pluripotent cells is reorganized into the compact structures seen in somatic cells as Nanog expression is lost at the onset of differentiation.
**Figure 15. Nanog expression level impacts chromatin organization in J1 ES cells.**

A Low magnification immunofluorescence of variegated Nanog expression in a J1 ES cell colony with DAPI shown to the right, scale bar 20 µm. B High-magnification and resolution representative image of a J1 ES cell field with high and low Nanog expressing cells (top left panel) and DAPI counterstain of a 0.3 µm z-stack series of these same nuclei. DAPI line scan analysis quantification is adjacent, 30 nuclei of each expression level were analyzed and asterisk represents significance of $P < 0.001$. White arrows indicate direction of line scan (shown above) through optimal DAPI z-stack. C Withdrawal of LIF from J1 ES cell culture results in the loss of typical smooth colony morphology of the ES cells (phase) and formation of compact chromocentres (ESI) within 72 hours. Scale bar, 0.5 µm.

To assess the relationship between Nanog expression and heterochromatin organization *in vivo* we took advantage of Nanog variegation that briefly occurs in the ICM of E3.75 blastocysts (Figure 16) (Yamanaka et al, 2010). Cells were binned according to Nanog expression levels, excluding lineage committed Gata6 and Cdx2 positive cells, which are no longer pluripotent (Figure 16A). Using a DAPI line scan analysis we found that Nanog levels correlate directly with heterochromatin compartmentalization and significant differences were observed between high and low Nanog expressing nuclei. Chromatin in both high and low Nanog positive nuclei were significantly less compartmentalized than the lineage committed Cdx2 or Gata6 positive cells (Figure 16B). Together these data indicate that in the absence of Nanog heterochromatin in lineage-committed cells during early development is organized in a compact structure. In contrast, high Nanog protein levels in pluripotent cells of the blastocyst directly correlate with their dispersed heterochromatin organization. These findings are consistent with the global analysis of chromatin structures associated with embryonic development and with the striking chromatin compaction of the ICM nuclei in the Oct4-null embryos (Ahmed et al, 2010). The implications of these data combined are two-fold: 1) The pluripotent transcription factor network regulates chromatin structure and dictates unique chromatin architecture of the pluripotent state and 2) chromatin reorganization precedes differentiation both *in vitro* and *in vivo*. 
Figure 16. Heterochromatin reorganizes in Nanog-high pluripotent cells *in vivo* in the early mouse blastocyst.

A Immunofluorescence of variegated Nanog (red), GATA-6 (green) and Cdx2 (blue) expression in E3.75 day early mouse blastocyst. Scale bar, 20 μm. B DAPI images of 3 optical sections above and below the optimal image plane, representative chromocentres in these cells are adjacent to IF image, z = 2.1 μm. White arrows indicate direction of DAPI line scans, shown below and quantified. Double-headed black arrows represent chromocentre DAPI peak height relative to nucleoplasmic background. Asterisk represents minimum significance to \( P < 0.005 \), from a minimum of 40 nuclei per category.

3.3 Discussion

We used LM/ESI to identify novel heterochromatin structure reorganization during iPS cell reprogramming. We demonstrate that constitutive heterochromatin in chromocentres transition from a very compact closed chromatin fibre domain in MEFs and partial iPS cells to a more open domain of loosely packed chromatin fibres in ES and full iPS cells. Although generally dispersed chromatin was previously observed in ES cells (Efroni et al, 2008; Hiratani et al, 2010), we show here that heterochromatin specifically enriched in H3K9me3 is composed entirely of 10 nm fibres in ES and full iPS cells. This is compatible with the general concept that pluripotent stem cells have more open chromatin structure to make the cells more responsive to
differentiation cues that they receive. We were surprised to observe 10 nm chromatin fibres in the very densely packed chromocentres of MEFs and partial iPS cells although we cannot exclude the possibility that 30 nm fibres can also be found in these structures. However, prevalence of 10 nm fibres in both compact and disrupted heterochromatin domains indicates the transition between open and closed chromatin domains involves, at least in part, transitions between closely packed and highly folded 10 nm chromatin fibres. This challenges the absolute requirement for transitions between 10 and 30 nm chromatin fibres in defining heterochromatin domains.

To directly demonstrate that heterochromatin reorganization specifically occurs in fully reprogrammed cells, we converted partial iPS cells into full iPS cells via 2i treatment. We thus confirm that 2i conversion is an effective means to complete reprogramming, and that during conversion heterochromatin reorganization coincides with other late events in reprogramming including retrovirus silencing and Nanog gene activation. These events are specific for ES and full iPS cells, and therefore are not induced by the combined expression of Oct4, Sox2 and Klf4 in the primary transduced MEFs or the partial iPS cells. This implies that heterochromatin reorganization is dependent on establishment of the endogenous pluripotency transcriptional network. Since endogenous Nanog expression is a reliable marker of mouse full iPS cells, we took advantage of its natural variegation in ES cells and blastocysts. These analyses show that high Nanog ES cells have the most dispersed heterochromatin, whereas low Nanog cells have less dispersed heterochromatin. These findings are confirmed in vivo during normal development, where heterochromatin is dispersed most in Nanog high cells and is more compartmentalized in Nanog negative differentiated cells. Thus, expression of Nanog, or its targets in the pluripotency network, correlates with late heterochromatin reorganization events in reprogramming. This finding is consistent with the known ability of Nanog both to activate pluripotent gene expression and to silence genes normally expressed in somatic cells. Comparing the gene expression differences between the partial EOS3F-24 iPS cells and the 2i-treated cells may in future identify candidate epigenetic pathways that participate in heterochromatin reorganization. Experimental manipulation of such pathways in partial iPS cells may direct chromocentre reorganization and establish the exact timing of heterochromatin disruption during the acquisition of the fully reprogrammed state.
Chapter 4

“Open” and “closed” domains in the mouse genome are configured exclusively as 10 nm chromatin fibres

Based largely on “Open and closed domains in the mouse genome are configured as 10 nm chromatin fibres” Fussner E*, Strauss M*, Djuric U, Li R, Ahmed K, Hart M, Ellis J and Bazett-Jones DP. Submitted and in editorial consideration at Nature. EF designed the experiments and collected and analyzed the tomography data.

Chromatin fibres, composed of DNA and histones, must accomplish the packaging of over $10^9$ base pairs of the mammalian genome while simultaneously participating in a regulatory role of keeping certain genetic loci in an “open” or transcriptionally competent state and other domains in a compact “closed” configuration. The conventional model is that compact chromatin domains consist of higher-order configuration of the 10 nm “beads-on-a-string” nucleosome fibre, beginning with the 30 nm chromatin fibre. Folding of 10 nm chromatin fibres into 30 nm fibres is commonly thought to regulate transitions between active euchromatin and repressed heterochromatin. Using three-dimensional chromatin imaging by combining electron tomography with electron spectroscopic imaging, we searched for in vivo evidence of 30 nm fibres in the most compacted heterochromatin domains in mammalian cells. Surprisingly, we observed only 10 nm chromatin fibres in mouse fibroblasts and tissues regardless of the degree of compaction in heterochromatin domains. We conclude that chromatin compaction depends on the degree of bending and spacing between 10 nm chromatin fibres. Given the absence of 30 nm fibres, genome organization and regulation does not involve transitions between 10 nm and higher-order chromatin fibre types.

4.1 Introduction

Prevailing models of genome organization use the 30 nm structure as the default fibre conformation in situ, of both silenced and bulk chromatin within the mammalian nucleus
According to these models, the 30 nm chromatin fibre is formed by either coiling 10 nm fibres into a solenoid (Kruithof et al., 2009; Robinson et al., 2006) or by folding 10 nm fibres into a zigzag conformation (Horowitz et al., 1994). Although a number of conventional models predict that the bulk of the genome is maintained in a structurally closed 30 nm chromatin fibre conformation (Belmont & Bruce, 1994; Gilbert et al., 2004; Naughton et al., 2010), *in situ* evidence for the 30 nm fibre is limited and largely restricted to non-mammalian cell-types (Fussner et al., 2011a; Maeshima et al., 2010; Tremethick, 2007; Woodcock, 1994). A definitive understanding of higher-order chromatin structure *in situ* has remained elusive, mainly due to limitations of direct imaging technologies of sufficiently high resolution. For instance, CTEM requires contrast-enhancing agents, which obscure the detail of individual chromatin fibres. ESI analysis on the other hand, of nuclear structures is particularly powerful as it provides both biochemical and structural information of *in situ* samples without the use of contrast enhancing reagents, which can obscure valuable fine structural detail. As an example, using ESI analysis, we were able to identify some 10 nm chromatin fibre structures even in regions of densely packaged chromatin fibres in MEF chromocentres (Chapter 3). However, previous studies using contrast enhancing reagents were unable to detect these fibre structures and resulted in measurements of chromocentre domains as exclusively higher-order chromatin fibre assemblies (Rego et al., 2008). In densely packed chromatin domains, resolving the absolute chromatin fibre configuration is often difficult or impossible, even when using ESI. This is due to the fact that ESI suffers from one of the same limitations of CTEM in which data is acquired as a projection image; i.e. multiple, often overlapping, chromatin fibres are projected onto a two-dimensional image plane. Although ESI has an exceptional capacity to resolve individual nucleosomes and intervening linker sequences, *in situ* it is limited to two-dimension image analyses. We were thus motivated to combine ESI (Ahmed et al., 2010; Ahmed et al., 2009; Bazett-Jones et al., 1999; Bazett-Jones & Ottensmeyer, 1981) with electron tomography (Aronova et al., 2007; Mastronarde, 2005), a technique that involves acquiring an image tilt series and computationally combining these data. This enables us to overcome the projection limitation of conventional ESI and yields three-dimensional (3D) images of chromatin fibres *in situ* (Fussner et al., 2011b) where nucleosomes and intervening linker regions of even overlapping 10 nm chromatin fibres can be resolved.
4.2 Results and Discussion

4.2.1 30 nm chromatin fibre assemblies of starfish sperm nuclei

Since previous studies have detected *in situ* higher-order chromatin structures in sperm nuclei of *Patiria miniata* (Bazett-Jones, 1992; Horowitz et al, 1994) we chose to use these nuclei to develop the combined ESI-tomography method. As expected, we observed exclusively 30 nm chromatin fibres within the mature sperm nuclei both by conventional ESI and by ESI-tomography. After tomographic reconstruction we observed that these 30 nm chromatin fibres are organized in domains where the fibres frequently exhibit parallel packing with a spacing of 20 nm between fibres (Figures 16 and 17).
Figure 17. 30 nm chromatin fibres of starfish sperm nuclei are detected by ESI and ESI-tomography

A Zero-tilt phosphorus ratio map of chromatin fibres in starfish sperm nucleus (white on black background). Arrows indicate location of protein-rich acrosome, found in mature sperm cells. Arrows indicate rotation between phosphorus map and SIRT reconstruction of tomogram, shown in B. B Tomogram of the entire field reveals the fibre organization, width and spacing in three-dimensions. Enlarged regions are shown in panels (boxes) C-E. In some regions, fibres are
parallel to the section plane. These regions were used to measure fibre width and fibre spacing, indicated by arrows and shown below in histogram. The scale bar is 0.5 µm in panels A and B and 125 nm in panels C-E.

4.2.2 MEF chromatin within both constitutive heterochromatin domains and surrounding chromatin are exclusively 10 nm

Since the 30 nm fibre and higher-order chromatin assemblies of starfish sperm are clearly discernable using ESI-tomography we sought to determine the structure and configuration of chromatin fibres in mammalian nuclei. The mouse genome provides a unique opportunity to assess the higher-order chromatin organization of closed chromatin domains. Major satellite pericentromeric repeat-sequences from multiple chromosomes cluster to form large, easily recognized, cytologically compacted chromocentres. These chromocentre domains, which are thought to be comprised of higher-order chromatin fibre assemblies, including 30 nm fibres (Rego et al, 2008) are enriched in the definitive marks of constitutive heterochromatin, such as H3K9me3 and H4K20me3 (Guenatri et al, 2004). These marks identify the chromatin domains for imaging by correlative immuno-fluorescence microscopy and ESI (Dellaire et al, 2004; Fussner et al, 2011b).

After tomographic reconstruction of MEFs we observed 10 nm chromatin fibres not only in the open domains, but surprisingly, within the closed heterochromatin chromocentre domains as well (Figure 18C-D and Figure 19). The chromatin fibres that make up these chromocentre domains are highly bent and an example of this severe bending is shown (Figure 19G). Fibres in both the open and closed domains, however, often cross over one another (Figure 19C-E), which in a two-dimensional ESI image may have been mistaken for higher-order or 30 nm chromatin fibres.
Figure 18. ESI-tomographic analysis detects both 30 and 10 nm chromatin fibres in situ.

A Phosphorus map of a starfish sperm nucleus, chromatin fibres (white on black background), and B central plane of the tomogram of an enlarged region in A (box). The arrow illustrates a rotation between the phosphorus map and the tomogram. The 30 nm chromatin fibres in the field (box) are oriented parallel to the plane of the section. C Phosphorus map and D central plane through the tomogram of a chromocentre and its surrounding field within the nucleus of a MEF. The chromocentre was identified by correlative fluorescence microscopy using an antibody against H3K9me3. The field represented in the tomogram (box) contains 10 nm
chromatin fibres that show a high degree of folding or bending. Scale bar represents 0.5 µm in panels A and C and 100 nm in B and D.
Figure 19. 10 nm chromatin fibres populate both the surrounding chromatin and the constitutive heterochromatin domains in MEFs.

A Zero-tilt phosphorus ratio map of chromatin fibres (white on black background) in a region containing a chromocentre. Fiducial arrows indicate orientation of the chromocentre in the phosphorus map and the tomogram in B. B central plane of the tomogram of the same field shown in A where individual chromatin fibres can be resolved, in contrast to the zero-tilt phosphorus maps. Enlarged regions of the tomogram (boxes) are shown in panels C-E. C Chromatin fibres outside the chromocentre (“open” chromatin) are all 10 nm fibres. D, E Chromatin fibres in the “closed” chromocentre domain are all 10 nm. F Three-dimensional rendering view (Chimera) of chromatin fibres shown in panel D with enlarged region (yellow box) showing a highly bent continuous fibre in G. The two panels show different orientations of the fibre to confirm its continuity and separation from other fibres. Scale bar is 0.5 µm in panels A and B and 170 nm in panels C-E.

To confirm that these fibres were in fact the canonical 10 nm fibre of the “beads-on-a-string” morphology we acquired higher-magnification data sets from MEF chromocentres where individual nucleosomes and even intervening linker DNA could be readily detected in three dimensions after SIRT tomographic reconstruction (Figure 20). At this resolution both end-on and side views of individual nucleosomes could be resolved in these data (Figure 21).
Figure 20. High-resolution images “beads-on-a-string” 10 nm chromatin fibres of a mouse embryonic fibroblast chromocentre.

Chimera enhanced phosphorus SIRT reconstructions from a tomographic slab shown in panels A and C with strings of nucleosomes shown in cartoon representations in panels B and D. Nucleosomes are false coloured yellow and intervening linker DNA in purple.
Figure 21. Higher magnification images of chromatin fibres in a mouse embryonic fibroblast chromocentre showing both side views and end-on views of nucleosomes in situ.

A Entire field of chromocentre reconstructed and rendered with Chimera (chromatin in white on black background). B-D phosphorus map reconstructions showing 10 nm chromatin fibres with individual nucleosomes and intervening DNA linker sequences showing examples where both side views (top panel) and end-on views (bottom panel) of individual nucleosomes are clearly resolved (boxes).

We considered the remote possibility that chemical fixation may have caused a disruption of 30 nm into 10 nm chromatin fibres. This was unlikely since the same fixation and embedding procedures were used in preparing the starfish sperm specimens, where 30 nm chromatin fibres were observed (Figure 17 and 18A) (Bazett-Jones, 1992). Regardless, we examined the chromatin structure of chromocentres in specimens that were cryo-preserved followed by freeze-substitution and resin embedding. These cyro-fixed MEFs also displayed 10 nm chromatin fibres both within the chromocentres and in the surrounding nucleoplasmic regions (Figure 22). Furthermore, we observed 10 nm chromatin fibres in the neighbouring somatic cell nuclei or the starfish testis.
Figure 22. Structures observed by cryo-fixation are consistent with chemical fixation.

A Phosphorus-rich chromatin (yellow) and nitrogen-rich protein (blue) maps of a field containing a cryo-fixed MEF chromocentre. Regions containing chromocentre (box) and field in which higher-resolution and magnification image is shown in B white arrows indicate regions where 10 nm fibres are readily detected. Scale bar is 0.5 µm in A and 0.2 µm in panel B.

4.2.3 Mouse tissue cells are comprised exclusively of 10 nm chromatin fibres

We then asked whether the genome of MEFs, comprised entirely of 10 nm chromatin fibres, was simply a consequence of culturing these cells in vitro. We thus analyzed a variety of mouse tissues by ESI (Figure 23) and selected both spleen lymphocytes and liver cells for tomographic analysis. The global chromatin organization of MEFs is primarily represented by highly dispersed 10 nm chromatin fibres, but also comprises discrete, compact domains, which include chromocentres and blocks of chromatin along the edge of the nucleolus and nuclear envelope. In contrast, lymphocytes have an exceptionally high representation of compact chromatin domains throughout the nucleus (Figures 23-25). We thus expected that this cell type would provide the best opportunity for observing 30 nm and higher-order chromatin fibre structures, assuming that such fibres are required for forming compact chromatin domains. After tomographic analysis we observed that these densely packed, compact chromatin domains were exclusively configured as 10 nm chromatin fibres (Figure 25 and Figure 26A-C). To confirm our image analysis we applied a Fourier analysis, enabling the global calculation of chromatin fibre sizes in three dimensions, if sufficiently represented within the image to be detected in Fourier space. The Fourier analysis supports the image analysis, showing a frequency peak correlating to 9.2 nm chromatin fibres in these very densely packaged domains.
Figure 23. Diversity of chromatin organization is evident in mouse tissues assessed by ESI.

A Phosphorus-rich chromatin (yellow) and protein-rich nitrogen (blue) maps of a mouse cardiac cell nucleus B mouse glial cell nucleus; C mouse lymphocyte cell from spleen tissue with the most compact chromatin domains observed in any of the cell types analyzed. D An example of a mouse liver cell nucleus. Both open (white arrow) and closed (red arrow) chromatin domains are well represented in these nuclei, scale bar in all image panels is 0.5 µm.
Figure 24. Chromatin fibres in mouse tissue lymphocyte nuclei are comprised exclusively of 10 nm chromatin fibres.

A Zero-tilt phosphorus map of a spleen lymphocyte nucleus. Compact domains (white arrowheads) are prevalent in these cells both at the nuclear envelope and within the nucleoplasmic space. A nuclear pore (arrows) is indicated in both the zero-tilt and the rotated tomogram shown in B. B a central plane of the tomogram of the same field shown in panel A where now each individual chromatin fibres can be resolved. RNA containing structures (arrowheads) in the tomogram are visible in the extra-chromosomal spaces. Enlarged regions (boxes) are shown in C-E where densely packed 10 nm chromatin domains are surrounded at their periphery with RNA containing structures. The scale bar is 0.5 µm in panels A and B and is 220 nm in panels C-E.
Figure 25. Compact chromatin domains in mouse tissues are organized in 10 nm chromatin fibre assemblies

A Phosphorus map of a lymphocyte nucleus from spleen tissue. Compact chromatin domain (box) is represented in a central plane of the tomogram B, fiduciary arrows indicates rotation between panels. B The fibres in these compact domains are entirely 10 nm. C A phosphorus map of a region of a mouse liver cell. Region of chromatin along the nuclear envelope (box) is presented as a slice through a central plane of the tomogram D. D The tomographic reconstruction reveals 10 nm chromatin fibres within the compact chromatin domain as well as dispersed surrounding 10 nm chromatin fibres. Scale bar, 0.5 µm in panels A and C and 150 nm in B and D.
In addition to tissue spleen lymphocytes we performed tomographic analysis of liver cells, where both open (dispersed chromatin) and closed (compact chromatin) domains are well represented (Figure 25 and 26). Measurements of tomographic reconstruction once again reveal that both the open and closed domains of the liver cell are comprised of 10 nm chromatin fibres (Figure 26C-E).

**Figure 26. Chromatin fibre organization in both open and closed chromatin domains of liver nuclei in mouse tissue are comprised of 10 nm chromatin fibres.**

A Zero-tilt phosphorus image (white on black background) of a liver nucleus from mouse tissue, open chromatin (red arrowhead) and a closed domain (white arrowhead) are highlighted. Closed chromatin domain (arrow) in both the zero-tilt and the rotated tomogram shown in B. B A
central plane of the tomogram of field shown in A. The chromatin fibres within both the open and closed domains are now distinguishable and found to be 10 nm. Enlarged regions (boxes) in panels C-E reveal 10 nm chromatin fibres in both open chromatin C and in closed domains D, E. White arrowheads indicate rough ER. The scale bar is 0.5 µm in panels A and B and 210 nm in panels C-E.

The absence of 30 nm chromatin fibres in somatic cells, while surprising, is consistent with recent studies by both the Dekker laboratory (Dekker, 2008; Lieberman-Aiden et al, 2009) and Eltsov and colleagues (Eltsov et al, 2008). In a whole-genome analysis using 3C molecular biology methodologies, the Dekker group found that yeast chromatin is best modeled by a 10 nm chromatin fibre (Dekker, 2008), and using the high-C technique, they found that both open and closed domains of a lymphocyte cell line were best modeled by a single chromatin fibre type rather than a hierarchy of fibre types (Lieberman-Aiden et al, 2009). In addition, cryo-EM images of cryo-preserved mitotic chromosomes obtained by Eltsov and colleagues revealed only 10 nm chromatin fibres (Eltsov et al, 2008). A formal possibility raised by the authors is that the 10 nm chromatin fibres populating the mitotic chromosomes could have arisen from “polymer melting”, in which similar frequencies of inter-fibre and intra-fibre contacts could disrupt the 30 nm fibre conformation. Our demonstration of only 10 nm chromatin fibres in chromatin domains of widely differing densities argues against the “polymer melt” explanation.

Despite the fact that the bulk genome is organized into 10 nm chromatin fibres, distinctive domains of compact, biochemically marked heterochromatin, and more open domains are evident – most striking in the MEF cells (Figure 19) – indicating that the genome does in fact organize into discrete, structurally defined, domains. However, what ultimately dictates how these domains are formed or maintained remains to be determined. The fibres within the more compact heterochromatin regions are more densely packaged and folded (Figure 19), indicating a role for factors that modulate fibre-fibre distances and fibre bending, two parameters that would establish local domains of compaction.

In the spleen lymphocytes, where we observed the highest levels of compact chromatin, we also observed the most pronounced fibre bending (Figures 24 and 25C). One might speculate that the
per cent volume of the nucleus occupied by chromatin may reflect the degree of fibre bending. However, in the relatively small volume of the ES cell, where it might be expected that chromatin compaction is required due to the low volume, no compact chromatin domains are apparent. Yet in the significantly larger volumes of the differentiated tissue nuclei, chromatin is highly compacted into discrete domains. Since large regions within these nuclei are depleted of chromatin, a primary role of chromatin condensation apparently is not to “fit” the genome into the cell nucleus, but serves to create distinct structural and functional domains. When carefully considered, these data are surprising in that the variation of chromatin organization differs to such a degree between cell types with essentially identical genomes.

4.2.4 Chromatin compaction versus condensation

If the highest order organization of chromatin in the interphase mammalian nuclei is solely contributed by 10 nm chromatin fibres, then how does the genome “fit” within the confines of the cell nucleus? Higher order fibres in the mitotic chromosome are thought to be required for the effective shortening of the DNA in the chromosome by a factor of approximately 10,000 (Alberts et al, 2008). Because the nucleosome provides a compaction ratio of 6-fold, additional higher order fibres, such as solenoids, are thought to be required (Table 3). Unfortunately, the term “compaction” is ambiguous in this context. In one sense, chromatin does not compact DNA. The volume of DNA in a diploid human genome is 6.4 µm³, only 5.7% of the nuclear volume. DNA complexed as nucleosomes, i.e. 10 nm fibres, actually takes up a greater volume, 20.6 µm³, because it is wrapped around a histone octamer core, a major component of the core particle’s volume. As 10 nm chromatin, DNA occupies 18.2% of the nuclear volume. Likewise, DNA organized as 30 nm chromatin occupies an even greater 28.5 µm³. Hence the role of chromatin is not to “compact” the DNA to fit into the nucleus. Rather its role is, first to balance the charge on the DNA, and second, to allow the DNA to become highly folded on itself through wrapping and bending. Although the 30 nm fibre could greatly increase the “folding” of DNA, the same degree of linear compaction could also be achieved by frequent bending and kinking of the 10 nm fibre.
I would argue that the degree of bending of 10 nm chromatin fibres together with mechanisms that permit 10 nm fibres to come into close proximity to each other, is sufficient to create both “open” and “closed” domains with widely differing chromatin densities. A highly bent or kinked linker DNA between nucleosomes would not be consistent with the persistence length of naked DNA measured \textit{in vitro} (Hagerman, 1988). However, Widom and colleagues have demonstrated that the persistence length of linker DNA is highly susceptible to modulation by salts and protein factors, thereby allowing consecutive nucleosomes to come into close proximity and contact each other (Yao et al, 1990). Thus in the crowded nuclear environment a 10 nm fibre could readily fold on itself to generate chromatin domains \textit{in vivo} that vary widely in their degree of compaction.
4.3 Concluding remarks

The primary role of histones is to allow DNA to be compacted into the cell nucleus. They do this by balancing the negative charge of the DNA and by decreasing the effective length of the genome (Fussner et al, 2011a). The conventional model of chromatin is that higher-order fibres, such as the 30 nm chromatin fibre, augment the shortening function of the nucleosome itself. Alternatively, frequent folding or bending of the 10 nm nucleosomal fibre could accomplish the same thing. Our data supports the latter. However, the protein factors which regulate transitions between 10 and 30 nm chromatin studied with in vitro model systems (Kan et al, 2009; McBryant et al, 2010) are very likely the same factors responsible for modulating the degree of folding and spacing of 10 nm chromatin fibres that we observed in both cultured MEF and mouse tissue cells. Future studies will address this question.
Chapter 5

5 Perspectives and Future Directions

5.1 Significant findings

Pluripotent stem cells are characterized by a unique “open” chromatin architecture. In this thesis I provided evidence that disruption of “closed” chromatin or compact domains is concurrent with the acquisition of the fully reprogrammed iPS cell state. This phenomenon may represent a significant biomarker of pluripotency, since partially reprogrammed iPS cell lines maintain compact constitutive heterochromatin and fail to disrupt these domains. The disruption of constitutive heterochromatin in pluripotent cells, however, is not dependent exclusively on the absence of epigenetic modifications classically associated with these domains. In all iPS cell lines, both partially and full reprogrammed cells, chromocentres contain H3K9me3 and H4K20me3 modified histones and in male cell lines the underlying DNA sequence is hypermethylated. Yet, only the fully reprogrammed iPS cells have disrupted chromocentre domains. The implications of these data are, at least in pluripotent cells, that specific epigenetic modifications alone are insufficient to dictate the degree of bending, folding and packing of chromatin fibres in pluripotent cells. The presence of these constitutive heterochromatin-associated modifications and associated proteins, such as HP1, indicate that pluripotent-specific factors or factors with pluripotent specific functions, are required to counterbalance the compaction effect that these modifications have on chromatin structures in somatic cells. These findings motivate the following questions 1) why is the pluripotent genome organized in an open configuration and 2) what factors are involved in maintaining these cells in this open chromatin state?

The lack of correlation between compaction and the canonical epigenetic modifications associated with constitutive heterochromatin may be unique to the pluripotent state. Since somatic cells contain both open and closed chromatin domains. Presumably these open and closed domains would be dependent upon specific epigenetic modifications. However, I would argue that the structural distinction between these two domains in somatic cells is not dependent on the formation of 30 nm and higher-order chromatin fibre assemblies, regardless of the underlying epigenetic modification. I demonstrated that chromatin fibre assemblies in
constitutive heterochromatin domains of both MEF chromocentres and in the most compact regions of lymphocytes from mouse spleen tissues were comprised exclusively as 10 nm chromatin fibre structures. This finding was made possible using a novel ESI-tomography approach. This methodology enables one to circumvent the projection image limitations of TEM while preserving the remarkable resolution of ESI, which can be used to visualize individual nucleosomes and intervening linker DNA, with a practical resolution limit below 2 nm. Thus, the unambiguous nature of chromatin fibre assemblies could be elucidated in situ for the first time in densely packed chromatin domains. I conclude that the formation of higher-order chromatin fibre assemblies, including the 30 nm chromatin fibre, is not required to condense the genome into the small confines of the nucleus, as was previously assumed. Rather, extensive bending and folding of 10 nm chromatin fibres may be needed to form the closed domains biochemically associated with heterochromatin in somatic cell types. These observations motivate questions regarding the precise relationship between biochemistry, DNA sequence, and chromatin structure.

5.2 Future directions

5.2.1 What are the factors that maintain chromatin in the characteristic open state in pluripotent cells?

A prominent and largely unanswered question arising from the study of disruption of constitutive heterochromatin domains during somatic cell reprogramming is what regulates or maintains this “open” signature chromatin structure? And is the regulation of this chromatin structure directly impacted by the pluripotent transcription factor network? Recently Wdr5, a component of the trithorax-group complex involved in gene-activation and histone acetylation was shown to associate and act in a functional complex with the pluripotent-associated transcription factor Oct4 (Ang et al, 2011). In addition, work in this thesis has provided some preliminary evidence that Nanog expression, a key-regulator of the pluripotent state (Silva et al, 2009) is correlated with chromatin disruption. Nanog expression levels are variegated in both the mouse embryo and in ES cells. Nuclei exhibiting the highest-levels of Nanog expression had the highest degree of disrupted chromatin associated with chromocentres in ES, full iPS and E3.75 pre-implantation mouse embryos, measured by relative DAPI intensity line scans. Taken together, these
observations implicate the pluripotent transcription-factor network, specifically Nanog, in directly regulating chromatin organization.

To test whether increased Nanog expression is sufficient to induce chromatin disruption, it could be exogenously expressed in the partially reprogrammed iPS cell line EOS3F-24. This partial iPS cell line maintains largely compact chromocentre domains reminiscent of the parental MEF cell line. The potential Nanog-dependent disruption of chromocentre domains could be monitored by both DAPI line scan analyses and by correlative LM/ESI experiments. If over expression of Nanog in the EOS24 iPS cell line did disrupt chromatin associated with chromocentres in these cells this would provide indirect evidence that Nanog was in fact involved in regulating chromatin organization in pluripotent cells. One of the confounding factors in these experiments is that as a transcription factor Nanog regulates several downstream target genes, which could ultimately be responsible for chromatin reorganization in pluripotent cells. Genome wide transcription profile microarrays could be used to monitor other potential epigenetic modifying proteins that are up or down regulated in EOS3F-24 cells in response to exogenous Nanog expression. The potential changes in gene-expression specifically related to epigenetic factors may garner insights into a putative Nanog-dependent network of chromatin regulators in ES cells.

In addition to the EOS3F-24 partial iPS cell line model system, it may also be possible to investigate the role of Nanog in regulating chromatin structure in the pluripotent state using an Epi Stem Cell (EpiSC) culture model system. EpiSCs are derived from day 5.5 mouse embryos, in contrast to ES cells, which are derived from the day 3.5 embryo. It has been previously demonstrated that these two cell types are significantly different in their global chromatin organization (Ahmed et al, 2010). Unlike ES cells that exhibit a completely disrupted global chromatin organization, EpiSCs have early signatures of somatic cell-like chromatin organization with visible chromatin domains and inter-chromatin spaces. EpiSCs have reduced expression of the pluripotent-associated transcription factors, notably Nanog. EpiSC cell lines have been generated that stably expressed Nanog, and preliminary ESI experiments of the global chromatin organization in these EpiSCs indicate robust Nanog expression is sufficient to disrupt the organized domains (Figure 27).
Figure 27. Nanog expression in Epi-Stem cells correlates with a disruption in chromatin

**Left panel** An ESI micrograph of a typical ES cell with dispersed uniform chromatin fibres throughout the nucleoplasmic space. In all images chromatin fibres are shown in yellow and protein-rich structures in blue. **Middle panel** shows an image of a typical EpiSC where chromatin fibres are beginning to form more discrete and clumpy domains. **Right panel** shows an EpiSC that is robustly expressing Nanog, chromatin fibres now appear more uniformly distributed throughout the nucleus. Scale bar is 0.5 µm.

I would speculate that Nanog expression is not merely influencing chromatin compaction by transcriptional regulation of downstream epigenetic modifying proteins. Instead I would hypothesize that Nanog participates directly in the organization of open chromatin domains in pluripotent cells. Although possible, albeit unlikely, Nanog could bind directly to constitutive heterochromatin and recruit a chromatin-remodeling complex to these regions. A more probable alternative would require an adaptor protein that binds directly to heterochromatin and would target Nanog, along with a putative epigenetic modifying complex to compact domains in pluripotent cells. This putative complex would then actively maintain heterochromatin in pluripotent cells in an open configuration. Nanog has been shown to be associated with several proteins associated with ATP-dependent chromatin remodeling complexes, including SWI-SNF-Brg1 (Kidder et al, 2009), and NuRD (Liang et al, 2008) complexes. Several components of an
ES cell specific SWI/SNF complex have been shown to have an impact on maintaining pluripotency and indirect evidence indicates that knockdown or knockout of these components have an impact on chromatin organization in ES cells (Lessard & Crabtree, 2010). Using correlative LM/ESI technologies, the impact of these SWI/SNF knockdowns on chromatin compaction could be directly assayed in the context of the mutant ES cell lines. The expectation being that SWI/SNF remodeling factors that are integral to the maintenance of the signature open chromatin state, when absent, would result in formation of compact domains in these cells. In addition, chromatin IP experiments using Nanog as a capture protein may yield important insights into the organization of Nanog-regulated genomic regions.

A promising candidate SWI/SNF component is the Baf-Brg1 complex. Knockdown-rescue experiments with these and other high probability candidate complex proteins could be performed to assess the role of these factors in pluripotent chromatin organizations. For instance, if the effect that Nanog is having on the EpiSC chromatin structures is a consequence of its interaction with Brg1 in the SWI/SNF complex, then knockdown of Brg1 in EpiSCs and induction of Nanog protein expression would result in a failure of the EpiSC chromatin to convert to the ES-like open chromatin state.

5.2.2 Is the pluripotent genome organized in an open configuration to provide genomic plasticity?

At this time, I can only speculate on the functional significance of maintaining chromatin in an open configuration in pluripotent cells. An attractive hypothesis is that this open configuration may confer an advantage to these unique cells by providing the needed structural plasticity to respond rapidly to differentiation cues. This open architecture would allow for the rapid reorganization of the genome into a tissue or lineage-committed state reflective of the unique transcriptional profile of the differentiating cell-type.

It may be possible to test this hypothesis indirectly by making use of a secondary iPS cell line, such as those generated in the laboratory of Dr. Andras Nagy (Woltjen et al, 2009) and investigate chromosome territory organization during cellular reprogramming. These cells, derived from a chimeric mouse, can be dox-induced to begin expressing the “Yamanaka”
transcription factors. The advantage of using the dox-inducible secondary iPS cell system are twofold: 1) cells that are generally more refractory to reprogramming, such as mature lymphocytes, can be studied more readily; and 2) the efficiency of these reprogramming experiments are much higher then *de novo* reprogramming experiments, thus allowing for pseudo-temporal chromatin reorganization studies.

Lymphocyte cells have robust pair-wise chromosome territory positioning propensities, where chromosomes 12, 14, and 15 tend to cluster together in these cells (Parada et al, 2002). In addition, lymphocytes have among the most compact chromatin organization observed to date (Figure 3), where clearly more than major satellite and associated DNA sequences are maintained in a highly compact state. If the hypothesis prevails that open chromatin provides the structural plasticity required to allow for rapid reorganization of chromatin during differentiation, then one would expect that as lymphocytes reprogram, these pair-wise associations would become less significant. However, if pair-wise chromosome territory organization, a proxy in this case for genome plasticity, were not a functional requirement of this open chromatin state, then the expectation would be that although the chromatin fibres within these territories would likely become more disrupted upon iPS cell generation, these lymphocyte-specific chromosome territory clusters would be maintained. To test this hypothesis, lymphocytes isolated from a chimeric mouse would be dox-induced to initiate the expression of the “Yamanaka” factors in these cells, and chromosome territory positioning would be monitored by using whole chromosome paints for chromosomes 12, 14 and 15. Alternatively, one could use a SKY 3D-FISH experiment, a technique that enables the simultaneous visualization of all chromosomes within a single nucleus. By monitoring chromosome territory pairing during the course of reprogramming, and by repeating these analyses over the 18 days required to generate a dox-independent fully reprogrammed cell in this system, the propensity for disruption of these lymphocyte-specific chromosome territory associations could be assessed.
5.2.3 What is the relationship between epigenetic modifications and chromatin structure?

Global chromatin architectures are reflected in the developmental stage and tissue types and, it is assumed, that epigenetic modifications regulate these specific chromatin structures. Remarkably, the signature open chromatin organization of the pluripotent cell is apparently uncoupled from the underlying epigenetic marks: whereby both chromatin with modified histones associated with heterochromatin and with the surrounding unmarked chromatin were often indistinguishable. In stark contrast, the chromatin structures enriched in heterochromatin modifications in MEFs and partially reprogrammed cells have easily distinguished compact chromatin domains, which directly correlate with these modifications. In tissues, such as mouse liver nuclei, distinct chromatin domains are also evident. Given these apparently disparate observations, it begs the questions: (1) What is the relationship between epigenetic modifications and chromatin structure and (2) is the lack of correlation between the two a unique feature of pluripotent cells? To begin to address this question, our ESI-tomography analyses, in combination with LM/ESI, imaging could be employed. Biochemically distinguishable chromatin features in several tissue types could be analyzed such as H3K9me3 constitutive heterochromatin domains, H3K27me3 facultative heterochromatin and H3K4me3 euchromatin regions can be indirectly labeled, within the same nucleus, with specific antibodies to these histone modifications. Tomographic reconstructions of these correlated regions may reveal important structural features that underlie the biochemistry of genome silencing. To elucidate these parameters, both inter-fibre distances, distances between nucleosomes, and the degree of fibre bending, would be measured in three-dimensional space using commercially available software. Based on visual inspection of the tomograms generated to date, I would speculate that the distance between chromatin fibres in compact domains would be both closer and less randomly organized and have increased propensities of regular fibre bending relative to those in more open chromatin domains. Given data from previous microccocal nuclease digest experiments of heterochromatin and euchromatin, the expectation would be that nucleosome spacing and density would vary between these two biochemical distinct types of chromatin.

It also may be the case that specific epigenetic modifications work in tandem to generate specific structural features. It may be fruitful to investigate constitutive, facultative, and euchromatin biochemical modifications in the background of specific epigenetic reader or writer null cell
lines. For example, in G9a null cells, one of the enzymes required to generate H3K9me2 modified histones may have an impact on H3K9me3 structures, whereby nuclesome spacing, or inter-fibre distances would be impacted in this null cell line. Again correlative LM/ESI-tomography could be used to measure these chromatin fibre structure parameters. To date several epigenetic null ES cell lines have been generated and could be differentiated or converted to EpiSC cell lines, in an attempt to address the affect of tandem epigenetic modifications on chromatin structure. Taken together, these experiments should reveal important insights into the relationships between biochemical modifications associated with specific gene profiles and chromatin structure.

In addition, it may be possible and interesting, to investigate chromatin structures, including inter-fibre distances and degrees of fibre bending, at specific genomic loci using a correlative FISH/ESI-tomography approach. Unfortunately, FISH techniques, which can be used to label specific genomic loci, have a profound effect on chromatin structure, likely due to the combination of heat and chemical denaturation required to anneal fluorescence DNA probes to chromatin. To circumvent this inherent problem in the FISH protocol, serial sections could be generated of a specimen of interest and alternate sections could be used for FISH experiments. The fluorescence signal could be mapped to the mirror section, where the underlying chromatin fibre structure is preserved, for ESI-tomography analysis. Using this approach it may be possible to finely map specific regions of the genome and assess the impact of differentiation or cell-type specificity of genomic loci of interest.
References


Meyer-Ficca M, Muller-Navia J, Scherthan H (1998) Clustering of pericentromeres initiates in step 9 of spermiogenesis of the rat (Rattus norvegicus) and contributes to a well defined genome architecture in the sperm nucleus. *J Cell Sci* **111 (Pt 10):** 1363-1370


