THE UNIVERSITY
of EDINBURGH

The Impact of DNA Sequence on the Structure and Dynamics of the Higher-order Chromatin Fibre *in vitro*

Hannah A Wheeldon

Thesis presented for the degree of Doctor of Philosophy

The University of Edinburgh

2019

MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, EH4 2XU, U.K.

Declaration of Originality

I hereby declare that except where specific reference is made to other sources, the work contained in this thesis is the original work of the author. It has been composed by myself and has not been submitted, in whole or in part, for any other degree, diploma, or other qualification.

Hannah Wheeldon December 2019

Acknowledgements

I am very grateful to Nick Gilbert for supervising this project, for supporting me in his lab, providing excellent ideas and advice throughout my PhD, and for his patience during my write up. I am also very thankful to Jim Allan for sharing his expertise in the field of chromatin, for his encouragement and for many useful and enjoyable discussions.

I would also like to thank the Gilbert lab for their advice and support over the last four and a half years. Adam, Catherine, Ryu-Suke, Lora, Issy, Waad, Elaine, Jamilla, Sam, Peter, Kate, Clarissa, Covi and Maria - thanks for teaching me new methods, answering many questions and for useful discussions during lab meetings. I am especially thankful to Davide for helping with the analysis of electron microscopy images.

I am grateful to my thesis committee: my second supervisor Andrew Wood, external supervisor Lynn Paterson and chair Richard Meehan for their input and advice at different stages of my project.

I am grateful to Alex Makarov for his help and expertise in platinum shadowing for my EM experiments, and to Stephen Mitchell for his assistance during microscopy.

I would like to thank Nikul Khunti, Nathan Cowieson and Rob Rambo at B21 at the Diamond Light Source for their help with small-angle x-ray scattering experiments.

I am also grateful to John van Noort for having me visit his lab in Leiden and for his excellent insight regarding the force spectroscopy experiments. I would also like to thank Artur for welcoming me to Leiden during my visits, and for all your help with using the tweezers and in analysing and interpreting these results.

I am very thankful to my family, for all their support over many years of study. I am grateful to have made many friends in Edinburgh: flatmates, dancers and my church family, who I thank for their encouragement and occasionally much-needed distractions from PhD stresses.

Finally, thank you to my husband Daniel. Thank you for your encouragement and support. Thank you for taking John out for long walks to give me time to write. Thank you for giving me perspective and reassuring me during many stressful moments. You are such a blessing to me, and this thesis would not have been possible without you.

Abstract

DNA in the eukaryotic cell is organised into a complex called chromatin, which protects the DNA from damage and allows the careful regulation of the genome. Precisely how the genome is used by the cell is dependent on the structure and regulation of this complex, and discovering how this is organised is therefore one of the principle challenges of modern biology.

DNA is wrapped around histone octamers to form arrays of nucleosomes, which are subsequently folded into a higher-order structure, speculated to be a 30-nm fibre. *In vitro* studies of higher-order chromatin fibre structure have provided valuable information about this structure, but it is not well understood how changes to the DNA sequence might affect the structure and dynamics of the complex. DNA sequence is known to affect nucleosome binding strength and positioning within an array, and I therefore hypothesised that DNA sequence changes are likely to impact higher-order chromatin fibre structure. Using an *in vitro* model of chromatin fibre structure, reconstituting purified DNA with purified core histones by salt dialysis, allowed me to isolate the effects of DNA sequence in the absence of confounding factors such as transcription factor binding.

I compared the higher-order chromatin structure and dynamics of the well-studied "601" DNA repeat sequence with two novel reconstitution templates which contain biologically-derived nucleosome positioning sequences. Sucrose gradient sedimentation of folded chromatin fibres suggested that non-*601* fibres may be as compacted as *601* fibres, but have more heterogeneous structures. However, non-*601* fibres were more easily perturbed under tension than *601* fibres, suggesting that such sequences might promote a more accessible chromatin environment. While repetitive *601* fibres were found to have a regular nucleosome repeat length by DFF digestion, non-repetitive, biologically-derived sequences had a more heterogeneous nucleosome repeat length, which I suggest is responsible for their increased accessibility.

I also found that the compacted higher-order structure of the *601* fibre is disrupted by the introduction of a single sequence with low affinity for the histone octamer. These structures can be separated by sucrose gradient sedimentation, and I suggest that this could be a useful method to examine the individual effects of a wider range of DNA sequences on higher-order chromatin fibre structure *in vitro*.

Lay Summary of Thesis

The genetic code required for life is contained within strands of DNA. In animal, plant and yeast cells, DNA is compacted into the complex of chromatin within the cell nucleus. The structure of this complex is of vital importance, as its compaction at different regions of DNA determines how accessible each part of the DNA is, and therefore how easily a cell can read the genetic code found in that region. Reading different parts of the genetic code allows cells with the same genome to form drastically different cell types. For example a muscle cell and a liver cell contain the same underlying genetic instructions, but different regions of DNA are accessible in each cell type due to the structure of the chromatin complex.

The precise structure of the complex, and how it is regulated has been well studied over recent years. There are many factors that play a role in compacting or decompacting chromatin structure, changing DNA accessibility. I have been researching the effects of the underlying DNA sequence on the structure of the folded chromatin fibre. In order to do this, I have used an *in vitro* system, creating chromatin from its purified components (DNA and histone proteins) outside of the cell. This has allowed me to assess the chromatin structures formed by different DNA sequences, in the absence of any other factors found *in vivo* that might confound my results.

I have found that non-repetitive DNA sequences derived from the sheep genome form a structure with a similar degree of compaction as the the artificial, repetitive DNA sequences that have been previously used to study chromatin fibre structure *in vitro*. However, when I pulled on individual chromatin fibres using magnetic tweezers I found that these biologically-derived sequences unfold more easily than the canonical artificial DNA sequence. These results suggest that chromatin in cells is likely to be more accessible than previous *in vitro* studies may have predicted.

Table of Contents

List of Abbreviations

°C	Degrees Celsius
µg/µl/µm/µM	1×10^{-6} grams/litres/metres/molar
A	Adenine
ACF	ATP-dependent Chromatin Assembly Factor
ATP	Adenosine Triphosphate
AUC	Analytical Ultracentrifugation
BAC	Benzalkonium Chloride
BLG	β-lactoglobulin
bp	Base Pairs
BSA	Bovine Serum Albumin
C	Cytosine
cm	1×10^{-2} metres
CE	Chicken Erythrocyte
ChIP	Chromatin Immunoprecipitation
ChIP-seq	ChIP Sequencing
CHO	Core Histone Octamer
CpG	Cytosine-guanine Dinucleotide (in which Cytosine is 5')
CTCF	CCCTC-binding Factor
CTD	C-terminal Domain
Da	Daltons
DFF/CAD	DNA Fragmentation Factor, a.k.a. Caspase Activated DNase
DNA	Deoxyribonucleic Acid
DNaseI	Deoxyribonuclease I
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid

EM	Electron Microscopy
EMSA	Electrophoretic Mobility Shift Assay
ENCODE	Encyclopedia of DNA Elements
FACT	Facilitates Chromatin Transcription Protein
FAIRE	Formaldehyde-assisted Isolation of Regulatory Elements
FISH	Fluorescent *in situ* Hybridisation
FOXA1	Forkhead Box Transcription Factor A1
FRET	Fluorescence Resonance Energy Transfer
G	Guanine
HATs	Histone Acetyl Transferases
HDACs	Histone Deacetylases
HMG	High-mobility Group Proteins
HP1	Heterochromatin Protein 1
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kbp	1×10^{3} Base Pairs
kDa	1×10^{3} Daltons
L	Litres
LH	Linker Histone
m	metres
M	Molar/Moles
Mb	1×10^{6} base pairs
me1/me2/me3	mono/di/tri-methylation
mESCs	Mouse Embryonic Stem Cells
mg/ml/mm/mM	1×10^{-3} grams/litres/metres/molar
min	Minutes
MNase	Micrococcal Nuclease
MW	Molecular Weight

MWCO	Molecular Weight Cut Off
NAP1	Nucleosome Assembly Protein
ng/nl/nm/nM	1×10^{-9} grams/litres/metres/molar
NPS	Nucleosome Positioning Site
NRL	Nucleosome Repeat Length
NTD	N-terminal Domain
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate-buffered Saline
PIC	Pre-initiation Complex
PMSF	Phenylmethyl Sulfonyl Fluoride
pN	Piconewtons
PTM	Post-translational Modification
RCC1	Regulator of Chromatin Condensation 1
RICC-seq	Ionizing Radiation-induced Spatially Correlated Cleavage of DNA with Sequencing
RNA	Ribonucleic Acid
RNA PolII	RNA Polymerase II
RPE1	Retinal Pigmented Epithelium (Cell Culture)
RPM	Revolutions per Minute
SAXS	Small-angle X-ray Scattering
SDS	Sodium Dodecyl Sulphate
SEC-SAXS	Size Exclusion Chromatography coupled with Small-angle X-ray Scattering
SELEX	Systematic Evolution of Ligands by Exponential Enrichment
SEM	Standard Error of the Mean
SMFS	Single Molecule Force Spectroscopy
SWI/SNF	SWItch/Sucrose Non-Fermentable Chromatin Remodelling Complex

T	Thymine
TAD	Topologically-associated Domain
TBE	Tris Borate EDTA Buffer
TE	Tris EDTA
TSS	Transcription Start Site
UV	Ultraviolet
V	Volts

List of Tables

List of Figures

Chapter 1. Introduction

Understanding how the genome is regulated and how this contributes to cell function in health and disease is one of the major challenges facing modern biology. While the primary sequence of DNA directly impacts the composition of the proteins created within a cell, the packaging of this sequence into a three-dimensional structure contributes to a complex system regulating the transcription, replication and repair of the genome.

The impact of the packaging and structure of the genome into chromatin on gene activity have become increasingly apparent. This packaging is dependent on many different factors that have been well documented by chromatin biologists including protein binding, protein modification and DNA modification. However, the intrinsic properties of the DNA sequence itself and how these impact its own regulation are challenging to study *in vivo*, due to the wide-ranging effects of changing the primary DNA sequence.

In vitro, DNA sequence has been found to play an important role in the structure of the 10-nm chromatin fibre by determining the positioning and binding strength of nucleosomes, and this positioning has been found to be strongly associated with that seen *in vivo* (Gencheva et al., 2006; Segal et al., 2006). However, the extent to which changes in the underlying DNA sequence impact the compaction and stability of the higher-order chromatin fibre have not been well studied.

1.1 Chromatin Structure and Function

Within a eukaryotic cell nucleus, DNA is packaged into the highly organised nucleoprotein complex of chromatin. The folding and compaction required to form this complex allows the 6 billion base pairs of DNA that form the human genome, a length of 2 m, together with approximately an equal mass of protein, to occupy a cell nucleus that is 20-50 μm in diameter (Radman-Livaja and Rando, 2010). The compaction of DNA into chromatin is a major regulator of DNA accessibility, and therefore of all DNA-dependent processes, crucially: gene transcription, DNA repair and genome replication. The understanding of these structures and their components, and the interplay between structure and function is therefore vital in understanding how the genome is regulated in development and disease.

Several factors influence the positioning of nucleosomes on any strand of DNA, including transcription and replication machinery, chromatin remodelling proteins, and histone variants and modifications. This thesis explores the influence of DNA sequence on nucleosome positioning, and crucially, how this affects the 30-nm chromatin structure. While a considerable amount is known about the sequence determinants of nucleosome positioning

and occupancy, less is known about how this impacts the structure of the higher-order chromatin fibre.

1.1.1 Components of Chromatin

Chromatin consists of DNA wrapped around histone proteins to create a chain of "beads on a string". Additional proteins then interact with this fibre to fold it into higher-order structures. Each of these components: DNA, histone proteins and additional proteins may be individually altered by various different mechanisms which allow tight but incredibly complex regulation of the genome.

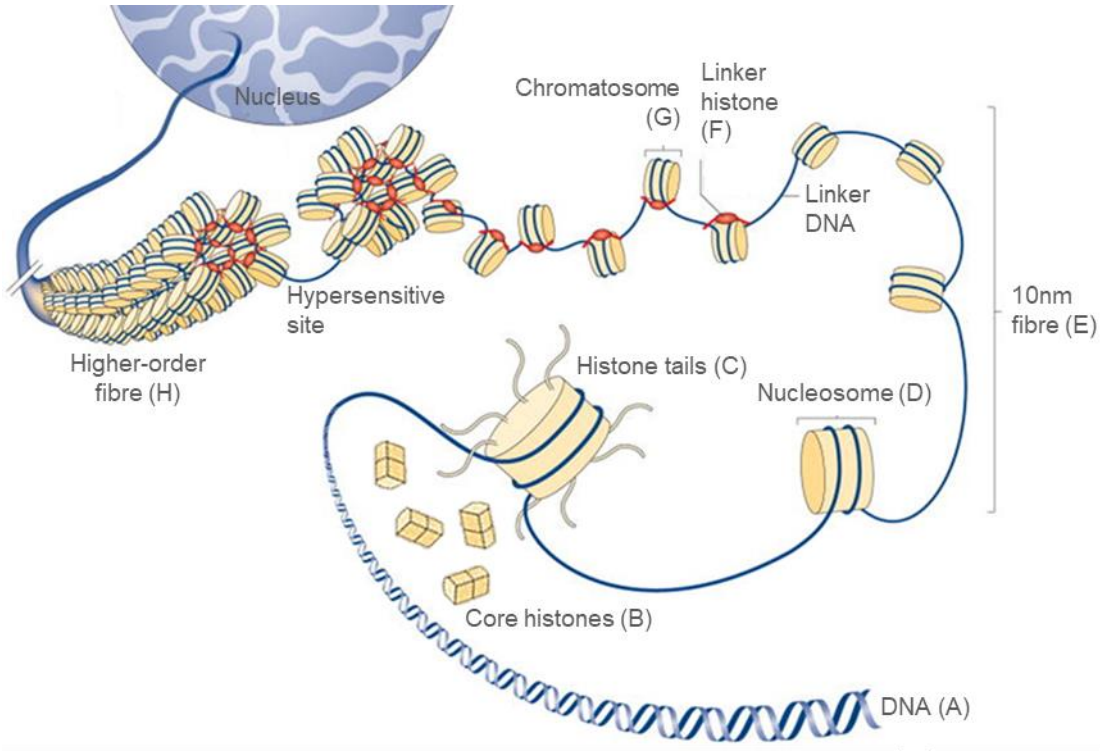

Figure 1.1. Chromatin Structure Within the Eukaryotic Cell Nucleus.
A) DNA. B) Core histones. C) Histone tails. D) Nucleosomes. E) The 10-nm fibre, or "beads on a string". F) Llinker histones. G) Chromatosomes. H) Higher-order chromatin fibre. Adapted from Fyodorov et al. (2018).

1.1.1.1 DNA

Deoxyribonucleic acid (DNA) consists of two polynucleotide molecules twisted around one another to form a double helix structure (Figure 1.1A). The sequence of nucleotide bases contained in each of these polymeric chains contains the genetic information required for a cell to manufacture the proteins which are necessary for life.

Nucleotide bases consist of a deoxyribose sugar molecule, containing five carbon atoms, with a phosphate group attached to the fifth carbon atom and a nitrogenous base attached to

the first carbon atom. These polymerise when phosphodiester bonds form between the phosphate group attached to the fifth carbon atom and the hydroxyl group attached to the third carbon atom of a neighbouring nucleotide. Therefore, one strand of DNA consists of a sugar-phosphate backbone, with varying nitrogenous bases attached to each sugar residue. The ionised phosphate groups within the DNA backbone cause the molecule to have a negative charge.

The two polynucleotide strands in the double helix are held together by hydrogen bonds between opposite nitrogenous bases. These bonds occur between the 4 different nitrogenous bases found in DNA: adenine (A); cytosine (C); guanine (G) and thymine (T). A always bonds with T in the opposite polynucleotide chain, C always pairs with G. Therefore, while the sequence of nitrogenous bases is different between the two polynucleotides, the two strands correspond with one another. It is this sequence of nitrogenous bases that encodes the information necessary to manufacture proteins and therefore controls the characteristics of an organism.

In addition to carrying the code needed to form proteins, the varying characteristics of each of the nucleotide base can allow genetic regulation through varying DNA structures other than the canonical double helix, varying flexibility of the double helix, and varying affinity for DNA-binding factors. Furthermore, proteins may bind to specific DNA sequences, influencing genetic regulation and genome architecture.

Modification of these nucleotides may change their properties. The best studied example of this is the methylation of the cytosine residue, which alters the mechanical properties of DNA reducing its flexibility (Ngo et al., 2016) and which promotes the binding of 5-methylcytosine specific proteins to DNA, which can further alter gene expression (Tate and Bird, 1993). Methylated DNA is associated with more compacted chromatin, and the methylation of CpG island promoters has been associated with the silencing of tumour suppressor genes in cancer (Feinberg and Tycko, 2004).

In addition, DNA may be supercoiled, where the double helix is either over or underwound, with effects on chromatin structure. Underwound DNA has been found to be associated with transcriptionally active regions of the genome, and is thought to promote a more disrupted, and therefore transcriptionally permissive chromatin structure (Naughton et al., 2013).

1.1.1.2 Core Histones

The components of the histone octamer are the proteins H2A, H2B, H3 and H4 (Figure 1.1B). Each histone protein contains a globular domain, which interacts with the other histones to form the core octamer structure, and a protruding N-terminal tail. The globular

domains of each core histone protein possess a histone-fold domain, consisting of three α helices linked by two loops, which interact to allow histones to dimerise. The octamer consists of one H3/H4 tetramer and two H2A/H2B dimers. These proteins have a positive charge due to the high content of basic amino acids lysine and arginine, in contrast to the negative charge found in DNA.

The histone-fold domains of the proteins interact with DNA in the nucleosome by several mechanisms (section 1.1.2.1). Each histone-fold pair is directly associated with 27-28 bp of DNA, with about 4 bp linking each of these sections (Luger et al., 1997). The wrapping of DNA around histones requires the DNA to bend sharply, and its ability to do this securely is dependent on the flexibility of the DNA nucleotide sequence.

Histone variants including H2A.Z and H3.3 may replace canonical histones, but have a different amino acid sequence which impacts the binding of histones to DNA and the properties of chromatin. For example, H2A.Z has an extended acidic patch promoting chromatin compaction (Fan et al., 2004), and is thought to stabilise the interaction between the H2A/H2B dimer and H3/H4 tetramer within the nucleosome (Kim et al., 2016). However, where H2A.Z is found in combination with the histone variant H3.3, nucleosomes have been found to be less stable and are associated with active promoter regions (Jin et al., 2009).

Post-translational modifications, primarily to the N-terminal tails of H3 and H4 (Figure 1.1C), also alter the structure and function of histones with the same underlying amino acid sequence. For example, histone acetylation is associated with chromatin decompaction. Acetyl groups remove a positive charge from the protein which may weaken the interactions between positively charged histones and negatively charged DNA (Hong et al., 1993), and are associated with more "open" and active chromatin. Acetylation of H4K16, within the N-terminal tail of the histone, inhibits the higher-order folding of chromatin *in vitro* (Robinson et al., 2008; Shogren-Knaak et al., 2006), whereas H3K56ac (which is closer to the globular core of the histone protein) has been found not to have an effect on chromatin folding, but increases DNA "breathing" within the nucleosome – suggesting a weaker interaction between the DNA and the histone octamer (Neumann et al., 2009). It is uncertain how important any direct effect of histone acetylation plays in chromatin structure *in vivo*, or whether any changes to chromatin structure are primarily dependent on the recruitment of effector proteins. Acetylated lysine residues are bound by bromodomains, often found in chromatin remodelling complexes such as SWI/SNF (Hassan et al., 2002).

In contrast to acetylation, lysine residues which are modified by methylation retain their positive charge and influence chromatin structure by the recruitment of effector proteins. For example, H3K9me3 binds heterochromatin protein 1 (HP1) (Bannister et al., 2001), which recruits histone deacetylases, promoting chromatin compaction, and may play a role in chromatin architecture by phase separation, causing heterochromatin to be sequestered within droplets of HP1 (Larson et al., 2017; Strom et al., 2017). Numerous other histone modifications including phosphorylation, ubiquitination and citrullination affect chromatin dynamics in various different ways, and may be combined within the same nucleosome in multiple different configurations, creating a complex regulation system.

1.1.1.3 Linker Histones

Linker histones do not form part of the core histone octamer but bind to DNA that is wrapped around the octamer at the dyad, the point of symmetry in the nucleosome where DNA enters and exits the complex, occupying an additional 10 bp of DNA at each end of the nucleosome and forming the chromatosome (Figure 1.1G, (Zhou et al., 2015)). The linker histone shows some sequence specificity for A/T rich regions of DNA. Linker histones consist of a globular domain, which binds to the dyad through its winged-helix domain, and an extended C-terminal tail which is positively charged and intrinsically disordered. The C-terminal tail binds to linker DNA between nucleosomes and plays an important role in stabilising the higher-order folding of the chromatin fibre (Allan et al., 1986). The linker histone is therefore a key regulator of transcription and other DNA-dependent processes. While the location of core histones upon the DNA is relatively stable, linker histones have been seen by FRAP to bind far more dynamically to chromatin, with H1 molecules being constantly exchanged between chromatin regions (Misteli et al., 2000).

In vitro, the compacting effect of linker histones on chromatin has been found to significantly reduce the transcription of a chromatinised DNA template (Laybourn and Kadonaga, 1991). In addition, linker histones have been found to interact with the H3K9 methyltransferase SUV39 *in vitro*, recruiting it to repetitive sites in Drosophila causing further compaction and gene repression by recruitment of HP1. H1 is itself associated with sites enriched for particular repressive core histone modifications, notably methylated H3K9 and H3K27 (Fyodorov et al., 2018). Furthermore, linker histone H1 is depleted in active promoter regions that are enriched for active histone marks such as H3K4me3 (Fyodorov et al., 2018).

Like core histones, there are several linker histone variants, with different underlying amino acid sequences; H1 is the canonical linker histone, H5 is found alongside H1 in avian

erythrocytes and is associated with highly condensed chromatin that is characteristic of these terminally differentiated and transcriptionally inactive cells. In addition, linker histones can be modified by phosphorylation, methylation and acetylation, though the effects of these modifications in linker histones are not so well-understood as they are for core histones (Fyodorov et al., 2018). H1 phosphorylation levels change throughout the cell cycle, becoming superphosphorylated during mitosis, when chromosomes are most condensed (Gurley et al., 1978). However, Turner et al. (2018) recently reported that phosphorylation of the H1 c-terminal tails reduces the affinity of the tail for the linker DNA, and suggest that this reduces chromatin condensation by reducing separation of the CTD/DNA complex into phase separated droplets, or "complex coacervates" where H1 acts as a liquid-like glue *in vitro*.

1.1.1.4 Other Factors

Multiple other proteins interact with the chromatin fibre and are required to maintain its structure. In addition to linker histones, CTCF, HP1 and transcription factors all play a role in stabilising and regulating chromatin compaction. High mobility group (HMG) proteins are important architectural factors in chromatin structure, with various roles regulating cell function. While these proteins may be considered integral components of the chromatin complex, throughout this thesis they will be referred to as additional or externally acting proteins.

RNA has also been suggested to form an integral part of the chromatin complex. Treatment with RNase A has been found to cause decompaction of chromatin (Rodríguez-Campos and Azorín, 2007). RNAs that are associated with chromatin may have multiple roles, notably the Xist RNA plays a crucial role in the silencing of the inactive X chromosome through the recruitment of epigenetic factors (Chow and Heard, 2009). However, Nozawa et al. (2017) have found that chromatin-associated RNAs interact with SAF-A, allowing this protein to oligomerise in the presence of ATP, forming a transcriptionally responsive mesh that maintains an open large-scale chromatin structure at gene rich regions. These chromatin-associated RNAs, thought to play a substantial role in maintaining chromatin structure, have recently been mapped in different cell types using next generation sequencing technology (Bell et al., 2018; Sridhar et al., 2017; Zhou et al., 2019a)

1.1.2 Organisation of Chromatin

A hierarchy of folding compacts the DNA with other chromatin components from a long strand into the cell nucleus. While there are several distinct levels of folding organisation, the degree of compaction within each will both have an effect on the ability of higher levels

of folding to effectively condense and will affect the accessibility of smaller scale structures. Each level of folding within this hierarchy is therefore interdependent.

1.1.2.1 The Nucleosome and the 10-nm Fibre

In forming chromatin, 146 base pairs of DNA wraps approximately 1.7 times around an octamer of histone proteins to form the nucleosome core particle (Figure 1.1E) (Luger et al., 1997), the fundamental unit of chromatin. The presence of core particles over DNA occludes access to the DNA by other proteins, inhibiting DNA-dependent processes such as transcription and DNA replication. Interactions between histone octamers and DNA are promoted by the opposing charges; DNA having a negative charge and histones having a positive charge. Using X-ray crystallography, Luger et al. (1997) identified the key interactions between DNA and histones in the core particle. (Table 1)

Interaction Description
The α helices of the histone-fold domains generate dipoles due to the positioning of the carbonyl groups within the peptide chain, generating a positive charge at the N-terminus which interacts with negatively charged phosphate groups in the DNA backbone, fixing them in position.
Hydrogen bonds form between the phosphate backbone of the DNA and amide groups within the protein chain.
Arginine side chains enter the minor groove of the DNA as it faces the octamer.
The peptides form non-polar interactions with the deoxyribose sugar in the DNA backbone.
Hydrogen bonds and salt bridges may form between phosphate groups and basic or hydroxyl-containing amino acid side chains.

Table 1. Description of Histone-DNA Interactions within the Nucleosome. Described by Luger et al. (1997).

The structure, stability and dynamics of each individual nucleosome is dependent on the underlying DNA sequence and modifications, histone variants and post-translational histone modifications, as well as external factors. Nucleosomes have been found to dynamically partially unwrap *in vitro*, with FRET experiments revealing that the DNA in a single "601" nucleosome may by partially unwrapped from the complex 2-10% of the time (Li and Widom, 2004; Zhou et al., 2019b), although it is possible that these dynamics will be different within the context of a compacted chromatin fibre. Furthermore, this unwrapping

may be promoted/inhibited by the presence of histone modifications or changes to the DNA sequence, altering protein access to nucleosomal DNA (North et al., 2012).

The formation of nucleosomes, repeated down the length of a DNA strand, can form a structure in low salt concentrations (eg. 10 mM NaCl) referred to as "beads on a string" (Figure 1.1F). The nucleosomes, 10 nm in diameter, form the beads, and are connected by linker DNA, represented by the string. These structures, known as 10-nm fibres, were first visualised *in vivo* by Olins and Olins (1974). The presence of monovalent or divalent cations will promote interaction between nucleosomes and lead to higher-order folding.

Linker histones bind to DNA at the point of exit/entry to the nucleosome, forming the chromatosome (Figure 1.1G). The presence of a linker histone may stabilise the nucleosome, restraining the movement of the core histones along the DNA (Pennings et al., 1994) and promotes further compaction of the nucleosome array in appropriate salt conditions (Robinson and Rhodes, 2006a; Thoma et al., 1979), although it may not be required to form 30-nm fibres from arrays with short NRLs *in vitro* (Dorigo et al., 2004; Robinson and Rhodes, 2006a). The binding of H1 is very dynamic compared to core histones, with linker histones continuously moving between different chromatin regions (Misteli et al., 2000).

The histone octamer protects the 147 base pairs of DNA from digestion by enzymes such as micrococcal nuclease. When a linker histone is present, an additional 20 bp are protected from digestion. Enzyme digestion can therefore be a useful indicator of primary chromatin structure, allowing the ubiquity of linker histones and the nucleosome repeat length (NRL) to be determined. The NRL may be as short as 154 base pairs in yeast (Szerlong and Hansen, 2011) and as long as 237 base pairs in sea urchin cells (Spadafora et al., 1976). Van Holde's metadata set of nucleosome repeat length analyses (1989) suggest that the repeat length of most vertebrate cells is between 177 and 207 bp, while that of cultured HeLa cells is 187-197 bp (Compton et al., 1976; Tate and Philipson, 1979). The 10-bp phasing seen in the nucleosome repeat lengths of different species (167, 177, 187 etc.) suggests that the linker length between nucleosomes, which occupy 147 bp of DNA is commonly a multiple of 10 bp, a complete turn of the DNA helix. *In vivo*, the nucleosome repeat length is influenced by DNA sequence, chromatin remodelling and protein binding, and long range *cis* interactions. Blank and Becker (1995) have found that electrostatic interactions within chromatin play a role in the nucleosome repeat length. The NRL has consequences for the structure and conformation of the higher-order chromatin fibre *in vitro* (Kruithof et al., 2009; Routh et al., 2008), and *in vivo* (Williams and Langmore, 1991).

Nucleosomes have been thought to generally repress transcription of chromatin templates, as DNA is occluded within the structure and is less accessible to transcription machinery (Laybourn and Kadonaga, 1991). However, the presence of a nucleosome array over the *PHO5* gene was recently found to increase RNA transcription *in vitro* (Nagai et al., 2017), in the presence of particular protein factors, highlighting the importance of the chromatin complex for effective transcription in eukaryotes.

1.1.2.2 Higher-order Chromatin Folding: A 30-nm Fibre?

In the presence of linker histones at physiological salt conditions, nucleosomes interact with one another, causing 10-nm arrays to fold further to produce a higher-order fibre (Figure 1.1H). The interaction between neighbouring nucleosomes occurs between an acidic patch of 7 amino acids within H2A and the N-terminal tail of H4 (Zhou et al., 2007). It has been shown by analytical ultracentrifugation that removal of the N-terminal tail of H4 prevents the formation of a higher-order chromatin fibre, displaying the important role that this interaction plays in the compaction of chromatin (Dorigo et al., 2003). In addition, mutating the H2A acidic patch and H4 tail (H2A E64C and H4 V21C) to include cysteine residues enables the creation of a disulphide bond which is able to stabilise chromatin compaction *in vitro* (Dorigo et al., 2004). As several chromatin-interacting factors, such as regulator of chromatin condensation 1 (RCC1) interact with the acidic patch on H2A (Makde et al., 2010), compaction of chromatin means that these proteins are unable to bind, as they must compete with the binding of the H4 tail.

Linker histones promote further compaction of a nucleosome array. The globular domain of the linker histone stabilises the nucleosome, sitting at the dyad between the entering and exiting strand of DNA, protecting approximately an additional 20 base pairs of DNA from digestion by micrococcal nuclease (Noll and Kornberg, 1977), and forming the chromatosome (Allan et al., 1980a). The C-terminal tail interacts with the linker DNA, aiding the compaction of the chromatin fibre by neutralising the charge of the DNA. Linker histones stabilise the intrinsic salt-dependent compaction of the nucleosome array (Carruthers et al., 1998) and allows increased compaction of the folded fibre *in vitro* (Robinson and Rhodes, 2006a; Routh et al., 2008), although this may not be required with fibres of a shorter nucleosome repeat length (Dorigo et al., 2004; Robinson and Rhodes, 2006a). *In vivo*, H1 has been found to be essential for mammalian development (Fan et al., 2003) and depleting H1 in mESCs has been found to have profound effects on chromatin structure, including reduction of the nucleosome repeat length, changes in histone modifications, and changes in higher-order chromatin structure resulting in less compact and more heterogeneous fibre conformations (Fan et al., 2005).

Compaction of a nucleosome array into a higher-order structure has been found to reduce the accessibility of nucleosomal DNA to restriction enzymes by 3-8 fold, and the accessibility of the linker DNA between nucleosomes by as much as 50 fold (Poirier et al., 2008). It is not well understood how various factors that impact the stability of the nucleosome may act differently on a compacted chromatin structure (Zhou et al., 2019b).

These inter-nucleosome interactions result in a fibre structure with a diameter of approximately 30 nm *in vitro*. As discussed below, the precise diameter, structure and ubiquity of this structure *in vivo* is questioned, and I will therefore refer to this level of organisation as the "higher-order fibre" rather than the "30-nm fibre" throughout this thesis. The canonical 30-nm fibre, was first identified by Finch and Klug (1976), who visualised folded fibres with a diameter of 30 nm and a pitch of 11 nm using electron microscopy. Assuming that this structure is a one-start solenoid (Figure 1.2A), where nucleosomes form a "stack" which twists to form a single helix, this 30-nm fibre contains 6 nucleosomes within each turn of the helix, allowing chromatin to be compacted with a nucleosome line density of 1.5-2.1nm/nucleosome (Robinson and Rhodes, 2006a). This would allow a 10-nm fibre with 50 bp linker DNA (a 197 bp repeat) become 12-18 times shorter upon compaction (Kruithof et al., 2009).

Several alternative models to the one-start solenoid have been suggested (Figure 1.2). Woodcock et al. (1984) first suggested the possibility of a two-start solenoid, where histones alternate between two "stacks", which then twist around in a helical ribbon (Figure 1.2 B). A second two-start twisted helical model with linker DNA localised in the centre of the fibre crossing between opposite nucleosome stacks (Figure 1.2C) has been suggested and visualised by *in vitro* studies. Utilising reconstituted nucleosome arrays, it has been suggested that fibres with a shorter nucleosome repeat length form a two-start helical structure, whereas those with a longer nucleosome repeat length form a one-start solenoid (Kruithof et al., 2009; Routh et al., 2008), however even these controlled studies have not reached a unanimous conclusion on the structure of the higher order fibre, as Robinson et al. (2006b) suggest based on EM measurements that fibres with a nucleosome repeat length between 177 bp and 237 bp form a one-start solenoid, whereas Song et al. (2014) find by cryo-electron microscopy that fibres with a repeat length of 177 or 187 bp form a two-start helix.

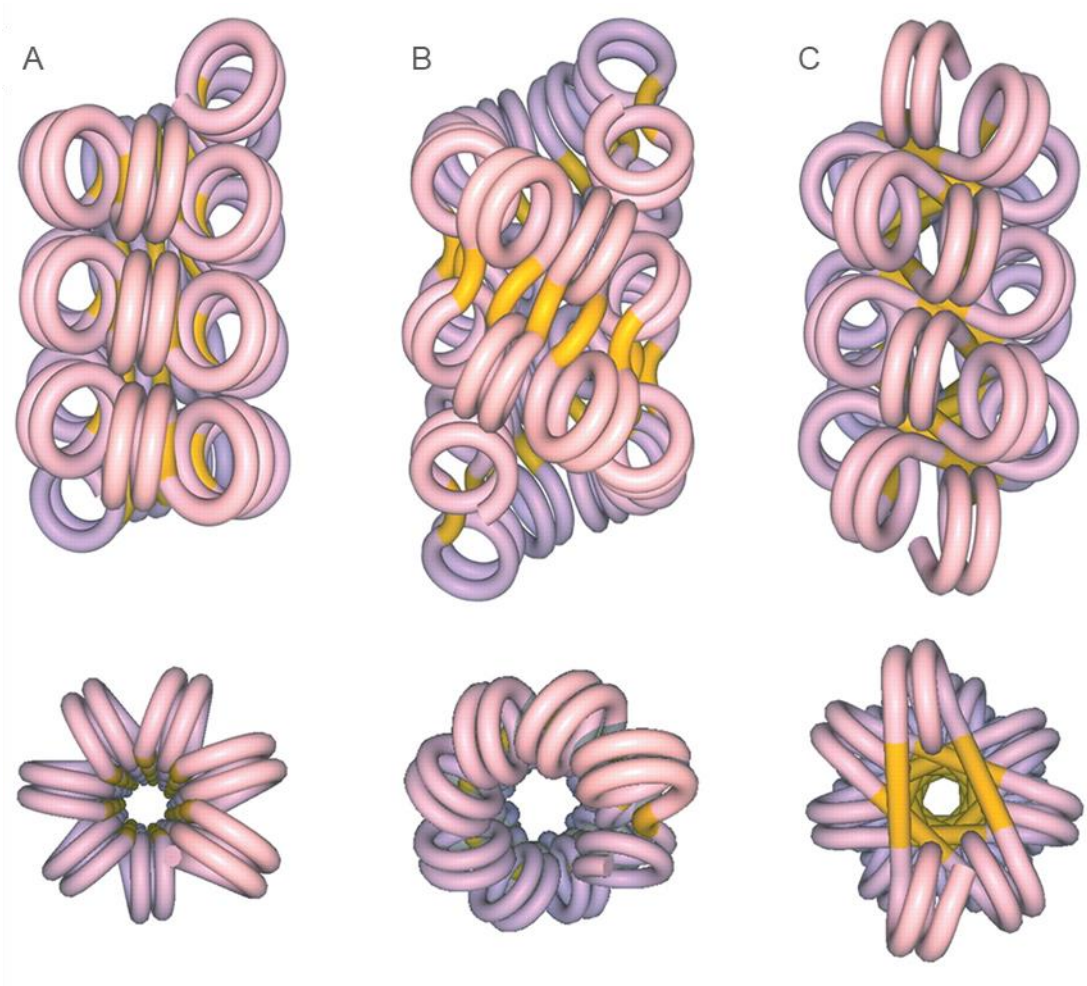

Figure 1.2. Various structures suggested for the higher-order "30-nm" chromatin fibre. A) One-start solenoid. B) Two-start helix. C) Two-start helix with crossed linker DNA. Modified from Dorigo et al. (2004).

Robinson et al. (2006b) suggest that nucleosome repeat length influences the diameter of the higher-order chromatin fibre *in vitro*, with two distinct structural classes with different diameters being formed by different repeat lengths (all repeat lengths of 10n bp). Langmore and colleagues (Athey et al., 1990; Williams and Langmore, 1991) found by SAXS and cryo-EM that different cell types of different nucleosome repeat length had wider higher-order fibres as NRL increased. However, Woodcock et al. (1984) state that there should be no variation in the higher-order fibre diameter of compacted chromatin with different nucleosome repeat lengths in a two-start solenoid configuration.

There is speculation around the exact structure and ubiquity of the 30-nm fibre within the eukaryotic cell nucleus. Nishino et al. (2012) and Joti et al. (2012) have not found 30-nm structures in HeLa cells by cryo-EM or small-angle x-ray scattering, during neither interphase nor metaphase. While 30 nm structures were evident by small-angle x-ray

scattering, this was lost upon removal of contaminating ribosomes. However, Langmore and colleagues (Langmore and Paulson, 1983; Langmore and Schutt, 1980) and Woodcock (1994) have measured 30-nm structures in chicken erythrocyte nuclei lacking ribosomes using these methods. The transcriptional inactivity of erythrocytes is likely to contribute to this structure and may allow it to achieve a greater compaction than transcriptionally active cells. Langmore and Paulson accounted for the confounding presence of ribosomes which might contribute to a 30 nm peak by subtracting the scattering of the cytoplasm alone, and moreover, found that treatment with DNase1 caused the loss of the 30 nm scattering peak.

Further attempts to understand localised chromatin structure *in situ* by Micro-C (Hsieh et al., 2015), RICC-seq (Risca et al., 2017), and Hi-CO (Ohno et al., 2019) have been thus far limited. Micro-C involves the crosslinking of neighbouring nucleosomes by formaldehyde, followed by micrococcal nuclease digestion of the genome. Fragments are then religated to those sequences in close spatial proximity to which they have been crosslinked and these fragments are then analysed by paired-end deep sequencing. In Hi-CO, chromatin is crosslinked similarly to in Micro-C, but the addition of a small adapter DNA sequence prior to the relegation of DNA allows nucleosome orientation as well as nucleosome positioning to be considered during analysis. However, it is possible that assays relying on formaldehyde fixation may be affected by artefacts caused by the crosslinking of sites that are not spatially correlated, but joined by a "bridge" of other cellular components. RICC-seq utilises ionising radiation to induce cleavage of spatially correlated DNA sites (within a 3.5 nm radius). These breaks are then repaired and religated to bring together sequences in close spatial proximity to create fragments that can be analysed by gel electrophoresis or high-throughput sequencing. Thus far, these methods have only succeded in analysing local chromatin structure on the tetra-nucleosome scale.

Other analyses have suggested that chromatin lacks a 30-nm higher-order structure *in vivo*. Several different alternative models to the 30-nm fibre have been suggested, including a packaged polymer-melt (Hansen et al., 2017; Shimamoto et al., 2017), nucleosome "clutches" (Ricci et al., 2015), or a disorganised chain with variable diameter (Ou et al., 2017). Ou et al. (2017) used DRAQ5 to visualise chromatin structure *in vivo* by ChromEMT. This suggested that in human osteosarcoma cell cultures, chromatin fibres form a disordered chain with a diameter between 5 and 24nm, and which pack together at different densities within the interphase nucleus. However, as this agent is a DNA intercalator, which will associate preferentially with negatively supercoiled regions of the genome, there is speculation that this method may be selective for more open and active regions of the genome which are associated with underwound DNA (Naughton et al., 2013).

It is known that there are variations in the structure of the higher-order fibre throughout the genome, where the structure is more "open" or "disrupted" around gene-rich regions, promoting transcription, and more compacted in gene poor-regions (Gilbert et al., 2004). Furthermore, the nucleosome repeat length varies throughout the genome, and reconstitution analysis by Routh et al. (2008) and single molecule force spectroscopy results from Kruithof et al. (2009) suggest that this may impact the configuration of the fibre. It is therefore credible that the conformation of the fibre might vary throughout the genome, and that depending on factors including DNA sequence, histone modifications and the action of external proteins, the chromatin fibre may form both a one-start and a two-start structure or various other structures in different genomic contexts. Wu et al. (2019; 2016) have suggested that chromatin may transition from a two-start to a one-start structure with longer linker lengths.

1.1.2.3 Large Scale Structure

Chromatin structures larger than the 30 nm scale are less well defined and appear to vary throughout the cell cycle. Belmont and Bruce (1994) visualised "chromonema" fibres with a diameter of 100-130 nm, which decondense into structures with a 60-80 nm diameter as cells enter S phase. During mitosis, when chromosome become most compact, fibres of 200-400 nm diameter have been visualised (Bak et al., 1977; Taniguchi and Takayama, 1986).

Chromatin fibres are structured within large scale loops and domains that contribute to the regulation of gene transcription. Within the eukaryotic cell nucleus, each chromosome occupies a distinct territory (Cremer et al., 1993; Lichter et al., 1988; Stevens et al., 2017). Within this, chromosomes are organised into evolutionarily conserved topologically-associated domains (TADs), which can contain hundreds of kilobases of DNA. TADs have been found to be associated with regulation of replication timing (Pope et al., 2014) and in gene expression. Nora et al. (2012) report co-expression of genes within TADs on the inactive X chromosome during mESC differentiation. The boundaries between these TADs are enriched for CTCF, which might enable chromatin loop formation and insulates chromatin domains.

Long-range interactions are thought to play an important role in gene regulation, in particular interactions between enhancers and promoters. Fluorescence *in situ* hybridisation (FISH) can analyse the colocalisation of specific DNA loci to analyse long range chromatin interactions, for example between the sonic hedgehog gene and the ZRS enhancer approximately 1 Mb away (Williamson et al., 2016). Chromatin conformation capture (3C) techniques use crosslinking to measure the spatial proximity of sequences across the genome

(Dekker et al., 2002). FISH and 3C have allowed the identification of TADs and chromatin loops, however in many cases these two techniques are not in agreement with one another (Williamson et al., 2014), possibly as a result of comparing the average interactions across a large number of cells captured by 3C techniques with those seen in a single cell using FISH, or due to extensive crosslinking of large complexes during 3C causing sequences that are spatially distant to appear close together.

The location of chromatin within the cell nucleus is also related to its transcription state, whereby the chromatin at the nuclear periphery may be transcriptionally repressed relative to the chromatin at the centre of the nucleus, at least in part due to the deacetylation of histones at the periphery of the nucleus (Finlan et al., 2008). Therizols et al. (2014) reported that the transcriptional activation of DNA located at the nuclear periphery caused it to be relocated to the centre of the nucleus.

1.1.3 Functions of Chromatin

Folding and compaction of DNA serves several purposes. In human cells, it enables the 2 meters of DNA found in a diploid human cell to fit into the cell nucleus, which measures approximately 20-50 μm in diameter (Radman-Livaja and Rando, 2010). This organisation of the genome also affects DNA accessibility, and therefore allows the regulation of gene transcription and DNA replication. This control of gene transcription enables cells with identical genomes to form different cell types, determined by varying gene expression patterns. Compacted "heterochromatin" is repressive to transcription, as it precludes necessary proteins, such as transcription factors and polymerases, from accessing the DNA, whereas less compacted "euchromatin" is permissible to transcription. Changes to chromatin structure, through histone and DNA modification, additional protein binding, and the introduction of supercoils, is therefore pertinent to such processes as transcription, DNA damage and DNA replication, which has implications for development and cancer progression.

There are multiple different methods of controlling the manufacture of proteins apart from the primary DNA sequence, which allows different cell types from the same organism to have very different characteristics despite having the same underlying DNA sequence. DNA may form alternative structures, be modified by methylation or contain positive or negative supercoils, histone octamers may contain variant histones or have post-translationally modified residues that affect folding, and external proteins may act by binding to specific chromatin structures, PTMs or DNA sequences to cause the change of any of these factors. The genome is therefore highly dynamic.

During important cellular processes, such as replication, transcription and repair, the two strands of DNA must temporarily separate. Chromatin obstructs this, and therefore 30-nm fibre structures must be disrupted and histone proteins removed for other proteins to access DNA. Chromatin may be made more accessible by this mechanism through two main pathways. Firstly, enzymatic modifications of histones including acetylation, methylation and phosphorylation of amino acid residues, particularly those within the tails protruding from the central core of the nucleosome, affect chromatin structure by changing the interactions between the histone octamer and DNA, and affecting the recruitment of additional proteins. Secondly, ATP-dependent chromatin remodelling complexes can displace histones, revealing the underlying DNA sequence to other proteins.

Compacting DNA into heterochromatin serves several purposes: recombination of repeat-rich sequences is supressed, unnecessary or cell-type inappropriate gene transcription is supressed and cell identity becomes fixed during development (Becker et al., 2017). Compaction into chromatin also protects DNA from physical damage, for example by radiation (Takata et al., 2013).

1.2 Chromatin Disruptions and Transcriptional Regulation

1.2.1 Chromatin Disruptions

Packing DNA into chromatin causes access to DNA sequence motifs by other proteins to be obstructed, therefore repressing DNA-dependent cellular processes including transcription. The location of a site within the nucleosome structure or within the 30-nm fibre may render a site inaccessible to transcription factors and refractory to transcription. Furthermore, the two nucleotide strands of DNA are unable to be separated while constricted within the structure of the nucleosome. A compacted chromatin fibre structure must therefore be disrupted in order to allow access to transcriptional machinery.

The presence of a nucleosome over promoter regions has been found to inhibit gene transcription *in vitro* (Lorch et al., 1987) and *in vivo* (Han and Grunstein, 1988). The removal of a nucleosome from the DNA sequence creates an accessible region, typically of around 200 bp, disrupting the chromatin structure and allowing the binding of nuclear proteins. These accessible regions of chromatin are associated with regulatory regions of DNA and are known as DNase1 hypersensitive sites as they are more prone to damage by nuclease digestion (McGhee et al., 1981; Weintraub and Groudine, 1976) as well as physical damage by techniques such as sonication (Auerbach et al., 2009) compared to bulk chromatin. The differences in chromatin structure between a compacted higher-order fibre and a chromatin region containing a hypersensitive site affects the buoyancy of such

samples, allowing them to be separated and analysed by sucrose gradient sedimentation (Caplan et al., 1987). Hypersensitive sites have been extensively mapped in particular cell types, identifying genome regulatory elements (Boyle et al., 2008; Thurman et al., 2012) and exploited in the analysis of regulatory DNA regions and of chromatin structures including the 30-nm fibre (Staynov, 2000) and large scale chromatin organisation in the nucleus (Ma et al., 2015). DNase-seq has been an important technique in analysing hypersensitive sites, but more recently ATAC-seq has been used to analyse accessible chromatin regions. Such regions are accessible to transposase enzymes including Tn5, as well as nucleases, allowing amplifiable tags for sequencing to be preferentially inserted into open chromatin regions. While there are likely to be some differences in the accessibility of different regions to different enzymes, ATAC-seq has some advantages over DNase-seq as there are fewer experimental steps and therefore far fewer cells are needed for analysis.

The occurrence of hypersensitive sites is correlated strongly with gene promoters, transcription factor binding sites and enhancers (Thurman et al., 2012) as the binding of proteins to these regulatory regions requires an accessible template. Hypersensitive sites over regulatory regions may be constitutive, promoting constitutive gene activation, or may be transiently generated in response to transcription factor binding. Proteins associated with transcription such as topoisomerases and polymerases have also been found to be associated with a subset of accessible sites (Gross and Garrard, 1988). DNA polymorphisms associated with disease tend to be concentrated at these accessible sites (Maurano et al., 2012), emphasising the fundamental importance of these regions in regulating proper cell function.

DNA sequence has been suggested to play a role in maintaining hypersensitive sites. Field et al. (2008) suggest that sequences rich in dA and dT create histone-depleted regions that may increase the accessibility of adjacent DNA to regulatory factors. These poly(dA:dT) elements are typically associated with maintaining stretches of DNA depleted of nucleosomes, for constitutive gene activation, as seen during the transcription of his3 and pet56 in yeast (Struhl, 1985). In mammalian cells, CpG islands over promoters are associated with nucleosome depletion and constitutive gene activation (Fenouil et al., 2012; Ramirez-Carrozzi et al., 2009). DNA supercoiling also appears to play a role in the generation of hypersensitive sites, which may be lost upon the removal of torsional stress (Villeponteau and Martinson, 1987; Villeponteau et al., 1984).

While the higher-order structure of chromatin will be markedly disrupted by the presence of a hypersensitive site caused by the loss of a nucleosome as described by Caplan et al. (1987) , a more subtle disruption may occur in the higher-order structure of chromatin, without the

loss of a nucleosome but allowing proteins increased access to linker DNA and to the surface of adjacent nucleosomes. Such a disruption in the higher-order structure may occur as a consequence of linker histone depletion or unstable underlying nucleosomes.

1.2.2 Transcription Factors and Transcription Initiation

Transcription initiation in eukaryotes by RNA PolII requires the binding of the large pre-initiation complex (PIC), which in yeast contains 58 protein subunits, to the promoter region of DNA (Kornberg, 2007). General transcription factors allow the binding of RNA PolII, mediator and other proteins to the promoter sequence and unwind the DNA helix allowing RNA PolII access to the sequence template. The binding of the large transcription initiation complex to DNA requires the absence of a nucleosome over the promoter region (Boeger et al., 2003; Lorch et al., 1987).

Transcription factors may be divided into two groups: "settler" transcription factors which can bind to pre-formed open chromatin architecture such as a hypersensitive site and recruit transcription machinery, and "pioneer" transcription factors, which can bind to nucleosomal DNA, destabilising and displacing nucleosomes from DNA, creating a hypersensitive site and allowing settler factors and other proteins to access their binding motifs. Whereas some hypersensitive sites are constitutive in a given cell type, allowing constitutive gene activation, pioneer factors are able to transiently induce gene activity in response to external factors (Choi and Kim, 2008; Tirosh and Barkai, 2008). Local chromatin environment and the stability of the chromatin fibre in the vicinity of pioneer transcription factor binding sites is likely to influence their binding and their ability to disrupt chromatin structure. Multiple complex factors including histone modifications, histone variants such as H3.3 and H2A.Z (Jin and Felsenfeld, 2007) and DNA sequence are likely to determine whether a chromatin region can be easily disrupted or if it is refractory to transcription.

FoxAI, the canonical pioneer transcription factor, binds to nucleosomal DNA at the nucleosome dyad in a similar manner to the linker histone via a winged-helix binding domain. By displacing linker histones it causes the decompaction of the higher-order chromatin fibre (Cirillo et al., 1998), providing access to additional proteins that promote decompaction and chromatin disruption. It also interacts with the H3/H4 tetramer, promoting nucleosome depletion (Cirillo et al., 2002). The removal of nucleosomes by pioneer transcription factors may involve the recruitment of further chromatin remodelling complexes such as SWI/SNF (Kadam et al., 2000), which reduce the interactions between DNA and histone proteins, promoting the formation of accessible nuclease sensitive sites (Bouazoune et al., 2009). In association with transcription factor binding, Owen-Hughes et

al. (1996) found that SWI/SNF causes the disruption of nucleosomes and creates hypersensitive sites allowing transcription to occur. Other pioneer factors are able to influence local chromatin architecture by binding to the linker DNA flanking the nucleosomes. Pho5 in yeast is upregulated in response to phosphate starvation, and involves the binding of the transcription factor Pho4 to the 70 base pair linker between nucleosomes and displacing surrounding histone octamers, exposing additional Pho4 binding sites and subsequently the TATA box (Reinke and Hörz, 2004). This process has been found to involve the recruitment of several different chromatin remodellers (Musladin et al., 2014).

1.2.3 Elongation Factors

Once RNA Polymerase has bound to the promoter, it moves along the gene reading the DNA code. This requires the destabilisation of nucleosomes from the DNA to allow the reading of the DNA code by Pol II, which is enabled by elongation factors such as FACT and the chromatin remodelling enzyme Chd1. Recently, cryo-EM has been used to visualise the movement of RNA PolII through the nucleosome structure (Kujirai et al., 2018). Histone-DNA contacts around the surface of the nucleosome appear to cause the pausing of RNA PolII as it transcribes through the nucleosome. Nucleosome remodellers such as the SWI/SNF complex promote the movement of RNA PolII through a nucleosomal template (Brown et al., 1996). Belotserkovskaya et al. (2003) suggest that the FACT complex appears to cause the transient destabilisation of a nucleosome by the temporary removal of one H2A/H2B dimer from DNA, whereas Schwabish and Struhl (2004) found that transcriptional elongation in the presence of FACT causes the eviction of core histones from the *S. cerevisiae GAL10* coding region, and that new core histones are recruited over the transcribed region within 1 min of the passage of PolII. Strong nucleosome positioning sequences are able to form a stable nucleosome through which Pol II may not transcribe *in vitro* (Bondarenko et al., 2006). However, such nucleosome positioning sequences have been found not to cause transcriptional pausing *in vivo* (Perales et al., 2011), demonstrating that factors other than DNA sequence have a profound effect on this property of the DNA sequence.

As Pol II transcribes through DNA, positive supercoils (overwound DNA) are generated ahead of Pol II while negative supercoils (underwound DNA) appear behind the transcription machinery. These supercoils are likely to have an effect on chromatin structure around sites of transcription, but may be relieved by topoisomerases (Kouzine et al., 2013). Naughton et al. (2013) used biotinylated-trimethylpsoralen to probe DNA structure as this is preferentially intercalated into underwound DNA. They found that domains of negative

supercoiling have a more disrupted chromatin structure, being enriched in hypersensitive sites and less compacted higher-order chromatin fibres, that may facilitate gene activation in these transcriptionally active regions. Kaczmarczyk (2019) found that higher-order chromatin fibres can accommodate a degree of positive supercoiling and that positive twist may stabilise chromatin fibres.

1.3 Effects of DNA Sequence on Chromatin Structure

The differing affinities of DNA sequences for the histone octamer impacts chromatin structure by influencing the positioning and binding strength of nucleosomes. *In vitro* studies show that chromatin assembles far more easily and regularly on particular sequences, such as the 5S RNA gene and the "601" nucleosome positioning sequence, than others ((Lowary and Widom, 1998), section 1.4). This variability in affinity for the histone octamer is based, at least in part, on the anisotropic flexibility of the DNA sequence, which allows nucleosomal DNA to wrap 1.7 times around the core octamer; a degree of flexibility that could not be achieved by free DNA without the presence of the positively charged protein complex.

The histone octamer is thought to have at least a 5000-fold difference in affinities between different DNA sequences (Gencheva et al., 2006). While there are several factors known to increase or decrease DNA sequence affinity for the histone octamer, the full range of factors affecting this variability is not completely understood. Histones interact with DNA at 14 crucial points around the nucleosome (approximately every 10 bp), where positive charges within the protein interact with phosphates within the DNA backbone. High-affinity DNA sequences are thought to be primarily recognised by the H3/H4 tetramer (Dong and van Holde, 1991).

Flexible dinucleotides influence DNA bending. Drew and Travers (1985) found that stretches of A/T nucleotides prefer to bend so that the minor groove of the DNA double helix forms the inside of a curve, whereas stretches of G/C nucleotides prefer to curve with the minor groove facing out. A periodicity of AA or TA dinucleotides occurring every 10 bp, where the minor groove of the DNA faces the histone octamer, has been shown to increase the affinity of the DNA sequence for the histone octamer (van der Heijden et al., 2012; Segal et al., 2006; Widom, 2001) as these sequences facilitate DNA bending. GC and CG dinucleotides are also found with a 10 bp periodicity, offset by 5 bp from AA or AT dinucleotides, where the minor groove faces away from the octamer (Figure 1.3). Pich et al. (2018) hypothesised that nucleosome locations define minor groove orientations, which has an effect on DNA damage/repair. This in turn reinforces the nucleosome-favouring 10-bp

periodicity, which may explain why the DNA of eukaryotes such as yeast has a higher affinity for nucleosomes than that of prokaryotes, which does not form chromatin. It has been suggested that the eukaryotic genome has evolved to accommodate the bending required to wrap around the histone octamer (Zhang et al., 2009).

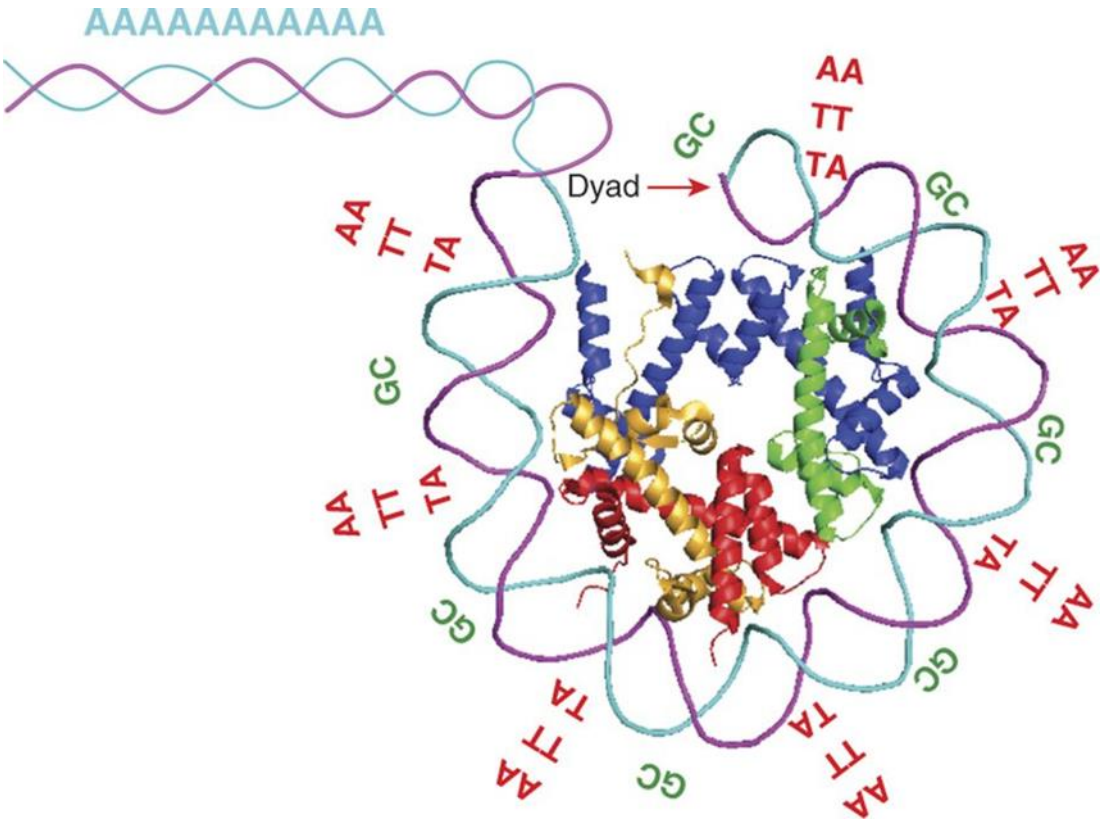

Figure 1.3. DNA Sequence and Nucleosome Positioning.
Poly(dA:dT) tracts are poor substrates for nucleosome binding, whereas sequences with a 10 bp periodicity of AA/AT dinucleotides, phased with GC dinucleotides have a high affinity for the histone octamer (Struhl and Segal, 2006).

In contrast, poly(dA:dT) tracts (more than 3 A/T nucleotides in succession) are unlikely to form nucleosomes. These have a stiff structure which is refractory to nucleosome positioning due to its sequence properties, and not to the binding of an external factors (Suter et al., 2000). These sequences and surrounding regions are significantly depleted on nucleosomes *in vivo* (Field et al., 2008; Yuan et al., 2005), and are unlikely to position nucleosomes *in vitro* (Anderson and Widom, 2001). Nucleosomes formed over these DNA sequences *in vitro* have a distorted structure which is less stable than a canonical nucleosome (Bao et al., 2006). As mentioned in section 1.2.1, these "anti-positioning" sequences may be exploited by certain classes of promoters, particularly prevalent in yeast genomes, creating nucleosome-free regions which are more accessible to transcription factors (Field et al., 2008; Struhl, 1985). The strong anti-positioning properties of these sequences are likely to

strengthen the positioning of nucleosomes on either side, which may form a barrier against which further nucleosomes may be packed.

This intrinsic positioning code has been reported to account for approximately 50% of nucleosome positioning *in vivo* (Segal et al., 2006), though this is disputed by the barrier model which suggests that only the +1 and -1 nucleosomes are specified by the DNA sequence and that these form a barrier against which further nucleosomes are packed (Mavrich et al., 2008). This has been tested by comparing the nucleosome occupancy of genomic DNA *in vivo* with that following reconstitution on DNA *in vitro* by salt dialysis. Kaplan et al. (2009) and Fraser et al. (2009) both find that nucleosomes reconstituted *in vitro* are correlated with those found in both yeast and mammalian DNA respectively. This method of reconstitution by salt dialysis relies on intrinsic histone-DNA interactions which are dependent on DNA sequence, whereas reconstitution using histone chaperones causes weaker nucleosome positioning (Zhang et al., 2009), suggesting that the effects of DNA sequence will be less dominant *in vivo*. Reconstituting genomic DNA with histones from foreign organisms indicates that DNA sequence plays a greater role in defining primary chromatin structure than histone amino acid sequence (Kaplan et al., 2009; Segal and Widom, 2009a; Truong and Boeke, 2017).

However, there are some significant features of chromatin structure *in vivo* that are not captured by histone-DNA interactions *in vitro*. Nucleosome positioning data (from MNase-seq or cuprous phenanthroline cleavage of chromatin followed by sequencing) show that the +1 nucleosome is often strongly positioned *in vivo* (Gencheva et al., 2006; Kaplan et al., 2009; Mavrich et al., 2008; Segal et al., 2006), and it may be that this feature requires transcriptional initiation by RNA pol I (Zhang et al., 2009). Elsewhere in the genome, chromatin structure is dependent on the binding of transcription factors that may compete with the histone octamer for various DNA sequences.

1.4 Chromatin Structure *in vitro*

Studying the structure of the higher-order chromatin fibre *in vivo* is challenging due to the occluded nature of chromatin within cells. The nucleus is impermeable to many chromatin-probes, and chromatin in the nucleus cannot be visualised by methods such as conventional microscopy. Releasing chromatin from the cell in order to study its conformation is likely to change its structure. Cryo-electron microscopy and small-angle x-ray scattering results have suggested that the higher-order structure of chromatin *in situ*, during both mitosis and interphase, is not the canonical 30-nm chromatin fibre, but forms an irregular structure (Joti et al., 2012; Nishino et al., 2012), however, these methods can only be used to visualise

general chromatin structure, and not to investigate specific loci, for example in active or heterochromatic regions. New sequencing-based techniques to probe localised chromatin structure in situ, as described in section 1.1.2.2, have thus far only identified structures on the tetra-nucleosome scale (Hsieh et al., 2015; Ohno et al., 2019; Risca et al., 2017).

In addition, the incredibly complex regulation of chromatin structure makes it difficult to measure the contribution of any one aspect of this regulation in the absence of confounding factors. *In vivo* there are multiple determinants of nucleosome positioning: chromatin remodelling enzymes, non-histone protein-DNA interactions and transcription of DNA. To analyse for example the contribution of DNA sequence to the structure of the complex, it would be difficult to determine what direct role the sequence plays, as changing the DNA sequence would possibly cause indirect effects, for example by recruiting transcription factors or histone modifiers.

To carefully study the structure (10-30 nm) of the chromatin fibre, *in vitro* models of chromatin structure have been developed, forming chromatin from its purified component outside of a cellular environment. To avoid confusion, throughout this thesis I will use the term *in vitro* to refer to chromatin explicitly in this form, when it is reconstituted from the purified components. I will not use the term *in vitro* to refer to chromatin inside cells or isolated nuclei, whether these are *in situ*, in tissue culture or otherwise isolated.

Previously, *xenopus* egg extracts, which contain a large excess of core histones have been used to reconstitute chromatin *in vitro*. These extracts also contain molecular chaperones that space nucleosomes in arrays (Almouzni and Wolffe, 1993). More recently, nucleosome arrays have been reconstituted from their purified components (DNA and core histones) *in vitro*. Mixing DNA and core histones together at high (2 M) salt concentrations overcomes the electrostatic interactions between the negatively charged DNA and the positively charged histones. Slowly reducing the salt concentration by gradient dialysis allows DNA to bind and wrap around histones without promoting precipitation. Wolffe (1998) states that reconstituting chromatin by salt dialysis causes nucleosomes to pack together on DNA rather than regularly spacing themselves, with one nucleosome found every 150-160 bp. Simpson et al. (1985) overcame this problem by reconstituting chromatin onto regularly spaced nucleosome positioning sequences (the 5S RNA gene) which have a high affinity for the histone octamer.

Nucleosome arrays can be compacted into folded and higher-order structures by varying the concentration of monovalent and divalent cations (Thoma et al., 1979). Schwarz and Hansen (1994) showed that chromatin *in vitro* can be compacted to the same degree that is achieved

in vivo, indicating that chromatin compaction is constrained by the charge of the DNA phosphodiester backbone, which can be shielded by these cations (Wolffe, 1998).

The artificial "601" sequence of DNA was identified by Lowary and Widom who used SELEX experiments to identify 147 bp DNA sequences with a high affinity for the histone octamer (1998). Arrays formed of repeats of this positioning site form regular spaced arrays and well-compacted chromatin fibres *in vitro*. "601" fibres have since been incorporated into constructs containing varying numbers of repeats, and with linker DNA of varying different lengths. Most studies done to date have used a repeat length of 147+10n bp (as this maintains multiple full twists of the DNA helix between nucleosomes). A DNA sequence containing regular repeats of the "601" positioning site can be reconstituted into a very regular higher-order chromatin fibre, which has been used to assess the structure of this 30-nm fibre (Robinson and Rhodes, 2006a; Robinson et al., 2006b; Routh et al., 2008; Schalch et al., 2005; Song et al., 2014). This template has been extremely useful for structural analysis of chromatin *in vitro*, particularly as the extreme regularity of the structure makes it possible to form crystals for x-ray analysis (Schalch et al., 2005). However, the repetitive nature of these arrays also limits the experimental possibilities as it is impossible to sequence the arrays and map fragments to a single location. Furthermore, the abnormal stability and regularity of these fibres makes them questionable models to represent the irregular structure that the human genome probably adopts.

Using a *601* repeat sequence, with repeat lengths between 177 bp and 237 bp, Robinson et al. (2006b) reconstituted chromatin fibres in the presence of linker histones to fold the arrays into higher-order structures and studied them by electron microscopy. They reported that the fibres folded into a one-start solenoid structure, and that the measurements of the compacted chromatin fibres were not compatible with the measurements of a two-start helix. Schalch et al. (2005) performed x-ray crystallography on tetranucleosomes with a 167 bp repeat, compacted in the presence of 20-60mM Mg^{2+}. They found that under these conditions, the chromatin fibre forms a two-start helical structure, with two stacks of two histones beginning to twist round one another. It is likely that arrays containing short nucleosome repeat lengths (167 bp) are constrained to form a two-start helix, but that those with a longer 197 bp repeat are able to form a one-start solenoid which is dependent on the linker histone for compaction (Routh et al., 2008). However, Song et al. (2014) used arrays of 12 nucleosomes with a 177bp or a 187bp repeat length and reported a similar two-start structure to Schalch et al. with these longer arrays, in disagreement with the results of Robinson et al. (2006b), showing that it is difficult to form a consensus on the higher-order structure of these fibres even under these controlled conditions.

Most studies *in vitro* have used a nucleosome repeat length satisfying the equation "147 + 10n bp", which allows a number of full turns of the DNA helix between adjacent nucleosomes. Brouwer et al. (in preparation) have found by single-molecule force spectroscopy that varying the linker length +/- 5 bp causes changes to the higher-order structure which impacts the unfolding dynamics of the complex.

1.5 Thesis Aims

The impact of DNA sequence on nucleosome positioning is well studied. It is likely that this has a significant impact on the structure and compaction of the higher-order chromatin fibre, but this is not so well understood. I set out to analyse the contribution of the underlying nucleotide sequence on the structure and dynamics of the higher-order chromatin fibre. I hypothesised that sequences with a lower affinity for the histone octamer or weaker nucleosome positioning properties than the well-studied "601" sequence might be expected to form less compacted higher-order structures which are more easily disrupted.

I first aimed to characterise *in vitro* reconstituted nucleosome arrays based on defined DNA sequences. Non-repetitive DNA templates containing biologically derived nucleosome positioning sites were expected to form heterogeneous nucleosome arrays that might be a more accurate representation of bulk chromatin structure *in vivo* than the "601". I aimed to examine the reconstitution efficiency of these sequences as well as the nucleosome positioning and the heterogeneity within each of these fibre populations.

I next examined the higher-order structure and dynamics of chromatin fibres based on these defined DNA sequences. Studies of bulk chromatin samples were first used to compare the structures and shapes of populations of chromatin fibres, however, population analyses were found to be limited due to the increased heterogeneity of non-601 chromatin fibres. I therefore used a single-molecule assay to probe the structure and dynamics of these different chromatin structures.

Finally, I aimed to uncover the impact of a single nucleosome disruption within a compacted higher-order chromatin structure. A single nucleosome positioning site can have profound effects upon higher-order chromatin structure, and this may be an important mechanism for gene transcription (Caplan et al., 1987).

Chapter 2. Materials and Methods

All chemicals were purchased from Sigma Aldrich and all restriction enzymes were purchased from New England Biolabs, unless stated otherwise.

2.1 Plasmid Cloning and Purification

2.1.1 Cloning of Plasmid Sequences

Standard molecular biology techniques were performed as described in Sambrook et al., (2001). A plasmid containing 26 repeats of a 197 bp "601" sequence in a pUC18 vector was provided by John van Noort. The repetitive nature of this plasmid meant that the number of repeats is inherently unstable; individual clones grown at 37°C following transfection into DH5α cells contained between 19 and 26 repeats of the "601" sequence, however, no plasmids containing 25 repeats were identified. It is possible that a single copy of the 197 bp repeat is never deleted due to constraints in the bending of the DNA required to achieve this. A plasmid containing 25 copies of the 601 sequence (*601*) and one with 24 copies of the "601" sequence and a central low affinity nucleosome (*601+LA*) were constructed as follows: A plasmid containing 24 repeats of the 197 bp "601" site was partially digested with *Ava*I, which cuts between each of the "601" nucleosome positioning sites (Figure 2.1A). Linearised sequences containing 24 "601" sites were isolated by gel extraction from NuSieve GTG Agarose (Lonza) with an E.N.Z.A. gel extraction kit (Omega), and the 5' base was dephosphorylated by Antarctic Phosphatase (New England Biolabs). 3' phosphorylated DNA primers were designed to be annealed into a 26 bp dsDNA insert (Figure 2.1B, described in appendix 1) with *Ava*I overhang sites and containing *Bsg*I restriction sites that would cut outwith the insert once it was ligated to the linearised plasmids and used to recircularise the DNA. These Plasmids, which contained insertion sequences at any of the *Ava*I sites between each 601 sequence were, transformed into Stbl2 competent cells (Thermo Fisher), clones were grown on agar plates and miniprepped (Quiagen). *Eco*RV and *Bsg*I digestion were used to identify clones where the insert appeared in the centre of the template, with 12 "601" sequences on each side (Figure 2.1C). These plasmids were cut with *Bsg*I, and linear dsDNA (GeneArt Strings by Thermo Fisher, see appendix 1) containing either "601" or a low affinity positioning sequence were cloned into the plasmid, such that the 197 base pair repeat was maintained, *Bsg*I restriction sites were lost, *Ava*I restriction sites between nucleosome positioning sites were reformed, and the inserted "601" sequence was indistinguishable from the other repeats (Figure 2.1D).

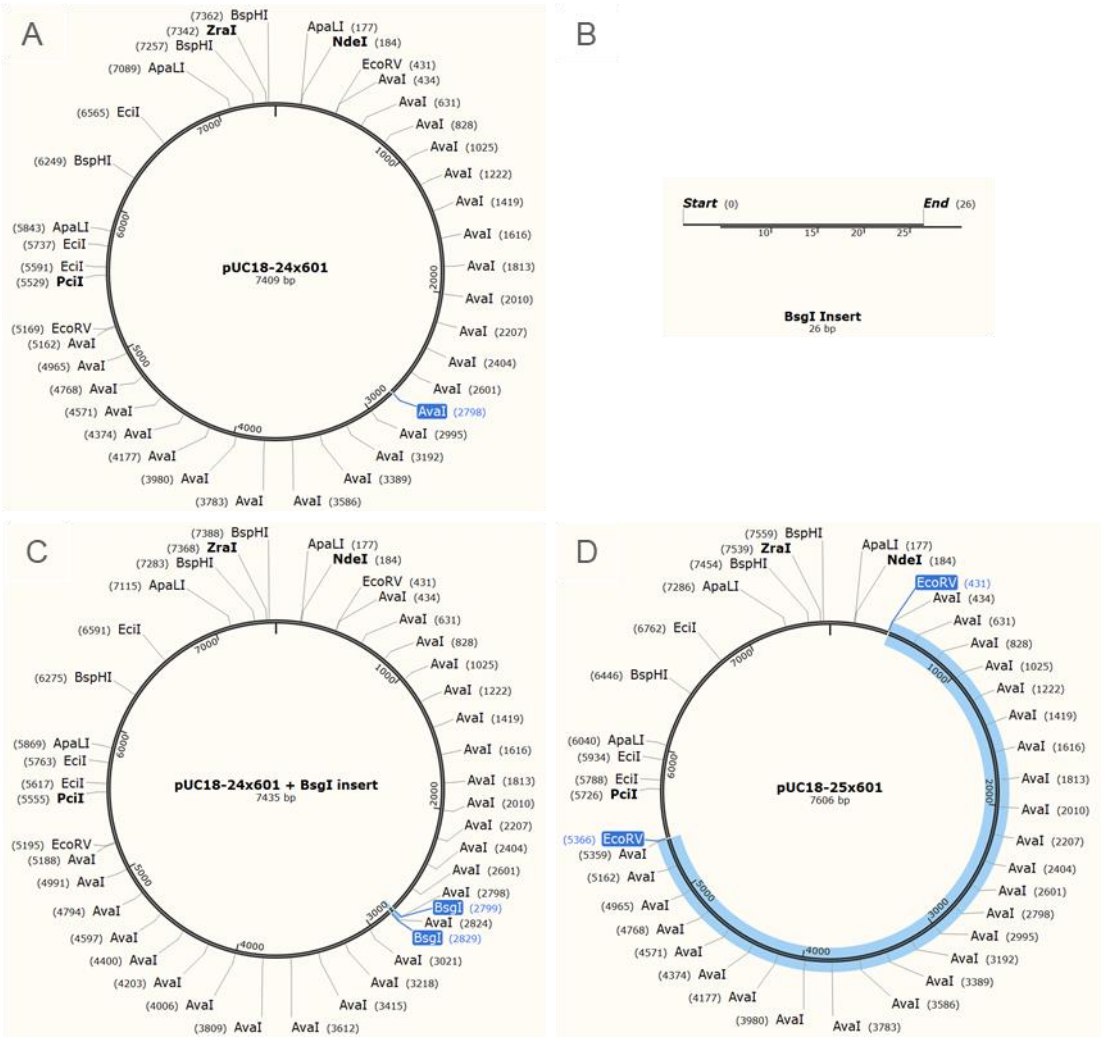

Figure 2.1. Cloning a 25 x 601 DNA template and a 601 template incorporating a low affinity nucleosome positioning site.
A) A 24 × 601 bp template within a pUC18 vector was partially digested with AvaI and the 5' base was dephosphorylated. B) A 26 bp sequence (formed from annealed primers, each with a phosphorylated 3' base) containing two *Bsg*I restriction sites which cut outwith the insert. C) A 26 bp insert cloned into the centre of a 24 × 601 template within a pUC18 plasmid, with 12 "601" repeats appearing either side of the insert. A clone with correct placement of the insert was identified by restriction digestion and gel electrophoresis. D) A plasmid containing a 25 × 601 template (*601*). The central "601" or "Low Affinity" sequence (found in appendix 1) was cloned into the plasmid following digestion with *Bsg*I, located in the centre of the template with 12 "601" repeats flanking the inserted site.

BLG and *601/BLG* sequences were designed by Nick Gilbert and Jim Allan and synthesised by GenScript (China) in a pUC57 based-vector. Maps of all vectors are found in Appendix 1.

2.1.2 Plasmid Preparation

To maintain repeat stability (Figure 2.2) "601" plasmids were grown at 30°C in LB supplemented with 100 µg/ml ampicillin. Plasmids were isolated from *E. coli* using the

Invitrogen PureLink HiPure Plasmid Maxiprep Kit according to the manufacturer's instructions, or by alkaline lysis followed by plasmid purification by caesium chloride sedimentation. 3 ml starter cultures were grown using 50 µl of glycerol stock or by picking single colonies from an agar plate, overnight at 30°C then were diluted 1/100 into fresh LB and grown for a further 16 hours at 30°C. 200 ml bacterial culture were pelleted by centrifugation at 5,000 rpm in an SLA-1500 rotor (Sorvall) for 10 min and resuspended in 10 ml 50 mM Tris-HCl (pH 8) and 10 mM EDTA supplemented with 10 µg/ml RNaseA (Invitrogen). An equal volume of 200 mM NaOH and 1% SDS were added, and cells were lysed at room temperature for 10 min. An equal volume of 3 M Potassium Acetate (pH 5.5) was added to precipitate genomic DNA. The precipitate was removed by centrifugation, and plasmid DNA was precipitated from the remaining solution using isopropanol. DNA was resuspended in TE buffer (10 mM Tris-HCl (pH 8) and 1 mM EDTA), CsCl was added to give a density of 1.6 g/ml and ethidium bromide was added to a final concentration of 1.2 mg/ml. Samples were centrifuged at 80,000 rpm overnight in a Ti100.3 rotor (Beckman), separating protein, RNA and DNA based on their different densities. Ethidium bromide bands containing plasmid DNA were extracted and EtBr was removed by washing with water saturated butan-1-ol. Plasmid DNA was precipitated using ethanol and resuspended in TE.

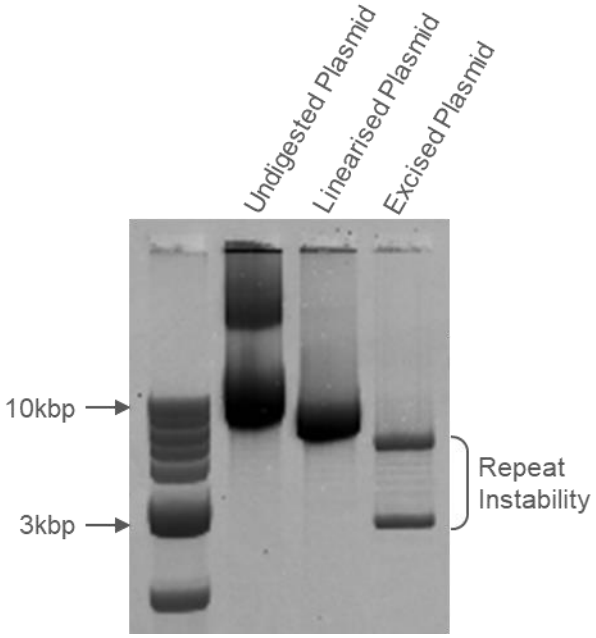

Figure 2.2. Repeat Instability of Repetitive Plasmid DNA Sequences.
Plasmid DNA (undigested, linearised by ZraI, or with the 26 x 601 template excised from the plasmid by EcoRV) analysed by agarose gel electrophoresis. Multiples of 197 bp repeats were deleted in a minority of plasmids, creating a ladder, most clearly seen when the DNA template was excised from the plasmid.

2.2 Chromatin Reconstitution

2.2.1 Materials

2.2.1.1 Template DNA

DNA templates (see Appendix 1) for chromatin reconstitution were excised from the plasmid by *Eco*RV (for the *601* and *601+LA* sequences) or *Xho*I (*BLG* and *601/BLG* sequences) restriction digestion to give an approximately 5 kbp fragment. The vector backbone (approximately 2.5 kbp), was digested into fragments smaller than 500 bp using *Apa*LI, *Bsp*HI and *Eci*I.

Digested plasmid was fractionated by gel filtration chromatography as described by Rogge et al. (2013). A column (65 cm height and 1.6 cm diameter) packed with sephacryl S 1000 (GE Healthcare) in a buffer containing 10 mM Tris-HCl (pH 7.5), 1 mM EDTA and 50 mM NaCl, was used with a flow rate of 0.6 ml/min. Digested plasmid DNA was loaded on to the column, and fractions were collected every 2.5 mins. Fractions containing template DNA were identified by agarose gel electrophoresis in 1 × TBE buffer.

Small amounts of DNA templates for single-molecule force spectroscopy were instead isolated by agarose gel extraction from SeaPlaque agarose (Lonza) using β-Agarase I (New England Biolabs).

2.2.1.2 Competitor DNA

Two different types of DNA competitor were used in the study, in order to prevent the oversaturation of template DNA by core histones. Firstly, a 147 bp fragment as described by Huynh et al. (2005) was included. This fragment was derived from a section of the vector backbone of the *601*-containing (pUC18) plasmid and was isolated by PCR amplification and gel extraction of the 147 bp fragment by Jim Allan. I performed subsequent PCR amplification of the competitor directly from the isolated 147 bp fragment, eliminating the need for gel extraction. Following PCR, excess primers and nucleotides were removed using a S-300 microspin column (GE Healthcare) according to the manufacturer's instructions.

Due to difficulties in purifying large quantities of template and 147 bp competitor DNA for large-scale studies, I also used a second strategy employing the vector backbone DNA following plasmid digestion as described in section 2.2.1.1. This fragment of DNA, approximately 2.5 kbp in length was sometimes included as a whole fragment, and was sometimes digested into fragments smaller than 500 bp using *Apa*LI, *Bsp*HI and *Eci*I restriction enzymes.

2.2.1.3 Histones

Recombinant *Xenopus* core histone octamers were kindly provided by Ramasubramanian Sundaramoorthy (Tom Owen-Hughes group, Dundee), having been prepared essentially according to Luger et al. (1999).

Chicken erythrocyte core histones were purified from chicken blood as described by Peterson and Hansen (2008) with minor modifications. Chicken erythrocyte chromatin is predominantly transcriptionally inactive and the cells are non-replicating so histones have relatively few post-translational modifications that might subsequently affect the structure of reconstituted nucleosome arrays (Peterson and Hansen, 2008). Erythrocyte nuclei were prepared from chicken blood by Jim Allan, and stored at –80°C. Core histones were isolated from nuclei as follows: nuclei were partially digested by micrococcal nuclease (New England Biolabs or Worthington Biochemical Corporation) and the soluble fraction of chromatin was isolated (Figure 2.3A). Linker histones H1 and H5 were removed from oligo-nucleosomes by Sepharose CL-4B (Amersham Biosciences) chromatography in 650 mM NaCl. The column was 35 cm in height and 2.6 cm in diameter with a flow rate of 2 ml/min.

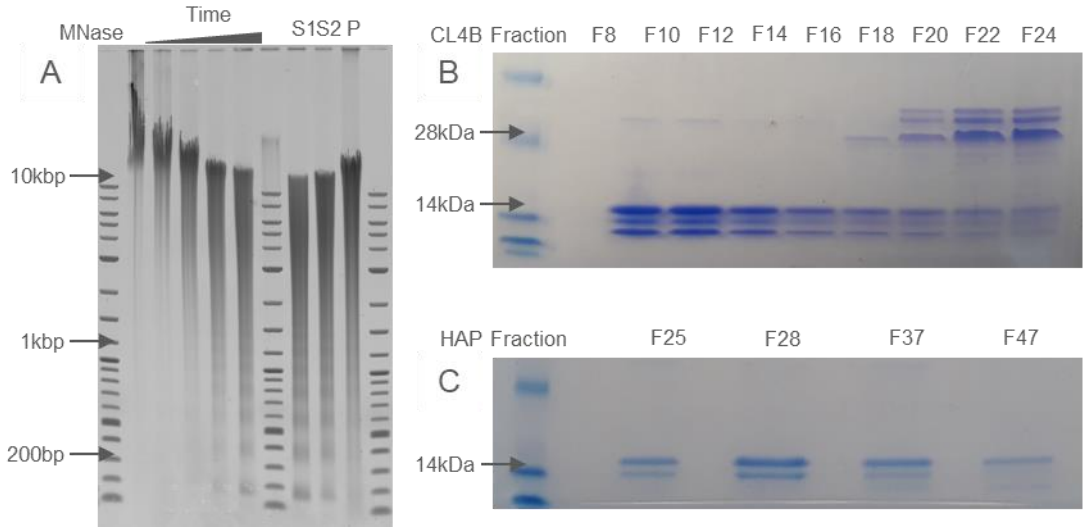

Figure 2.3. Isolation of Core Histones from Chicken Erythrocyte Nuclei.
A) Digestion of chicken erythrocyte chromatin by micrococcal nuclease over 30 min and release of soluble chromatin. Soluble fractions S1 and S2 were used to make core histones, insoluble fraction P was discarded. B) Fractions from Sepharose CL-4B chromatography to separate core and linker histones. Core histones were eluted from fraction 10, linker histones began to be eluted from fraction 18. Small amounts of H3 dimers are seen in fractions 10 and 12. Fractions 9-16 were pooled for HAP chromatography. C) Fractions from HAP chromatography to remove core histones from DNA. Only small amounts of protein were loaded on the gel so not all histone types could be observed. Fractions 25-41 were pooled based on A280. Following concentration, all histone types could be visualised after polyacrylamide gel electrophoresis and staining (See Figure 3.3). Samples were fractionated on 12% polyacrylamide gels alongside SeeBlue Plus2 protein standards.

Fractions were collected every 2.5 minutes and examined on a polyacrylamide gel to identify fractions containing core and linker histones (Figure 2.3B). Core histones were removed from DNA by hydroxyapatite chromatography. Hydroxyapatite was prepared according to the protocol described by Miyazawa and Thomas (1965). Oligo-nucleosomes were bound to HAP in a column 35 cm in height and 2.6 cm in diameter. Core histones were eluted using 2 M NaCl at a flow rate of 2 ml/min and fractions were collected every 2.5 minutes. Fractions containing core histones were identified by polyacrylamide gel electrophoresis (Figure 2.3C). Histone octamers were concentrated using Amicon Ultra Centrifugal Filters with 10 kDa molecular weight cut off (Millipore) and stored at -20°C.

Linker histone H5 was isolated from chicken erythrocyte chromatin by Jim Allan by hydroxyapetite chromatography of native chromatin to isolate linker histones, followed by amberlite chromatography to separate linker histones H1 and H5, according to Allan et al. (1980b).

2.2.1.4 BSA

Bovine serum albumin (New England Biolabs) was included in chromatin reconstitutions to prevent chromatin from adhering to membranes and storage tubes. This was not included in samples for small-angle X-ray scattering analysis.

2.2.2 Reconstitution by Salt Dialysis

Template DNA and histones were combined in the presence or absence of competitor DNA in 2 M NaCl, with 10 mM Tris-HCl (pH 7.5), 0.2 mM EDTA, with 1 mg/ml BSA.

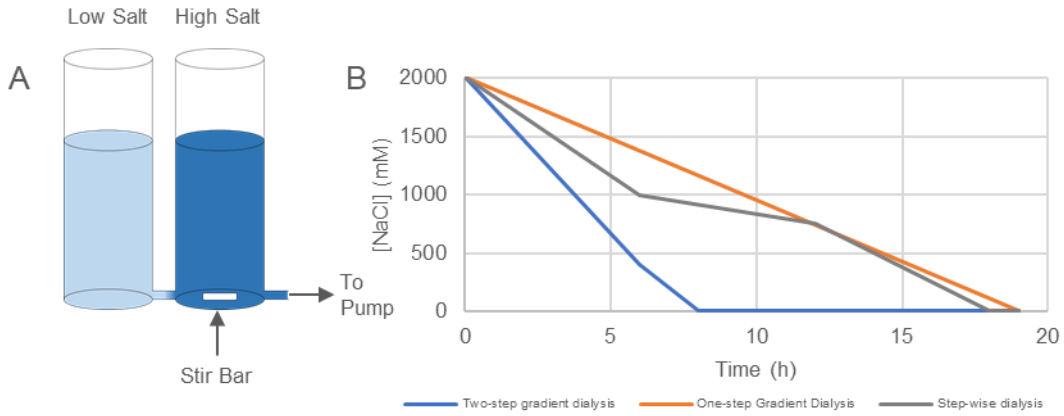

Figure 2.4. Chromatin Reconstitution by Salt Dialysis.
A) Linear gradient maker used to dialyse chromatin from 2 M to 400 mM NaCl. B) Rate of dialysis for three different methods of reconstitution.

Reconstitutions were routinely dialysed from high salt to low salt by two-step gradient dialysis in Slide-A-Lyzer Mini Dialysis Units with 10 kDa molecular weight cut off (Thermo Scientific). Dialysis caps were placed in a linear gradient maker (Figure 2.4A) at 4°C for approximately 6 hours, over which time the NaCl concentration decreased from 2 M to 400 mM. Dialysis units were removed from the gradient maker and dialysed directly into low salt buffer (2.5, 10 or 80 mM NaCl) overnight.

For single-molecule force spectroscopy experiments, chromatin was prepared as described by Kaczmarczyk et al. (2017). Chromatin was dialysed in mini dialysis units in a linear gradient from 2 M to approximately 10 mM NaCl over 19 hours (one-step gradient dialysis).

For some initial SAXS experiments, step-wise salt dialysis was performed as described by Rogge et al. (2013), where chromatin is dialysed from 2 M to 1 M NaCl over 5-6 hours, 1 M to 750 mM NaCl overnight, then to 2.5 mM NaCl over 5-6 hours in dialysis tubing. The rate of change of NaCl concentration for of each of these three dialysis methods are shown in Figure 2.4B.

Chromatin reconstituted using any of these techniques was not torsionally constrained, and therefore no positive or negative supercoils would be expected to form during the assembly of the chromatin.

2.2.3 ATP-dependent Chromatin Assembly by NAP-1 and ACF

Assembly of chromatin *in vitro* using the purified histone chaperone NAP1 and chromatin assembly and remodelling factor ACF is described by Fyodorov and Kadonaga (2003). Chromatin assembly was achieved using a Chromatin Assembly Kit (Diagenode), which included recombinant human core histones, NAP1 and ACF. In brief, core histones were combined with NAP1 in a chromatin assembly buffer. DNA, ATP and ACF are subsequently added, and chromatin was assembled over 3 hours with incubation at 30°C.

Due to the large amounts of chaperone protein required to assemble small amounts of chromatin, this method was not suitable for structural analyses such as SAXS, therefore salt dialysis was primarily used to prepare chromatin.

2.3 Gel Electrophoresis

2.3.1 Agarose Gel Electrophoresis

Unless otherwise stated electrophoresis was performed in 1% agarose (Biogene) gels in 0.5 × TBE (44 mM Tris base, 44 mM Borate, 2 mM EDTA) in the presence of 500 ng/ml ethidium

bromide. Gels were imaged using a geldoc (Syngene) or scanned with a laser scanner (Fuji FLA-5100) and lanes were analysed using Aida Image Analyser.

Unless stated otherwise, NEB 2-Log DNA Ladders were used.

2.3.2 Electrophoretic Mobility Shift Assay

EMSA was used to measure a band shift in chromatin compared to DNA, as reconstituted chromatin runs more slowly on an agarose gel than unreconstituted DNA or partially reconstituted nucleosome arrays. As described by Huynh et al. (2005), chromatin was fractionated on 0.7% agarose gels in 0.25 × TBE (22 mM Tris base, 22 mM Borate, 1 mM EDTA) in the absence of ethidium bromide, for 3 hours at 100 V. Gels were subsequently stained in 500 ng/ml ethidium bromide and scanned using a laser scanner (Fuji FLA-5100).

Unless stated otherwise, NEB 2-Log DNA Ladders were used.

2.3.3 Polyacrylamide Gel Electrophoresis

Protein samples were fractionated in NuPage LDS sample buffer (Invitrogen). Samples were heated to 95°C for 5 min with 50 mM DTT to denature proteins and break disulphide bonds. NuPage 12% bis-tris gels or 4-12% bis-tris gradient gels (Invitrogen) were ran at 100 V in MOPS running buffer (Invitrogen). Gels were subsequently stained with coomassie (50% Methanol, 10% Glacial Acetic Acid, 0.25% Brilliant Blue R250) shaking at room temperature for 1 hour. Gels were destained (50% Methanol, 10% Glacial Acetic Acid) and photographed.

Unless stated otherwise, Thermo Fisher SeeBlue Plus2 protein standards were used.

2.4 Electron Microscopy

2.4.1 Preparation of Samples

Samples for electron microscopy were prepared by dialysing chromatin into 10 mM triethanolamine-HCl (pH 7.5), 0.2 mM EDTA and sodium chloride (2.5, 10 or 80 mM NaCl). If required, magnesium chloride was added following dialysis. Chromatin samples taken from a sucrose gradient (approximately 5 ng/µl) or chromatin taken directly from reconstitution (25 ng/µl) were fixed with 0.1% glutaraldehyde at 4°C for approximately 18 hours.

Chromatin was adhered to 3.05 mm copper grids with a 200 mesh formvar/carbon support film (TAAB). Chromatin was coated in benzalkonium chloride (BAC) by incubating in 2 × 10^{-4} % BAC for 1 hour at room temperature and was applied to grids for 5 mins. Grids were

washed twice with deionised water, dehydrated with 90% ethanol and blotted dry before shadowing.

2.4.2 Platinum Shadowing

For contrast enhancement, grids were rotary shadowed with 3 nm platinum at an angle of 7° using a Leica ACE600 vacuum evaporator with a Pt-loaded electron gun. Jim Allan optimised the techniques to prepare these slides and Alex Makarov and Charles Dixon assisted me in the use of this equipment.

2.4.3 Imaging

Samples were examined in a JEOL JEM-1400Plus Transmission Electron Microscope at 80 kV at a magnification of ×20,000-30,000. Alex Makarov and Stephen Mitchell assisted me in the use of this microscope at the Centre Optical Instrumentation Laboratory, King's Buildings, Edinburgh.

2.4.4 Image Analysis

ImageJ scripts for counting nucleosomes in unfolded chromatin and the area measurement of folded chromatin fibres were written by Davide Michieletto. Images were cropped to isolate individual chromatin fibres, then blurred to obscure the background. Nucleosomes were identified using a Phansalkar thresholding approach as this copes well with low contrast and non-uniform images (Phansalkar et al., 2011). Initially, overlapping nucleosomes were counted as single objects and then a watershed segmentation was used to separate individual nucleosomes (See Figure 3.6). ImageJ's built-in watershed segmentation calculates the Euclidian distance map of the foreground areas found by thresholding (measuring the distance from the centre to the edge of each object), then dilates each of the ultimate eroded points in the map to form a "dam" separating overlapping objects.

2.5 Digestion of Chromatin

2.5.1 Restriction Digestion

Chromatin was digested with the restriction enzymes *Ava*I, *Psi*I and *Pfl*23II (Thermo Fisher Scientific) in RB50 buffer (10 mM HEPES–KOH, pH 7.6, 50 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM EGTA) as described by Maier et al. (2008).

100 ng of chromatin was digested with 5 units of *Ava*I or *Psi*I or 3 units of *Pfl*23II overnight at 4°C. *Psi*I and *Pfl*23II digests were added to genomic lysis buffer and proteinase K to dissociate and degrade proteins, DNA was then purified by phenol/chloroform extraction and precipitated with ethanol. *Ava*I digests were analysed as native chromatin (to see the fraction

of mono-nucleosomes vs unreconstituted “601” sites) on a 1.1% agarose gel, as described by Huynh et al. (2005).

2.5.2 Micrococcal Nuclease Digestion

Digestion by micrococcal nuclease (New England Biolabs) was performed in the presence of 10 mM Tris-HCl (pH 7.5), 10 mM NaCl, 0.2 mM EDTA, 2 mM $CaCl_2$ with 0.1 mg/ml BSA. 0.75 units of MNase was used to digest 2 μg of chromatin in 150 μl buffer at 37°C. Fractions were removed at different time points up to 30 mins and added to an equal volume of 2 × genomic lysis buffer (300 mM NaCl, 20 mM EDTA, 1% SDS) to stop digestion. Proteinase K was added, and samples were incubated at 55°C for 15 mins. DNA was then purified by phenol/chloroform extraction, precipitated with ethanol and analysed by electrophoresing in a 1% agarose gel in TBE buffer.

2.5.3 Preparation of DFF/CAD Nuclease

Constructs for the expression of recombinant DFF/CAD nuclease, engineered to replace the site of caspase-3 cleavage with a TEV protease cleavage site, were provided to the Gilbert laboratory by William Garrard and Fei Xiao. These were produced in E. coli and purified essentially as described by Xiao et al. (2007).

Both the human and the mouse isoforms of the nuclease have previously been cloned into a pRSF-Duet1 (Novagen) expression construct. These were transformed into BL21 (DE3) *E. coli* (Novagen) and colonies were picked from agar plates. 3 ml starter cultures were grown from individual colonies overnight, then 1 ml was diluted into 100 ml fresh LB and grown at 30°C in the presence of 25 μg/ml chloramphenicol and 50 μg/ml kanamycin until it had reached an A600 of 2. IPTG was added to 1 mM to induce expression of DFF, and the culture was then grown at 16°C overnight.

Bacteria were collected by centrifugation, washed in PBS, and resuspended in binding buffer (300 mM sodium chloride, 15 mM imidazole, 50 mM Tris-HCl (pH8), 10% glycerol, 10 mM 2-merceptoethanol, PMSF and ETDA). The cell slurry was incubated on dry ice for 30 min, thawed and lysozyme was added to a final concentration of 1 mg/ml. The culture was then incubated on ice for a further hour. Samples were sonicated for 15 min and triton-X 100 was added to 0.1%. The cell debris was removed by centrifugation at 16,000 g. The soluble fraction was added to nickel-agarose beads and incubated at 4°C overnight.

Beads were washed 3 times in wash buffer (300 mM NaCl, 20 mM imidazole, 50 mM Tris-HCl (pH8), 10% glycerol, 10 mM 2-mercaptoethanol, PMSF) then transferred to a column and washed again with 2 volumes of wash buffer. Protein was eluted with 10 ml elution

buffer (identical composition to wash buffer but containing 250 mM imidazole) and 0.4 ml fractions were collected. Fractions containing DFF were identified by polyacrylamide gel electrophoresis (Figure 2.5A and B) and pooled. These were buffer exchanged into 2 × storage buffer (100 mM KCl, 20 mM Tris-HCl (pH8), 0.2 mM EDTA, 2 mM DTT, 10% glycerol) using a G-25 sephadex column. BSA was added to 200 µg/ml and glycerol was added to 50%. Proteins were stored at -20°C.

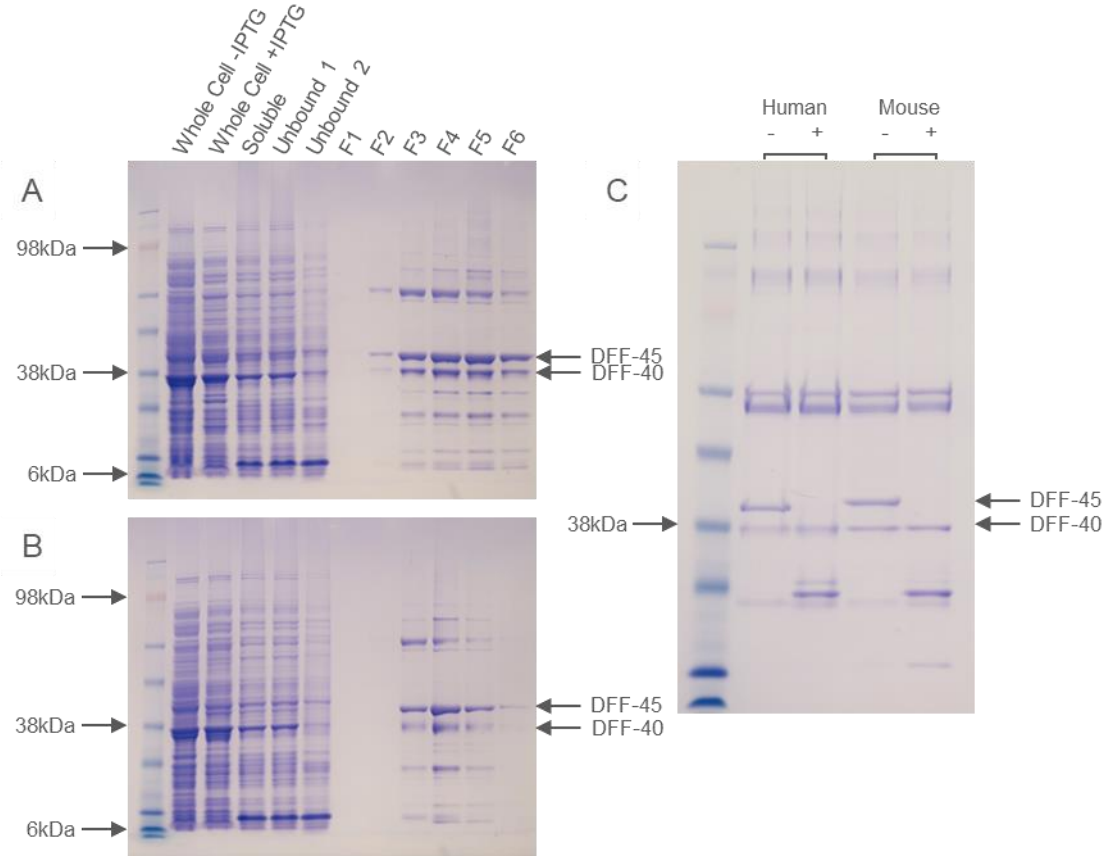

Figure 2.5. Preparation of Recombinant DFF/CAD.
Stages of purification of human (A) and mouse (B) DFF. Whole cell extracts with and without induction of DFF by IPTG, soluble protein, fractions not bound to Ni-Ag beads, fractions eluted from Ni-Ag beads. Human DFF was found in fractions 3-6, mouse DFF was found in fractions 3-5. C) Activation of human and mouse DFF by cleavage of the 45 kDa inhibitory subunit by TEV protease for 30 mins at 37°C. Protein was fractionated on a 4-12% polyacrylamide gel.

2.5.4 Digestion of Chromatin by DFF/CAD Nuclease

DFF/CAD nuclease was activated by TEV protease (Invitrogen), which cleaves the DFF-45 inhibitory subunit in the manner of caspase-3, releasing the DFF-40 active subunit. DFF was diluted 1:1 with deionised water, then 1 µl TEV protease was added for each 20 µl of diluted DFF and incubated at 37°C for 30 mins. Cleavage was confirmed by polyacrylamide gel electrophoresis (Figure 2.5C).

Activated DFF was added to chromatin in 10 mM Tris-HCl (pH 7.5), 10 mM NaCl, 0.2 mM EDTA and 3 mM $MgCl_2$, and digested at 4°C or 37°C. Fractions of the sample were removed at various time points and added to 2 × genomic lysis buffer (300 mM NaCl, 20 mM EDTA, 1% SDS) to stop digestion. Proteinase K was added and samples were incubated at 55°C for 15 min. DNA was purified by phenol/chloroform extraction, precipitated with ethanol and analysed on a 1% agarose gel in TBE buffer.

2.6 Sucrose Gradient Sedimentation

Chromatin was fractionated on 6-40% isokinetic sucrose gradients (Noll and Noll, 1989), which were prepared in 12 ml ultraclear centrifuge tubes (Beckman) using an isokinetic gradient maker. Up to 200 µl of chromatin (typically 10-20 µg) was layered on top of the gradient using a P1000 Gilson pipette. Gradients were centrifuged in a swing-out SW41 rotor (Beckman) at 41,000 rpm at 4°C for 5 hours. Gradients were fractionated by upward displacement at 1 ml/ml through a UV monitor and logged using an analogue recorder (Anachem). Fractions were collected every 30 sec.

2.7 Caesium Chloride Gradient Sedimentation

Chromatin density was measured by caesium chloride isopycnic gradient sedimentation according to Gilbert and Allan (2001). Chromatin fractions were taken from a sucrose gradient and were dialysed into 10 mM TEA-HCl buffer (pH 7.5) containing 80 mM NaCl and fixed using 0.1% formaldehyde overnight at 4°C. Caesium chloride (Boehringer Mannheim) was added to 48.5% w/w, to a density of 1.55 g/ml. Samples were spun at 20°C in a vertical TV865 rotor (Sorvall) at 50,000 rpm for approximately 50 hours. Gradients were fractionated from bottom to top using a pump speed of approximately 1 ml/min, collecting fractions every 15 sec. The refractive index of each fraction was measured using an Abbe refractometer (Bellingham and Stanley) to determine the CsCl concentration and calculate the density of each fraction.

2.8 Small-angle X-ray Scattering

Small-angle X-ray scattering was performed at Diamond Light Source B21, Harwell Science and Innovation Campus, Didcot, with assistance from Nathan Cowieson, Nikul Khunti and Rob Rambo. Nikul Khunti performed some experiments when prepared samples were mailed directly to the facility. B21 operates in a fixed camera length configuration (4.014 m), suitable for particles with a radius of gyration under 200 Å, at 12.4 keV. B21 can measure a resolution range from 0.0031 to 0.38 $Å^{-1}$ using a flux of ~1012 photons per second and is optimized for solution state SAXS experiments (Diamond Light Source, 2018).

Chromatin was prepared by salt dialysis either using a gradient maker as described in section 2.2.2, or using a step-gradient as described by Rogge et al. (2013), where chromatin was dialysed from 2 M to 1 M NaCl over 5-6 hours, 1 M to 750 mM NaCl overnight, then to 2.5 mM NaCl over 5-6 hours in dialysis tubing. Linker histones or magnesium were added after reconstitution to fold nucleosome arrays. Samples were analysed by SEC-SAXS using a Shodex KW405-4F column in 10 mM Tris-HCl (pH 7.3), 2.5 mM NaCl and 0.2 mM EDTA pumped at 0.25 ml/min, or by direct injection onto the SAXS beam. For direct injection, samples were loaded into a temperature-controlled quartz capillary cell and exposed for 3 min. Data was corrected using a matched buffer and analysed using Scatter (BIOISIS) and Sasview (NSFDANSE) software.

2.9 Single-molecule Force Spectroscopy using Magnetic Tweezers

Single-molecule force spectroscopy was performed according to Kruithof et al. (2009) at the Leiden Institute of Physics, with assistance from Artur Kaczmarczyk and John van Noort.

2.9.1 DNA Preparation

DNA labelled with biotin and digoxigenin was prepared by digesting plasmid DNA with *Nde*I and *Pci*I, generating a fragment containing 25 nucleosome positioning sites plus a total of around 600 bp of additional DNA at the ends of the 5 kbp template. The *Nde*I restriction site was filled in with dATP and labelled with biotin-dUTP (Roche) using DNA Polymerase I Large (Klenow) Fragment. Excess nucleotides were removed by spinning DNA through G-50 microspin columns (GE Healthcare). The *Pci*I restriction site was then filled in with dATP, dCTP and dGTP and labelled with dig-dUTP (Roche) using Klenow. Labelled template DNA was separated from the vector backbone and excess nucleotides by gel extraction from SeaPlaque Agarose (Lonza) using β-Agarase I (New England Biolabs).

Biotin and digoxigenin labelling of DNA was confirmed by dot blot. Protran BA85/20 nitrocellulose membranes were prepared by washing in deionised water for 5 mins followed by 20 min in 20 × SSC, and dried on blotting paper. Samples were spotted onto the membrane with 500-50 femtomol/µl dilutions of a standard primer. DNA was air-dried and membranes were crosslinked by UV. Membranes were then blocked by 30 mg/ml BSA in 100 mM Tris-HCl (pH 7.5) and 150 mM NaCl. Anti-digoxigenin-AP or Streptavidin-AP (Roche) were then added at 0.75 U/ml in 100 mM Tris-HCl (pH 7.5) and 150 mM NaCl and incubated with shaking for 1 h at room temperature. Membranes were washed twice with 100 mM Tris-HCl (pH 7.5) and 150 mM NaCl for 15 min, and a final rinse with 100 mM Tris-HCl (pH 9.5) for 5 mins. Membranes in 5 ml 100 mM Tris-HCl (pH 9.5) were

developed using BCIP/NBP Alkaline Phosphate Substrate Kit (Vector Laboratories) and photographed.

2.9.2 Chromatin Preparation

Chromatin was prepared by salt dialysis as described in section 2.2.2, or by one-step gradient dialysis described by Kaczmarczyk et al. (2017), where chromatin is dialysed from 2 M to 10 mM NaCl over approximately 19 hours. Chromatin was reconstituted using recombinant xenopus core histones in the presence of a 147 bp monomer competitor DNA fragment. Chromatin was then diluted to approximately 20 ng/ml in ESB (+) buffer (10 mM HEPES (pH 7.6), 100 mM NaCl, 2 mM magnesium acetate, 10 mM sodium azide, 0.1% (v/v) Tween-20 and 0.02% (w/v) BSA) before injection into the flow cell.

2.9.3 Flow Cell Preparation

Flow cells were prepared according to Kaczmarczyk et al. (2017). A coverslip was coated with 0.1% nitrocellulose in amylacetate, dried, then mounted onto a flow cell containing a flow channel of approximately 10 × 40 × 0.4 mm. Anti-digoxigenin (Roche) was diluted to 3 μg/ml in deionised water, injected into the flow cell and incubated at 4°C for at least 2 h. The flow cell was then incubated overnight at 4°C with 4% (w/v) BSA and 0.1% Tween-20.

Flow cells were washed with ESB (+) buffer before chromatin was injected at a concentration of approximately 20 ng/ml. Flow cells were incubated with chromatin at 4°C for at least 10 min, then ESB (+) buffer containing 2.8 μm-diameter streptavidin-coated M270 magnetic beads (Invitrogen) were injected into the flow cell. Beads were incubated for a further 10 min at 4°C before untethered beads were washed out with ESB (+) buffer.

2.9.4 Single-molecule Force Spectroscopy

The van Noort laboratory's home-built multiplexed magnetic tweezer set up was used to measure the extension of chromatin fibres under force (Brouwer et al., in preparation; Kaczmarczyk et al., 2017). A continuously increasing force is applied by moving the magnet towards the flow cell. Subsequently, the magnet trajectory is reversed to decrease force. The force exerted was calculated using a double exponential function calibrated prior to the experiment. The change in height of the magnetic beads (and therefore extension of chromatin molecules) was measured in real time at a frame rate of 30 Hz with a digital camera (CMOS Vision Condor).

2.9.5 Data Analysis

Data analysis and curve-fitting was performed using custom software written in Labview by the van Noort laboratory. The force-extension curve for each individual chromatin fibre was

fitted to the van Noort laboratory model of chromatin unfolding by first offsetting the curve so that the extension at high force (above 60 pN, after the final rupture event) to a worm-like chain of a known contour length (5542 bp for *601*, 5595 bp for *BLG* and *601/BLG*) with a persistence length of 50 nm and a stretch modulus of 1200 pN, in order to correct for off-centre attachment of the chromatin fibre to the magnetic bead. Rupture events at high force (6-60 pN) were identified with a *t*-test step-finding algorithm with a 5-point window. At forces below the first nucleosome-rupture event, force extension curves were manually fitted to the statistical mechanical model described by Meng et al. (2015) by specifying the nucleosome repeat length (197 bp), the number of nucleosomes (variable), the length of folded nucleosomes (1.5 nm), the stiffness of a folded fibre (variable), the number of unfolded nucleosomes, the length of DNA unwrapping from nucleosomes at low force (56 bp), the ΔG of three force transitions (variable) and the degeneracy (degeneracy of 1 was used, indicating a one-start model of chromatin structure consistent with a 197 bp repeat (Meng et al., 2015)).

Artur Kaczmarczyk performed analysis of high force rupture probability and provided MATLAB scripts to display force-extension curves of chromatin unfolding.

Chapter 3. Characterisation of Chromatin Arrays

3.1 Introduction

DNA sequence is known to play a role in nucleosome positioning due to differences in the bendability of different DNA sequences (Struhl and Segal, 2013), as described in section 1.3. Stiff poly(dA:dT) tracts have a low capacity to position nucleosomes (Field et al., 2008; Segal and Widom, 2009b), whereas sequences with a helical periodicity of flexible dinucleotides (AT, TA), occurring approximately every 10 base pairs, have a high affinity for the histone octamer (Lowary and Widom, 1998).

Nuclease digestion of chromatin is an important tool in the study of chromatin structure, particularly in the analysis of nucleosome positioning. Most nucleases are unable to access DNA wrapped around the histone octamer; instead they attack the accessible DNA between nucleosomes, leaving a footprint of DNA sequences bound to the histones which can be used to analyse nucleosome positions or to identify nucleosome-free regions. DNaseI has been used to identify regulatory regions which are depleted of histone proteins, identifying DNaseI hyper sensitive sites (Boyle et al., 2008; McGhee et al., 1981; Weintraub and Groudine, 1976). Micrococcal nuclease is used preferentially to map nucleosome positions by isolating and sequencing individual mono-nucleosome sequences both *in vitro* and *in vivo* (Fraser et al., 2009; Kaplan et al., 2009). These enzymes are suited to these different tasks as DNaseI can be thought of as a sharp cutter and is more penetrative than the "blunt-nosed" MNase.

In vivo, multiple factors play a role in nucleosome positioning, including the underlying DNA sequence, chromatin remodelling enzymes, and gene transcription. Sequences at some gene promoters have been found to maintain DNA accessibility by disfavouring nucleosome formation (Field et al., 2008; Sekinger et al., 2005). When *in vivo* nucleosome occupancy data derived from MNase digestion of the genome has been compared with *in silico* predictions of high nucleosome occupancy, it appears that approximately 50% of nucleosome positioning *in vivo* can be attributed to the intrinsic properties of the underlying DNA sequence (Segal et al., 2006).

When chromatin is reconstituted onto genomic DNA *in vitro* by salt gradient dialysis, the nucleosome positioning is well correlated to the native nucleosome position *in vivo*. This has been shown by comparing micrococcal nuclease digestion and sequencing of yeast mono-nucleosomes derived from cells and following *in vitro* reconstitution of the yeast genome (Kaplan et al., 2009) and of the ovine β-lactoglobulin gene region (Gencheva et al.,

2006). In addition, promoter sequences that enhance DNA accessibility *in vivo* have been found to poorly bind to nucleosomes *in vitro* (Kaplan et al., 2009; Segal and Widom, 2009b; Sekinger et al., 2005).

The "601" sequence is a chemically synthetic DNA sequence selected from a random pool of sequences by SELEX experiments (Lowary and Widom, 1998). It is known to position nucleosomes strongly following reconstitution *in vitro* and repeats comprised of this sequence form regular nucleosome arrays where nucleosomes are precisely spaced. Reconstituted fibres can be folded into higher-order structures; depending on the repeat length they have been observed to have a diameter of 30-nm *in vitro*, which have been analysed by electron microscopy (Robinson et al., 2006b), X-ray crystallography (Schalch et al., 2005) and cryo-electron microscopy (Song et al., 2014). While the conformation of the fibres may vary depending on the nucleosome repeat length of arrays, Robinson et al. found that fibres with a nucleosome repeat length between 177 bp and 237 bp form a one-start helical structure, while Schalch et al. and Song et al. find that fibres with a nucleosome repeat length between 167 bp and 187 bp form a two-start helix with a tetranucleosomal substructure. While arrays of the "601" sequence have been extremely important as a model system, their repetitive nature and strong affinity for the histone octamer make them a questionable model for the chromatin structure of the majority of the human genome. Thåström et el (1999) found that artificial nucleosome positioning sequences including the "601" have approximately a six-fold higher affinity for the histone octamer than the strongest biological nucleosome positioning sequences examined, and suggest that the eukaryotic genome has not evolved to position nucleosomes so precisely as the "601" is able to.

To investigate the role of DNA sequence on chromatin structure, the Gilbert laboratory has designed novel DNA sequence templates using strong biologically-derived nucleosome positioning sequences within a non-repetitive template. In this chapter I have optimised the reconstitution of these DNA templates into nucleosome arrays and analysed the nucleosome positioning of each of these fibres by nuclease digestion. Ensuring that chromatin fibres are correctly saturated and understanding this primary level of chromatin structure will be essential to study the higher-order structures and dynamics of these different sequences.

3.2 Design and Construction of DNA Templates for Reconstitution

The laboratory of Jim Allan (Fraser et al., 2009) mapped sites of high nucleosome affinity *in vitro* across the ovine β-lactoglobulin gene. Essentially, linearised plasmid DNA encoding the ovine β-lactoglobulin DNA sequence (approximately 11 kbp) was reconstituted by salt

gradient dialysis, using a low histone:DNA ratio of approximately one nucleosome per DNA molecule so there is no nucleosome-nucleosome competition. The resulting chromatin was digested to mono-nucleosomes by MNase, leaving DNA sequences that are protected by the presence of a histone octamer. Protected fragments were purified by gel electrophoresis, sequenced, and mapped to the DNA sequence to determine nucleosome coverage (Figure 3.1A). This indicated where nucleosomes are most likely to form and protect DNA regions from nuclease digestion. Jim Allan analysed this data to generate dyad maps (Fraser et al., 2009) assuming that the particle size was 149 base pairs (Figure 3.1B), although it is possible that particles of different sizes (159 or 169 base pairs) could exist, which would affect the identification of the dyads. Using these dyad positions, 25 nucleosome positioning sites with high nucleosome occupancy were selected to develop a template for chromatin reconstitutions (Figure 3.1, grey).

In addition to experimental methods, algorithms have been developed to predict nucleosomes postions (Field et al., 2008; Kaplan et al., 2009; Segal et al., 2006). To compare experimentally determined nucleosome coverage maps generated using *Xenopus laevis* core histone octamers (Fraser et al., 2009) with predicted nucleosome positions the algorithm developed by van der Heijden et al. (2012) was used (http://bio.physics.leidenuniv.nl/~noort/cgi-bin/nup3_st.py). This algorithm uses the periodicity of various different flexible dinucleotides; the position of TA, TT, AA and GC dinucleotides is calculated, with other dinucleotides weighted in order to normalise the probability of nucleosome occupancy. The algorithm is parameterised by three parameters: The probability amplitude which captures the sequence specificity for nucleosome binding by specific dinucleotides, the periodicity of dinucleoide distributions within the nucleosome, and the window of DNA that defines the binding of histones during the reconstitution. The chemical potential is also specified, defining the average affinity of the histone octamers for DNA, which is dependent on their relative concentration and allows the possibility that bound histones will block occupied binding sites or statistically position nucleosomes alongside themselves to be captured by the model. Using an amplitude of 0.2, a periodicity of 10.1 bp, a window of 147 bp and a potential of -1.0 kT, the probability of nucleosome occupancy across the ovine β-lactoglobulin gene was measured (Figure 3.1C, yellow). While van der Heijden et al. state that a window of 74 bp best captures the strong nucleosome positioning properties of the “601” sequence, a 147 bp window was used here as this was found to best capture the nucleosome occupancy of the yeast genome reconstituted by salt gradient dialysis *in vitro* by Kaplan et al. (2009).

Variability in the strength of the nucleosome positioning sites between experimental (using xenopuscore histones) and theoretical data was observed (Figure 3.1C). The algorithm shows generally less variation between points of high positioning probability and low probability and seems to show more peaks suggesting that there are more sites that are statistically likely to position a nucleosome, based on their sequences, than are captured in

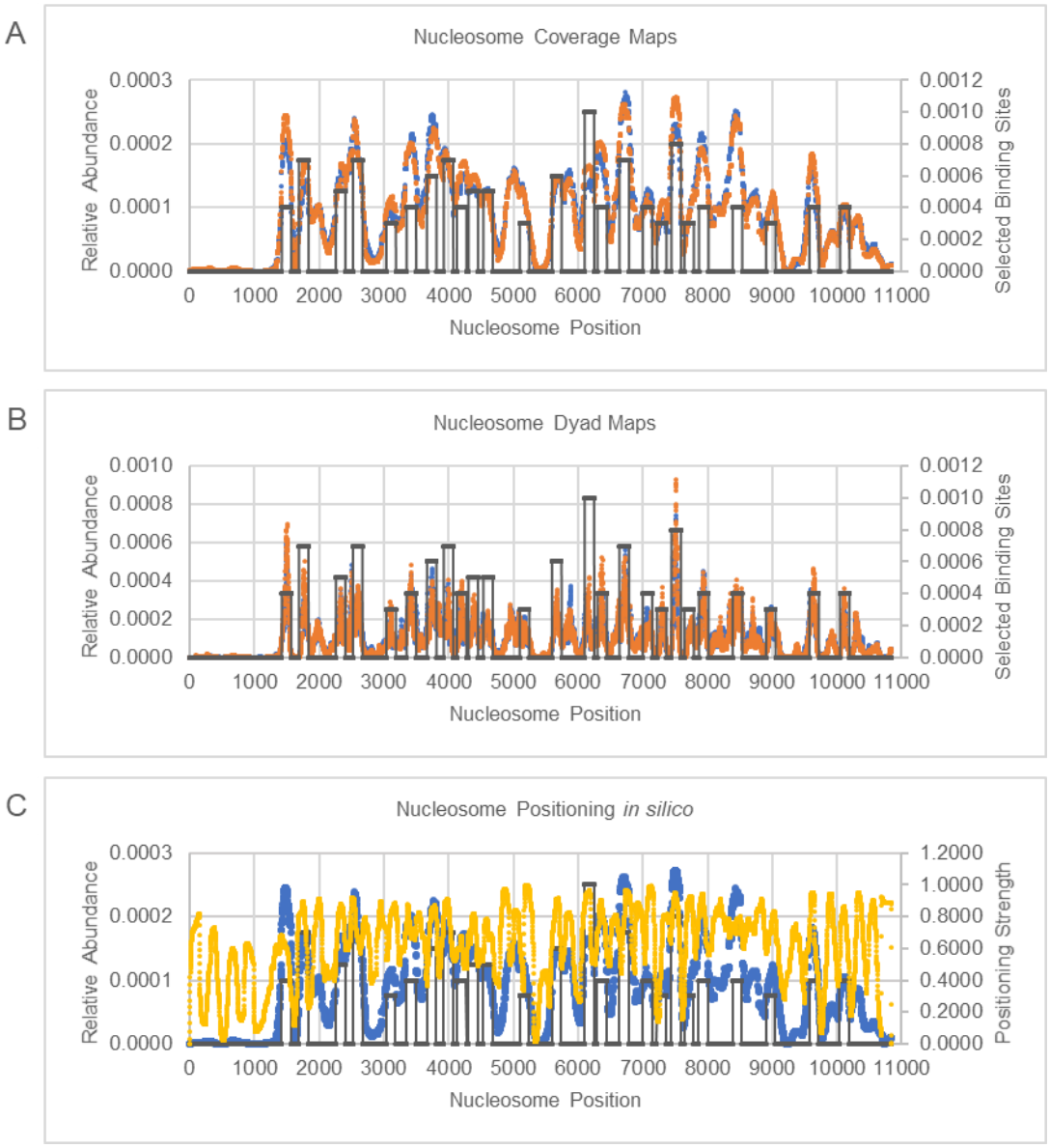

Figure 3.1. Nucleosome Positioning Maps of the Ovine β-lactoglobulin Gene.
A) Nucleosome coverage maps for the ovine β-lactoglobulin gene reconstituted using chicken erythrocyte (orange) or recombinant Xenopus (blue) core histones. Sequencing data from Fraser et al. (2009) replotted. B) Dyads calculated from coverage maps based on a 149 bp window by Jim Allan. C) Nucleosome coverage map (blue – using xenopus histones as in part A) compared with in silico nucleosome positioning analysis (yellow – right axis) (van der Heijden et al., 2012). Grey plateaus indicate 25 selected 197 bp binding sites with increased affinity for the histone octamer.

the experimental data. The algorithm also suggests that nucleosomes would expect to be positioned over regions of DNA (such as the region around 0-1300 bp) where no nucleosomes were found experimentally. However, there appears to be a degree of correlation within the datasets. For example, at approximately 5300 bp is a region where very few nucleosomes were experimentally identified, and the algorithm also suggests is very unlikely to position a nucleosome. The predicted histone occupancy *in silico* was compared with the positioning of the 25 selected nucleosome positioning sites that have the high affinity for the histone octamer *in vitro* (grey). A peak in the probability of histone occupancy in the *in silico* data was identified within every one of the 197 bp sites derived from the *in vitro* data, with the average distance between the peak maximum and the dyad of the selected positioning site being 29.84 bp. In 2 of these 25 positioning sites, the peak of predicted histone occupancy *in silico* aligned precisely with the dyad calculated from the experimental data. This suggests that there are DNA sequences whose histone occupancy seems to be well captured by algorithms to predict nucleosome positioning, but that there is much still unknown about the affinity of the histone octamer for different DNA sequences that means that histone occupancy *in vitro* still cannot be perfectly be captured.

3.3 Novel Templates for Chromatin Reconstitution

To investigate the relationship between underlying DNA sequence and the structure and dynamics of the higher-order chromatin fibre two new DNA templates with different properties were developed.

A non-repetitive and a partially repetitive sequence for the reconstitution of chromatin *in vitro* were created from high affinity nucleosome positioning sites identified in the ovine β-lactoglobulin gene. The 25 highest affinity octamer positioning sequences were selected (section 3.2) with the central base pair assumed to be the dyad (Figure 3.1B). 197 bp sequences centred on each dyad position were synthesised together (in the order that they appear within the β-lactoglobulin gene) to create a 4925 base pair DNA sequence, termed *BLG*. A second DNA sequence was synthesised where every other BLG nucleosome positioning sequence was replaced with a canonical 197 bp "601" repeat. This DNA template was denoted *601/BLG* (Figure 3.2A).

These novel DNA templates were compared to a "601" sequence containing 25 repeats of a 197 bp sequence comprised of the 147 bp "601" core and a 50 bp linker sequence. In this thesis, where *601* is referred to in italics, I refer to this specific DNA template containing 25 repeats of the 197 bp "601" monomer.

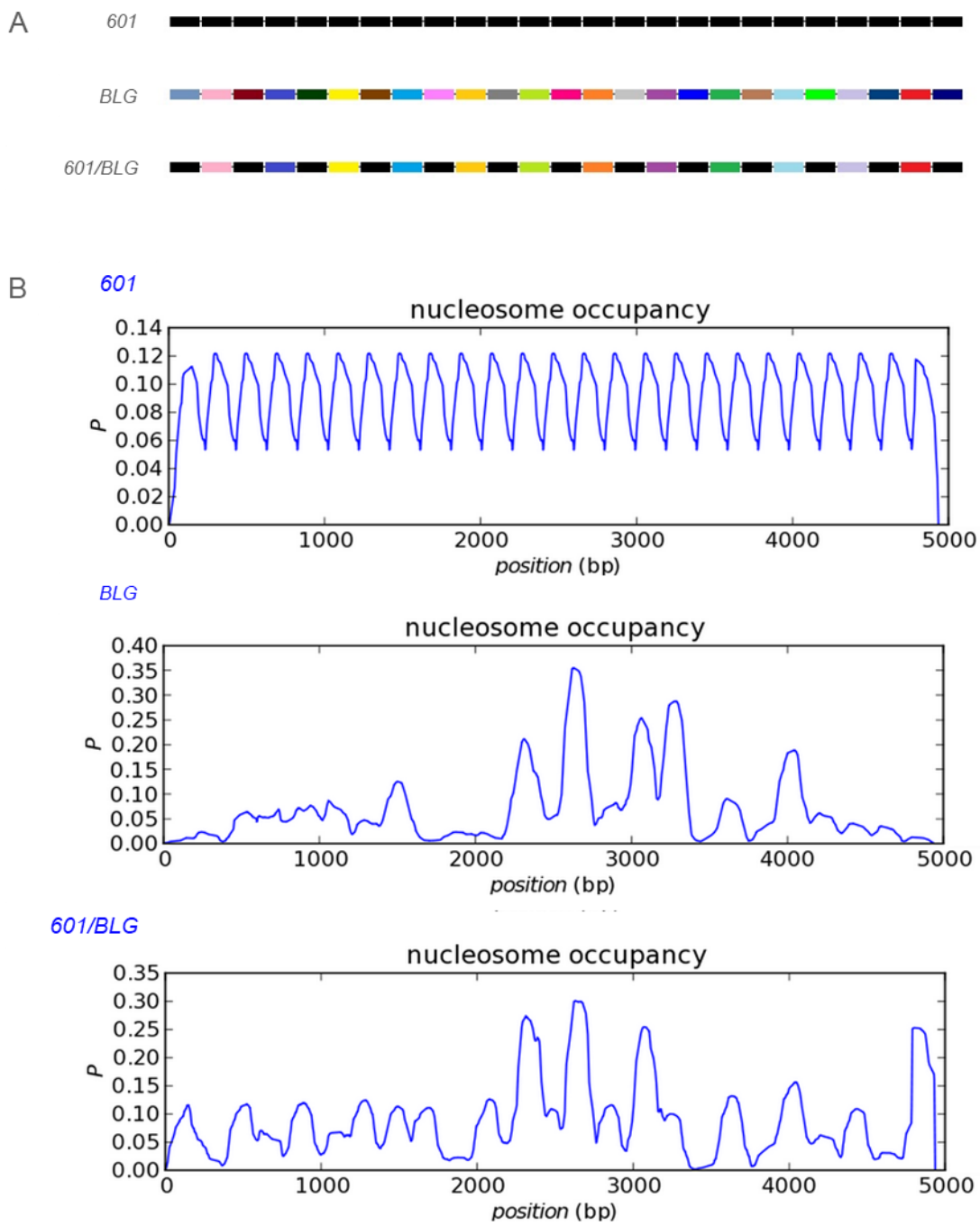

Figure 3.2. DNA Templates for Chromatin Reconstitution.
A) Schematic of three DNA templates: the repetitive 601, BLG containing 25 unique nucleosome positioning sites from the β-lactoglobulin gene, 601/BLG containing alternating "601" and unique nucleosome positioning sites. Each nucleosome positioning site consists of 197 bp DNA. B) In silico nucleosome positioning analysis (van der Heijden et al., 2012) of three reconstitution templates. P = predicted occupancy.

Each of the sequences differs slightly in their G/C content. *BLG* has the highest G/C content at 60.6%, *601* has 57.3% G/C, and *601/BLG*, being a mixture of the other sequences has a G/C content between these two, of 57.7%. G/C rich regions of the genome, such as those

found within gene bodies are associated with higher levels of nucleosome occupancy *in vitro* (Valouev et al., 2011).

The three different DNA templates each contain 25 × 197 base pair nucleosome positioning sites but as the DNA sequences are different, the affinity of the histone octamer for each site is likely to vary. The algorithm described by van der Heijden et al. (2012) was used to determine the nucleosome positioning properties of the *601*, *BLG*, and *601/BLG* (Figure 3.2B).

The algorithm predicts strong, regular positioning of nucleosomes by the *601* sequence. In contrast, the *BLG* sequence contains some sites which appear to have a greater affinity for the histone octamer than the "601", but also contains regions which do not appear to position nucleosomes well. In the *601/BLG* sequence, the presence of the "601" repeats may force nucleosomes to be positioned less variably upon the "BLG" sites in between them.

As each of the non-"601" nucleosome positioning sites have been experimentally mapped following reconstitution by salt gradient dialysis, it is expected that these templates should form nucleosome arrays containing 25 histone octamers when reconstituted. However, the positioning of these nucleosomes may vary from the very regular *601*, where nucleosomes are known to position strongly over the 147 bp core of each repeat. These differences in the position and strength of nucleosome positioning would then be expected to have an impact on the structure of the folded chromatin fibre.

3.4 Reconstitution of Chromatin by Salt Dialysis

To form nucleosome arrays from DNA templates and core histone octamers *in vitro*, chromatin was reconstituted by salt dialysis. DNA and core histone octamers were combined in high salt (which inhibits interactions between the negatively charged DNA and positively charged histones), and slowly dialysed to low salt, allowing histone octamers to bind to preferential sequences within the DNA and subsequently fold into nucleosome arrays. To form folded higher-order chromatin fibres from nucleosome arrays, linker histones and/or magnesium ions were added after salt dialysis (see section 4).

Throughout I used a linear gradient maker to dilute 2 M sodium chloride to 400 mM over a 6-hour period (section 2.2.2). Upon reaching 400 mM sodium chloride histone octamers will be bound to and positioned on the preferred DNA sequences. Nucleosome arrays in 400 mM sodium chloride were then dialysed into low salt (2.5, 10 or 80mM NaCl) overnight; at 80 mM salt they represent chromatin found in the nucleus, but in lower salt they are likely to be an unfolded state and suitable for analysis of nucleosome arrays.

Competitor DNA sequences can be added to reconstitutions to minimise oversaturation of template DNA sequences by histone octamers, as excess histone octamers will bind to the competitor. "601" sequences have such a high affinity for the histone octamer that they become saturated with nucleosomes before any protein binds to the competitor, excess histones will bind to the competitor rather than oversaturate the "601". Previously, a 147 bp competitor DNA has been used to prevent the oversaturation of "601" nucleosome arrays (Huynh et al., 2005), whilst reconstitution of *601* in the absence of a competitor requires careful titration of histone octamers to ensure that nucleosome arrays do not become oversaturated.

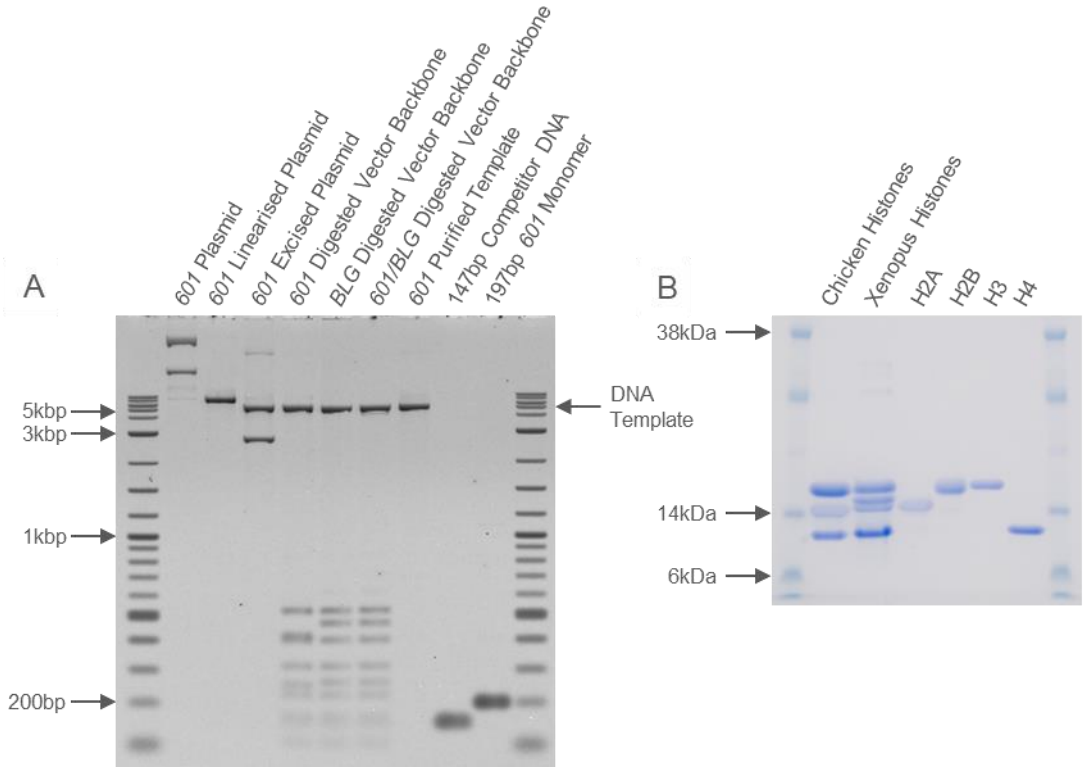

Figure 3.3. Components of Chromatin Reconstitution.
A) DNA templates for chromatin reconstitution. B) Chicken erythrocyte and recombinant Xenopus histone octamers fractionated alongside individual purified recombinant human histones (New England Biolabs). Samples were analysed on a 12% polyacrylamide gel.

Two different types of competitor were used during my studies (Figure 3.3A). Firstly, the 147 bp mono-nucleosomal competitor sequence derived from the pUC18 vector backbone, was used in reconstitutions as described by Huynh et al (2005). This was used for small scale experiments including single-molecule force spectroscopy (section 4.7), however, it was difficult to purify sufficient competitor and template DNA for large scale experiments. As an alternative, for some experiments including sucrose gradient sedimentation (section

4.4), the *601*, *BLG*, and *601/BLG* templates were excised from the plasmid vector backbone, the backbone was digested to fragments of approximately 500 bp or smaller, and these fragments were used as competitor DNAs. For experiments including SAXS where competitor DNA would affect the results, purified template was reconstituted in the absence of any competitor, as described by Rogge et al. (2013).

Core histone octamers were purified from chicken erythrocytes (see section 2.2.1.3). For defined small scale experiments, such as single-molecule force spectroscopy, recombinant *Xenopus* core histones were provided by Tom Owen-Hughes and Ramasubramanian Sundaramoorthy (Figure 3.3B).

3.5 Heterogenous DNA templates require additional histones for saturation

To determine the histone:DNA ratio to correctly saturate DNA templates with 25 nucleosomes in the presence and absence of a competitor, DNA was titrated with histone octamers and the nucleosome arrays were analysed by an electrophoretic mobility shift assay (EMSA). Reconstituted chromatin migrates more slowly through an agarose gel than unreconstituted DNA as a result of the decreased overall charge and increased mass of the reconstituted chromatin. Furthermore, the topology of the molecule is likely to impact its migration speed, in a similar manner to a supercoiled plasmid migrating more quickly than linear DNA.

Previously, Huynh et al. (2005) suggested that "601" fibres reach a maximum mobility shift when they become saturated with histone octamers (in this case, 25 octamers bound to each DNA fibre).

Chromatin fibres in the presence and absence of competitor DNAs were reconstituted at different histone:DNA ratios at 40 ng/µl (template DNA concentration) and analysed on 0.7% agarose gels in 0.25 × TBE, in the absence of ethidium bromide. Figure 3.4 shows the three chromatin templates, titrated with varying amounts of core histones in the absence of a competitor (A), or the presence of an undigested plasmid backbone as competitor (B) (the size of the competitor fragment when undigested is approximately 2.5 kbp). The chromatin fibres migrate at different rates relative to the DNA markers (Figure 3.4C). However, as chromatin runs differently to DNA depending on gel conditions (specifically agarose concentration, see Figure 4.2), the maximum shift cannot be compared between each of the different templates which are ran on different gels (see below), but for each individual fibre,

this enables the point where each template reaches a plateau indicating fibre saturation to be determined.

In the absence of competitor, *601* appears to be fully saturated at a histone:DNA ratio of 1.4:1 (Figure 3.4A – indicated by *). In contrast, the non-*601* templates require a slightly higher amount of histones to achieve full saturation in the absence of a competitor, corresponding to a 1.6-1.8:1 histone:DNA ratio for the *BLG* and *601/BLG* respectively. This is consistent with the lower affinity of the non-"601" sequences for the histone octamer.

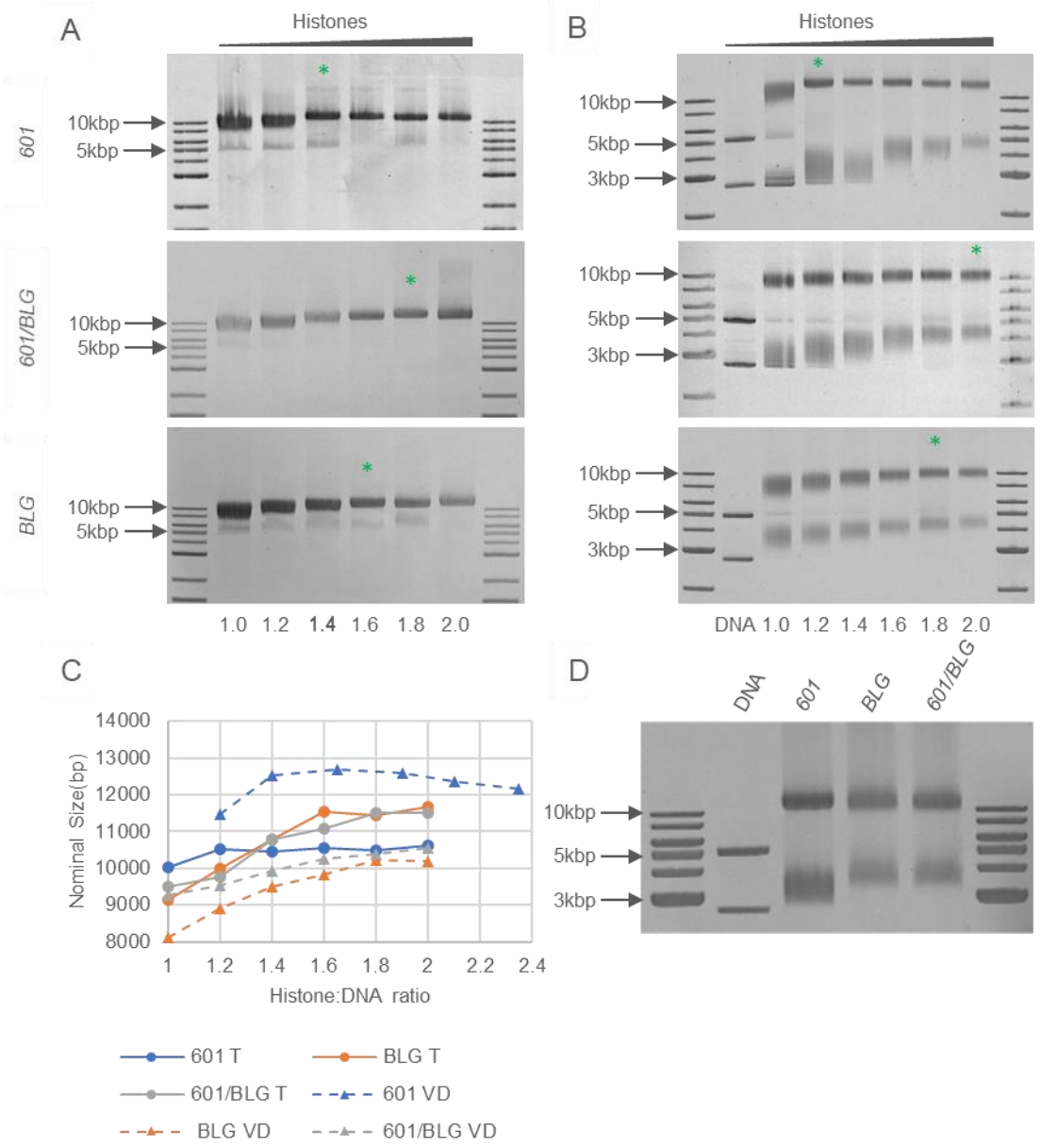

Figure 3.4. Electrophoretic Mobility Shift Assay of DNA Templates Reconstituted with Varying Amounts of Core Histone Octamer.
(Full legend found on following page)

A) Titration of purified template DNA with chicken erythrocyte core histones. Green asterisk indicates the histone:DNA ratio at which the shift in gel migration plateaus, calculated from the nominal fragment size. B) Titration of templates with *Xenopus* core histones, using the undigested plasmid backbone as a competitor. Note that following remeasurement of the *601* DNA sample concentration, histone:DNA ratios used were found to be higher than those in the corresponding *601/BLG* and *BLG* titrations. See panel C. C) Shift measured by calculating the fragment size relative to the DNA ladders. Chromatin runs differently on different percentage gels relative to DNA, and therefore the maximum shifts between gels cannot be compared, but these can be used to visualise the plateau at which chromatin fibres are thought to become fully saturated in each case. True fragment size is not calculated, but used as a marker for mobility shift. For the template reconstituted in the absence of competitor, these results are the product of two technical replicates (samples ran on two separate gels), in the presence of the vector competitor, but data from only one gel is shown. T (solid line) = template reconstituted in the absence of competitor (A), VD (dotted line) includes a competitor of the digested vector backbone. D) Fully saturated samples (*601* at 1.7:1, *BLG* and *601/BLG* at 2:1) including a vector backbone competitor are compared on the same gel and the relative positions of migration were measured as 11388bp, 11377bp and 11058bp respectively.

In the presence of competitor, the *601* template appears to achieve full saturation at a very similar histone:DNA ratio (1.4:1) as without competitor, and there is very little concomitant mobility shift in the vector backbone (Figure 3.4B). The large difference in affinity between the template and the competitor appears to allow the *601* to be reconstituted as if the competitor was not present up to a histone:DNA ratio of 1.4:1. At histone:DNA ratios higher than this, once the template becomes fully saturated with 25 histone octamers, additional histones bind to the competitor as described by Huynh et al (2005), preventing oversaturation of the template. At ratios higher than 1.4:1, there is a mobility shift in the competitor DNA, showing more histones binding at increasing histone:DNA ratios (Figure 3.4B).

For both the *BLG* and *601/BLG* samples, more histones are required to saturate the templates in the presence of a competitor, as histones bind to the competitor at sub-saturating ratios of histone:DNA. This confirms that these templates do not have such a high affinity for the octamer as the *601*. To achieve saturation, approximately a 1.8-2:1 histone:DNA ratio is required for the *BLG* and *601/BLG* respectively. Intriguingly, in both the absence and presence of competitor, the *601/BLG* appears to require slightly more histones than the *BLG*, although this would be expected to have an overall higher affinity for the histone octamer due to the presence of the "601" repeats. It is possible that for the *601/BLG* template the histones bind to the "601" repeats earlier during the reconstitution, and that this makes it more difficult for nucleosomes to subsequently assemble on the "BLG" sites between them.

The fact that a plateau is reached in the shift of the fibres, even in the absence of a competitor DNA (where excess histones might be expected to bind to the template), suggests that excess histones are either not bound to the template, or that a gel mobility-based assay cannot discriminate well between a saturated and an oversaturated chromatin fibre. However, where *601* has been over-titrated in the presence of a competitor, the template appears to run faster as the histone:DNA ratio is increased above 1.6:1. This might indicate a structural change at this level of saturation that causes the chromatin to run faster than a fibre with 25 nucleosomes, however if this were the case, it would also be expected to occur in the absence of a competitor, when the template would become oversaturated more easily, indicating there might be limitations to the assay.

Although recombinant *Xenopus* core histones were used to titrate the DNA in the presence of a competitor (Figure 3.4B), many other experiments used chicken erythrocyte core histones for reconstitutions. A comparison between the reconstitution efficiency of chicken and xenopus core histone octamers of a *601* template can be found in Appendix 3.

As chromatin migrates slightly differently depending on agarose gel concentration it is most informative to run all samples on the same gel. To confirm the optimal reconstitution conditions the three DNA templates (*601* at 1.7:1, *BLG* and *601/BLG* both at 2:1), were reconstituted under optimal conditions in the presence of a vector backbone competitor and analysed together to reveal a similar degree of mobility shift (Figure 3.4D).

3.6 Electron Microscopy

To assess the extent of the nucleosome array saturation and to examine fibre structure, chromatin was reconstituted using chicken erythrocyte core histones in the absence of competitor DNA and viewed by electron microscopy. As described in section 2.4, chromatin samples were fixed using formaldehyde, attached to carbon grids, platinum shadowed, and imaged at 20-30,000 × magnification using a JEOL JEM-1400Plus Transmission Electron Microscope. A *601* sample reconstituted at a 1.6:1 histone:DNA ratio (Figure 3.5A) at low resolution showed variation in the structures across the population of molecules. At high resolution, individual fibres from the different templates are easily discernible (Figure 3.5B). At a histone:DNA ratio of 1.2:1 the samples appear slightly undersaturated, but at 1.6:1 fibres look well saturated. At a ratio of 2:1, some fibres appeared very folded so that individual nucleosomes were not easily discernible, indicating oversaturation. In some views condensed particles were visible (Figure 3.5C); it is not clear what these correspond to and so were excluded from the analysis.

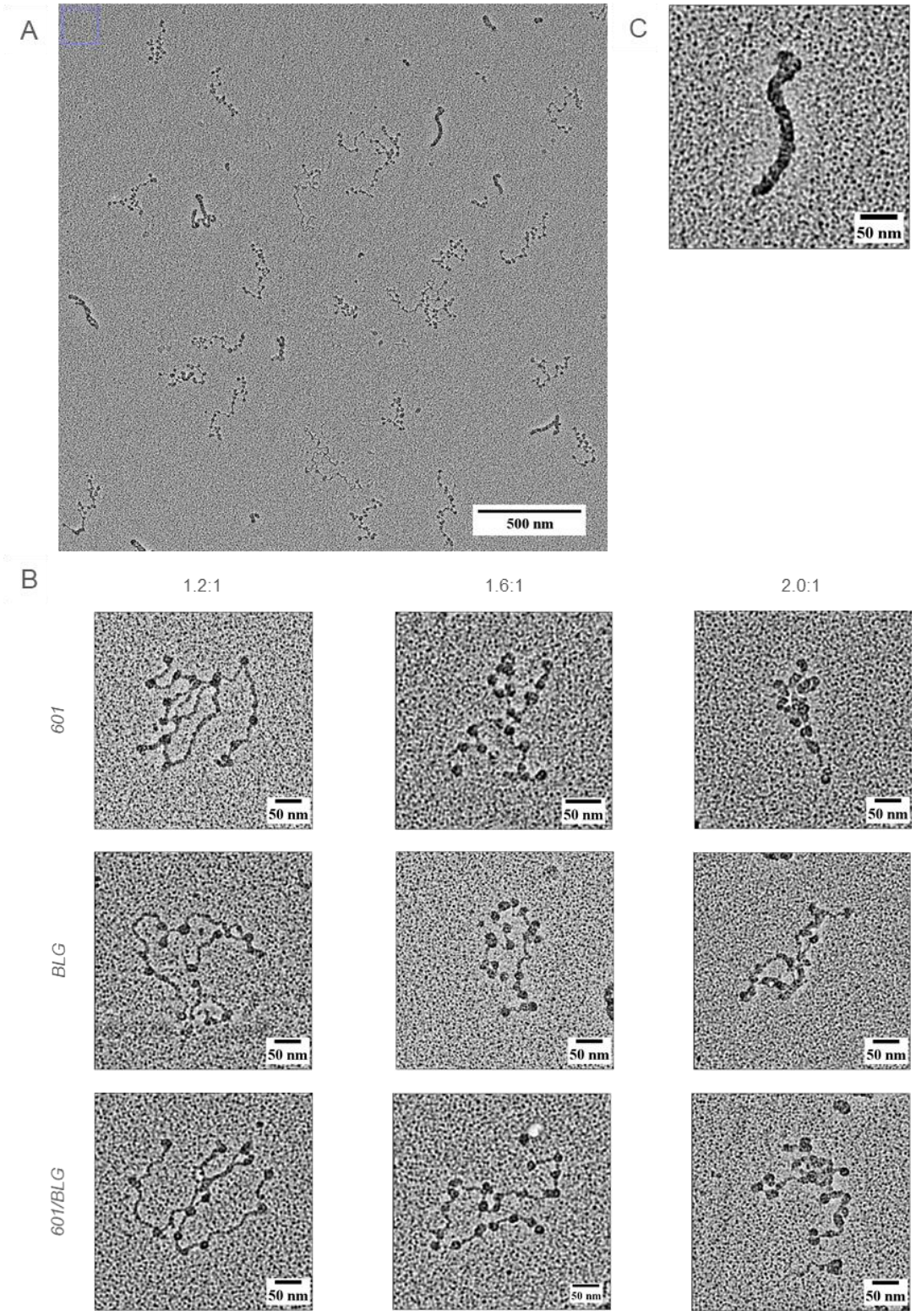

Figure 3.5. Electron Microscopy of Nucleosome Arrays.
A) Image field from an electron micrograph of 601 chromatin fibres reconstituted at 1.6:1 (histone:DNA). Sample is magnified 20,000× and scale bar is 500 nm. B) Typical fibres observed for each DNA sequence template reconstituted at different histone:DNA ratios. C) Example of a condensed particle which appears at similar amounts in each sample and was excluded from further analysis.

To count the number of nucleosomes and to measure the area occupied by each nucleosome array within an EM image, ImageJ scripts were written by Davide Michieletto (see section 2.4.4). Images of individual chromatin fibres were blurred to obscure background and nucleosomes were identified using a Phansalkar thresholding approach. A watershed segmentation was added to separate overlapping nucleosomes which would otherwise be counted as single objects on a more saturated chromatin fibre (Figure 3.6A).

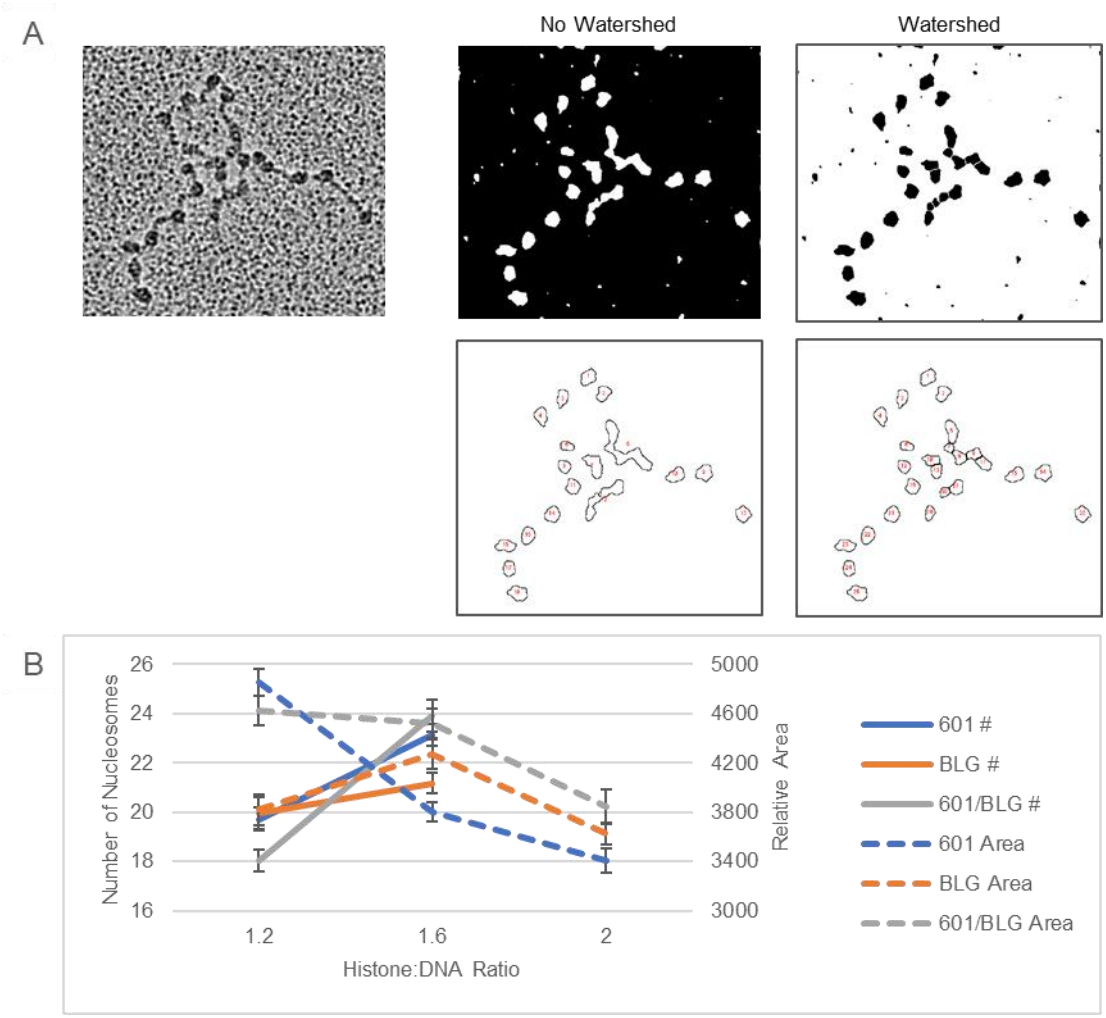

Figure 3.6. Counting Nucleosomes in Electron Microscopy Images.
A) Automatic counting of nucleosomes with and without watershed separation of nucleosomes. A 601 fibre reconstituted at 1.6:1 has 18 nucleosomes counted without watershed separation, and 25 nucleosomes with watershed separation. B) Number of nucleosomes and area of chromatin fibres measured by automated nucleosome counting algorithms, +/- SEM.

For each sample, nucleosomes on 20 fibres were manually counted, and the radii used to blur and threshold the images were optimised for each individual sample. These scripts were then applied to a wider selection of chromatin fibres to count the number of nucleosomes and

measure the area covered by each fibre (Fibre 3.6B). It was not possible to manually count individual nucleosomes on fibres reconstituted at a 2:1 histone:DNA ratio, but the area was measured by the algorithms optimised for fibres reconstituted at 1.6:1.

At a 1.2:1 histone:DNA ratio, an average of 19.7 nucleosomes were counted on *601* fibres (N=102), 20.0 nucleosomes were counted on *BLG* fibres (N=43), and 18.1 nucleosomes were counted on *601/BLG* fibres (N=78). At 1.6:1, an average of 23.1 nucleosomes were counted on each *601* chromatin fibre (N=82), 21.1 nucleosomes were counted on each *BLG* chromatin fibre (N=85) and 23.9 nucleosomes were counted on each *601/BLG* chromatin fibre (N=70). At a histone:DNA ratio of 1.6:1, a *601* fibre might be expected to be slightly oversaturated, a *BLG* fibre correctly saturated, and a *601/BLG* fibre slightly undersaturated, based on the results of the EMSA (Figure 3.4A). While the chromatin fibres reconstituted under each of these conditions all appear to be undersaturated it is surprising that *601* and *601/BLG*, which are expected to be respectively oversaturated and undersaturated from the band shifts (Figure 3.4A), they have a similar number of nucleosomes while the *BLG*, expected to be correctly saturated, has fewer. It is possible that nucleosomes are less evenly spaced within the *BLG* templates, causing more cases of nucleosome overlapping which can not always be separated by manual counting or the watershed segmentation. Similarly, when *601* became oversaturated with histone octamers, increased nucleosome overlapping may restrict nucleosomes being counted. When reconstituted at a 2:1 ratio, this algorithm was typically only able to identify 15-17 nucleosomes on each chromatin fibre due to this overlapping. This might suggest that nucleosomes cannot be counted to accurately determine fibre saturation following EM, but this might still be a useful method to compare fibre saturation between different samples.

The area occupied by the chromatin fibre also decreases as fibres become saturated, possibly as a consequence of increased folding and nucleosome overlapping. The area of *601* fibres, expected to become fully saturated at a histone:DNA ratio of 1.4:1, drops steeply between histone:DNA ratios of 1.2:1 and 1.6:1, but only moderately between 1.6:1 and 2:1. Conversely, the area of *601/BLG* fibres, which are expected to become saturated at 1.8:1, drops moderately between 1.2:1 and 1.6:1, then steeply between 1.6:1 and 2:1. This suggests that the area of reconstituted fibres drops considerably at the point when fibres become oversaturated, possibly as a result of increased fibre folding in these conditions. *BLG* fibre area was maximal at a ratio of 1.6:1, at which point they should be correctly saturated, but dropped steeply as the histone:DNA ratios was increased to 2:1.

Unfortunately, electron microscopy could not be used to study saturation levels of chromatin reconstituted in the presence of a competitor without first removing the competitor DNA from the sample (see section 4.4).

3.7 Density Gradient to Measure Reconstitution Efficiency

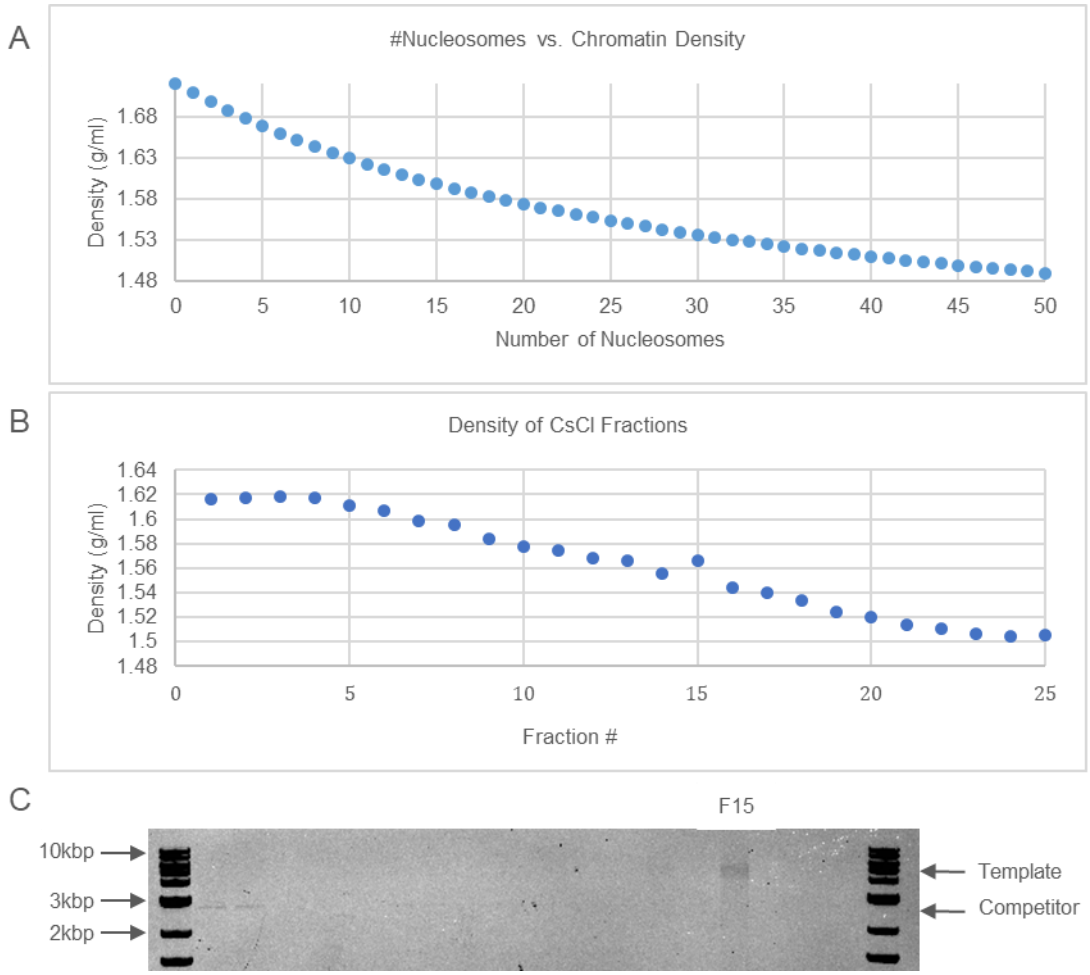

Figure 3.7. Density Gradient of 601 chromatin.
A) Calculated density of a 25 x 197 bp DNA template onto which a given number of nucleosomes have been reconstituted. 25 nucleosomes will have a density of 1.55 g/ml. B) Density of fractions of a CsCl gradient, calculated from the refractive index and corresponding CsCl concentration. C) CsCl fractions (fractionated from bottom to top) analysed on 1% agarose gel. Template DNA appears in fraction 15 of the gradient.

To quantitatively analyse the number of nucleosomes bound to different DNA templates, the density of the chromatin fibres was measured using caesium chloride isopycnic gradients. Chromatin was reconstituted in the presence of a competitor (the vector backbone), and therefore required the separation of competitor and template chromatin fragments by sucrose gradient sedimentation. *601* reconstituted at 1.7:1 and *BLG* and *601/BLG* reconstituted at 2:1 histone:DNA ratios (Figure 3.4D) were centrifuged on a 6-40% isokinetic sucrose

gradient for 5 hours to isolate reconstituted chromatin (Figure 4.3), which was then extensively fixed with formaldehyde. Cross-linked chromatin was centrifuged on a CsCl gradient with a density of 1.55 g/ml. As DNA has a density of 1.72 g/ml and protein has a density of 1.35 g/ml, this is the calculated density of a chromatin fibre with one histone octamer bound per 197 bp DNA (Figure 3.7A).

601 chromatin sedimented in fraction 15 of the gradient (Figure 3.7C) consistent with the density of a saturated nucleosome arrays of 1.55 g/ml (Figure 3.7B). In contrast, the sucrose gradient sedimentation profiles showed *BLG* and *601/BLG* appeared to be more heterogeneous in either reconstitution efficiency (with different numbers of nucleosomes reconstituted on each fibre) or in fibre structure (see Figure 4.3B). As a result, these samples were more diluted when removed from the sucrose gradient, and therefore could not be seen on a gel following caesium chloride sedimentation and fractionation, removal of crosslinks and isolation of DNA.

When crosslinking and subsequent caesium chloride sedimentation were attempted without first separating the template by sucrose gradient sedimentation, template DNA appeared throughout the entire gradient, suggesting that the fibre could be quite heterogeneous. It is also possible that the chromatin fibre density could also be affected by template chromatin being crosslinked to undersaturated competitor fragments.

3.8 Restriction Digestion to Assess Chromatin Reconstitution Efficiency and Nucleosome Positioning

To assess chromatin reconstitution efficiency and the positioning of underlying nucleosomes, chromatin was digested with restriction enzymes. Restriction enzyme recongnition sequences within the core of each nucleosome positioning site would be expected to be occluded if a nucleosome was positioned there, while motifs within the linker regions would be expected to be accessible to endonucleases, enabling digestion of chromatin arrays into mono-nucleosomes.

Chromatin was digested overnight at 4°C by *Ava*I, *Psi*I or *Pfl*23II, in the presence of 1.5 mM magnesium chloride (Figure 3.8-3.10). DNA from *Pfl*23II and *Psi*I digests was isolated and analysed by electrophoresis, whereas *Ava*I digests were analysed as native chromatin on a 1.1% gel to see the fraction of reconstituted monomers, as described by Huynh et al. (2005).

3.8.1 *601* Reconstitution Efficiency is Determined by Inaccessibility to Nucleosomal DNA by Restriction Enzymes

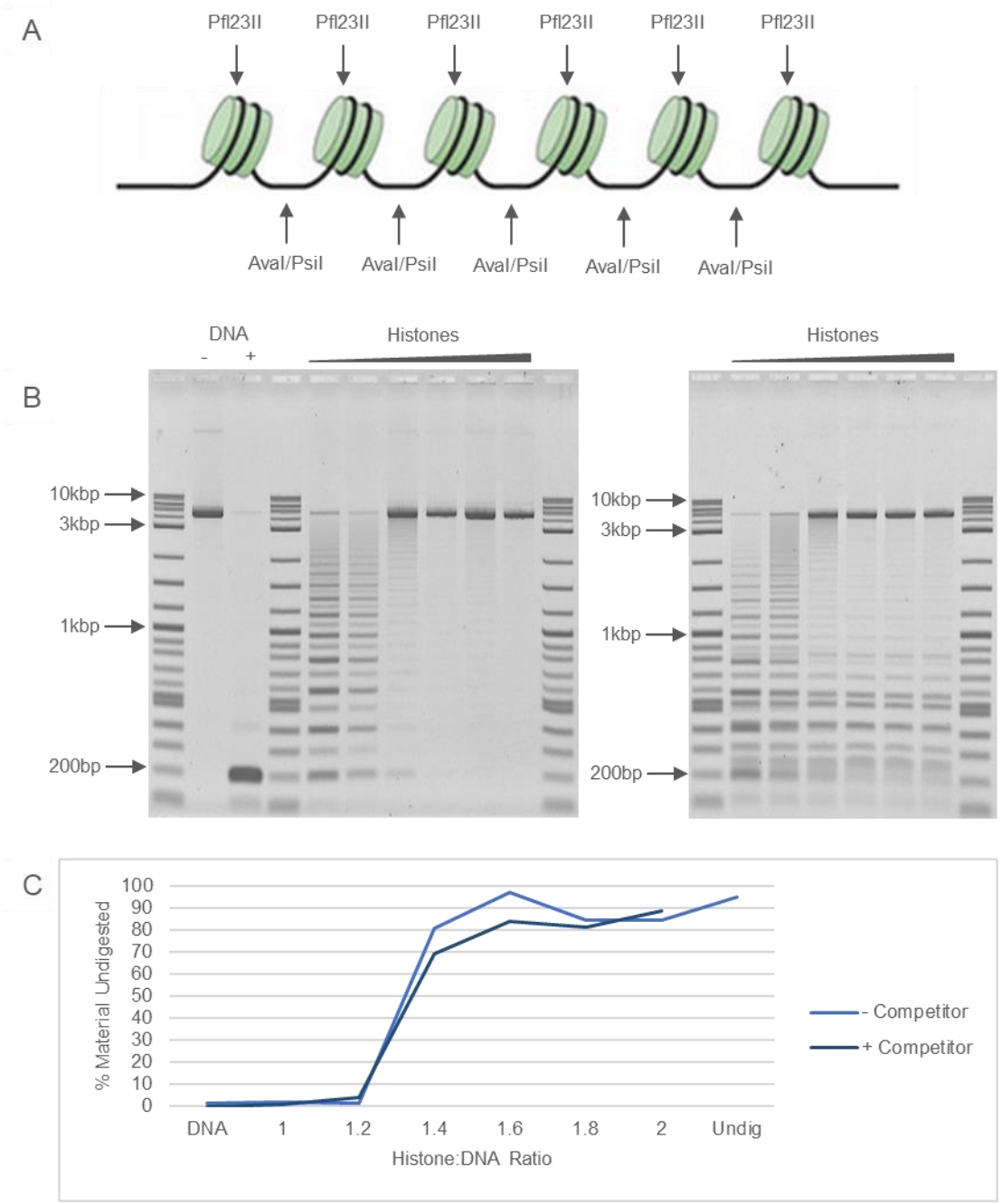

Figure 3.8. Restriction Enzyme Digestion of Chromatin by Pfl23II.
A) Map of restriction enzyme sites within reconstitution templates. 601 fibres contain Pfl23II sites within the nucleosome core and AvaI sites within linker regions. BLG and 601/BLG contain PsiI sites within linker regions. B) Digestion of 601 nucleosome arrays by Pfl23II, reconstituted at different histone:DNA ratios without (left) and with (right) a competitor DNA. C) Quantification of material digested by Pfl23II. Nucleosome arrays reconstituted at ratios of 1.4:1 and higher have minimal amounts of digestion, suggesting complete saturation.

Within the 147 base pair core of the "601" repeat there are several restriction sites, which can be utilised to assess reconstitution efficiency. When nucleosomes position over the restriction site, endonucleases are unable to recognise and cleave the sequence motif. As a result, properly reconstituted *601* chromatin will not be digested by enzymes such as *Pfl*23II (Figure 3.8A).

To assess chromatin reconstitution efficiency, *601* template was reconstituted at a concentration of 20 ng/µl with CE core histones in the presence and absence of competitor DNA (using digested vector backbone competitor fragments) at histone:DNA ratios of 1.0-2.0:1, and then digested with enzyme (see section 2.5.1). Unreconstituted *601* DNA was almost completely digested by the enzyme under these conditions (Figure 3.8B) whereas partial digestion was seen for chromatin reconstituted at lower histone:DNA ratios. DNA template that remained undigested was expected to be saturated with histones. The percentage of material that remained undigested plateaued at around 85% at ratios higher than 1.4:1 in the absence of competitor (Figure 3.8C). While 15% of the material was digested, it does not follow that 15% of the nucleosome positioning sites are unreconstituted, but that in 15% of the fibres, at least one of the 25 positioning sites is not occupied by a histone octamer. It is possible that there may be some enzyme activity that can move nucleosomes to access the restriction sites that causes 15% to be partially digested, or that these 15% of sites are unoccupied.

When reconstituted in the presence of a competitor, the competitor DNA (digested plasmid backbone) accounts for 33% of the total DNA, which is discounted from the analysis. Around 80% of the template DNA remained undigested in these samples, though at a histone:DNA ratio of 1.4:1 the proportion of undigested template was slightly lower than this, suggesting that slightly more histones may be required to achieve the maximum level of saturation.

3.8.2 *601* Reconstitution Efficiency is Analysed by Digestion to Mono-nucleosomal Material

*Ava*I sites are located between each "601" repeat of the *601* sequence and are within the linker region of these chromatin fibres. In the *BLG* and *601/BLG* sequences, this is replaced by the *Psi*I restriction site. In a regularly reconstituted chromatin fibre, where nucleosomes are positioned over their cognate binding sites and not over the linker regions, these positions should be accessible to endonucleases (Figure 3.8A).

To assess chromatin reconstitution efficiency, *601* chromatin fibres were reconstituted in the presence and absence of competitor as described above and digested with *Ava*I. As the

"601" sequence has strong nucleosome positioning properties, the restriction sites would be expected to be accessible in each nucleosome array, whilst if the fibre was oversaturated digestion would not be complete. Nucleosomal fragments can then be analysed by native agarose gel electrophoresis (Figure 3.9A).

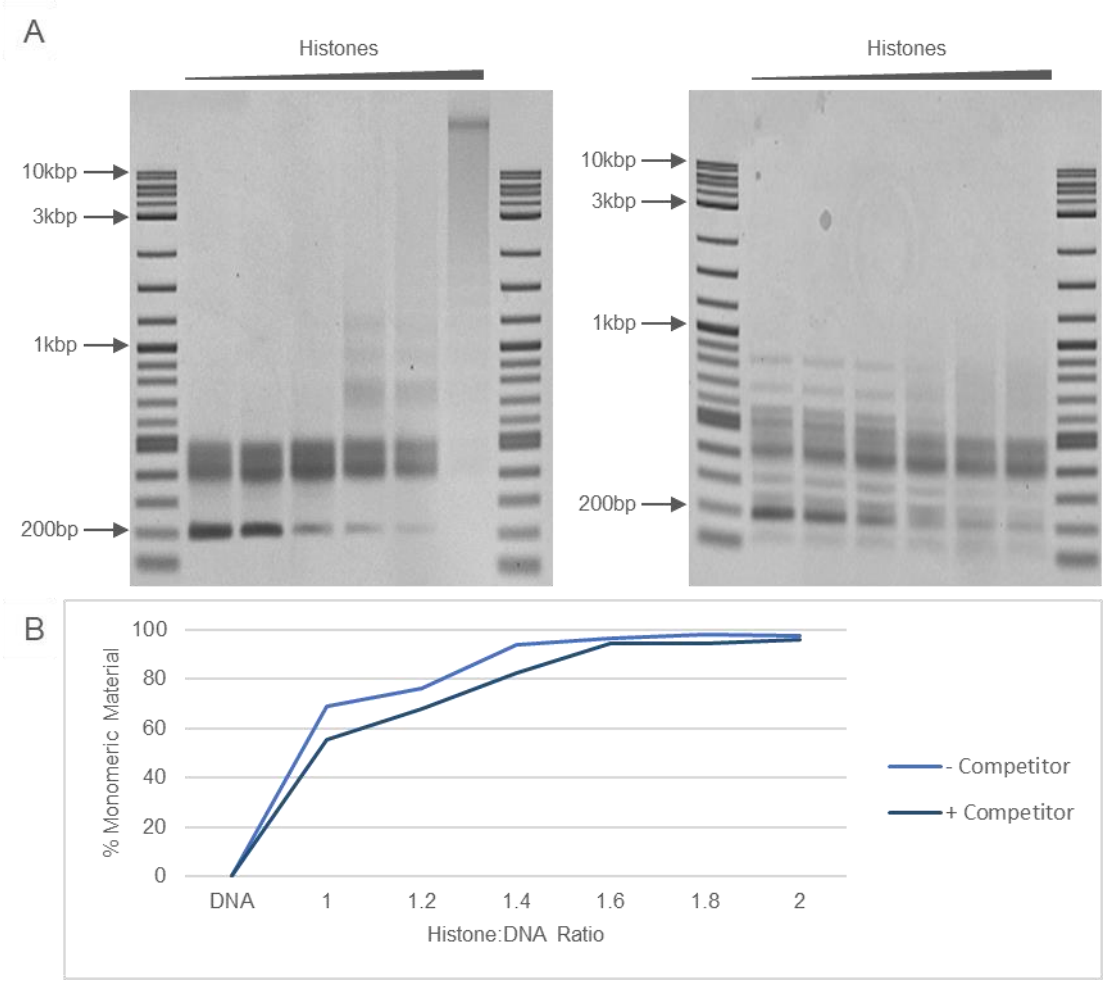

Figure 3.9. Restriction Digestion of Chromatin by *Ava*I.
A) Digestion of *601* nucleosome arrays by *Ava*I, reconstituted at different histone:DNA ratios without (left) and with (right) a competitor DNA. B) Quantification of mononucleosomal material compared to monomer DNA. Nucleosome arrays reconstituted at ratios of 1.4:1 and higher have almost 100% of nucleosomes reconstituted. In the absence of a competitor DNA, digestion is limited at ratios above 1.6:1, suggesting that chromatin becomes oversaturated.

The proportion of mono-nucleosomal material (which runs close to the 400 bp DNA marker) compared to monomer DNA (197 bp) was measured. Based on the results of Paoletti et al. (1977), the intensity of the nucleosomal material was adjusted to account for the fact that chromatin only intercalates 60% as much of the ethidium bromide dye as naked DNA. Without competitor, this was found to plateau at ratios higher than 1.4:1 where around 94% of material appears mono-nucleosomal and 6% remained as unreconstituted. When a

competitor was included, a 1.6:1 ratio was required to achieve this level of saturation, with only 82% of sites being reconstituted at 1.4:1.

In the absence of competitor, undigested species can be seen at ratios higher than 1.6:1, suggesting these samples were oversaturated. This does not appear to be an issue in the sample where a competitor DNA was used, demonstrating that the addition of this molecule seems to effectively prevent oversaturation. Based on one experiment, it is difficult to comment whether the protective effect the competitor has over the reconstitution is statistically significant at histone:DNA ratios of 1.6:1 or 1.8:1. Furthermore, undigested material will affect the measurement of the proportion of reconstituted nucleosomes, as an unreconstituted dimer migrates at a similar speed to a mono-nucleosome, but this would only be expected to cause a minor effect.

The results from each of these restriction digestions (both using *Pfl*23II and *Ava*I) suggest that the *601* template, when reconstituted without competitor, becomes saturated at histone:DNA ratios of 1.4:1, in agreement with the electrophoretic mobility shift assay (Figure 3.4). When a competitor DNA is used, it appears that slightly more histones may be required to saturate, due to a small fraction binding to the competitor DNA, but at histone:DNA ratios higher than 1.6:1 the reconstitution efficiency of the template appears to remain constant to ratios of at least 2:1, confirming that at 1.7:1 the nucleosome array remains correctly saturated, in agreement with the caesium chloride density gradient (Figure 3.7).

3.8.3 *BLG* and *601/BLG* Arrays have Heterogeneous Nucleosome Positioning, Limiting the Analysis of Reconstitution Efficiency by Restriction Digestion

In *BLG* and *601/BLG* DNA sequence templates, a *Psi*I restriction site is located in between nucleosome positioning sequences, in an analogous position to *Ava*I (Figure 3.8A). Templates were digested with this enzyme and DNA was extracted from the chromatin before separation by electrophoresis (Figure 3.10A), in contrast to Figure 3.9 where native nucleosomes are ran on the gel. Digestion of the non-repetitive templates with *Psi*I gave a partial digestion, even when chromatin was reconstituted at relatively low histone:DNA ratios, suggesting that histone octamers are blocking these sites and protecting them from digestion, and that nucleosomes are not forming centrally over the nucleosome positioning sites. This is unsurprising, considering the *in silico* predictions of nucleosome positioning shown in section 3.3, as the probability of nucleosome positioning over the centre of the positioning site is not always significantly higher than that over the surrounding linker

regions. This lack of complete digestion means that the analysis of reconstituted vs. unreconstituted sites as described in section 3.8.2 cannot be performed on these templates.

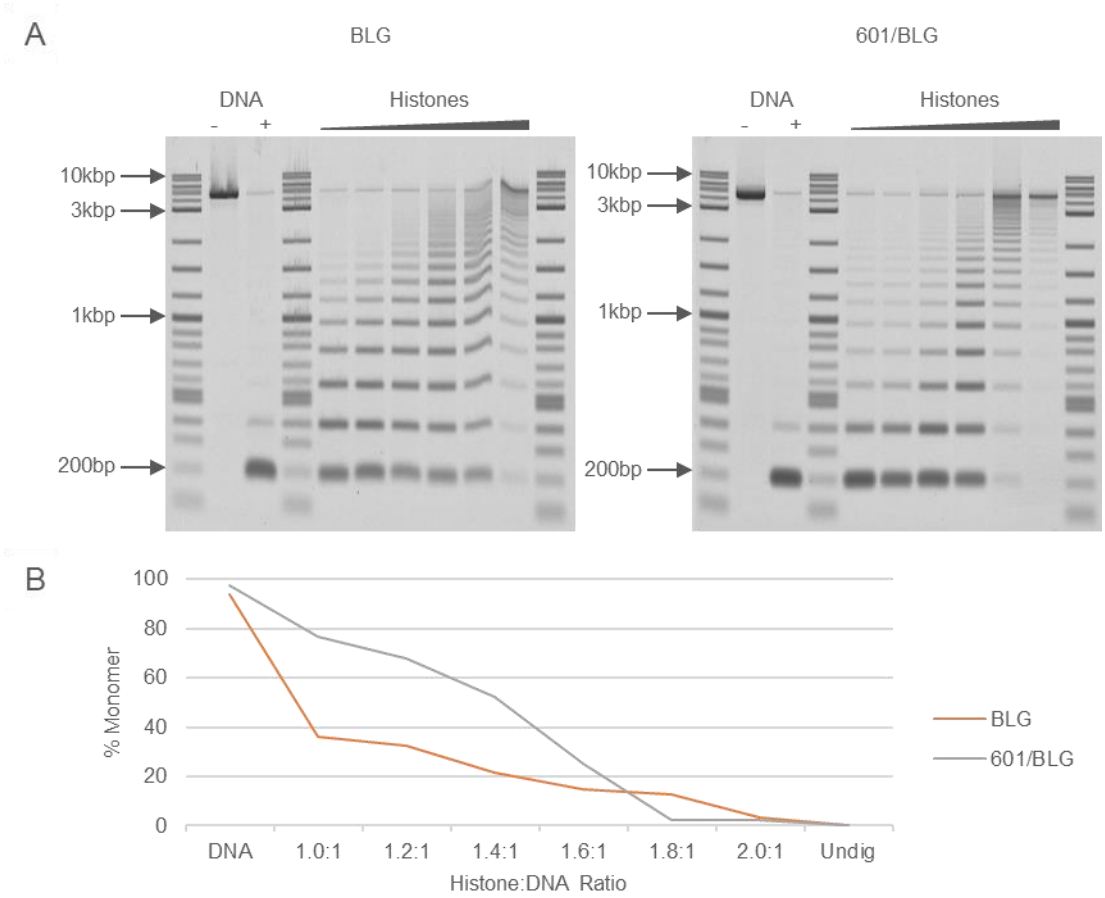

Figure 3.10. Restriction Digestion of Chromatin by PsiI.
A) Digestion of BLG (left) and 601/BLG (right) nucleosome arrays by PsiI, reconstituted at different histone:DNA ratios in the absence of a competitor DNA, and extracted from chromatin prior to agarose gel separation. B) Quantification of monomer DNA compared to DNA that has not been completely digested. Both remain incompletely digested, even at low histone:DNA ratios which would not be expected to saturate templates, suggesting that histones are positioned over linker regions.

Nevertheless, when *601/BLG* or *BLG* DNA was reconstituted, digested with *Psi*I and the DNA analysed, the resulting DNA ladders can provide information about the nucleosome positioning capacity of these fibres (Figure 3.10A). To compare the nucleosome positioning properties of these templates, the efficiency of digestion to monomer DNA by *Psi*I was measured. At histone:DNA ratios of 1.0:1 to 1.6:1, *601/BLG* is digested more efficiently, having a higher proportion of completely digested material, than the corresponding *BLG* fibres.

EMSA suggests that *BLG* becomes saturated in the absence of competitor at a histone:DNA ratio of 1.6:1, whereas *601/BLG* becomes saturated at 1.8:1 (section 3.5). At these ratios, *BLG* surprisingly appears to position nucleosomes slightly better than *601/BLG*. This is unexpected, as the "601" sites within the *601/BLG* are more likely to be occupied across the positioning sequence, and it was expected that this might force nucleosomes to be positioned more accurately across the interspersing "BLG" sites.

It should be noted that chromatin digestion enzymes (restriction enzymes, micrococcal nuclease or DFF) require divalent cations to function (either Mg^{2+} or Ca^{2+}). Restriction digestions were therefore performed in 1.5 mM magnesium chloride, which will induce folding of nucleosome arrays. It is possible that this folding may restrict enzyme access to some *Ava*I or *Psi*I restriction sites even where a nucleosome is not positioned directly over the site, but the site is nonetheless occluded within the folded structure. Poirier et al. (2008) found that folding nucleosome arrays *in vitro* using 5-10 mM $MgCl_2$ reduced the accessibility of linker DNA by as much as 50-fold compared to bare DNA. This may affect the efficiency of the enzymes, accounting for some degree of sample indigestion, and may cause a perfectly saturated nucleosome array to be slightly underdigested compared to a DNA control. Furthermore, if different reconstitution templates achieve different levels of chromatin compaction, or form different higher-order structures under these conditions, this could differentially impact each of the nucleosome arrays. If *601/BLG* forms a more compacted higher-order structure under these conditions than *BLG*, this might cause the lower level of digestion at saturating histone:DNA ratios despite more accurate nucleosome positioning over the 25 sites.

3.9 Micrococcal Nuclease Digestion of Chromatin is Affected by the Underlying DNA Sequence

In chromatin, micrococcal nuclease cleaves linker DNA leaving nucleosomal DNA intact. It is an important tool for studying primary chromatin structure, and recently has been paired with next-generation sequencing to map nucleosome positions. Complete digestion of cellular chromatin yields mono-nucleosomes and can be used to study nucleosome positions, while partial digestion will produce oligo-nucleosomes, and has been used to identify open regions of chromatin (Zhuo et al., 2017) – though DNaseI-seq and ATAC-seq are more widely used for this purpose. Micrococcal nuclease exhibits some sequence specificity towards A/T rich regions of chromatin (Dingwall et al., 1981) but this has not been found to substantially bias digestion of reconstituted chromatin on a typical genomic template (Allan et al., 2012).

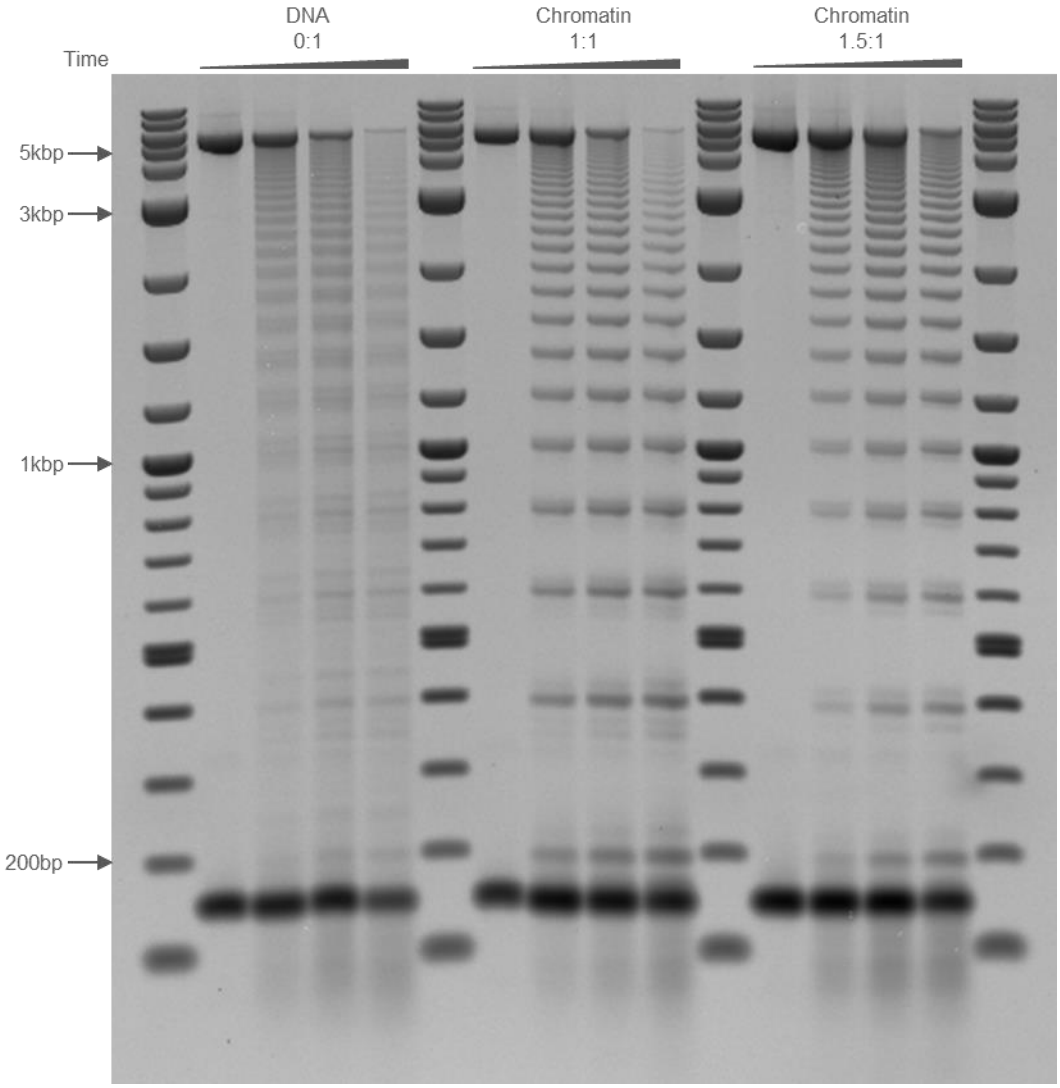

Figure 3.11. Digestion of DNA and Chromatin by Micrococcal Nuclease.
601 DNA and reconstituted chromatin digested by micrococcal nuclease. A nucleosomal ladder indicated chromatin reconstitution, but this ladder is also seen, albeit less clearly, upon digestion of the DNA sample, suggesting that the sequence bias of micrococcal nuclease contributes to this digestion pattern.

To analyse the primary structure of reconstituted fibres, ie. the nucleosome positioning and nucleosome repeat length, assessing the heterogeneity of these within each template, arrays were digested with MNase. A fibre with nucleosomes regularly spaced every 197 bp would be expected to reveal a 197 bp ladder on an agarose gel following digestion and DNA purification. A differently spaced ladder would suggest a different repeat length, whilst a sharply defined DNA ladder would indicate more regularly spaced and homogeneous population of fibres. Despite MNase not having sequence specificity on other DNA templates (Allan et al., 2012) it appears to be affected by the DNA sequences found in the "601" repeats. Micrococcal nuclease exhibits some specificity for A/T rich regions of DNA,

and is therefore more likely to cleave within the 50 base pair linker regions of this sequence, even in the absence of histones. When naked *601* DNA was digested by MNase, there was evidence of a 200 bp ladder, similar but less clear than the ladder seen following the digestion of *601* chromatin (Figure 3.11). For my purposes MNase was therefore not suitable to analyse nucleosome positions on the 601 template.

3.10 Digestion by DFF/CAD Nuclease Reveals Differences in Primary Chromatin Structure

3.10.1 DFF/CAD Nuclease

As micrococcal nuclease showed a high degree of sequence specificity when cleaving 601 DNA (Figure 3.11) I decided to use another enzyme called DFF/CAD (DNA Fragmentation Factor/Caspase Activated DNAse). It its inactive form, DFF/CAD is a heterodimer composed of a 40 kDa endonuclease subunit and a 45 kDa inhibitor subunit. Caspase cleavage during apoptosis cuts the inhibitory subunit, releasing the active endonuclease, which is able to cleave chromatin between nucleosomes in a similar manner to micrococcal nuclease. Structurally, DFF/CAD looks like a pair of "molecular scissors" as shown in Figure 3.12A (Allan et al., 2012; Samejima and Earnshaw, 2005). This structure likely contributes to its reduced sequence specificity in comparison to micrococcal nuclease, which binds within the minor groove of the DNA helix (Figure 3.12B). DFF also lacks exonuclease activity and it cuts within nucleosomes far less than MNase (Widlak and Garrard, 2006).

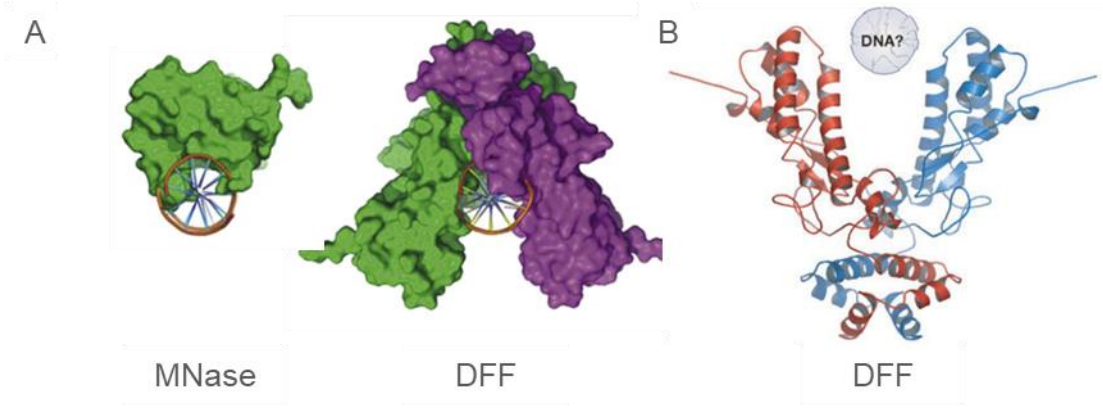

Figure 3.12. Structure of micrococcal nuclease and DNA fragmentation factor.
A) Comparison of the binding of micrococcal nuclease (left) and DFF (right) to DNA. Micrococcal nuclease binds to the minor groove, preferentially selecting for sites with higher A/T content, whereas DFF acts like a pair of “molecular scissors” with less sequence specificity and is unable to cut DNA within the nucleosome structure (Allan et al., 2012). B) DFF structure: A pair of molecular scissors (Samejima and Earnshaw, 2005).

As the enzyme is not commercially available I prepared the recombinant enzyme (section 2.5.3). While native DFF nuclease is activated through cleavage by caspase-3 or caspase-7 during apoptosis, a recombinant protein where the caspase cleavage site is replaced by a TEV protease cleavage site is described by Xiao et al. (2007). I purified this protein essentially as described by Xiao et al. in order to digest chromatin in a less sequence specific manner. Activation by TEV was always performed immediately prior to the experiment, as it is unknown how long the protein may remain active once cleaved *in vitro*.

3.10.2 DFF is Less Sequence Specific than MNase

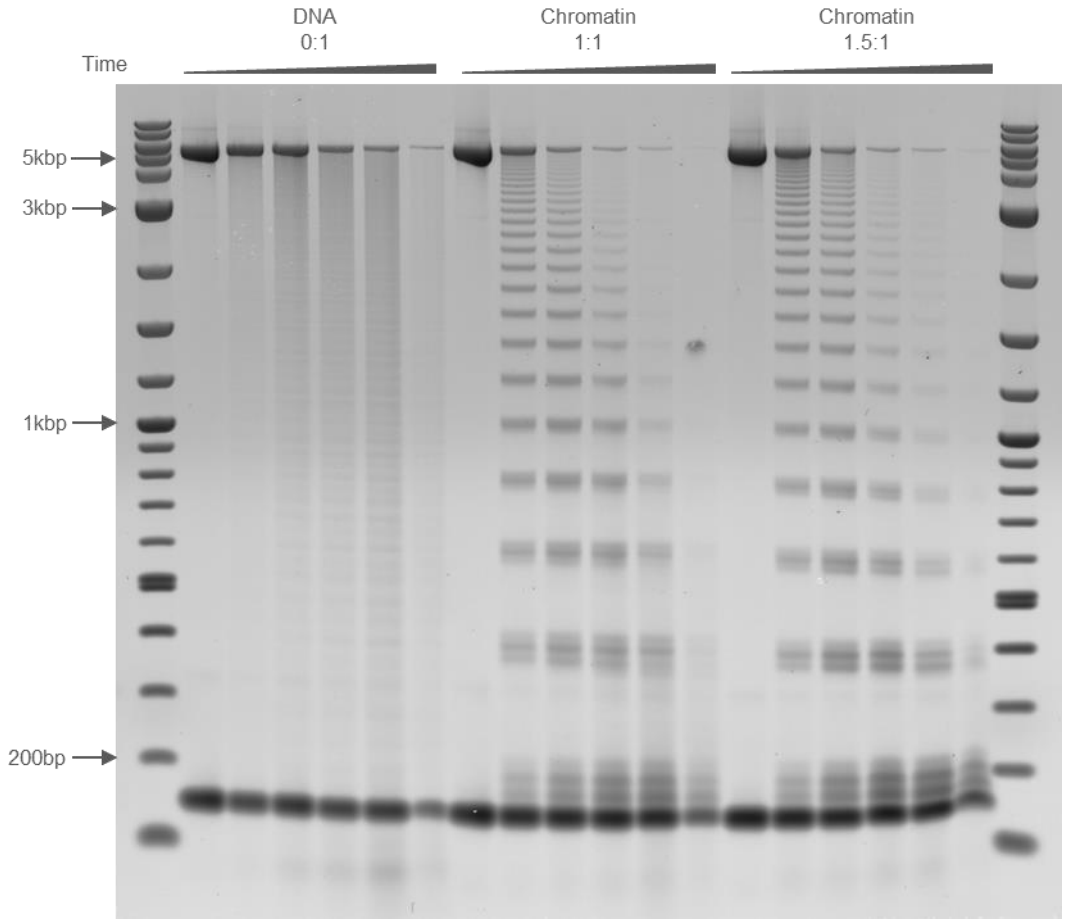

Figure 3.13. Digestion of 601 DNA and chromatin by DFF.
Gel showing 601 DNA and Chromatin reconstituted at 1:1 or 1.5:1 and digested by DFF for 0, 2, 4, 8, 16 or 32 minutes. A nucleosomal ladder is seen in reconstituted chromatin samples (1:1 and 1.5:1), but is not apparent when DNA alone is digested.

To confirm that DFF/CAD lacks sequence specificity compared to MNase, *601* DNA and *601* reconstituted chromatin were digested by the enzyme at 37°C. When unreconstituted DNA is digested and analysed by agarose gel electrophoresis, it lacks the characteristic digestion pattern of chromatin (Figure 3.13). Comparing the digestion of DNA by DFF/CAD in Figure 3.13 and by micrococcal nuclease in Figure 3.11, there is a pronounced

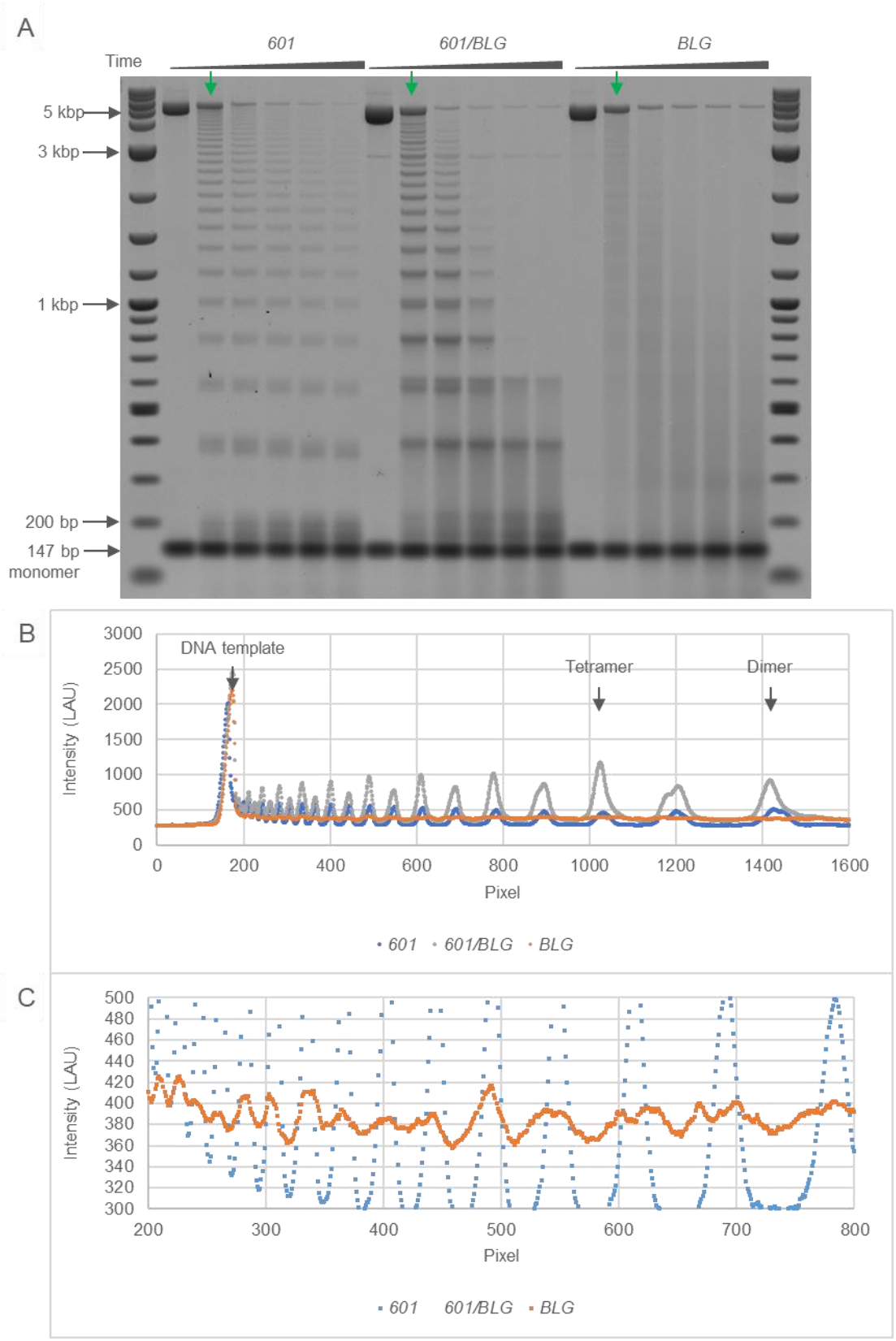

Figure 3.14. Digestion of Nucleosome Arrays by DFF/CAD at 37°C.
(Full legend found on following page)

A) Agarose gel showing protected DNA fragments after digestion of *601*, *BLG*, and *601/BLG* arrays reconstituted at 1.7:1, 2:1 and 2:1 respectively in the presence of a monomer DNA competitor. Nucleosomal ladders after 0, 2, 5, 10, 20 or 30 minutes digestion. B) Densitometry along the gel lane digested for 2 mins for the *601*, *601/BLG* and *BLG* templates (highlighted by the green arrow on gel). C) Densitometry along the gel lane digested for 2 mins for the *601* and *BLG* templates, zoomed in to show the faint ladder within the *BLG* corresponding with the ladder produced by the *601*.

difference with DFF/CAD not showing a nucleosomal ladder. In contrast chromatin is digested to leave a clear 200 bp ladder. There are several bands visible every 200 bp which appear to be approximately 157 bp, 177 bp and 197 bp as a monomer and approximately 360 bp, 380 bp and 400 bp as a dimer. This suggests that there might be three different sites within the "601" linker region where the enzyme preferentially digests, creating fragments of slightly different lengths, though it is surprising that there are no fragments larger than 200 or 400 bp observed.

3.10.3 Non-"601" Nucleosome Arrays are Less Stable at 37°C

To assess the differences in primary chromatin structure between the three nucleosome arrays, *601* was reconstituted at a histone:DNA ratio of 1.7:1 and *BLG* and *601/BLG* were reconstituted at 2:1 with a 147 bp monomer competitor. Chromatin was digested with DFF/CAD over 30 mins at 37°C, then DNA was extracted and fractionated on a 1% gel (Figure 3.14A). While a strong nucleosomal ladder is apparent for the *601* template and an intermediate ladder seen in the *601/BLG*, there is no clear ladder seen following digestion of the *BLG* for 2 mins (Figure 3.14B). There is also no evidence of a mono-nucleosomal fragment, suggesting that histone octamers are not protecting nucleosomal DNA from digestion by DFF/CAD. At later time points (20-30 min) a faint band appears suggesting digestion to a 286 bp fragment, however closer analysis of gel lane following 2 mins digestion reveals a faint nucloesomal ladder in the *BLG* that corresponds with the nucleosomal ladder seen following digestion of the *601* (Figure 3.14C). Surprisingly, the ladder produced by all three templates following 2 minutes digestion for fragment sizes above 3 kbp is approximately 240 bp, which might reflect the preferential cutting of DFF at different points within the linker. For *601* and *601/BLG* the ladder observed below 3 kbp has a clear 200 bp periodicity, but at this time point it is unfortunately not possible to see smaller fragments in the *BLG* to calculate whether this follows the same pattern as the *601* and *601/BLG* or if the ladder in fact reflects the 286 bp band which is seen in later time points. It is possible that *BLG* chromatin is less stable than the other chromatin samples at these temperatures and that nucleosomes are either removed or are easily able to slide along the DNA, allowing digestion of previously protected DNA regions resulting in this 286 bp band.

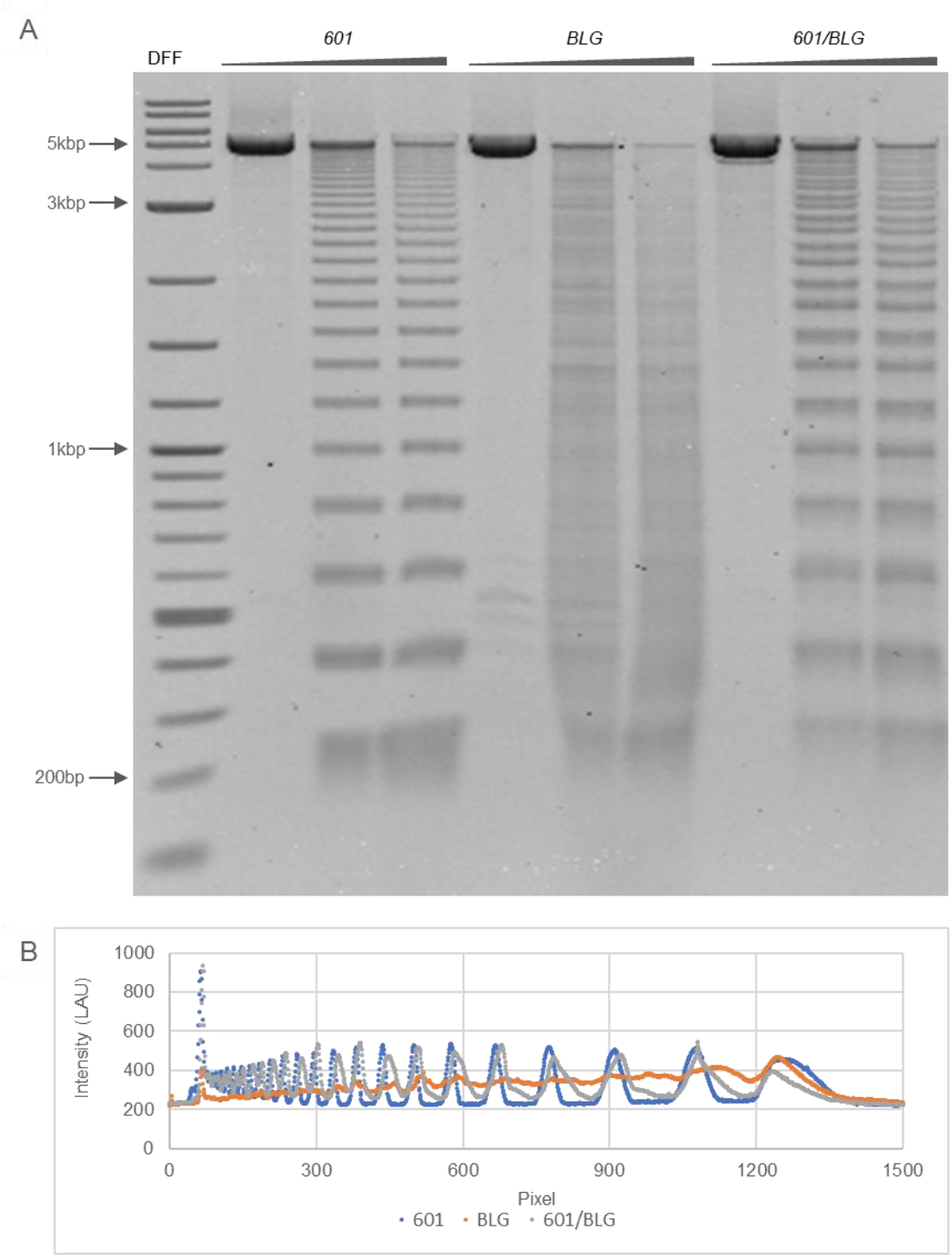

Figure 3.15. Digestion of Nucleosome Arrays by DFF/CAD at 4°C.
A) Digestion of *601*, *BLG*, and *601/BLG* arrays, reconstituted as described in section 4.5, in the absence of competitor. Samples digested with 1 or 4 units of enzyme for 40 min. B) Densitometry for samples digested with 4 units of enzyme, corrects for variations in sample migration across the gel.

Interestingly, there appears to be a bimodal digestion ladder seen in the *601/BLG*, with a clearer band appearing at every other "rung" of the ladder, suggesting that the enzyme is

more likely to cut at one end of the "601" repeat. The DFF has some sequence specificity based on purine and pyrimidine bases, particularly for sequences RRNYRNYY (5'-purine-purine-any-pyrimidine-purine-any-pyrimidine-pyrimidine-3') as described by Widlak et al. (2000). Two such sites are present within the "601" repeat which might contribute to this bimodal digestion.

3.10.4 Digestion by DFF at 4°C Yields a Mono-nucleosomal Ladder

To investigate whether temperature affected nucleosome mobility in these different DNA templates the samples were digested with DFF at 4°C. Chromatin was reconstituted in the absence of a competitor as described in section 4.5 and was digested for 40 min with 1 or 4 units of DFF. DNA was isolated by phenol/chloroform and analysed on a 1% agarose gel (Figure 3.15A) and quantified by densitometry (Figure 3.15B). Under these conditions, a mononucleosome approximately 200 bp in size is observed following digestion of the *BLG* fibre, though the ladder is not as sharp as either the *601* or the *601/BLG*, indicating either that *BLG* is still unstable at 4°C, or that *BLG* chromatin has a more heterogeneous structure than the *601* and *601/BLG* templates.

The digestion pattern of the *BLG* array, with a weak nucleosomal ladder, suggests that nucleosomes are often irregularly positioned over the 197 bp repeats and not centred on the dyad. This variability may be due to heterogeneous positioning of nucleosomes along individual chromatin fibres, or due to heterogeneous array structures throughout the population of fibres (or both). Sequencing of the monomers (and possibly the dimers in the case of the *601/BLG*) would distinguish between these possibilities and enable me to see whether there is variability between individual chromatin fibres, or if variability of the nucleosome repeat length occurs exclusively within individual chromatin fibres. However, it was not possible to isolate sufficient monomeric and dimeric material to perform this experiment.

3.11 Reconstitution by Salt Dialysis is more Efficient than ATP-dependent Chromatin Assembly

There are several methods of reconstituting chromatin onto DNA *in vitro*. ATP-dependent chromatin assembly using the core histone chaperone NAP-1 and the chromatin remodeller ACF as described by Fyodorov and Kadonaga (2003) may exhibit different sequence specificities to salt dialysis and may be a more effective method of reconstitution. As this partially mimics chromatin formation *in vivo*, it is possible that this approach might yield more biologically relevant nucleosome arrays.

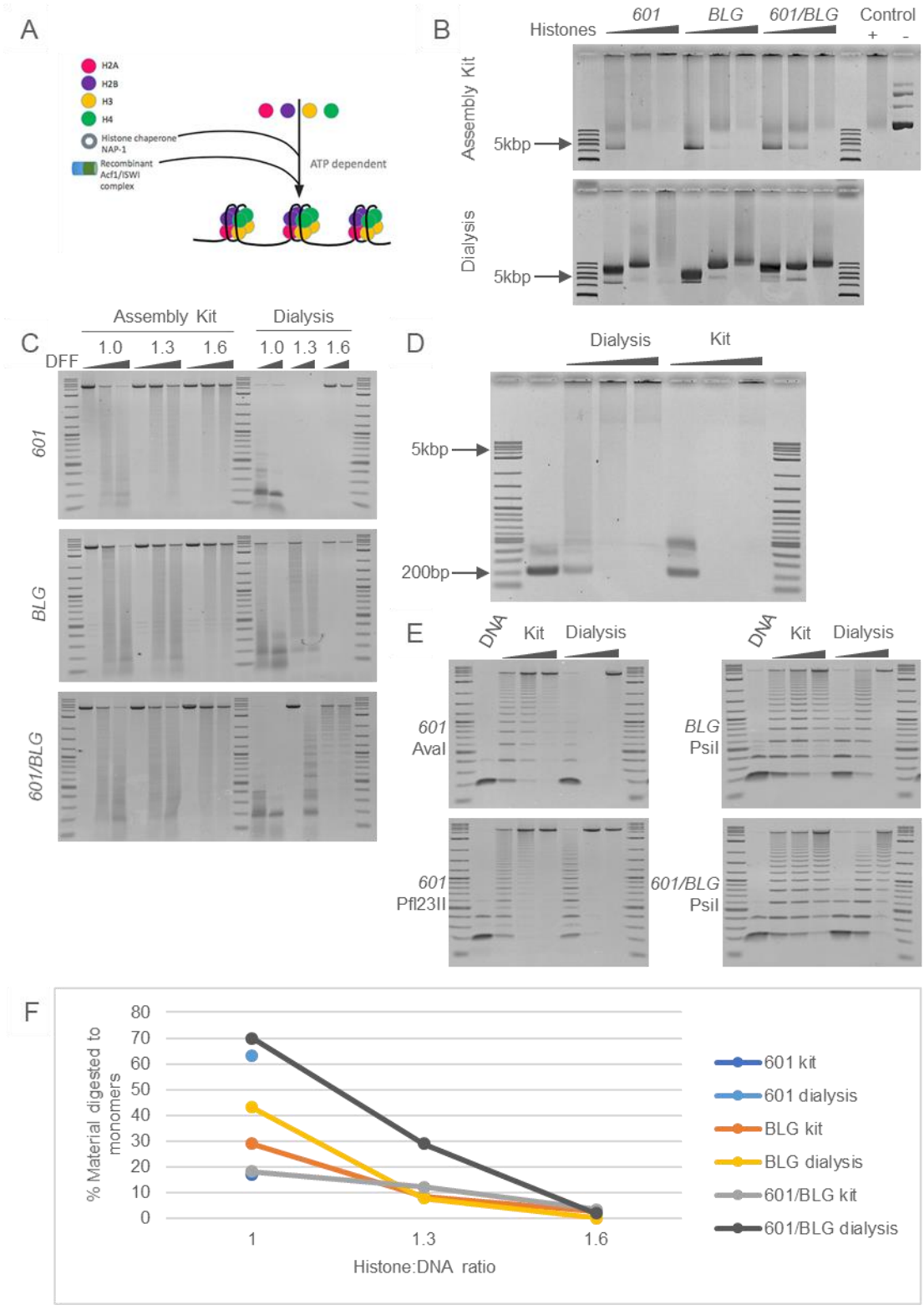

Figure 3.16. Analysis of Chromatin Reconstituted by ATP-Dependent Chromatin Assembly.
(Full legend found on following page)

A) Mechanism of chromatin assembly from DNA and recombinant histones using NAP-1 histone chaperone and ACF chromatin remodeller (Diagenode, 2015). B) Electrophoretic mobility shifts of samples reconstituted using a chromatin assembly kit compared to those reconstituted by salt dialysis. C) DFF/CAD digestion at 4°C of nucleosome arrays reconstituted using a chromatin assembly kit compared to those reconstituted by salt dialysis. D) *Ava*I digestion of *601* arrays reconstituted using a chromatin assembly kit or by salt dialysis to compare the amount of mononucleosomal and monomer DNA material. E) Restriction digestion of chromatin samples, with purified DNA separated by electrophoresis. *601* digestion by *Ava*I (top left), *601* digestion by *Pfl*23II (bottom left), *BLG* digestion by *Psi*I (top right), *601/BLG* digestion by *Psi*I (bottom right). F) Quantification of chromatin digested to monomers by AvaI or PsiI.

To compare reconstitution by salt dialysis with an ATP-dependent chromatin assembly, DNA templates were reconstituted using a Diagenode Chromatin Assembly kit according to the manufacturer's instructions, but using 1 µg of linear DNA and varying the amount of recombinant histone added. This was compared to chromatin reconstituted by salt dialysis using the recombinant histones included in the kit. Circular DNA, provided with the kit, was also reconstituted as a control. ATP-dependent chromatin assembly had the benefits of not losing material during dialysis (when chromatin presumably sticks to the dialysis membrane and around 20% of material may be lost) and of the sample volume not increasing during dialysis.

The chromatin prepared was analysed by a mobility shift assay and by digestion with restriction endonucleases and DFF/CAD. The mobility shift assay (Figure 3.16B) showed that when *601*, *BLG* and *601/BLG* were reconstituted using the assembly kit at histone:DNA ratios of 1:1 to 1.6:1, the resulting nucleosome arrays appeared to be more heterogeneous than those reconstituted by salt dialysis, with the bands appearing less sharply on a 0.7% agarose gel. These samples also precipitatedmore easily, as there was a considerable amount of material remaining in the well, even when reconstituted at a low histone:DNA ratio. At a 1.6:1 ratio, samples created by salt dialysis were also beginning to precipitate, which did not occur with chicken erythrocyte core histones (Figure 3.4A).

Chromatin was also digested with DFF (Figure 3.16C). This revealed some evidence of mono-nucleosomes, but very little evidence of a nucleosomal ladder when chromatin was assembled using a kit, suggesting that nucleosomes are not regularly spaced along these fibres, despite the presence of the chromatin remodeller ACF. When reconstituted by salt dialysis, mono-nucleosomes are seen at a 1:1 histone:DNA ratio, where the samples have been almost completely digested, and a nucleosomal ladder is seen at ratios of 1.3:1, though this is less clear for the *BLG* as the *601/BLG* (unfortunately, there was insufficient material for *601* reconstituted at 1.3:1 to be included in this analysis). At 1.6:1, there was no

digestion of *601* or *BLG*, suggesting that these are becoming oversaturated, but there is some digestion of the *601/BLG*, confirming that this template was less readily saturated with core histones.

When digested with *Ava*I and directly analysed by electrophoresis (Figure 3.16D), monomer DNA and mono-nucleosomal material are visible in the sample reconstituted using the kit with a 1:1 histone:DNA ratio. However, undigested chromatin and precipitated material are also visible in each lane. Chromatin does not appear to be well digested, suggesting oversaturation or incorrect placement of histones, at ratios of 1.3:1 and above. When the sample is reconstituted by dialysis at a ratio of 1:1, appears partially reconstituted, consisting of 61% monomeric material and 39% DNA (comparable to data following reconstitution with chichen erythrocyte core histones in figure 3.9). At a ratio of 1:1 the sample reconstituted by the kit appears to contain 56% monomeric material and 44% DNA. Unfortunately, not enough material reconstituted at 1.3:1 was recovered to perform this experiment. Furthermore, the DNA sample does not appear to have been completely digested by the enzyme, and the presence of dimer material in the chromatin samples will confound the results. The sample reconstituted at 1.6:1 is completely undigested, but appears to have come out of solution in the presence of 1.5 mM $MgCl_2$ (though this did not appear to occur when chromatin was reconstituted with CE core histones (Figure 3.9A), where chromatin appeared to be digested to fragments appearing as tetra-nucleosomes or smaller).

Chromatin samples were also digested by restriction enzymes *Ava*I and *Pfl*23II (*601*) and *Psi*I (*BLG* and *601/BLG*) and the digested DNA was purified and analysed by electrophoresis (Figure 3.16E). Digestion by *Ava*I reveals incomplete digestion of all three samples reconstituted by the kit, suggesting that even at low histone:DNA ratios, histones are occupying *Ava*I restriction sites in the linker regions, and not positioning over the "601" sites as they would following salt dialysis (Figure 3.16F). At a 1:1 histone:DNA ratio 17% of the material reconstituted using the kit was digested to monomers, compared with 63% of the material reconstituted by salt dialysis (Unfortunately, not enough chromatin reconstituted at 1.3:1 by salt dialysis was recovered to be able to perform this analysis on chromatin expected to be saturated). This suggests that chromatin assembly using ACF and NAP1 has little sequence specificity or a different specificity to that seen following salt dialysis.

Similar results were observed when *BLG* and *601/BLG* samples were digested with PsiI. Following reconstiton of *BLG* and *601/BLG* with the kit at 1:1 ratios and digestion by *Psi*I 43% and 70% of the material was found to be digested to monomers respectively (Figure

3.16F). Following reconstitution by ATP-dependent chromatin assembly, 29% of the *BLG* material and 18% of the *601/BLG* material had been digested to monomers, suggesting that salt dialysis positions nucleosomes more accurately over the designed nucleosome positioning sites than ATP-dependent chromatin assembly.

When digested by *Pfl*23II, which cuts within the "601" 147 bp core, *601* chromatin reconstituted by salt dialysis appeared to be digested when reconstituted at a 1:1 ratio, but is undigested at 1.3:1 and 1.6:1 ratios, suggesting that all the nucleosome positioning sites are occupied by nucleosomes at these ratios (with chicken erythrocyte core histones this was found to occur at a ratio above 1.4:1). When reconstituted using a chromatin assembly kit, there is still digestion when a 1.3:1 histone:DNA ratio is used, suggesting that the template is not fully saturated at this point, or that the nucleosome positioning is such that the *Pfl*23II site in the centre of the "601" core are not blocked by their occupancy. At 1.6:1, almost all of the *Pfl*23II sites appear to be occluded.

Salt dialysis seemed to produce more efficiently saturated and homogeneous chromatin samples where nucleosomes were relatively well positioned over the 25 nucleosome positioning sequences within each template. Therefore, all experiments in this thesis were conducted using chromatin reconstituted using salt dialysis. In addition, reconstitution using a chromatin assembly kit may not be easily scaled up to reconstitute chromatin at a high concentration, and the presence of additional proteins NAP1 and ACF may interfere with downstream analyses such as small angle X-ray scattering (section 4.5). The Diagenode Chromatin Assembly kit indicates that 5 µg of NAP1 (48kDa) should be used to reconstitute 1 µg of DNA; these proteins are likely to dominate structural analyses unless they are removed. It is possible that home-made enzymes may reconstitute chromatin more efficiently than those supplied within this kit.

3.12 Summary

In summary, three different DNA sequences with varying affinities for the histone octamer (Figure 3.2A) have been reconstituted to form nucleosome arrays. Different reconstitution methods have been trialled and salt dialysis was found to efficiently saturate DNA templates and create relatively homogeneous fibre populations (Figure 3.16). Several experiments have been conducted to confirm the reconstitution efficiency of each template and to identify the correct histone:DNA ratio to use to generate properly saturated nucleosome arrays for each, resulting in 25 nucleosomes formed on each template.

The *601* template appeared to require approximately a 1.4:1 histone:DNA ratio to be reconstituted in the absence of a competitor DNA, as measured by mobility shift assays

(Figure 3.4A) and digestion assays by *Pfl*23II and *Ava*I (Figures 3.8 and 3.9, respectively). In the presence of a competitor DNA, mobility shift assays indicate that the template still becomes saturated at a similar ratio (Figure 3.4B), although *Ava*I and *Pfl*23II digestion suggest that slightly more histones might be required to saturate all nucleosome positioning sites, as a small proportion of histones will bind to the competitor DNA (Figures 3.8 and 3.9). When additional histones are added, they appear to bind to the competitor; caesium chloride density gradients (Figure 3.7), and restriction digestion (Figure 3.9) show that template chromatin reconstituted in the presence of a vector backbone competitor at ratios as high as 1.7:1 are not oversaturated.

It is more difficult to assess the reconstitution efficiency of the novel *BLG* and *601/BLG* due to increased heterogeneity of the fibre population and the lack of restriction sites within the nucleosome binding sequences. *Pfl*23II and *Ava*I digestion, which are useful methods for determining the reconstitution efficiency of *601* arrays cannot be used on these templates. *Psi*I digestion reveals a reduced regularity in the nucleosome positioning properties of these sequences compared to the *601* (Figure 3.10). Nuclease digestion by DFF also suggested a population of chromatin fibres with more heterogeneously positioned nucleosomes (Figure 3.15). Caesium chloride gradients on these templates were not successful, likely as a result of their heterogeneity and therefore the reduced concentration that could be added to CsCl gradients. Electrophoretic mobility shift assays suggested that in the absence of a competitor these templates would require slightly higher histone:DNA ratios to achieve full saturation; *BLG* should be saturated at 1.6:1 and *601/BLG* should be saturated at 1.8:1 in the absence of a competitor. Band shifts also suggest that in the presence of a competitor extra histones are required compared to *601* due to the reduced affinity of these sequences for the histone octamer, approximately a 2:1 ratio.

Achieving full saturation for each of these templates is crucial as any variation in the level of saturation is likely to affect the folding of the arrays, and therefore without being confident that arrays are identically saturated, it is impossible to say that differences in the folding are purely a consequence of variations in the DNA sequence. It would be ideal if a more quantifiable measurement of the saturation of each fibre were possible, especially for the non-*601* fibres. Analytical ultracentrifigation or multi-angle light scattering could be used to calculate the saturation of each of these fibres from the sedimentation velocity and the molecular weight respectively, but it was not possible to create sufficient material to perform these analyses on a range of samples with different histone:DNA ratios.

The fact that it is impossible to internally control for chromatin saturation between different templates is the main challenge for studying each of these novel templates. However, I am confident that the *601* is becoming correctly saturated under the conditions described, based on its density and on restriction digestion, and that *BLG* and *601/BLG* are becoming similarly saturated based on their mobility shift.

Chapter 4. The Effect of DNA Sequence on Higher-Order Chromatin Fibre Structure

4.1 Introduction

The conformation of the higher-order chromatin fibre, often referred to as the "30-nm" fibre, is not well understood *in vivo* (section 1.1.2.2). The existence of the "30-nm fibre" seen in x-ray scattering experiments (Langmore and Paulson, 1983) and visualised by electron microscopy following the release of chromatin from nuclei (Thoma et al., 1979) has recently been disputed by Joti et al. (2012). Joti et al. state that while 30 nm structures may exist in transcriptionally inactive cells such as chicken erythrocytes, they did not find evidence of such a structure in the interphase nuclei of HeLa cells by SAXS, and that previous data was confounded by the presence of ribosomes in nuclei samples. It is known that the structure of the chromatin complex changes throughout the genome, with fibres of a more "open" higher-order structure being enriched in gene-rich domains whereas closed, compacted higher-order fibres are found in both heterochromatin and euchromatin (Gilbert et al., 2004). Ou et al. (2017) have recently described chromatin fibres with a variable fibre diameter between 5 and 24 nm, though it is possible that ChromEMT selects for negatively supercoiled regions of the genome which have a more open chromatin structure (Naughton et al., 2013) and can more readily bind the DRAQ5 drug.

In vitro, chromatin fibres have been visualised with a diameter between 25 and 45 nm, though the exact conformation of this fibre at varying repeat lengths is unclear. Robinson et al. (2006b) found by electron microscopy that "601" fibres form a one-start helical structure which has a diameter of approximately 35 nm when the nucleosome repeat length is between 177 and 207 bp, but increases to approximately 45 nm when the nucleosome repeat length is between 217 and 237 bp. Using a "601" fibre with a 167 bp repeat, Schalch et al. (2005) found by X-ray crystallography that the chromatin fibre formed a two-start helical structure with a diameter of 24-25 nm, which they expected could increase in diameter to accommodate a longer stretch of linker DNA. "601" arrays with a repeat length of 177 or 187 bp were found by Song et al. (2014) using cryo-electron microscopy to form a two-start helix with a diameter of 27.2 nm when the repeat length was 177 bp, increasing to 29.9 nm when the repeat length was increased to 187 bp. This is closely consistent with calculations by Athey et al. (1990) that fibre diameter might be expected to increase by 0.23 nm per base pair increase in the linker length. While it is difficult to reconcile results described by Robinson et al. (2006b) with those by Song et al. (2014), it is clear that fibre diameter and possibly conformation is dependent on the repeat length of the nucleosome array *in vitro*.

Kruithof et al. (2009) using single molecule force spectroscopy suggest that a "601" fibre with a 167 bp repeat has a two-start helical structure, while a 197 bp repeat has a one-start helical structure. Routh et al. (2008) suggest that the shorter DNA linker length of 167 bp repeat arrays forces nucleosomes to stack in a zig-zig arrangement, and therefore form a two-start helix upon compaction, whereas for 197 bp repeat arrays, the nucleosome arrangement within the higher-order structure is determined by the linker histone. Although experiments have been performed on fibres *in vitro* with varying nucleosome repeat lengths, these repeat lengths have been consistent within the individual fibres and no model has examined the effects of varying the nucleosome repeat length within a single chromatin fibre. Varying the DNA sequence of a chromatin fibre, rather than using repeating sites with strong nucleosome positioning properties is likely to affect the nucleosome spacing within individual fibres (section 3.10.4), which may have an effect on the regularity of such fibres. Athey et al. (1990) discuss the potential effects of heterogeneous nucleosome spacing within chromatin fibres, and suggest that local heterogeneity may give rise to local disorder, causing potential artefacts in studies of bulk chromatin populations. Furthermore, Ricci et al. (2015) suggest that variable nucleosome length might contribute to a "nucleosome clutch" structure *in vivo*, rather than a canonical 30-nm fibre. Since the discovery of the "601" nucleosome positioning sequence, studies of chromatin structure have become more focused on homogeneous populations of chromatin fibres while the effects of DNA sequence and variation in the length of the nucleosome linker within individual fibres upon the higher-order structure have not been well studied.

When studying the effects of DNA sequence on the structure of the folded chromatin fibre, it is likely that differences will arise primarily as a consequence of differential nucleosome positioning and heterogeneous nucleosome spacing, though the stability of individual nucleosome structures may also play a role. It is hypothesised that fibres with variable DNA sequences, including biological positioning sequences which do not have such a high affinity for the histone octamer or such strong positioning properties as the "601", will have a disrupted fibre structure. To assess the structure and dynamics of folded chromatin fibres, nucleosome arrays described in section 3.3 have been reconstituted and folded into higher-order fibres using divalent cations and linker histones. The structures were analysed by sucrose gradient sedimentation and small-angle x-ray scattering, while the dynamics of individual chromatin fibres with known saturation have been analysed by single-molecule force spectroscopy.

4.2 Folding Nucleosome Arrays with Linker Histones Does Not Affect Gel Mobility

When comparing the structure of folded chromatin fibres, it is first necessary to determine the amount of linker histone required to fold 10-nm nucleosome arrays into compacted structures. Linker histones bind to the nucleosome dyad and neutralise the charge of linker DNA to stabilise the salt-dependent folding of nucleosome arrays. Chromatin under-titrated with linker histones will not be fully compacted, but chromatin over-titrated with linker histones will become insoluble. The amount of linker histone within nuclei varies between cell types, with chicken erythrocytes containing 1.4 molecules of H5 per nucleosome (Kowalski and Pałyga, 2011) while other somatic cells contain between 0.4 and 0.8 molecules of H1 per nucleosome (Woodcock et al., 2006). The lower ratio of linker histones:nucleosomes in transcriptionally active cells is consistent with a role for linker histones in gene repression.

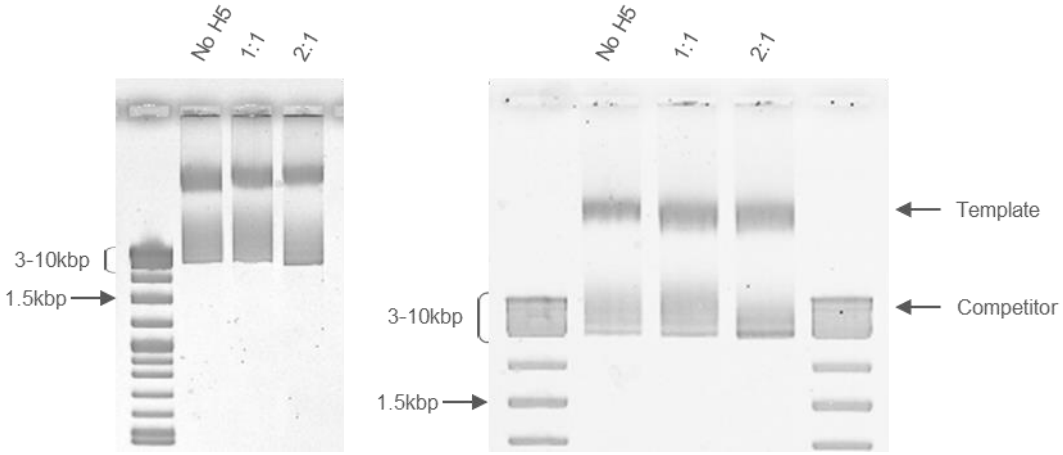

Figure 4.1. Electrophoretic Mobility Shift Assay of Nucleosome Arrays titrated with Linker Histone H5.
Two technical replicates of *601* chromatin titrated with linker histones. Chromatin was reconstituted in the presence of a vector backbone competitor at 1.7:1 core histones:DNA. Linker histone was added at ratios of 1 or 2 linker histone molecules per nucleosome. Nominal molecular weights of the template were calculated at 31.0 kbp, 32.0 kbp and 32.5 kbp (right gel) respectively, and 20.2 kbp, 19.6 kbp and 19.6 kbp (left gel). Upon adding more than 1:1 linker histone ratio, the difference in mobility appears to be primarily in the competitor DNA.

Huynh et al. (2005) describe a gel mobility-based assay to assess the binding of linker histones to nucleosome arrays *in vitro*. In 0.2 × TBE running buffer on 1.4% agarose gels, nucleosome arrays migrated faster after the addition of linker histones, possibly as a result of the structural changes induced. After "601" arrays became fully titrated, as confirmed by a

plateauing of the speed of migration, it was assumed that excess linker histones bound to competitor DNA fragments.

In order to assess the binding of linker histones to *601* arrays, chromatin was reconstituted in the presence of a vector backbone competitor at a core histone:DNA ratio of 1.7:1. Fibres were dialysed into 80 mM NaCl, a salt concentration at which nucleosome arrays would be expected to form folded structures, with interactions between individual nucleosomes, but not a true higher-order structure in the absence of linker histones. Following reconstitution, DNA concentration was measured by UV spectrophotometry, linker histone H5 was added at 1 or 2 molecules per 197 bp of template DNA and chromatin was analysed on 1.4% agarose gels (Figure 4.1).

Insignificant changes to the mobility of chromatin in the gel were measured following the addition of linker histones to *601* chromatin. A very small difference was seen between unfolded chromatin and chromatin with a 1:1 ratio of linker histones:nucleosome repeats. There was no difference between fibres titrated with a 1:1 and a 2:1 ratio of linker histone, but there was a difference in shift between the competitor fragments in these samples. In some cases, chromatin migrated slightly faster after the addition of linker histones (right gel), as described by Huynh et al. but in other experiments chromatin migrated slightly slower following compaction (left gel), likely due to very small differences in the agarose gel concentration of each replicate experiment. It was therefore concluded that this was not a sufficiently discriminative assay to measure linker histone titration.

4.3 Solubility of Nucleosome Arrays in the Presence of Linker Histones

The solubility of chromatin fibres after the addition of linker histones was measured to identify an appropriate ratio of linker histones:nucleosomal DNA, without causing the chromatin fibres to become insoluble.

Three DNA sequence templates were reconstituted in the presence of a vector backbone competitor. *601* was reconstituted at 1.7:1, *BLG* and *601/BLG* were reconstituted at a 2:1 ratio. Chromatin was titrated with linker histone H5 up to a 3:1 molar ratio of linker:nucleosomes in the presence of 80 mM NaCl. Chromatin was briefly centrifuged, and the supernatant was analysed on a 0.7% agarose gel, with (left) or without (right) purification of DNA (Figure 4.2A). The proportion of template DNA remaining in solution compared to a sample with no linker histones added was quantified from the DNA lanes (Figure 4.2B). *601* appears to remain most soluble in the presence of linker histones, with 84.4% remaining

soluble at a 1:1 ratio of linker histones:nucleosomes. This drops to 59.1% and 48.4% in *601/BLG* and *BLG* respectively. Assuming that these samples are equally saturated, surprisingly this suggests that fibres with a more heterogeneous structure do not retain solubility upon folding as well as regularly spaced nucleosome arrays.

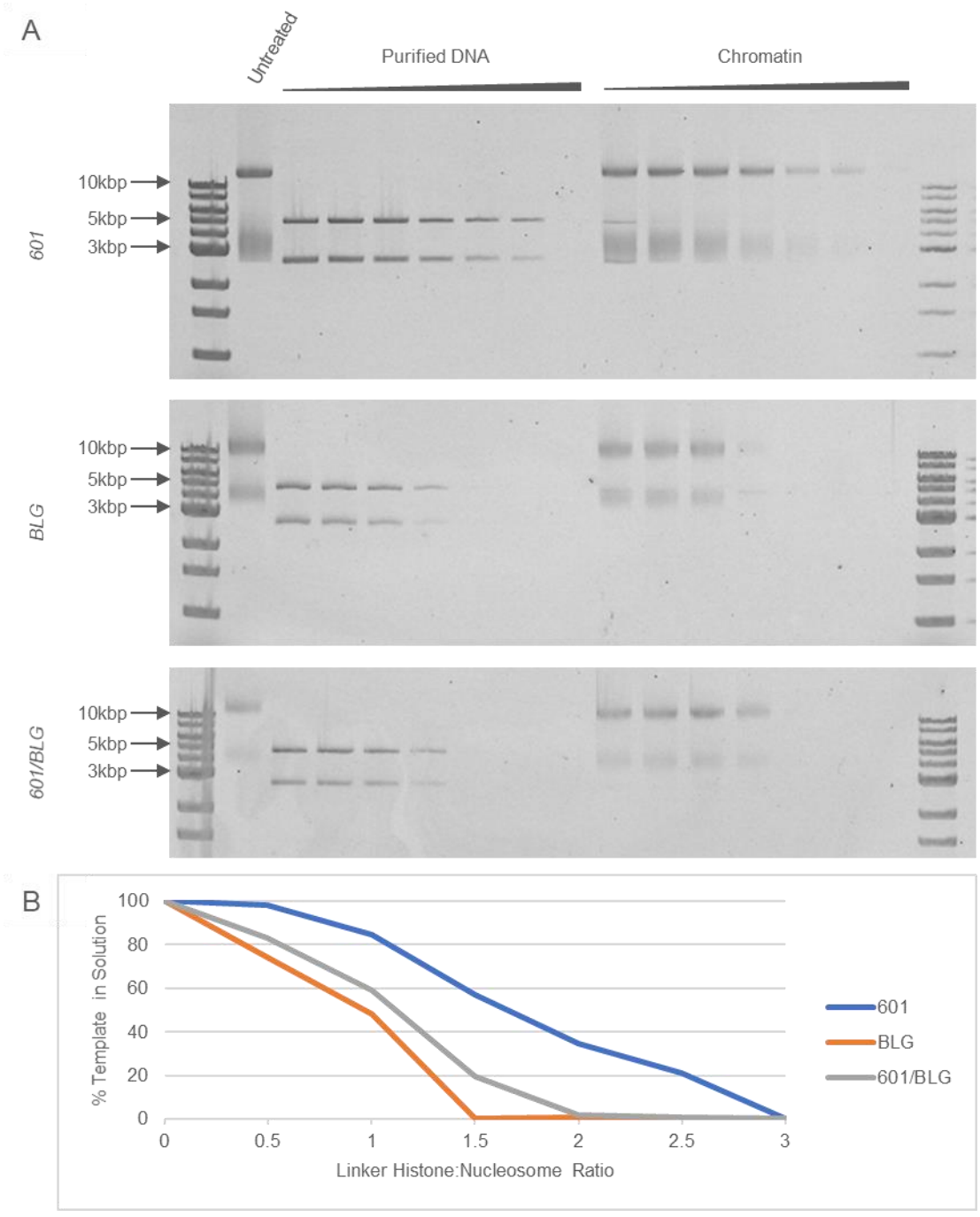

Figure 4.2. Titration of Nucleosome Arrays with Linker Histone.
A) Agarose gel analysis of DNA extracted from precipitated chromatin (left) or chromatin (right) following titration, incubation and centrifugation in the presence of linker histone H5 up to a H5:nucleosome ratio of 3:1. B) Quantification of template DNA that remained in solution after titration with linker histone.

It is difficult to control for the presence and nature of the competitor during this experiment. As the competitor is more reconstituted in both the *BLG* and *601/BLG* this may affect the affinity of linker histones for the competitor, which may affect the amount of linker histone binding to the template chromatin fibres. If linker histones were to bind more easily to the chromatinised competitor, less would bind to the template, causing it to appear comparatively more soluble. However, as the competitor DNA appears to precipitate at similar levels to the template in each case, this seems unlikely to be having a significant impact on the results.

4.4 Folded Fibres have a Similar Sedimentation Rate but Non-*601* fibres have a more Heterogeneous Structure

Sucrose gradient sedimentation separates macromolecules based both on their mass and their hydrodynamic shape. A heavier particle (for example, a chromatin fibre that is more saturated with core histones) will sediment faster than a less massive particle. Assuming that two particles have the same mass (such as two chromatin arrays of equal size and saturation), a more compacted particle will sediment more quickly than an unfolded molecule (Figure 4.3A). Heterochromatin that has a more compact structure has been shown to sediment more quickly than bulk chromatin of a similar fragment size (Gilbert and Allan, 2001), whereas chromatin that contains discontinuities to its higher-order structure will sediment more slowly (Caplan et al., 1987; Naughton et al., 2010). This method is therefore an informative measure of chromatin fibre compaction following the addition of linker histones to nucleosome arrays.

To measure the compaction of chromatin fibres following the addition of linker histones, *601* chromatin was reconstituted at a 1.5:1 histone:DNA ratio, *BLG* chromatin was reconstituted at a 2:1 ratio, and *601/BLG* chromatin was reconstituted at a 2:1 or a 1.8:1 ratio in the presence of a vector backbone competitor which had been digested into small fragments. Electron microscopy was performed on each of these samples, which appeared well-saturated, but unfortunately the average number of nucleosomes could not be counted due to the concentration of chromatin and the presence of salt or sucrose in the samples following sedimentation (See Figure 4.4). Linker histones were added at ratios of 1:1, 1.4:1 or 1.8:1 in 80 mM NaCl and sedimented on isokinetic 6-40% sucrose gradients. In an unfolded state, fibres might be expected to sediment at a similar rate, but upon folding differences in the compaction of chromatin fibres with different underlying DNA sequences might be revealed by a difference in the sedimentation rate of each fibre.

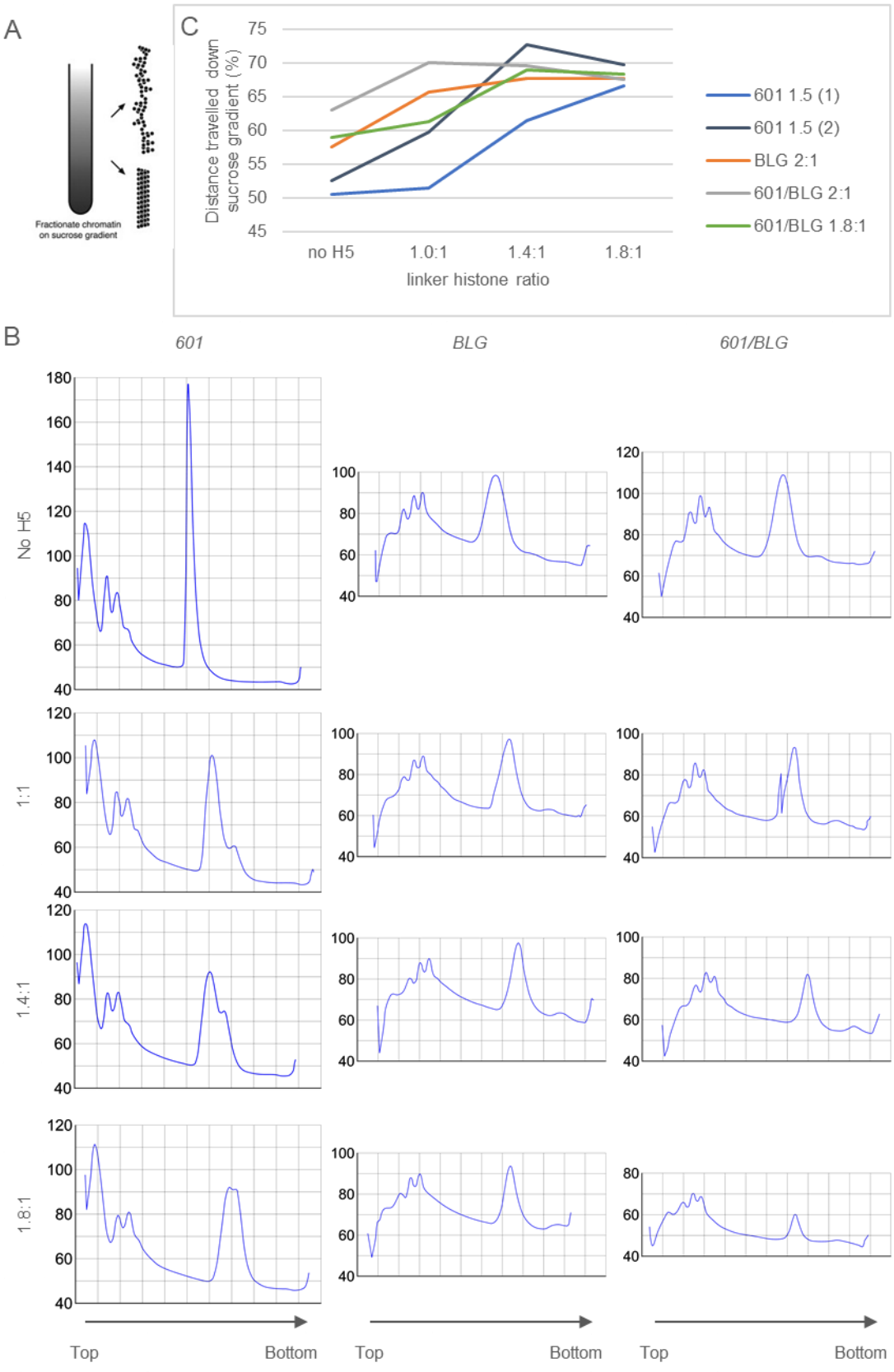

Figure 4.3. Sucrose Gradient Sedimentation of Fibres Titrated with Linker Histone.
(Full legend found on following page)

A) Sucrose gradient sedimentation separates chromatin based on both size and structure. A "compacted" chromatin fibre will sediment faster than "open" or "disrupted" chromatin fibres of equal mass (Gilbert et al., 2004). B) Sedimentation profiles of *601*, *BLG*, and *601/BLG* reconstituted at 1.5:1 (*601*) or 2:1 (*BLG* and *601/BLG*) histone:DNA ratios, sedimented in the absence of linker histones and titrated with linker histones at ratios of 1:1 to 1.8:1. C) Quantification of chromatin sedimentation (measured as a proportion of the length of the gradient profile) for three chromatin fibres. *601/BLG* was also reconstituted at 1.8:1 histones:DNA as a large amount of material became insoluble following reconstitution at 2:1.

The sedimentation profiles of each fibre show significant variation in the sedimentation velocity of the unfolded chromatin fibres (Figure 4.3B, C). Surprisingly, *601/BLG* and *BLG* both sediment faster than *601* arrays in 80 mM NaCl in the absence of linker histones. If these are assumed to be saturated with core histone octamers to a similar degree, this suggests that the heterogeneity in the nucleosome positioning of these arrays is having a demonstrable impact on their structure. It is possible that non-*601* fibres are more flexible and allow more nucleosome-nucleosome interaction than *601* fibres in 80 mM NaCl, causing them to have a more compacted structure, or that nucleosomes are able to be more closely spaced within regions of the non-*601* fibres, affecting their overall shape and mobility. The *BLG* and *601/BLG* arrays also display much broader peaks in their sedimentation profiles, as a result of their heterogeneity either in the number of nucleosomes bound to each chromatin fibre or in the folded structures that these arrays form in 80 mM NaCl.

Upon the addition of linker histones, fibre sedimentation velocity increases as a result of increased mass and compaction of their structures (Figure 4.3B). There is a considerable increase in the sedimentation rate of the *601* fibres, which sediment approximately 52% of the distance of the gradient profile in the absence of linker histones and 68% of the distance of the gradient with a 1.8:1 ratio of linker histones added (Figure 4.3C). The difference in the sedimentation of the non-*601* fibres after the addition of linker histones is considerably smaller. *BLG* and *601/BLG* fibres reconstituted at 2:1 core histones:DNA migrated approximately 59 and 63% of the sucrose gradient respectively. Upon the addition of linker histones to a 1.8:1 ratio, the sedimentation increases to approximately 68% of the distance of the gradient, a similar extent of sedimentation to that achieved by *601* fibres. If there are disruptions in the structure of the non-*601* fibres, ie. regions of naked DNA that cannot be compacted by linker histones, this might limit the folding of the fibre upon the addition of linker histone, while having a comparatively small effect on the dynamics of the fibre in 80 mM NaCl.

Width ½ height	*601*	*BLG*	*601/BLG*
No H5	3mm	10mm	9mm
1:1	6.5mm	9mm	6.8mm
1.4:1	11mm	7mm	5.5mm
1.8:1	9.5mm	6mm	4.8mm
2.2:1	7mm		

Table 2. Width at ½ Height of Chromatin Sedimentation Peaks.
Where peaks are divided into two distinct peaks, the larger peak is measured.

Interestingly, when linker histones were added to *601* chromatin, two clear species of fibres structures were seen by sucrose gradient sedimentation. These two species were also visualised by Huynh et al. (2005) in electrophoretic mobility shift assays, and may indicate that linker histones bind cooperatively to defined chromatin fibres. Linker histones bind cooperatively to naked DNA (Clark and Thomas, 1986; Thomas et al., 1992), and comparable interactions have been thought to exist in chromatin (Wolffe, 1998). The two species are likely to have different sedimentation velocities as a fibre with more linker histone bound will have a greater mass and be more compacted. It is possible that differences in core histone saturation in the population of arrays might cause differences in structure that can only be observed upon folding, however this is unlikely as the distribution of the two peaks changes depending on the amount of linker histone added. While a similar phenomenon may occur in the *BLG* and *601/BLG* fibres, the intrinsic heterogeneity of these fibres may mean that separate species cannot be resolved by sucrose gradient sedimentation. When two peaks were seen in sedimentation profiles, the sedimentation distance of the major peak was measured for analysis in Figure 4.3C and Table 2. As these peaks overlap, it was not possible to image them separately by electron microscopy to visually determine differences in structure.

Electron microscopy was performed following the sedimentation of chromatin (section 2.4). Prior to fixation, chromatin was dialysed into 80 mM NaCl buffer, removing the sucrose from samples, using mini dialysis units. While insufficient numbers of individual fibre images could be isolated for a complete analysis, images of typical chromatin fibres within each sample are shown in Figure 4.4. Unfortunately, no fibre images could be isolated of *601/BLG* titrated with a 1.8:1 ratio of linker histones. In the absence of linker histones, individual nucleosomes can be seen within *601* fibres, and these fibres appear to be less compacted than *BLG* and *601/BLG* fibres, consistent with them having a lower sedimentation rate. Upon the addition of 1.8:1 linker histones, *601* fibres appear far more

compacted, though there is still considerable variation in the lengths of the compacted fibres. The *BLG* fibres appear less rod-like and more disorganised at this histone:DNA ratio.

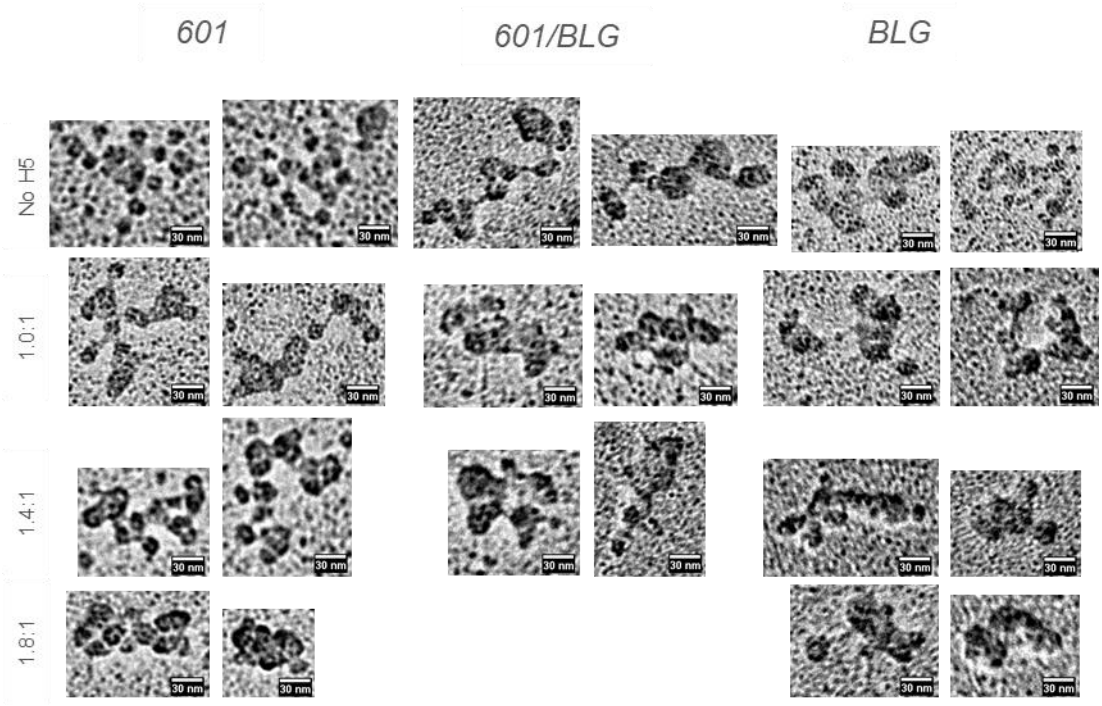

Figure 4.4. Electron Microscopy of Sucrose Gradient Samples.
601 reconstituted at 1.5:1 and *601/BLG* and *BLG* reconstituted at 2:1 in the presence of a competitor DNA, titrated with up to 1.8:1 ratios of linker histone H5 and separated by sucrose gradient sedimentation. Peak fractions were dialysed into 80 mM NaCl then fixed and imaged as described in section 2.4.

4.5 Small-angle X-ray Scattering Reveals Differences in the Folded Shape and Volume of Chromatin Molecules

Small-angle X-ray Scattering (SAXS) detects small differences in the electronic density of macromolecules, allowing their size and shape to be measured in solution. Analysing the structures of these fibres in solution, rather than by electron microscopy, does not rely on a fixative, which might change the structure of a molecule during the period of fixation.

X-ray scattering has revealed 30-nm structures in samples of "601" nucleosome arrays folded by divalent cations and linker histone H1 (Maeshima et al., 2016), though it is argued by this group that chromatin in solutions of high Mg^{2+} concentration better reflect chromatin structure *in situ*, forming globular structures that lack a 30-nm fibre. 30-nm structures have also been detected in samples of cellular chromatin (Langmore and Paulson, 1983). It has been suggested that these results were confounded by the presence of ribosomes in the samples (Joti et al., 2012; Nishino et al., 2012), however treatment of chromatin by DNase1

caused the loss of the 30-nm peak, suggesting that it is indeed chromatin that contributes to this structure.

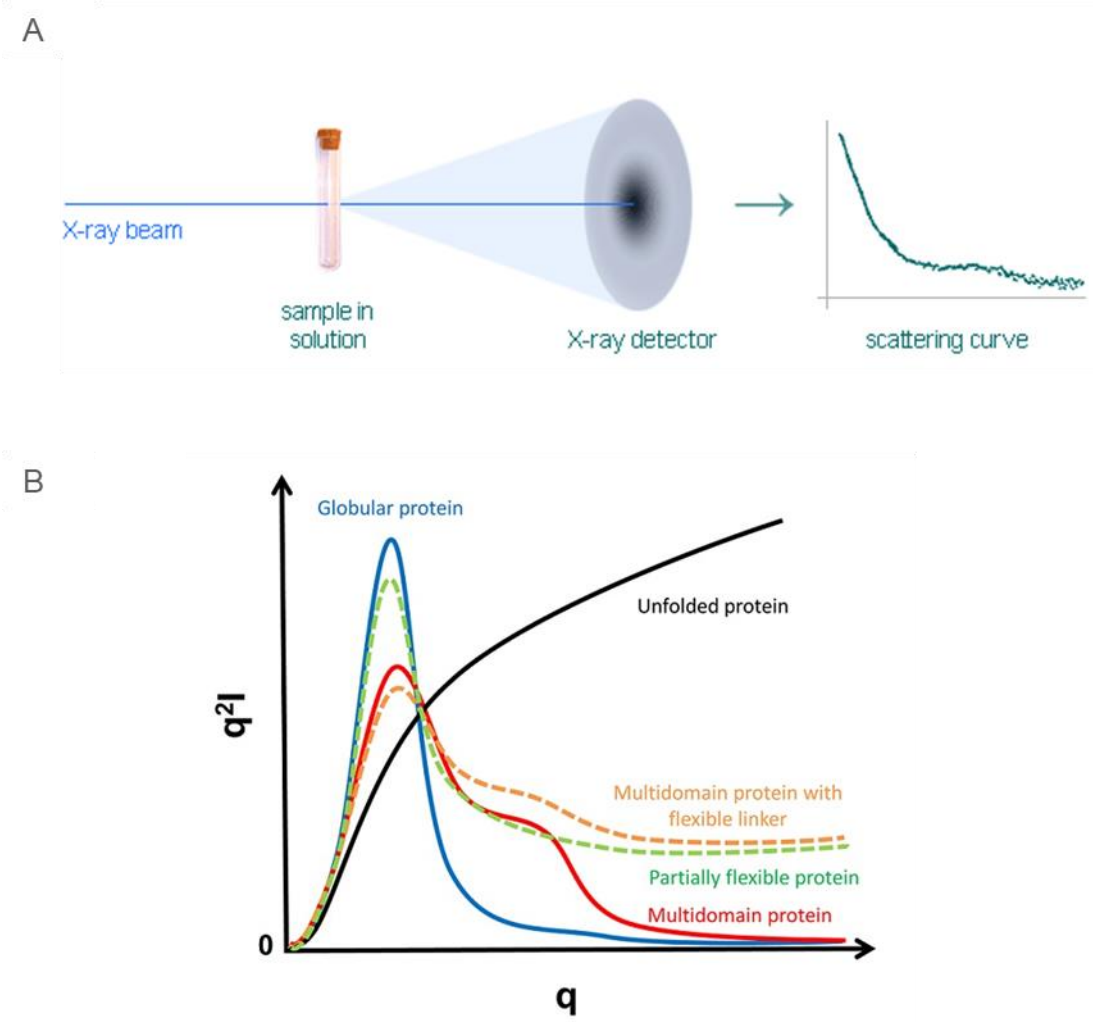

Figure 4.5. Small-angle X-ray Scattering to Analyse Fibre Structure in Solution.
A) Small angle x-ray scattering. An x-ray beam is directed through a solution containing the chromatin sample and the scattered radiation is registered by a detector. A scattering curve comparing intensity with scattering angle provides information about the size and shape of macromolecules. (BIOSAXS, 2018). B) Kratky analysis, comparing scattering angle with the square of the scattering angle × scattering intensity, reveals features in the shapes of the molecules in the sample. I = scattering intensity, q = scattering angle. (SSRL SLAC, 2017).

In order to analyse structural differences between chromatin arrays containing different DNA sequences in solution, SAXS was performed at Diamond Light Source B21 with assistance from Nikul Khunti, Nathan Cowieson and Rob Rambo. Samples were irradiated and the scattering intensity at various angles was measured (Figure 4.5A).

The scattering intensity at different angles (scattering vector = q) provides information about the size (low q), shape (intermediate q) and internal structure/volume (high q). The molecule shape is most easily assessed using a Kratky plot (Figure 4.4B) which divides out the decay of the scattering at larger angles, enhancing features of the scattering pattern to better analyse shape.

As SAXS takes an average measurement across all macromolecules in a sample, reconstituted chromatin must be isolated from the competitor so that this does not impact the assessment of the fibre structure. SAXS can be coupled with size-exclusion chromatography to separate different species within a sample and allow SAXS measurement of each individual peak eluted from a column. Initially, chromatin was reconstituted in the presence of competitor (vector backbone digested with *Apa*LI, *Bsp*HI and *Eci*I to fragments less than 500 bp) and SEC-SAXS was used to separate the template chromatin from the competitor fragments (Figure 4.6). A Shodex KW405-4F column was able to separate reconstituted chromatin from competitor fragments, with the template being eluted after 8.2 minutes and the largest competitor fragment being eluted after 10.2 minutes, running at 0.25 ml/min. However, separation diluted the chromatin extensively, with chromatin injected at a concentration of 0.36 mg/ml being diluted to 0.03 mg/ml upon elution. This concentration is far too low for accurate SAXS analysis, Grishaev (2012) suggests that concentrations of at least 0.1 mg/ml are required for analysis of protein complexes (concentrations as high as 10 mg/ml of macromolecules may be measured, depending on the nature of the sample). Whether the column eluate was directly analysed by SAXS or whether fractions were collected for direct injection onto the SAXS beam no significant scattering patterns were detected.

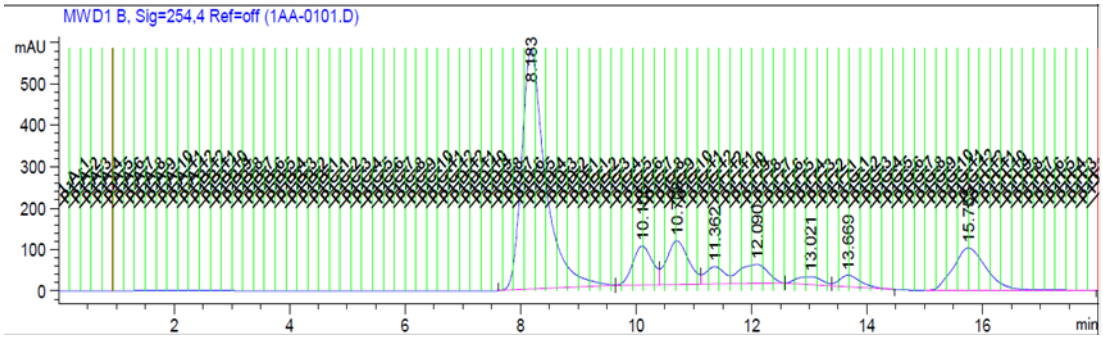

Figure 4.6. SEC-SAXS to Separate Competitor from Template Arrays.
UV trace of *601* chromatin separated from competitor fragments using a Shodex KW405-4F column. Template was eluted after 8.2 min and completely separated from digested vector competitor, first eluted after 10.1 min, but final concentration of chromatin, as measured by UV absorbance was approximately 0.03 mg/ml (diluted from 0.36 mg/ml), much lower than required for SAXS analysis.

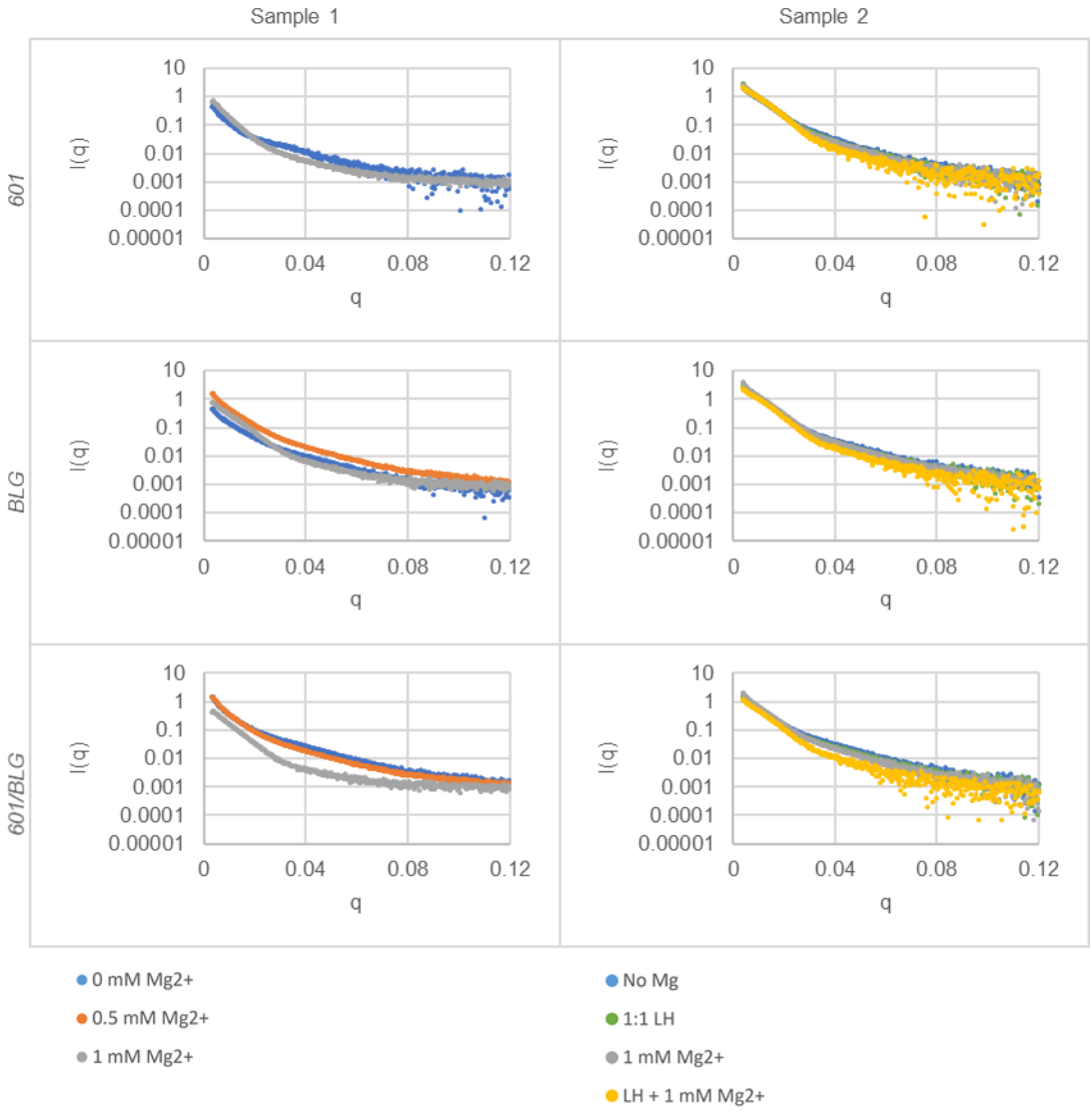

Figure 4.7. Log Intensity Curves of Scattering Data Derived from Fibres Titrated with Magnesium and Linker Histones.
Experiment 1 - left: *601* reconstituted at 1.2:1 and *BLG* and *601/BLG* reconstituted at 1.4:1 titrated with 0.5 or 1 mM $MgCl_2$. Unfortunately no data was collected for 601 at 0.5 mM $MgCl_2$. Experiment 2 – right: *601* reconstituted at 1.12:1 and *BLG* and *601/BLG* reconstituted at 1.3:1 titrated with 1 mM $MgCl_2$, 1:1 ratio of linker histones or both $MgCl_2$ and linker histones.

Chromatin was therefore reconstituted in the absence of competitor. While the histone:DNA ratios necessary to reconstitute chromatin in these conditions at low concentrations were determined in section 3.5, much higher concentrations of chromatin (approximately 0.2 mg/ml) were required for SAXS analysis. Chromatin reconstituted at the ratios described in section 3.5 (*601* at 1.4:1, *BLG* at 1.6:1 and *601/BLG* at 1.8:1) became insoluble and "clumpy" during dialysis, indicating that the histone:DNA ratio was too high. DNA was

therefore titrated with core histones and the highest ratio at which chromatin was found to remain in solution during dialysis at high concentrations was used for SAXS analysis.

For SAXS analysis, *601* chromatin was reconstituted at 1.2:1 (Experiment 1) or 1.12:1 (Experiment 2), *BLG* and *601/BLG* were reconstituted at 1.4:1 (Experiment 1) or 1.3:1 (Experiment 2). These samples were analysed by electron microscopy to confirm fibre saturation (see section 4.6). For Experiment 1, the average number of nucleosomes/fibre was found to be 23.18, 23.28 and 23.97 for *601*, *BLG* and *601/BLG* arrays respectively. For Experiment 2, the average number of nucleosomes/fibre was counted at 26.08 (n=37), 25.61 (n=36) and 25.16 (n=38) for *601, BLG* and *601/BLG* arrays respectively. In order to keep salt concentrations low to achieve the best SAXS results, chromatin was folded with a low concentration of magnesium ions, rather than 80mM NaCl. Chromatin in 2.5 mM NaCl was titrated with 0.5 mM or 1 mM $MgCl_2$ and/or a 1:1 ratio of linker histone H5. Dorigo et al. (2003) showed that "601" chromatin fibres achieved maximal compaction at 1 mM Mg^{2+}, and this remains unchanged up to Mg^{2+} concentrations of 100 mM. The scattering intensity (I) was measured at different scattering vectors (q) and the scattering of a matched buffer was subtracted to produce scattering intensity plots (Figure 4.7). Multiplying the intensity by q^2 to create a Kratky plot corrects for the random orientation of chromatin fibres relative to the x-ray beam that results in increased forward scattering (Langmore and Paulson, 1983).

In a Kratky plot (Figure 4.8), the scattering decay is multiplied out and peaks indicating structural features become more apparent. The shape of the scattering curve seems to indicate a multi-domain protein with a flexible linker in the absence of any folding agents (Figure 4.5B), shifting to a more globular structure upon the addition of magnesium or linker histones. Peaks in the data reflect periodicities within the shape of molecules in the sample. As described by Langmore and Paulson (1983), a periodicity of D Å in a sample is calculated by $2\pi/q$. A peak at 0.035 q would therefore indicate a periodicity (D) of 18 nm.

In analysing the scattering produced by unfolded chromatin fibres in experiment 1 (blue), a peak in seen in the *601* data that indicates a periodicity of 18 nm (q=0.035). This shifts slightly to 16-17 nm for the *BLG* and *601/BLG* (though the peak is not very clear in the *BLG*). I suggest that this peak is the result of the 50 bp linker DNA between the nucleosomes, which would be expected to measure approximately 17 nm. An 18 nm scattering peak is seen by Langmore and Paulson (1983) in nucleated erythrocytes, but not in other cell types. My data also show a relatively small peak indicating a distance of 57 nm (q=0.011) in the unfolded *601* data. In the *BLG* and *601/BLG* data a peak instead appears indicating a distance of 39-42 nm (q=0.015-0.016) and is relatively much larger than the 57

nm peak seen in the *601* data. Maeshima et al. (2016) also see a small shoulder in scattering intensity at approximately 40 nm, which becomes a true peak when chromatin is folded in 1 or 2.5 mM magnesium. This likely reflects a disordered but self-interacting structure in the presence of 2.5 mM NaCl, with more compacted folded forming in non-*601* fibres in the absence of linker histones or magnesium, which is consistent with these fibres having an increased sedimentation velocity relative to unfolded *601* fibres in 80 mM NaCl (section 4.4).

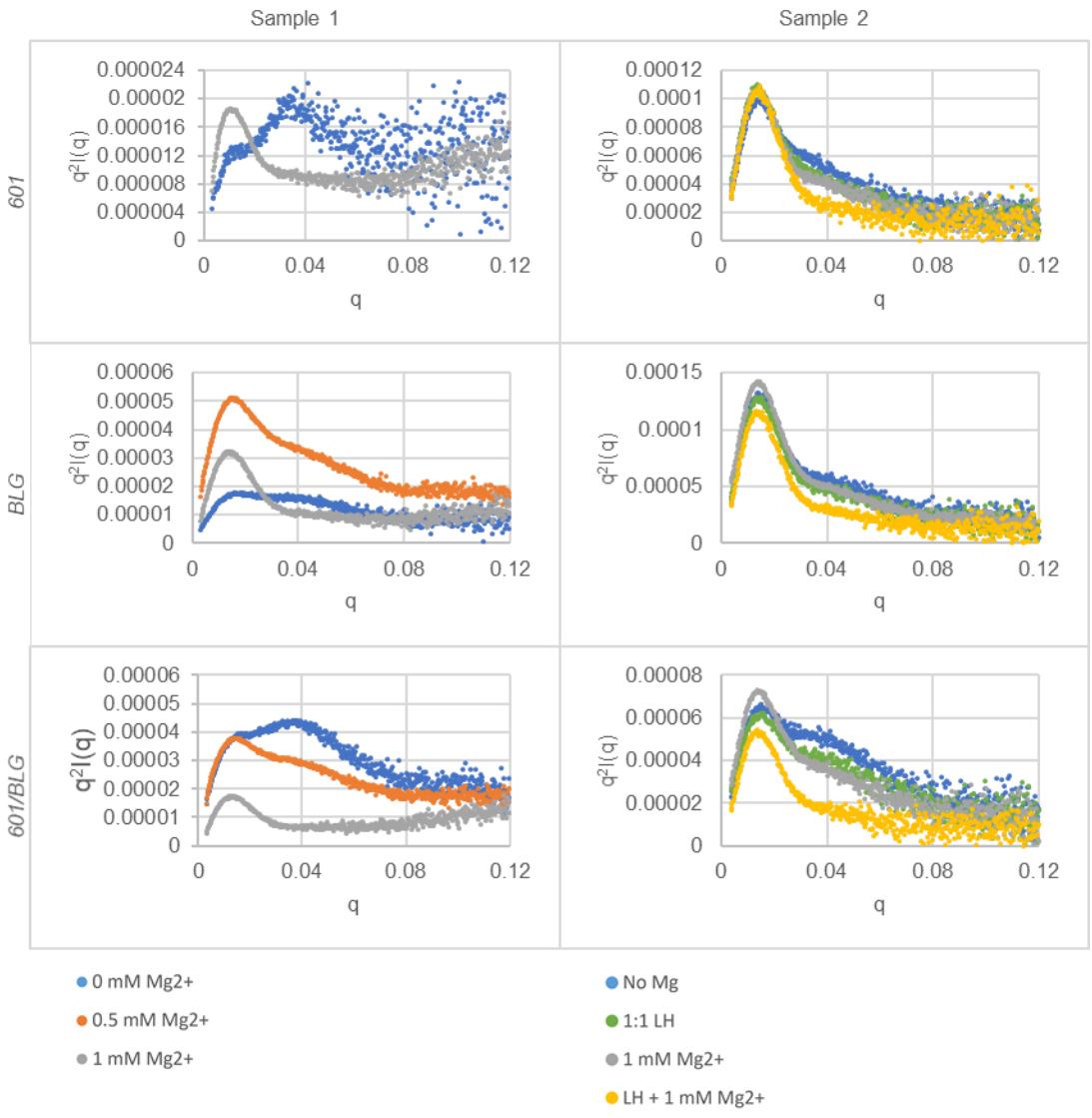

Figure 4.8. Kratky Plots of Scattering Data Derived from Three Fibres Titrated with Magnesium and Linker Histones.
Experiment 1 - left: *601* reconstituted at 1.2:1 and *BLG* and *601/BLG* reconstituted at 1.5:1, titrated with 0.5 or 1 mM $MgCl_2$. Unfortunately no data was collected for 601 at 0.5 mM $MgCl_2$. Experiment 2 - right: *601* reconstituted at 1.12:1 and *BLG* and *601/BLG* reconstituted at 1.3:1 titrated with 1 mM $MgCl_2$, 1:1 ratio of linker histones or both $MgCl_2$ and linker histones.

Upon folding with divalent magnesium cations (0.5 mM – orange, 1 mM – grey), there is a relative reduction in the 16-18 nm peak. If this indeed reflects the distance between nucleosomes in an unfolded fibre, this would be consistent with the compaction of the array. The samples compact into a more globular structure with only one dominant peak at 1 mM $MgCl_2$. In *601*, this peak represents a periodicity of 57 nm (q=0.011), in *BLG* this represents a periodicity of 45 nm (q=0.014), and in *601/BLG* this reflects a periodicity of 48 nm (q=0.013). Assuming, based on the sedimentation velocity of folded *601* fibres relative to *BLG* and *601/BLG*, that *601* fibres are becoming similarly compacted in the presence of linker histones to the *BLG* or *601/BLG* it is possible that these distances indicate differences in fibre shape, as more compacted fibres with longer linker lengths have been found to have larger diameters but shorter lengths *in vitro* (Robinson et al., 2006b). However, the folding conditions are different between the SAXS and sedimentation experiments described here, which may impact the structure of the different fibres.

While linker histones and magnesium ions have both been used to fold nucleosome arrays into higher-order fibres, there may be differences in the structures when different folding agents are used. Routh et al. (2008) find that the linker histone is essential for achieving complete compaction of "601" arrays with a 197 bp repeat. The 40 nm scattering peak seen by Maeshima et al. (2016) became more prominent upon addition of linker histone H1, but did not seem to shift in periodicity. To discover if this was true for our arrays, including those with varying underlying DNA sequences, chromatin was reconstituted at 1.12:1 (*601*) of 1.3:1 (*BLG* and *601/BLG*). While fibres in experiment 1 appear to be well saturated, SAXS analysis indicated that fibres became aggregated at 1 mM Mg^{2+} (Table 3), therefore slightly lower histone:DNA ratios were used in the repeat experiment. In experiment 2, there is much less evidence of nucleosome spacing before the addition of linker histones or magnesium. It is possible that these fibres are slightly undersaturated compared to experiment 1 and that there therefore is no discrete peak. The peak at q=0.015, thought to indicate a feature of nucleosome-nucleosome interaction in 2.5 mM NaCl, dominates and it is therefore not possible to analyse nucleosome spacing using the pearl necklace model in SasView. There is a peak in periodicity in all samples at approximately 42 nm, which was also seen in undersaturated fibres during trial experiments. It is not clear what this corresponds to. A "shoulder" to the right of the peak, particularly visible in the *601/BLG* sample, suggests a second periodicity comparable to that seen in unfolded fibres from experiment 1, but this was not prominent enough to be accurately measured. This shoulder was lost upon the addition of magnesium and linker histones.

Peaks at q=0.05 and q=0.1 would reflect a periodicity of 11 nm and 6 nm, which would reflect the dimensions of the diameter and height of the nucleosome respectively. These were seen by Maeshima et al. (2016) but are not apparent in my data.

Further analysis of SAXS experiments on folded chromatin fibres is limited, as the data in the Guernier region (low q), relating to the average size of particles in the sample, of intensity plots is of poor quality due to the attractive interactions between neighbouring chromatin molecules and possible aggregation within the sample. This prevents analysis of the precise average shape of molecules, for example calculating the radius of gyration and the maximum dimension of particles. Valuable information can still be gained from the high q region of the data, such as the Porod invariant which is dependent on particle volume (but not on its form, which the Rg is).

Sample		**Rg**	**Porod Volume**	**Porod Exponent**
601	No Magnesium	80.6	1481839	2.6
	0.5mM Mg^{2+}	-	-	-
	1mM Mg^{2+}	89.1	3662629	2.2
BLG	No Magnesium	61.3	904794	3.1
	0.5mM Mg^{2+}	45.5	449614	2.9
	1mM Mg^{2+}	242.3	17255855	2.5
601/BLG	No Magnesium	74	1235185	2.9
	0.5mM Mg^{2+}	47.7	513031	2.8
	1mM Mg^{2+}	221.0	19055389	2.1

Table 3. Porod Analysis of 601, BLG and 601/BLG fibres at Varying Mg2+ Concentrations.
Radius of gyration (nm) (automatically estimated by Scatter software) Porod Volume ($Å^3$) and Porod exponent of each sample in experiment 1.

The Porod exponent, which is related to the average particle flexibility and compaction and average particle volume are displayed in table 3. A higher Porod exponent indicates a more compacted particle. In the absence of magnesium ions, the *601* fibres appear to occupy the largest volume and are the most flexible. This is consistent with the observation that they have the lowest sedimentation velocity, as they are the least compact of the three fibres. In contrast, the *BLG* is the most compact and least flexible by SAXS measurement, although *601/BLG* was found to have the highest sedimentation velocity in an unfolded state (Figure 4.3), although the fibres examined by sucrose gradient sedimentation are in a higher monovalent salt concentration.

Upon the addition of 0.5 mM magnesium chloride to induce the folding of arrays into higher-order structures, the estimated radius of gyration of the *BLG* is reduced by 26% and that of the *601/BLG* is reduced by 35.5%. The average volume of the *BLG* particles concurrently reduces by 50.3% while that of the *601/BLG* molecules reduces by 58.5%. However, in 0.5 mM Mg^{2+}, *BLG* fibres maintain a smaller volume and Rg than *601/BLG* fibres. Unfortunately, the quality of the data for *601* fibres in 0.5 mM $MgCl_2$ was insufficient for analysis. At 1 mM MgCl, all three fibres seem to greatly increase in volume and Rg. This suggests that the samples are biased by large aggregates forming in the presence of this level of Mg^{2+}. This is more prominent for the *BLG* and *601/BLG* fibres than for the *601*, consistent with data shown in Figure 4.2, that these heterogeneous fibres become insoluble in Mg^{2+} more easily than the relatively homogeneous *601*.

4.6 Electron Microscopy of Chromatin Fibres

In order to analyse the level of nucleosome saturation and to get a clearer idea of the structure of folded higher-order fibres, electron microscopy was performed on these samples before and after the folding of the arrays.

The number of nucleosomes reconstituted onto arrays was measured by fixing chromatin in 10 mM NaCl and counting nucleosomes (section 3.6). Typical images of chromatin fibres reconstituted in the absence of competitor as described in section 4.6 and used for SAXS analysis are shown in Figure 4.9A. Unfortunately the density of fibres seen during imaging was relatively low, and more fibres would be required for a more thorough analysis. The average number of nucleosomes counted on each fibre in experiment 1 was 23.18±1.02 (n=31) for *601*, 23.28±0.77 (n=38) for *BLG* and 23.97±0.72 (n=38) for *601/BLG*. In experiment 2, the average number of nucleosomes counted within each sample was 26.08±0.58 (n=37) for *601*, 25.61±0.71 (n=36) for *BLG*, and 25.16±0.94 (n=38) for *601/BLG* (Figure 4.9A and 4.9C). As discussed in section 3.6, while a similar number of nucleosomes counted might indicate a similar degree of saturation for each of the different templates within each sample, there may in fact be more nucleosomes per fibre within sample 1, as fewer nucleosomes might be counted on fully saturated fibres due to the effects of nucleosome overlapping in electron micrographs on the qualification of nucleosomes within these images.

These samples were fixed in the presence (1 mM, 0.5 mM) or absence of magnesium chloride, or 1 mM magnesium chloride with a 1:1 ratio of linker histone H5 to induce varying degrees of folding of the chromatin fibres as described (section 4.5). Samples were fixed at low chromatin concentrations (25 ng/µl) in contrast to SAXS experiments which are

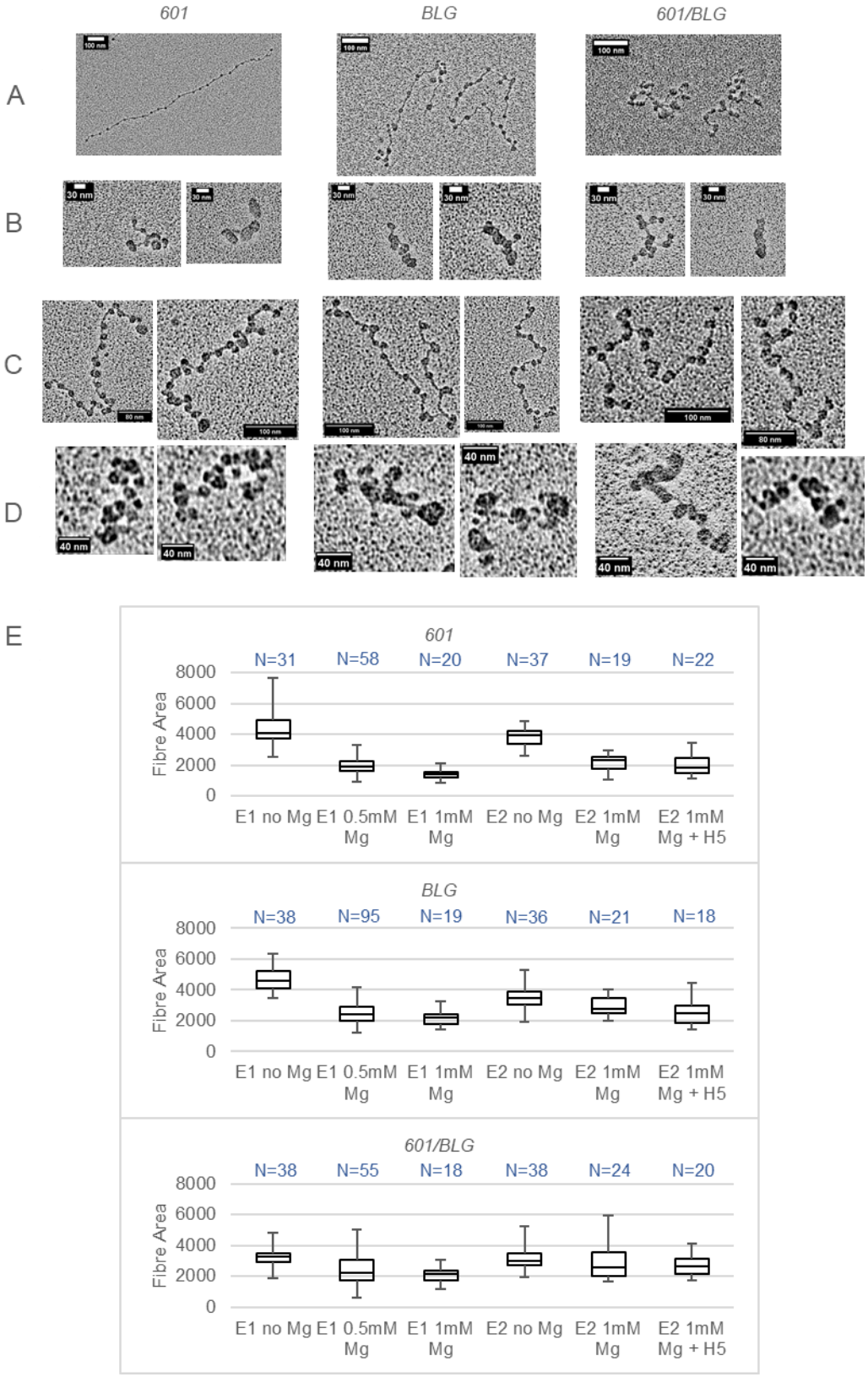

Figure 4.9. Electron Microscopy of Fibres Titrated with Linker Histones and Magnesium Ions.
(Full legend found on following page)

A) Experiment 1 in the absence of magnesium. 601 reconstituted at 1.2:1, 23.18 nucleosomes/fibre (N=31). BLG reconstituted at 1.5:1, 23.28 nucleosomes/fibre (N=38). 601/BLG reconstituted at 1.5:1, 23.97 nucleosomes/fibre (N=38). B) Experiment 1 arrays in the presence of 1 mM MgCl2. C) Experiment 2 in the absence of magnesium and linker histones. 601 reconstituted at 1.12:1, 26.08 nucleosomes/fibre (N=37). BLG reconstituted at 1.3:1, 25.61 nucleosomes/fibre (N=36). 601/BLG reconstituted at 1.3:1, 25.16 nucleosomes/fibre (N=38). D) Experiment 2 folded in the presence of 1 mM MgCl2 with a 1:1 ratio of linker histones. E) Measurement of fibre area from EM images

performed at very high chromatin concentrations, which may impact fibre structure. The area occupied by individual chromatin fibres in EM images at 20,000 × magnification was measured using an algorithm written by Davide Michieletto (Figure 4.9E). In the absence of magnesium, the area was measured was 4102 pixels for *601* fibres, 4591 for *BLG* fibres and 3287 for *601/BLG* fibres. The area occupied by the fibres in EM images might be expected to correlate with fibre volume, so it is surprising that the area occupied by *BLG* fibres is greater than that occupied by *601* or *601/BLG* fibres, as the *BLG* fibres appeared to have the smallest volume measured by SAXS (Table 3). Upon addition of magnesium to 0.5 mM, the area occupied by *601* fibres decreases by 53.6%, that of *BLG* fibres decreases by 48.4% and that of *601/BLG* fibres decreases by 33%. Upon addition of 1 mM $MgCl_2$, the area decreases compared to the no-magnesium control by 66.2% in the *601* sample, 52.8% in the *BLG* sample and by 35.7% in the *601/BLG* sample. The lower concentration of chromatin fibres in this experiment likely prevented the aggregation of fibres at 1 mM magnesium chloride that was observed during SAXS experiments. While the *601* fibres seem to compact significantly further between 0.5 mM and 1 mM $MgCl_2$ ($p < 5 \times 10^{-7}$), this difference is less significant in the *BLG* and *601/BLG* samples ($p < 7 \times 10^{-3}$ and $p < 6 \times 10^{-2}$ respectively). While the variation in the areas of the population of fibres is reduced at 1 mM Mg^{2+}, the average value reduces by only a small amount, which might suggest that the ability of these fibres to further compact is limited compared to the *601* (though they are more likely to become insoluble as described in section 4.3).

4.7 Chromatin Unfolding Dynamics by Single-molecule Force Spectroscopy

Experimental methods such as sucrose gradient sedimentation and small angle x-ray scattering provide valuable information about the average properties of a population of chromatin fibres. The heterogeneity in the number of nucleosomes within each chromatin sample however, including the *601* samples makes it difficult to analyse chromatin structure using population-wide approach. For example, when two peaks are observed in SAXS analysis, it is unknown whether each fibre contains structures that cause the two periodicities

measured, or whether there is a population of fibres (for example, undersaturated fibres) causing one peak, and another population (for example, properly saturated fibres) causing another.

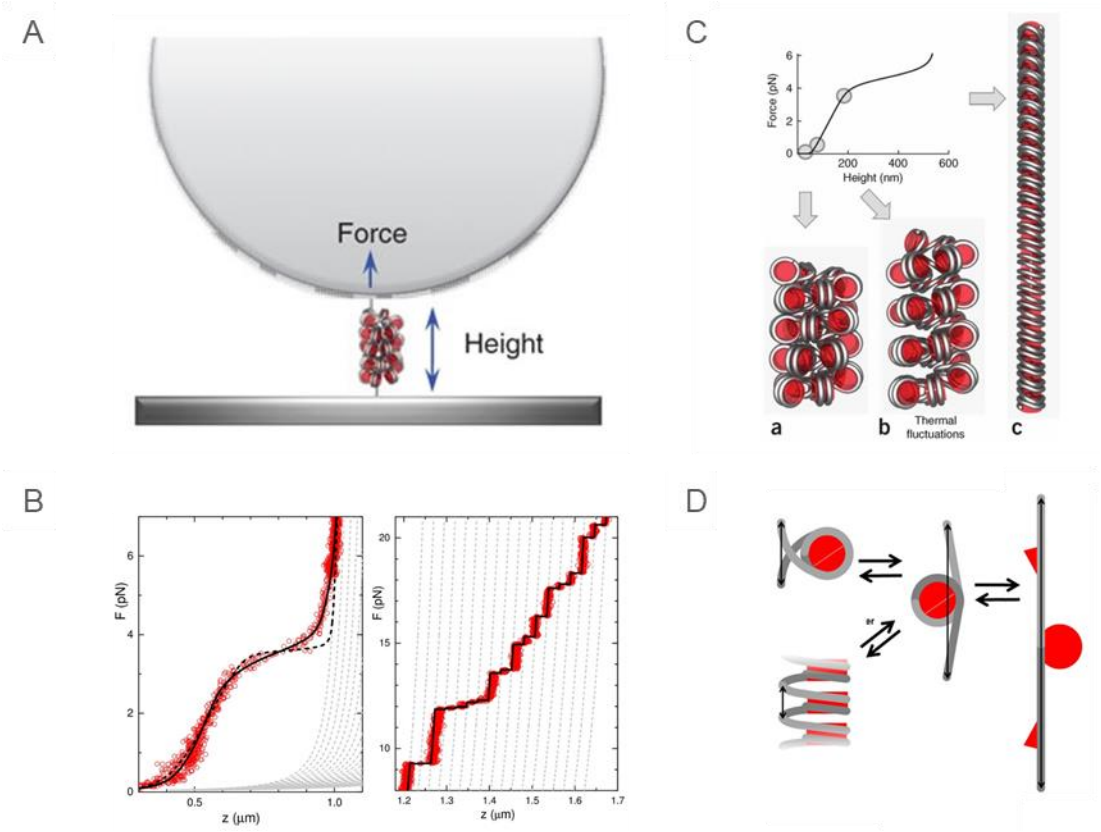

Figure 4.10. Magnetic Tweezers for Single Molecule Force Spectroscopy.
A) Schematic of magnetic bead pulling on individual chromatin fibre (Kruithof et al. 2009). B) Low force and high force regions of "601" fibre unwrapping under force. At low forces (left) the 30-nm fibres come apart and one turn of DNA (56 bp) unwraps from each nucleosome under approximately 3.5 pN of tension. At high forces (right), the second turn of DNA unwraps stochastically from each nucleosome, causing "steps" of extension to appear in the force-extension curve (Kruithof et al. 2009). C) Suggested unwrapping path of the 30-nm chromatin fibre. Thermal fluctuations at low forces lead to the unwrapping of the solenoid at forces lower than 3.5 pN, though nucleosomes are still interacting at this stage (Meng et al., 2015). D) The unwrapping process between a higher-order fibre (or a single nucleosome) and a chain of nucleosomes with a single DNA wrap (low force transition) is reversible. The high force transition (DNA completely unwrapping from the octamer) is irreversible (Meng et al., 2015).

Single molecule studies have allowed the study of defined particles, where the characteristics of one molecule can be examined without bias or interference from the surrounding population. Single molecule methods are particularly useful in studying populations of heterogeneous samples where one parameter, in this case the number of nucleosomes per chromatin fibre, is likely to affect the average patterns when examining an entire sample.

While electron microscopy can be used to study the shape of individual chromatin fibres, it is not possible to analyse both the folded structure and the degree of saturation for one single chromatin molecule, as the samples are fixed.

Single-molecule force spectroscopy, using either magnetic or optical tweezers, has allowed the analysis of the dynamics of individual chromatin fibres under tension (Chien and van Noort, 2009). To assess chromatin structure using magnetic tweezers, a magnetic bead is attached to one end of a DNA molecule, which is secured to an immobile surface at its other end (Figure 4.10A). By measuring the change in height of the bead as force is applied to pull the bead upwards, it is possible to determine the unfolding dynamics of single nucleosome arrays. Under such tension, it is possible to identify fibres upon which are reconstituted the correct number of nucleosomes, and to limit the analysis of chromatin structure and dynamics to these fibres. The variable number of nucleosomes per fibre is therefore controlled between the three DNA sequences in this experiment.

Kruithof et al. (2009) described the dynamics of the unfolding and unwrapping of individual chromatin fibres under tension using magnetic tweezers. The first level of unwrapping (at 3.5 pN) is in equilibrium and is reversible (Figure 4.10B, left). This extension reflects the unwrapping of the higher-order chromatin fibre structure (Figure 4.10C) and the unwrapping of one turn of DNA (56 bp) from each nucleosome within the chromatin fibre. If chromatin is not pulled at forces higher than ~6 pN (ie. if no histones octamers are removed) then the fibre will refold by following the same force-extension path as the force is removed. If pulled again, it will produce the same force-extension dynamics. Following this, there is a small intermediate unwrapping phase causing an extension of approximately 5 nm per nucleosome, which possibly reflects the dissociation of H2A/H2B dimers from the H3/H4 tetramer. At forces above 6 pN (Figure 4.10B, right) step-like rupture events occur as individual nucleosomes become completely unwrapped and histones become dissociated from DNA. The removal of histones octamers at forces above ~6 pN is stochastic and irreversible, and once histones have been removed the chromatin fibre will not refold. However, these rupture events and the length of fibre extension at forces above ~6 pN allow the measurement of fibre saturation for individual chromatin fibres. Upon release of the chromatin fibre, the force-extension curve returns to 0 nm extension by following the path of a worm-like chain, in the manner of a bare DNA molecule.

The forces used in this experiment to unfold chromatin fibres are highly physiologically and relevant to nuclear function. RNA Polymerase has been shown using optical traps to exert a

force of 14-25 pN on DNA (Wang et al., 1998; Yin et al., 1995) while other proteins likely exert a smaller force.

4.7.1 Sample Preparation

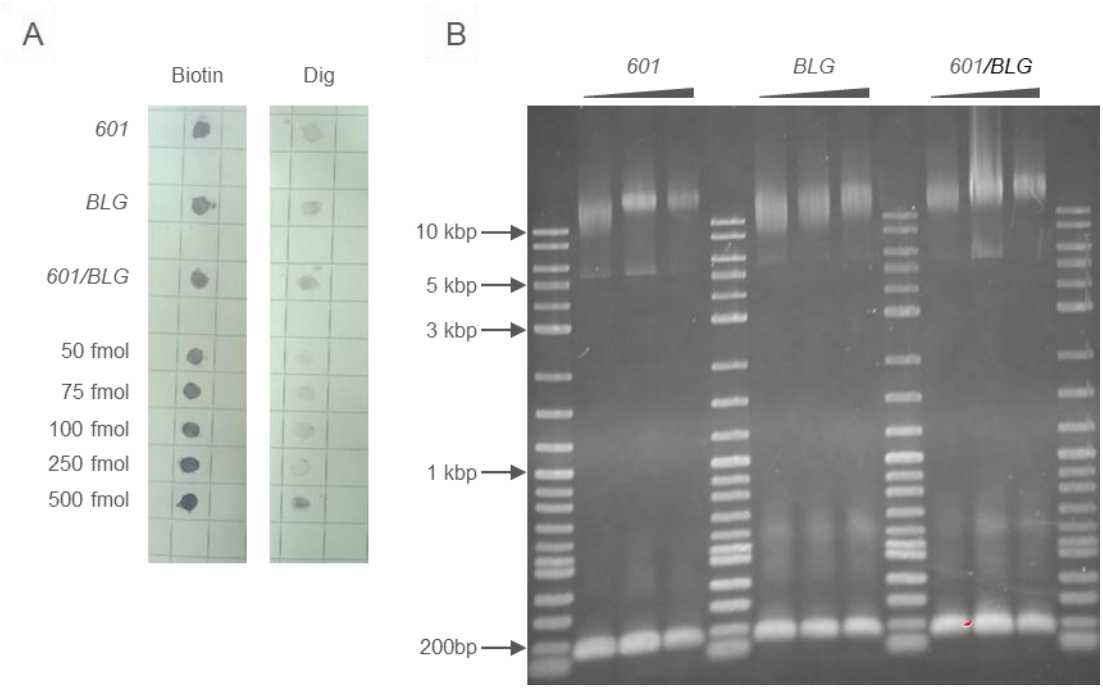

Figure 4.11. Sample Preparation for Single-molecule Force Spectroscopy.
A) Dot blot showing the labelling of DNA templates with biotin and digoxigenin, compared to standards. B) EMSA confirming reconstitution of DNA templates with recombinant *Xenopus* core histones and 147 bp monomer competitor. *601* reconstituted at 1.3:1, 1.5:1 and 1.7:1, appears to be reconstituted by 1.5:1. *BLG* and *601/BLG* reconstituted at 1.8:1, 2.0:1 and 2.2:1. *601/BLG* appears to be reconstituted by 2.0:1, *BLG* does not reach a plateau in its shift.

Plasmid DNA containing *601*, *BLG* or *601/BLG* templates (see appendix 1 for vector maps) was digested by *Pci*I and *Nde*I, rather than *Eco*RV or *Xho*I, creating template DNA fragments that have ~600 bp of additional DNA "handles" on the ends of each of the 5 kbp templates allowing the chromatin fibres to attach to the glass slide and the magnetic bead. In the manner of competitor DNA, on a *601* template these "handles" might be expected to remain unreconstituted when a nucleosome array is formed. However, due to the larger amount of histone octamers required to saturate *BLG* and *601/BLG* templates in the presence of competitor DNA (section 3.5), these "handles" might form nucleosomes on these templates. *Pci*I and *Nde*I restriction sites were used to label each DNA template with biotin-dUTP at one end and digoxigenin-dUTP at the other in order to attach the chromatin fibre to the microscope slide and to the magnetic bead. Biotin and digoxigenin labelling was confirmed by dot blot (Figure 4.11A). 115 ng of DNA was spotted on the membrane, this

corresponds to approximately 30 fmols of each template. Densitometry suggests that on the biotin dot blot, the intensity of the dot is approximately the same as the 50 fmol standards, but for the digoxigenin, each sample appears to be between 250 and 500 fmol, approximately 10 times higher than the amount of DNA on the dot blot. It is possible that these standards have not worked well (especially the digoxigenin, which has very low intensity dots until 500 fmol), but it appears that the templates are well-labelled with both biotin and digoxigenin.

Template DNA was separated from the vector backbone and equal mass of a 147 bp DNA molecule was instead used as competitor. Chromatin was reconstituted at 1.3:1, 1.5:1 and 1.7:1 (*601*) or 1.8:1, 2:1 or 2.2:1 (*BLG* and *601/BLG*) histone:DNA ratios using a one-step salt gradient dialysis (section 2.2.2). Chromatin samples were analysed by EMSA (Figure 4.11B), which suggested that *601* chromatin reached a plateau in its migration, indicating complete saturation, at a ratio of 1.5:1, *601/BLG* appears to plateau at a 2:1 ratio, but *BLG* chromatin samples do not appear to plateau.

Following reconstitution, chromatin was diluted into ESB(+) buffer, containing 100 mM NaCl and 2 mM $MgCl_2$. Kaczmarczyk et al. (2017) showed that the H4 tails and H2A acidic patches of neighbouring nucleosomes in *601* fibres under these conditions are interacting, indicating that the nucleosome arrays are forming higher-order folded structures.

4.7.2 Chromatin Unfolding Dynamics vary between Different DNA Sequences

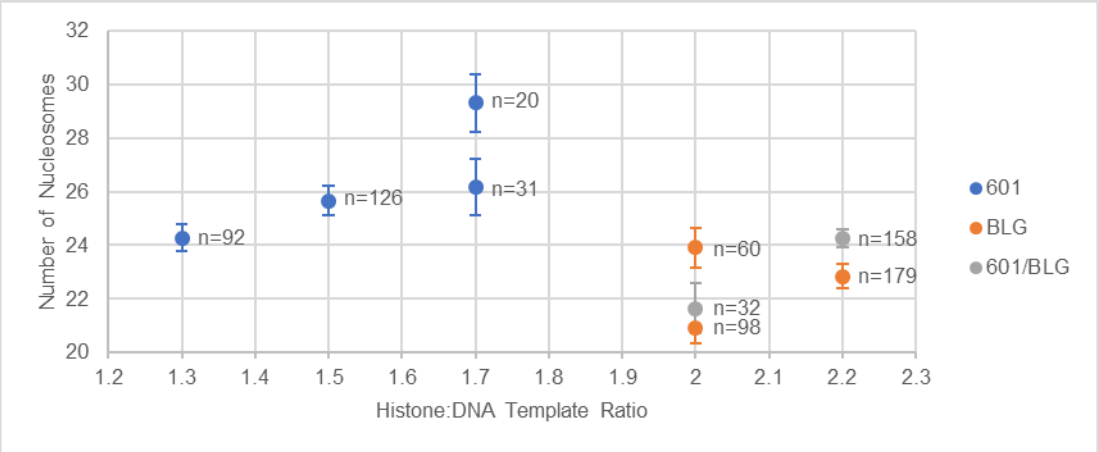

Figure 4.12. Number of Nucleosomes Counted on each Chromatin Fibre.
Graph showing the number of nucleosomes, counted by measuring the fibre extension at forces between 6 and 50 pN. +/- SEM.

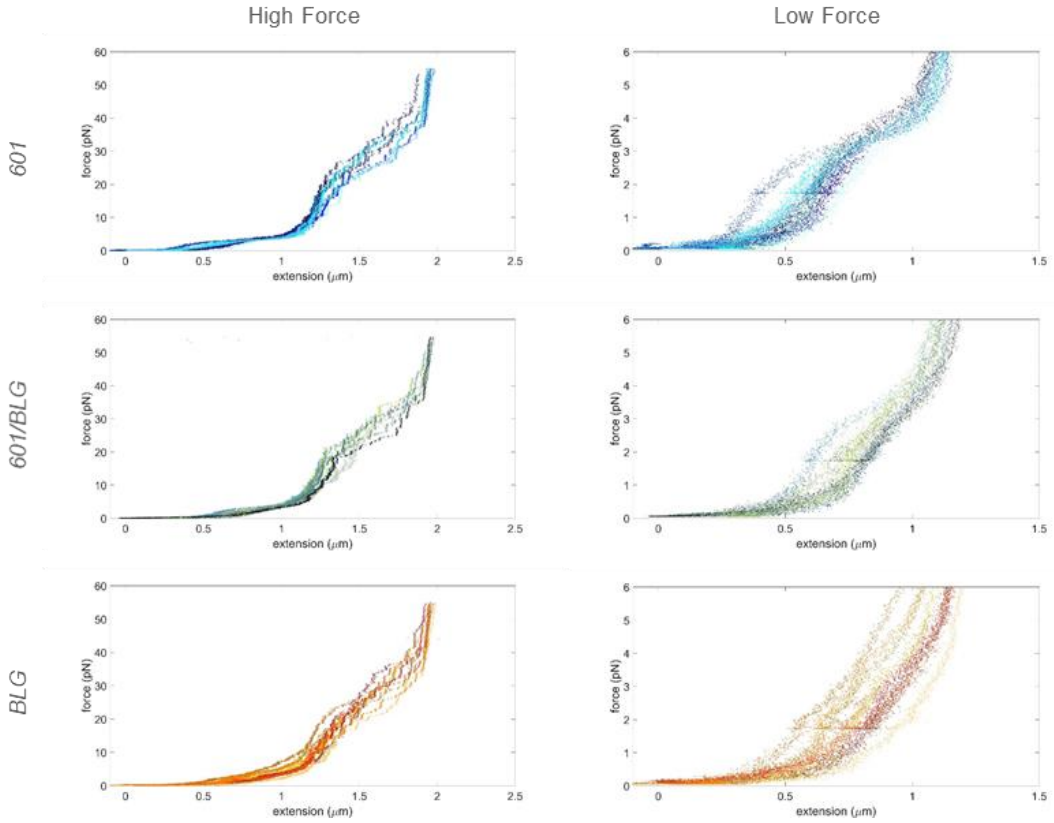

Figure 4.13. Force-extension Curves Showing Fibre Unwrapping of 601, BLG and 601/BLG Arrays.
Shown at a range of forces up to 50 pN (left) and zoom in to show low force region (right). 10 individual fibres are shown in each graph.

Chromatin is added to flow cells covered by a glass slide coated in anti-digoxigenin. The digoxigenin-labelled end of the chromatin fibre binds to the base of the flow cell, and streptavidin-coated magnetic beads are added and bind to the biotin label at the opposite end of the chromatin fibre. Magnetic tweezers pull on the magnetic bead and the change in bead height is measured as the force changes, allowing the unfolding of a chromatin fibre to be examined (Figure 4.10A). Multiplexed tweezers were able to measure the pulling of hundreds of beads at one time, increasing the throughput of the technique.

Firstly, the high-force regions of individual force-extension curves (Figure 4.13, left) were examined to calculate the number of nucleosomes within each fibre. The extension and the number of "steps" counted between ~6 and ~50 pN were analysed to determine the number of individual nucleosomes unwrapping (Figure 4.12). At a 1.3:1 ratio, the average number of nucleosomes loaded on a *601* fibre was slightly below 25 (n=92), whereas at 1.5:1 (n=126), there were slightly more than 25 nucleosomes per fibre, confirming that a 1.4:1 ratio is the correct saturating ratio for fibres reconstituted at these concentrations (20 ng/µl). The *BLG* and *601/BLG* appear to be slightly undersaturated, with an average of 23-24

nucleosomes at a 2.2:1 ratio (n=179 and n=158 respectively). It is unsurprising that slightly more histones might be required to fully saturate the fibre in the presence of a monomer competitor, as the amount of competitor DNA is higher than when the vector backbone was used. Sufficient fibres with 25 nucleosomes were identified for subsequent analysis.

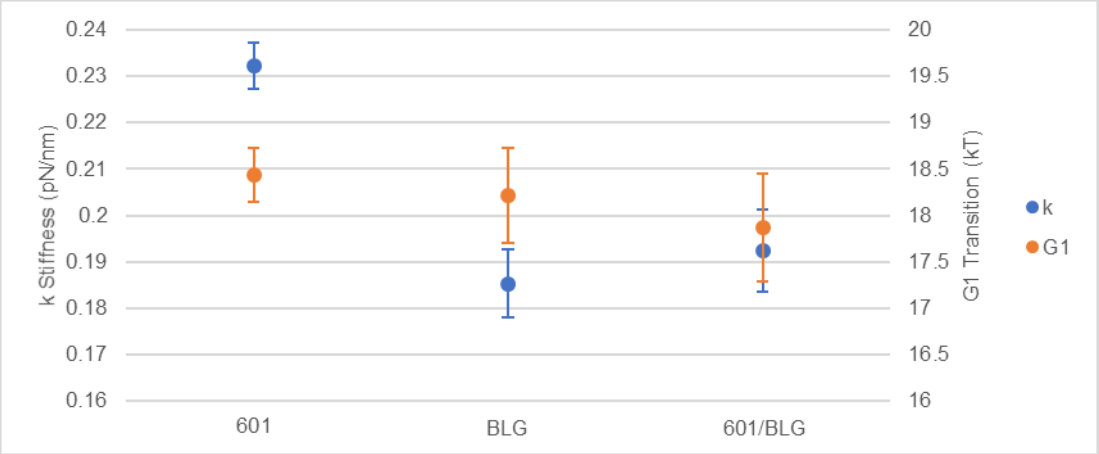

Figure 4.14. Analysis of Force-Extension Curves.
Fitting to the model described by Kruithof et al. (2009) allows various parameters to be calculated. The stiffness of each fibre (k) and ΔG of the first force transition are shown on the left and right axis, respectively. All fibres analysed have 24-27 nucleosomes. +/-SEM. 48 *601* fibres were analysed, 44 *BLG* fibres were analysed and 24 *601/BLG* fibres were analysed.

The unfolding patterns of 10 fibres containing 25 or 26 nucleosomes are shown in Figure 4.13. Fibres containing 26 nucleosomes are expected to have formed a nucleosome over one of the DNA handles that allow the labelling of the template by biotin or digoxigenein, and are therefore not oversaturated. The entire force-extension curve, including the high force region which is used to count the number of nucleosomes is shown on the left, then this is zoomed in to examine the low force regime, the unfolding of the higher order fibre and first turn of DNA from each nucleosome, more clearly on the right. There are clear differences in the unfolding dynamics at low forces between the *601* and the other two fibres. The *601*, as has been previously described, undergoes an initial extension at very low forces (<0.5 pN) as the DNA handles at the ends of the template become extended. At approximately 3-4 pN, a plateau of extension occurs, indicating the unfolding of the higher-order fibre and the first turn of DNA beginning to unwrap from each of the individual nucleosomes. In *BLG* and *601/BLG* fibres, there is very little evidence of an extension plateau at 3.5 pN, but there appears to be much more of a gradual extension in the length of the fibre between 0.4 and 1.0 µm at forces between 1 and 4 pN. Some fibres in the *601/BLG* sample have a short plateau, but others lack this. These differences can be described by fitting the force-extension curves to the force-extension model developed by the van Noort laboratory and

described by Meng et al. (2015). This allows us to fit various parameters of the DNA and chromatin fibres (section 2.9.5 and Meng et al. (2015)). The stiffness (k), number of fully folded nucleosomes, and three force transitions of the fibre (kT) were manually fitted to this model for fibres containing 24-27 nucleosomes. The primary difference between the three samples was found to be in the stiffness of each fibre, which measures the rate at which the fibre is pulled apart, breaking the interactions between nucleosomes and unwrapping the DNA from the octamer, in pN/nm.

The *601* sample has an average stiffness of 0.232±0.035 pN/nm (n=48), which is comparable to what has been previously published by Meng et al. (2015). As no sequence apart from the *601* has been examined using this apparatus before, it is unknown what differences might be expected of other DNA sequences. The *BLG* was found to have an average stiffness of 0.185±0.048 pN/nm (n=44), and that of the *601/BLG* was slightly higher at 0.193±0.043 pN/nm (n=24) (Figure 4.14, left axis). The ΔG of the first force transition (the force at which the unfolding of the higher-order fibre and unwrapping of the first turn of DNA occurs) was found to be very similar between the three fibres (Figure 4.14, right axis), but the differences between the samples were found to be entirely captured by the stiffness. The number of unfolded nucleosomes in each sample was also found to be the same on average between the three samples. These unfolded nucleosomes have been suggested to be tetrasomes, or nucleosomes with dissociated dimers. The van Noort group has found that these almost always need to be accounted for when analysing data from 197 bp repeat sequences, but not with 167 bp sequences.

This assay has limitations; I cannot be sure whether the lower force required to extend non-601 chromatin fibres is a result of weaker or absent interactions between nucleosomes within a higher-order chromatin fibre structure, or because the first turn of DNA becomes more easily unwrapped from the histone octamer. Experiments were attempted in the absence of magnesium, in order to compare the dynamics of an unfolded nucleosome array with a compacted fibre. However, this experiment did not seem to be possible under these conditions, as long, unfolded nucleosome arrays appear to “stick” to the surface of the slide and unstick several nucleosomes at a time, causing jumps in the force/extension curve. This means that a smooth unfolding trace cannot be seen in these curves as it can in the presence of magnesium, so it was not possible to dissect these two aspects of the chromatin fibre structure.

4.7.3 Individual Nucleosome Stability is Influenced by DNA Sequence

The step-size reflecting the unwrapping of a second turn of DNA from each individual nucleosome at forces between 6 and 40 pN was approximately 79 bp for *601* and *601/BLG*, consistent with results previously reported for the *601* (Kaczmarczyk et al., 2017), but increased to 84 bp for *BLG* indicating that more DNA may remain interacting with the H3/H4 tetramer following the unwrap of one turn of DNA from these fibres (however, the accuracy of the measurement is smaller than this distance, so it is not possible to draw any strong conclusions). The height that these steps occur, i.e. the force at which the DNA becomes fully dissociated from the histone octamer, varies between the samples. Artur Kazmarckyk analysed the forces at which these steps occur in saturated chromatin fibres and found that the cumulative probability of rupture was lowest for the *601* and highest for the *BLG* at any given force between ~6 and ~40 pN (Figure 4.15). This indicates that DNA unwraps slightly more easily from the sequences within the *BLG* fibre. If this phenomenon also applies to the unwrapping of the first turn of DNA at forces below 4 pN, this might contribute to the reduced stiffness of the non-*601* chromatin fibres described above.

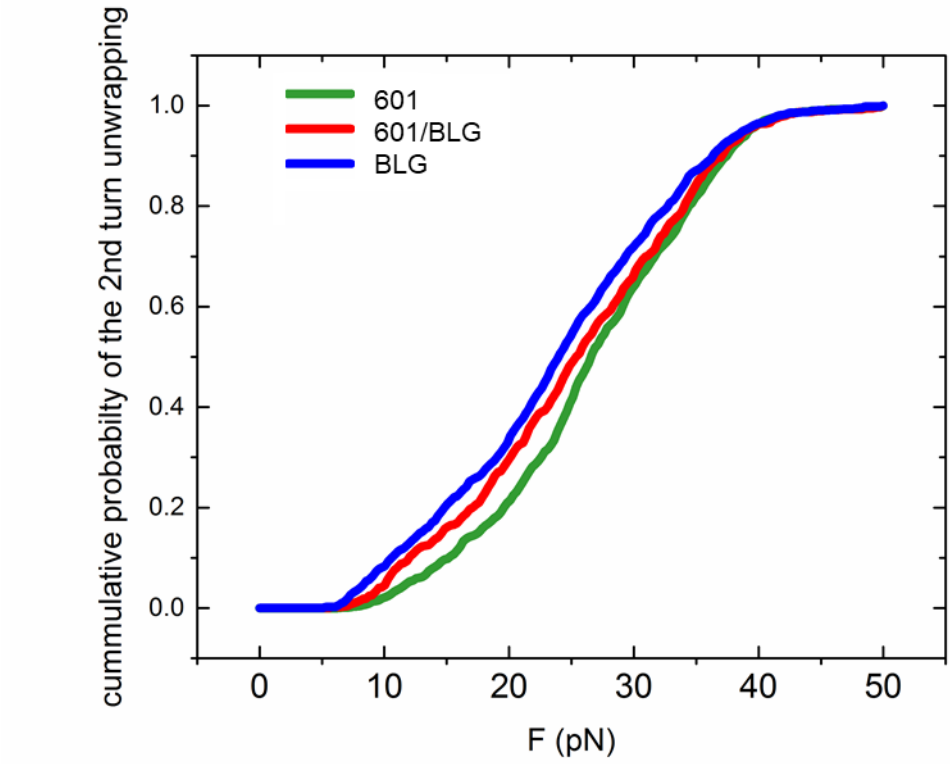

Figure 4.15. Step Distribution Analysis at High Forces.
The probability of individual nucleosome rupture was calculated from a selection of well-saturated fibres. BLG nucleosomes were more likely to rupture at lower forces than 601 nucleosomes.

4.7.4 Chromatin Unfolding, but not Folding, is Affected by Neighbouring Chromatin Fibres.

Labelled chromatin fibres must be diluted to extremely low concentrations (0.04 ng/µl) to perform single molecule force spectroscopy, as a high concentration of labelled fibres will enable too many magnetic beads to bind and cause "double tethers", where two or more chromatin fibres bind to the same streptavidin-coated magnetic bead, affecting the force-extension dynamics of the whole complex. However, chromatin concentration is known to affect fibre stability; it is possible that chromatin fibres might fall apart more easily at low concentrations.

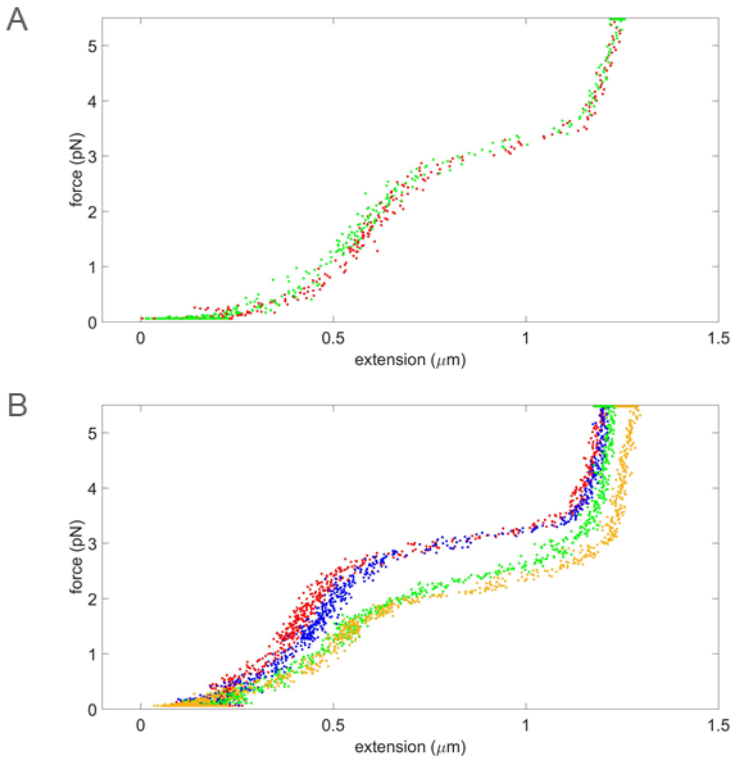

Figure 4.16. Refolding Chromatin at Different Concentrations.
A) Unfolding (red) and refolding (green) of a 601 fibre analysed at low chromatin concentrations. B) Unfolding (red and blue) and refolding (green and orange) of a 601 nucleosome array pulled twice at high concentration.

To test this additional unlabelled *601* chromatin fibres were included in the buffer during pulling, increasing the total chromatin concentration to 4 ng/µl (100 fold higher than previously described). Chromatin fibres were pulled at low force, which at low chromatin concentrations allow the fibre to completely refold, following the same force-extension path as its unfolding (Figure 4.16A). When additional, unlabelled chromatin was added to the flow cell, there was not a quantifiable difference in unfolding dynamics (though high force

pulls could not be used to confirm complete saturation), but the refolding pathway was altered in several cases, with chromatin not refolding until lower forces, around 2.5 pN had been reached. The chromatin fibre analysed in Figure 4.16B requires a force of approximately 3.5 pN to unfold (red), but as the tension is decreased it does not refold until the force decreases to approximately 2.5 pN (green). When this force is reapplied to this fibre it follows the same unfolding pathway (blue), suggesting that the structure of the fibre has not changed, but again does not refold until the force is decreased. These experiments were performed before multiplexed tweezer setups were available, and therefore it was not possible to confirm saturation of these fibres by high-force measurements or to measure a significant number of fibres for a complete analysis. While it was not possible to pull these fibres at high force to count the number of nucleosomes, the low force region of the curves shown for this fibre fits with a model that may have 24-27 nucleosomes. Curves that appeared to be undersaturated with nucleosomes (indicated by a shorted plateau at 3.5 pN) appeared to have a smaller difference between the unfolding and refolding pathways.

McDowall et al. (1986) suggested that adjacent chromatin fibres in the nucleus (where concentration is extremely high) are interdigitated with one another. Maeshima et al. (2016) have more recently suggested a polymer-melt structure, lacking a 30-nm fibre, where chromatin interacts with neighbouring fibres in the nucleus. This interaction would be highly likely to influence folding dynamics of chromatin.

4.8 Summary

Surprisingly my results suggest that although *BLG* and *601/BLG* arrays in 80 mM NaCl are more heterogenous they have a more compacted structure than *601* fibres. Under these ionic conditions, nucleosomes might be expected to be interacting within arrays, but not to form a 30-nm structure in the absence of linker histones. The heterogeneity of nucleosome positioning on the *BLG* and *601/BLG* templates may cause different types of folding, eg. between very closely positioned nucleosomes, or within longer nucleosome-free regions of DNA, allowing nucleosomes to interact more easily than within a *601* fibre, leading to a more compact structure as seen in their higher sedimentation velocity (Figure 4.3) and their smaller volume as indicated by SAXS analysis (Table 3). However, when analysed by electron microscopy, *BLG* arrays were found to occupy a larger area, suggesting a larger volume, than *601* fibres. This might indicate that *BLG* fibres are more likely to become decompacted at low concentrations than *601* or *601/BLG* fibres.

Upon the addition of linker histones or Mg^{2+} ions to induce folding of nucleosome arrays into higher-order fibres, *601* fibres appear to become more compact as evidenced by their

increased sedimentation velocity in the presence of linker histones (Figure 4.3) and their reduced fibre area measured by electron microscopy in the presence of Mg^{2+} (Figure 4.9). While non-*601* arrays become more compacted in low levels of magnesium, they appear to become insoluble more easily in the presence of magnesium or linker histones (Figure 4.2, Table 3), preventing them from compacting more extensively. This may be due to irregular spacing of nucleosomes; possibly there are sections of the template where nucleosomes space very close together with short linkers between, and these could promote array aggregation.

As it is impossible to control for fibre saturation during reconstitution of the three samples, there are likely to be small variances in array saturation that impact fibre folding. In addition, it is difficult to dissect the structural features of an average population of fibres, which may contain different folded structures.

Single-molecule force spectroscopy allows the folding dynamics of individual chromatin fibres of known saturation to be studied. *601* fibres were found to have very similar unfolding dynamics to those previously described (Figure 4.13, 4.14) with interactions breaking between nucleosomes and the first of turn of DNA becoming disassociated with the histone octamer at approximately 3.5 pN. *BLG* and *601/BLG* fibres were found to unfold more gradually, which might reflect the heterogeneity of the nucleosomes within the structures and their reduced stability. The second turn of DNA unwrapping occurred at lower forces in the *BLG* compared to the *601*, further evidencing the reduced stability of these nucleosomes under tension (Figure 4.15).

The main difficulties found with these experiments were in confirming nucleosome array saturation prior to folding. Without being sure that arrays are equally saturated, it is impossible to be confident that all differences are a product of the underlying DNA sequence, as missing or additional histone octamers are likely to cause changes in the folding of the fibre. While single-molecule experiments have successfully allowed us to study the dynamics of fibres which have a known number of nucleosomes, a method to study DNA sequence effects on the average fibre structure where the nucleosome array saturation could be internally controlled is necessary.

Chapter 5. Introducing Sequence Disruptions into Saturated Fibres

5.1 Introduction

While comparing the properties of *601*, *BLG* and *601/BLG*, the variable affinities of the 25 unique nucleosome positioning cause a variable and heterogeneous fibre (Chapter 4) that makes it difficult to attribute changes in fibre folding to a specific DNA sequence, as the surrounding chromatin landscape is likely to be different in each fibre within a sample.

A different approach is therefore required to study the properties of any particular DNA sequence. To address this and to study the impacts of different DNA sequences within a controlled chromatin landscape, novel templates were devised in which a single nucleosome positioning site varies within a stable, homogeneous "601" template.

Varying the DNA sequence of a single nucleosome positioning site might cause a point of disruption within an otherwise compacted chromatin fibre. Chromatin from the active β-globin locus of the chicken erythrocyte genome sediments more slowly through a sucrose gradient than bulk chromatin, due to a single nucleosome disruption (Caplan et al., 1987; Fisher and Felsenfeld, 1986; Kimura et al., 1983). Caplan et al. (1987) showed that a 6.2 kbp fragment (containing 29.2 nucleosomes on average) of the β-globin locus sedimented at approximately the same rate following the removal of linker histones. However, following the refolding of the fibres with linker histones this difference again became apparent, suggesting that the difference in the sedimentation velocity of the fibres is due to a single unoccupied nucleosome site which cannot be accommodated within the higher-order fibre structure, causing a disruption. It was reasoned that a nucleosome positioning site with a low affinity for the histone octamer, cloned into the centre of a reconstitution template would cause a similar disruption *in vitro* that could be identified by sucrose gradient sedimentation.

Figure 5.1. Sequence-dependent Nucleosome Disruptions in Chromatin Structure.
Two possible fibre conformations adopted by fibres with differing central nucleosome positioning sites (Caplan et al., 1987).

A 197 bp sequence with a low affinity for the histone octamer within a regular nucleosome array may form a structure where the 30-nm chromatin fibre is interrupted by a "hypersensitive site" (Figure 5.1), or it may be that the strong positioning of the "601" nucleosomes forces a nucleosome to fill the gap between them, depending on the DNA

sequence. In this case, a "disruption" in the chromatin fibre may be formed where there is a difference in the higher-order structure that might cause the DNA to be more accessible.

Kubik et al. (2017) suggest that "fragile" nucleosomes may exist *in vivo*, across promoters which are more accessible to transcription factor binding. Furthermore, a nucleosome-free region is often found upstream of active transcription start sites, generated by transcription factor binding. Some of these sequences have been found to have reduced affinity for the core histone octamer, including poly(dA:dT) tracts (Field et al., 2008). The lack of a nucleosome and disruption of the condensation of the higher-order chromatin structure may enhance the chromatin accessibility of this region to transcription machinery or other cellular factors.

5.2 Design of Template DNA Sequences

To study the effects of the sequence of a single nucleosome positioning site on chromatin structure, DNA constructs were designed containing 25 nucleosome positioning sites (section 2.1.1). A single 197 bp DNA sequence with a low affinity for the histone octamer was cloned into the centre of the template, with 12 × "601" sequences positioned on either side and compared with a template containing 25 canonical 197 bp "601" sites (*601*). It was reasoned that a sequence with a high affinity for the histone octamer would allow a regular nucleosome array to form, which could fold into a compacted chromatin fibre, while a sequence with a low affinity for the histone octamer would either not form a nucleosome (causing a hypersensitive site) or form a nucleosome which is less stable than the surrounding array. This unstable nucleosome might form a single "disruption" in the higher-order chromatin structure when folded.

A 197 bp nucleosome positioning site with low affinity for the histone octamer was identified from the sequencing of nucleosomes reconstituted onto the ovine β-lactoglobulin gene *in vitro* (Fraser et al., 2009) by Jim Allan. The low abundancy of this site suggests that nucleosomes rarely form over this sequence *in vitro* and is easily digested by micrococcal nuclease following reconstitution (Figure 5.2A, approx. 5400 bp). This site was cloned into the centre of a "601" DNA template with surrounding high-affinity nucleosome positioning sequences and the predicted nucleosome positioning properties of the sequence were compared with the 25×601 sequence (*601*) using the algorithm described by van der Heijden et al. (2012) and in section 3.3 (Figure 5.2B). There is a clear difference between the predicted nucleosome occupancy of the central low-affinity sequence and the surrounding "601" sites, with the low-affinity sequence having almost no measurable affinity for the

histone octamer. The presence of the low-affinity site does not seem to affect the positioning of nucleosomes over the adjacent "601" nucleosome positioning sites.

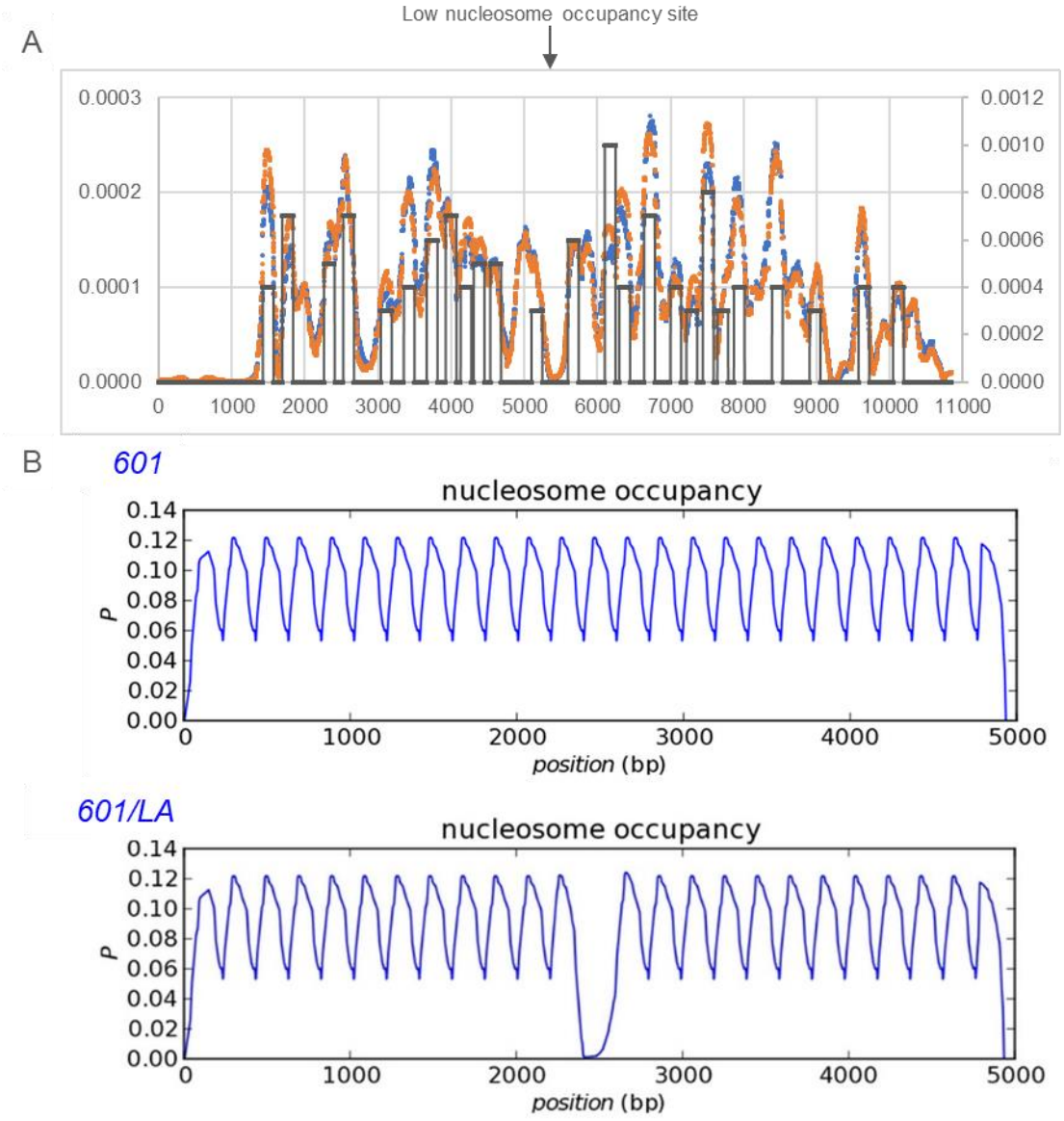

Figure 5.2. A Site with Low Affinity for the Histone Octamer.
A) Nucleosome occupancy maps of the β-lactoglobulin gene, highlighting a site with low affinity for the histone octamer. B) Comparison of predicted nucleosome occupancy of 601 and 601+LA according to van der Heijden et al. (2012).

It should be easier to control the saturation level of these templates, as unlike *601*, *BLG* and *601/BLG* they were expected to achieve a similar level of saturation when reconstituted at similar histone:DNA ratios. This method also allows a competitor to be easily used (except in the case of analyses such as SAXS) as the competitor would also be reconstituted to the same degree in each sample. As more histones were required to saturate the *BLG* and *601/BLG* templates and this caused the competitor to be more chromatinised, there was the

possibility that when linker histones were added these might bind differently to the template and the competitor fibres in each case. This would not be expected to be an issue between two fibres which vary in only one nucleosome positioning site.

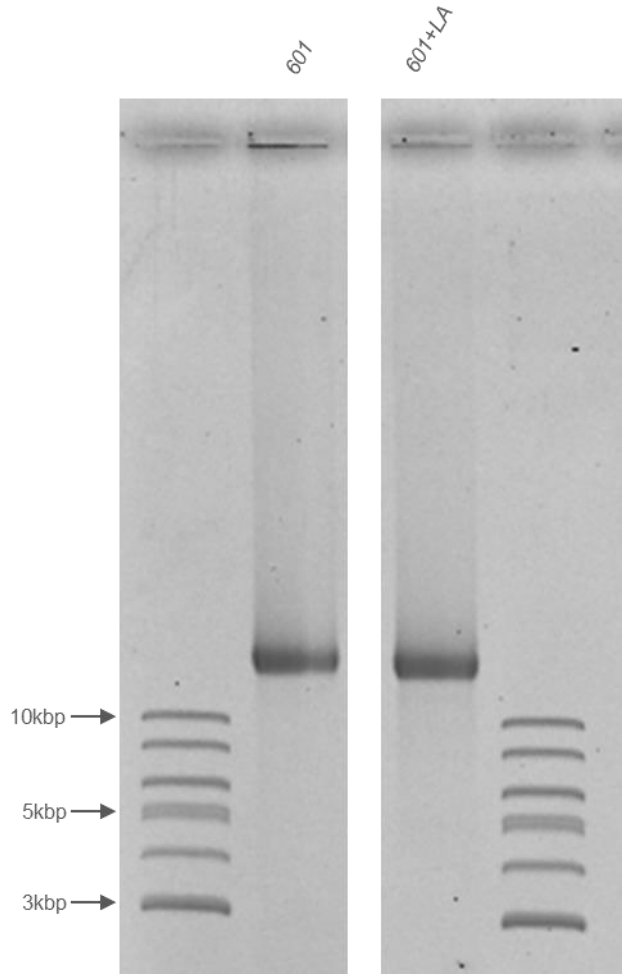

Figure 5.3. Electrophoretic Mobility Shift Assay of 601 and 601/LA.
0.7% gel showing shift of purified DNA reconstituted at histone:DNA ratios of 1.12:1. Nominal fragment sizes: *601* = 13657 bp, *601+LA* = 13848 bp.

Furthermore, studying the dynamics of *BLG* and *601/BLG* fibres only, allows us to study the average dynamics of each of the templates, rather than the composite effect of each nucleosome positioning sequence within it. For example, if the positioning of one nucleosome affects the properties of a neighbouring nucleosome, it is difficult to distinguish these. Using an approach where only one nucleosome positioning site is varied in each template removes the effects of differential nucleosome positioning, as surrounding "601" sites are likely to force a nucleosome, if one forms, into the designated site, rather than allowing it to sit over the linker regions, as commonly occurs within the *BLG* and *601/BLG*

(section 3.8.3). This removes the effect of the surrounding DNA sequence and enabled me to examine the properties of the one DNA sequence placed in the centre of the fibre.

The "601" template with a low-affinity central nucleosome positioning site derived from the β-lactoglobulin gene was termed *601+LA*. The *601* and *601+LA* sequences were reconstituted by salt dialysis (section 2.2.2). While histone octamers are likely to bind to the "601" sites first, it is possible that nucleosomes will form over the central nucleosome positioning sequences before octamers bind to the competitor DNA, even if sections of the competitor might have a higher affinity for the histone octamer, due to the cooperative binding mechanism of histones to DNA (Rubin and Moudrianakis, 1972). In the absence of a competitor, *601* and *601+LA* achieved a similar degree of saturation as measured by an electrophoretic mobility shift assay when reconstituted at a histone:DNA ratio of 1.12:1 in the absence of competitor (Figure 5.3).

5.3 A Low Affinity Nucleosome Binding Site is Sensitive to Nuclease Digestion

To analyse the sensitivity of each chromatinised template to nuclease, templates were reconstituted at histone:DNA ratios of 1.12:1 (see section 5.4) in the absence of competitor, and digested with DFF nuclease (section 3.10.4). A difference in the sensitivity of the central nucleosome site to DFF nuclease would indicate differential positioning of nucleosomes in these fibres, or the absence of a nucleosome positioned over the central low-affinity site in the *601+LA* template.

Following digestion for 40 min with 1 or 4 units of DFF nuclease at 4°C, chromatin was incubated in genomic lysis buffer with proteinase K, DNA was purified by phenol/chloroform extraction, precipitated and analysed by electrophoresis on a 1% agarose gel (Figure 5.4A). While the *601* showed a regular 200 bp nucleosomal ladder indicating that nucleosomes are well positioned over each 197 bp repeat, there was a region of increased digestion observed between the 12^{th} and 13^{th} bands of the *601+LA* template, which corresponds to the location of the low-affinity nucleosome positioning site. When the profiles of these lanes (after digestion with 4 units of enzyme) were compared (Figure 5.4B), it was observed that while there is increased digestion across this central nucleosome site in the *601+LA*, there is comparatively less material digested to 12mers or 13mers, likely as a consequence of increased digestion over the low-affinity nucleosome positioning site inhibiting the enzyme from cutting at the linker regions either side. Interestingly, there is also an increased frequency of digestion over the 11^{th} and 14^{th} linker regions, which are positioned one "601" nucleosome away from the low-affinity sequence. It is difficult to

analyse the quantity of digestion within the low-affinity nucleosome positioning site compared with the linker regions either side, as some digestion in the linker regions will be the result of the template being digested into more than two fragments.

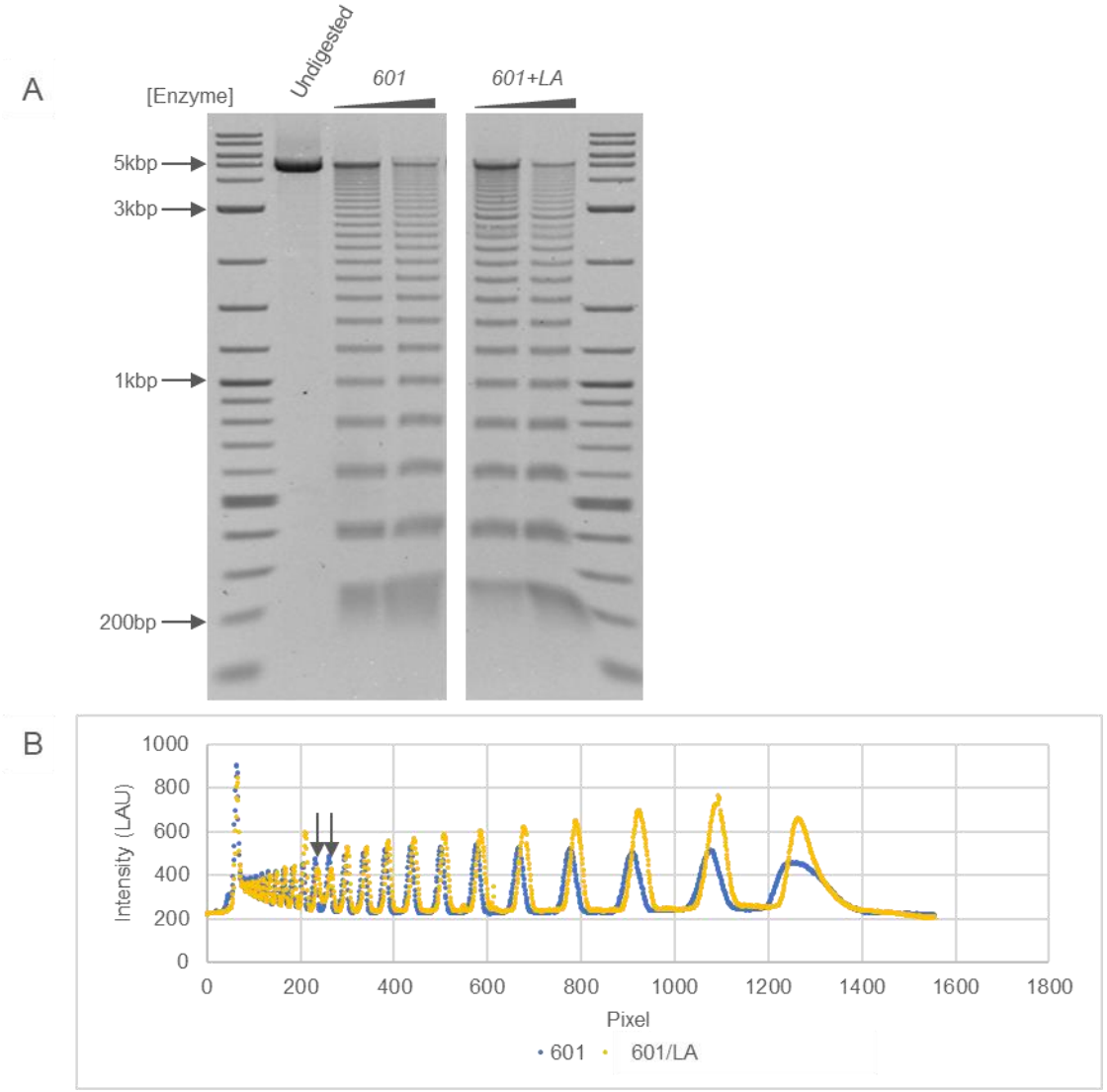

Figure 5.4. Digestion of 601 and 601+LA by DFF.
A) Digestion of *601* and *601+LA* reconstituted at 1.12:1 digested by 0, 1 or 4 units of DFF nuclease to nucleosomal ladders. B) Quantification of DNA following digestion with 4 units of DFF nuclease. See "spike" at 11n and 14n nucleosomes in *601+LA* and "blur" between 12 and 13n. Adjusted to account for gel curvature.

This data suggests a nucleosome can form over the LA site, but that DNA is more accessible or the histone octamer is more easily removed by the nuclease. Alternatively, increased nuclease sensitivity of this region might suggest a nucleosome may not be positioned over this low-affinity site in every fibre in the population of reconstituted arrays.

5.4 Small-angle X-ray Scattering

To analyse the shapes of *601* and *601+LA* chromatin fibres folded by linker histones and magnesium ions, I attempted to analyse the scattering properties of fibres by SAXS (section 4.5). Template DNA was reconstituted in the absence of a competitor at a ratio of 1.12:1 (see section 4.5).

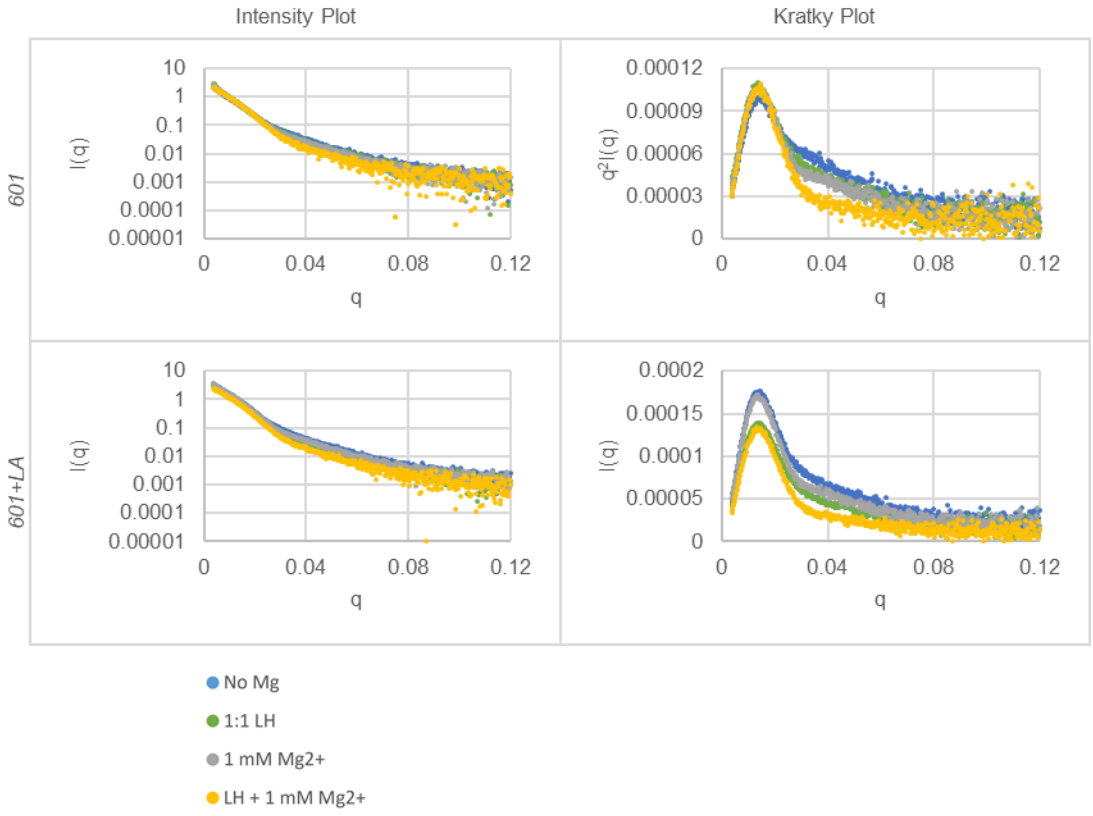

Figure 5.5. Scattering Intensity and Kratky Plots of 601 and 601+LA.
Plots show scattering patterns of *601* and *601+LA* reconstituted at 1.12:1, and titrated with 1 mM magnesium chloride and/or a 1:1 ratio of linker histone H5.

Intensity and Kratky plots (Figure 5.5) from the SAXS lack a prominent peak at 17-18 nm (0.035q) periodicity, as seen in Figure 4.8 thought to indicate the distance between individual nucleosomes, in the absence of linker histones or divalent cations. This may indicate the samples are undersaturated, similar to experiment 2 described in section 4.5. A prominent peak, however, is seen at 0.13-0.14 q suggesting a repeating distance of 45-48 nm. As linker histones and magnesium ions were added, a shoulder which might reflect a periodicity of 17-18 nm is lost but there is no change to the 45-48 nm peak. A reduction in the peak at 45-48 nm in the *601+LA* samples after the addition of linker histones might reflect a structural change or might suggest that a small amount of these arrays become insoluble under these conditions.

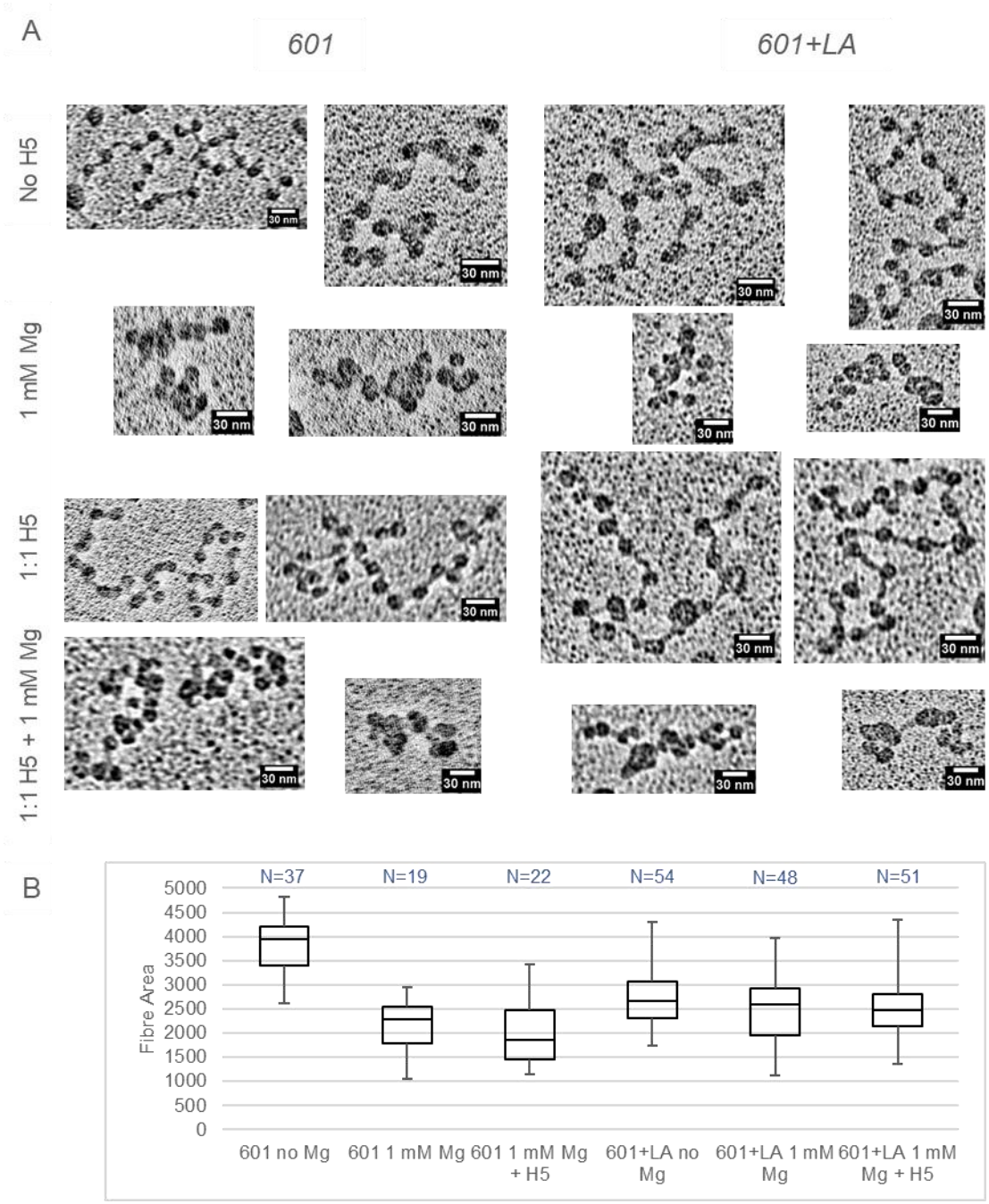

Figure 5.6. Electron Microscopy of 601 and 601/LA titrated with Linker Histones and Magnesium Ions.
A) *601* and *601+LA* reconstituted at 1.12:1 in the absence of competitor, titrated with linker histone H5 and/or 1 mM magnesium chloride. B) Area of individual fibres measured following electron microscopy.

Electron microscopy analysis suggested that these fibres were differently saturated, with *601* fibres having 26.08±0.57 nucleosomes/fibre (N=37) and *601+LA* having 21.68±0.63 nucleosomes/fibre (N=54). This demonstrates that even when two templates should have very similar affinities for histone octamers and become saturated at similar ratios, it is

difficult to control for reconstitution efficiency, likely due to small differences in DNA concentration prior to reconstitution.

The area of the *601* fibres is significantly higher than that of the *601+LA* fibres in the absence of magnesium or linker histones, but this is likely accounted for by the difference in fibre saturation (Figure 5.6B). Despite these differences, both fibres seem to compact well in the presence of magnesium ions, *601* compacted to a smaller volume despite initially having a higher volume. Surprisingly, the presence of linker histones in the absence of any magnesium had very little effect on the compaction of the fibres, and in the presence of magnesium caused only a modest reduction to fibre area.

Considering the difference in saturation, it is surprising that the samples both compact efficiently and that there is so little difference in the scattering profiles revealed by small-angle X-ray scattering.

5.5 Impact of Chromatin Disruption on Sucrose Gradient Sedimentation

A disruption in higher-order chromatin structure is expected to have a significant impact on the sedimentation velocity of the chromatin. To analyse the effect of a single DNA sequence disruption within the "601" templates on the sedimentation velocity of the chromatin arrays, *601* and *601+LA* were reconstituted in the presence of a vector backbone competitor at a histone:DNA ratio of 1.5:1 or 1.6:1 and dialysed into 80 mM NaCl. Using competitor DNA should allow us to more easily reconstitute two equally saturated samples, and any small differences might be expected to be seen only in the saturation of the competitor molecule. Linker histone H5 was added following the reconstitution up to a ratio of 2.2:1 and chromatin was centrifuged through 6-40% isokinetic sucrose gradients (section 4.4).

The sedimentation profiles of *601* and *601+LA* chromatin reconstituted at 1.5:1 or 1.6:1 histones:DNA (Figure 5.7A) show that *601+LA* fibres sediment slightly faster that *601* fibres in the absence of linker histones, but not as quickly as the *BLG* and *601/BLG* fibres. This is consistent with a hypothesis that fibres with less regular and precise nucleosome positioning are able to adopt more compacted structures in 80 mM salt in the absence of linker histones. Adding linker histone H5 to a linker:nucleosome ratio of 1.4:1 causes a considerable shift in the sedimentation velocity of the *601* and a comparatively smaller shift in the sedimentation velocity of the *601+LA*, so that the major peak of the *601* fibres spin faster than *601+LA*. At H5:nucleosome ratios higher than 1.4:1, *601* chromatin travels further through the gradient

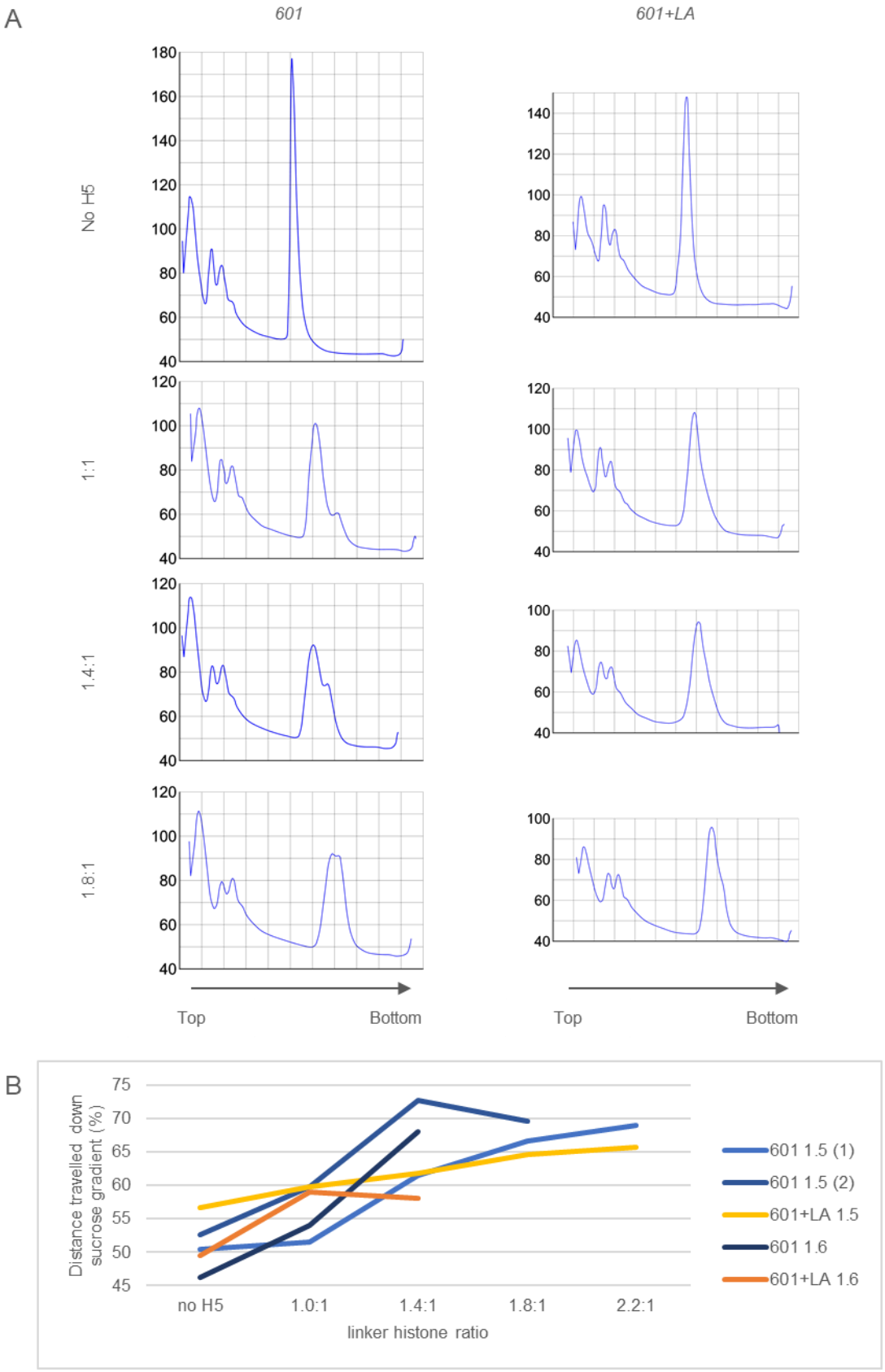

Figure 5.7. Sucrose Gradient Sedimentation of *601* and *601+LA* Titrated with Linker Histones. (Full legend found on following page)

A) Sedimentation profiles of *601* and *601+LA* chromatin reconstituted at 1.5:1 core histones:DNA and titrated with up to 1.8:1 H5 molecules per nucleosome. B) Distance sedimented by each sample relative to the length of the gradient.

than *601+LA*, suggesting that it adopts a more compact structure. As described in section 4.4, the profile of the *601* sedimentation shows two peaks suggesting the cooperative binding of linker histones. While this is less clear for the *601+LA* fibres, there is evidence of a "shoulder", especially at a 1.8:1 linker histone ratio. This suggests that the *601+LA* fibres are less heterogeneous than the *BLG* and *601/BLG* fibres, and therefore distinct species are visible. The height to peak width ratios are shown in Table 4.

Width ½ height	601	627
No H5	3mm	4mm
1:1	6.5mm	5.8mm
1.4:1	11mm	8mm
1.8:1	9.5mm	8mm
2.2:1	7mm	9mm

Table 4. Width at ½ Height of Chromatin Sedimentation Peaks.
Where peaks are divided into two distinct peaks, the larger peak is measured.

Following sucrose gradient sedimentation, fibres were imaged by electron microscopy. While insufficient individual fibre images could be isolated for a comprehensive analysis, typical fibres are shown in Figure 5.8. Fibres seem to adopt a fairly similar structure in the absence of linker histones, and became more compacted as linker histones were added. In a number of compacted *601+LA* fibres, it was possible to see two distinct sections to the compacted fibre, connected by a short length of DNA. This might indicate a disruption in the centre of the fibre which prevents it from folding into a continuous structure. As a 200 bp nucleosome-free region is expected to be approximately 66 nm in length, which is much longer than these disruptions appear to be, they may indicate a discontinuity caused by a weakly positioned nucleosome, or it may be that a longer stretch of DNA is occluded by the surrounding chromatin fibre in these images. These discontinuities only appear in a subset of fibres, suggesting that the low-affinity DNA sequence often positions a nucleosome which can be folded into a higher-order structure which cannot be easily distinguished from *601* fibres.

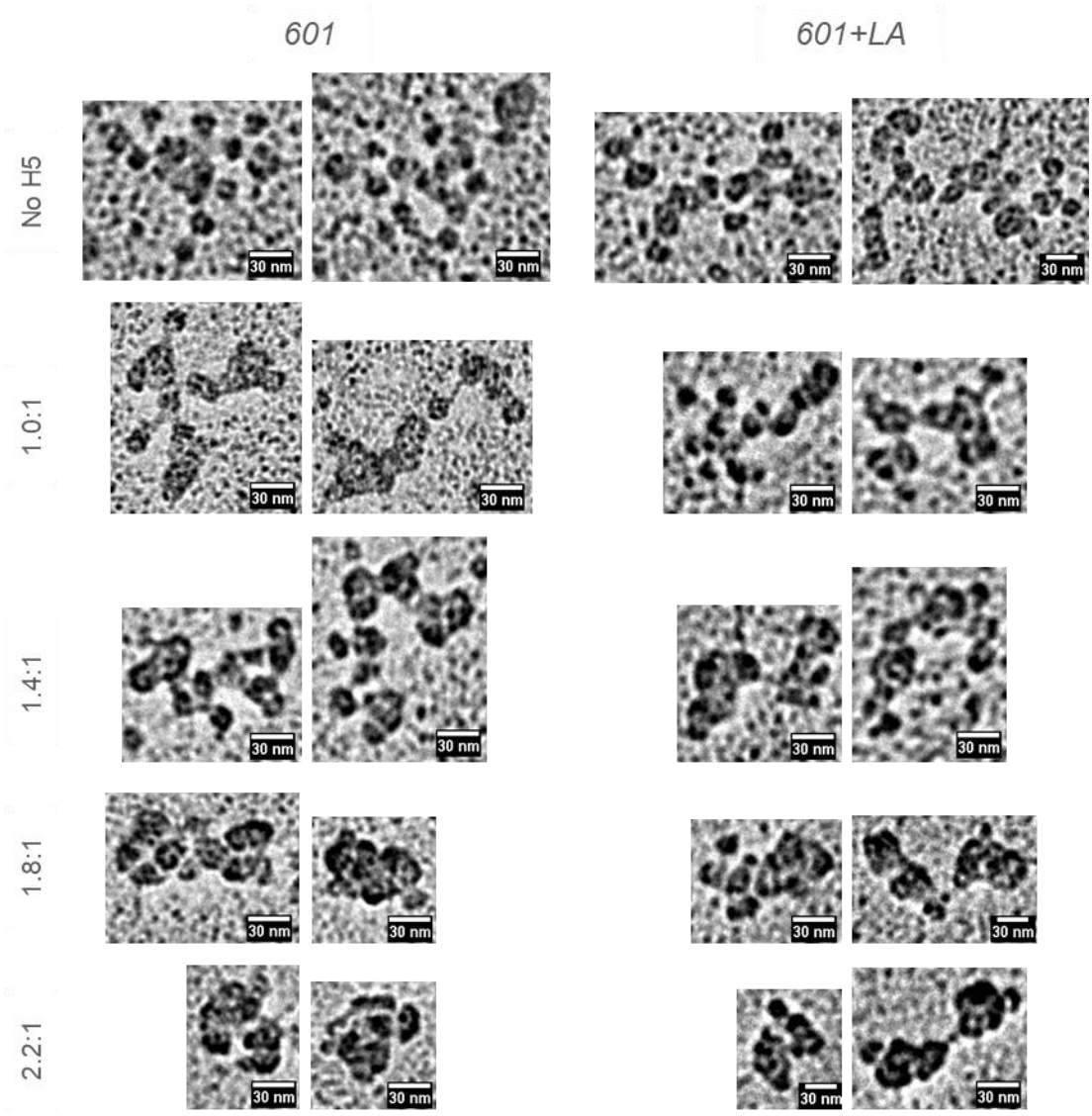

Figure 5.8. Electron Microscopy of 601 and 601+LA titrated with Linker Histones.
Typical images of each fibre at different linker histone ratios, reconstituted at 1.5:1 and titrated up to 2.2:1 H5 molecules per nucleosome.

5.6 Summary

To study the contributions of individual nucleosome positioning sequences in regulating the structure of higher order chromatin fibres, a template containing a single low-affinity nucleosome positioning sequence in the centre of a repeating "601" array was made (Figure 5.2). Electrophoretic mobility shift assay indicated that these fibres have a similar mobility when reconstituted at the same histone:DNA ratio in the absence of a competitor (Figure 5.3). Due to the similar overall affinity of each template for the histone octamer, these templates would also be expected to reach saturation at a similar histone:DNA ratio in the presence of a competitor, unlike the *BLG* and *601/BLG*.

When digested by DFF nuclease, there appeared to be increased digestion across the central nucleosome positioning site (Figure 5.4), suggesting that this site is either lacking a nucleosome in some or all of the fibre population, or that a nucleosome is positioned across this site, but is more "fragile" and allows the nuclease greater access to the underlying DNA.

Analysis of the scattering properties of these fibres provided limited information about their different structures when folded by linker histones or divalent cations, due to differences in the saturation of these samples (Figure 5.6).

Sucrose gradient sedimentation of arrays in 80 mM salt in the absence of linker histones revealed that arrays containing a low-affinity nucleosome positioning site have a slightly higher sedimentation rate than regular *601* fibres (Figure 5.7). This is consistent with my previous observation (section 4.4) that arrays with less regular nucleosome positioning are more compacted in the absence of linker histones. Upon the addition of linker histones, *601* fibres become more compacted than those containing a low-affinity positioning site, at linker histone:nucleosome ratios above 1.4:1. These differences in compaction may be due to a disruption in the higher-order structure of the *601+LA* fibres, caused by a missing or fragile nucleosome over the central low-affinity site. While insufficient fibres could be isolated by EM images for a complete analysis (less than 10 of each sample), such a "disruption" could be seen in a number of *601+LA* fibres, while *601* fibres generally adopted a more continuous structure (Figure 5.8). This model might be a useful method of studying the influence of a variety of DNA sequences on the higher-order chromatin fibre structure.

Chapter 6. Discussion

The higher-order structure and dynamics of the chromatin fibre are thought to play an important role in the regulation of the eukaryotic genome, but much remains unknown about the structure of this complex and how it is regulated. *In vitro* studies based on repeating units of strong nucleosome positioning sequences have provided insight into how nucleosomes within an array are compacted into a 30-nm fibre, with sequences containing a long nucleosome repeat (197 bp) forming a one-start solenoid and those with a short repeat length (167 bp) forming a two-start helical structure (Kruithof et al., 2009; Routh et al., 2008). The unfolding dynamics of these structures have been studied by single-molecule force spectroscopy, confirming that longer DNA lengths appear to form a one-start structure which unfolds at similar forces but is less stiff than a two start structure (Meng et al., 2015). However, as a limited number of DNA sequences, in particular the *601* sequence, have dominated these studies, it is not well understood what role the DNA sequence plays in defining this structure and the dynamics of chromatin unfolding.

I set out to investigate the structure and dynamics of chromatin fibres with different underlying DNA sequences. I hypothesised that non-repetitive DNA sequences containing strong, biologically-derived nucleosome positioning sequences might not form such a compacted higher-order chromatin structure as a repetitive "601" DNA sequence, which has a far higher affinity for the histone octamer than any biological nucleosome positioning sequence *in vitro* (Thåström et al., 1999), and that these structures might be more easily disrupted than *601* fibres. A fibre containing a combination of "601" and biologically-derived nucleosome positioning sequences was hypothesised to form a less compacted structure that is more easily perturbed than 601 fibres, but a more compacted structure than a non-repetitive sequence. I also investigated the properties of a higher-order fibre based on a "601" repeat with a single biologically-derived DNA sequence site with a low affinity for the histone octamer in the centre. I hypothesised that this sequence might create a disruption in the higher-order structure of the chromatin fibre, either by not binding a histone octamer or by weakly positioning a nucleosome.

6.1 The Impact of Nucleosome Positioning on Higher-order Folding

My results of DFF digestions show that biologically-derived non-repetitive DNA sequences containing strong nucleosome positioning sites form nucleosome arrays with a more heterogeneous nucleosome repeat length than *601* fibres (Figure 3.15). I hypothesised that this would impede their ability to fold into compacted higher-order structures to the same degree as the *601* fibre. However, I found by sucrose gradient sedimentation of chromatin

fibres folded by linker histones in 80 mM NaCl that biologically-derived chromatin fibres with strong nucleosome positioning properties are able to achieve a similar degree of compaction as *601* fibres, though they appear to form more heterogeneous structures than the *601* (Figure 4.3). Counterintuitively, in the absence of linker histones, these fibres were achieved a more compacted structure than *601* fibres.

I suggest the heterogeneous nucleosome positioning within these fibres creates a more flexible population of fibres which can achieve greater compaction in the absence of linker histones, however in the presence of linker histones they are still able to create a compacted higher-order structure. Fraser et al. (2006) hypothesised that the translational and rotational positioning of nucleosomes within an array would affect the capacity of nucleosomes to form the interaction between the acidic patch of the H2A globular domain and the H4 tail which is intrinsic to the compaction of the chromatin fibre (Dorigo et al., 2003). However, as the DNA sequences do not position nucleosomes as strongly as the "601", it is possible that nucleosomes slide along the DNA strand in order to accommodate a more energetically favourable higher-order structure.

Reconstituting a "601" fibre containing variable nucleosome repeat lengths could provide information about whether a nucleosome array with an irregular repeat can form a compacted higher-order chromatin structure, or whether nucleosome sliding might be required to achieve this, as nucleosomes would be unlikely to slide from the strongly positioning 147 bp "601" core sequence. Differences in the repeat length of the higher-order fibre are known to have an impact on structure, with long nucleosome repeat lengths having a wider diameter fibre and shorter nucleosome repeat lengths adopting a two-start helical structure rather than a one start solenoid (Routh et al., 2008). The majority of research into the structure of "601" arrays also use linker lengths increasing by 10 bp (167 bp, 187 bp, 197 bp etc.) which is consistent with the pitch of the helical turn of DNA, but very little research has been performed using nucleosome repeat lengths between these, where there may be different rotational positioning of nucleosomes. Brouwer, T., Kaczmarczyk, A., and van Noort, J. (personal communication) have found by performing single-molecule force spectroscopy on "601" chromatin with linker lengths varying by 5 bp that this positioning affects the higher-order structure of the chromatin fibres, as nucleosome phasing is out of synchrony with DNA helical rotation.

6.2 Relationship Between Structure and Function

The strength and regularity of nucleosome positioning are at least partially encoded in the DNA sequence, and it is possible that this has an impact on genome structure and function as

a result of the effects on the dynamics of the higher-order chromatin fibre. I have found by single-molecule force spectroscopy that biologically-derived non-*601* fibres which have a heterogeneous nucleosome repeat length are more easily perturbed under tension than regular *601* chromatin (Figure 4.12). RNA polymerase has been found to be capable of generating 14-25 pN of force on DNA, which is sufficient to disrupt its chromatin structure under these conditions.

Nucleosome repeat length has been found to be more irregular downstream of active promoters than of silent promoters (Baldi et al., 2018), but it is not known if this might be the consequence of RNA polymerase activity, or if this creates a permissive chromatin environment that favours gene transcription, where chromatin can be more easily perturbed. When the chromatin remodelling enzyme ACF is mutated in Drosophila embryo extracts, a less regular nucleosome repeat is observed downstream of promoters with low expression levels, which are subsequently upregulated (Scacchetti et al., 2018). This suggests that a regular nucleosome repeat length might contribute to a repressive chromatin state.

My results suggest that a fibre that has heterogeneous nucleosome positioning as a result of its DNA sequence, and which have a lower affinity for the histone octamer than the "601" sequence, are more easily disrupted than "601" repeat sequences. Strong nucleosome positioning sequences however, may still be considered markers of permissive environments for chromatin activation as strong nucleosome positioning sites may flank a nucleosome-free region that is important for transcriptional activation (Mavrich et al., 2008). Strong nucleosome positioning sequences in a regular repetitive pattern could form a repressive environment for transcription, while weaker nucleosome positioning sequences that are easily perturbed or strong positioning sequences that establish a nucleosome free region could form a permissive environment for transcription.

6.3 A Single Nucleosome Positioning Site can Cause a Disruption in Higher Order Chromatin Structure

I have also shown that a single sequence with low affinity for the histone octamer in the centre of a compacted fibre can cause a disruption in the chromatin fibre structure that affects its sedimentation rate. A significant difference is seen in the sedimentation profiles of *601*and *601+LA* chromatin in the presence of linker histones (Figure 5.7). This may be caused by a missing nucleosome or by differences in the flexibility of the higher-order structure caused by the different central nucleosome positioning sequence.

This may prove a useful model to analyse the contribution of nucleosome positioning sequences to the higher-order chromatin structure in a controlled manner. By cloning a library of nucleosome occupancy sites derived from micrococcal nuclease or DFF nuclease digestion of the genome into the centre of a "601" fibre, the higher-order structures formed by each different central sequence could be determined by sucrose gradient sedimentation. Sequences that position a nucleosome and fold into a compacted higher-order chromatin structure would be expected to sediment faster, and structures that do not position a nucleosome, or that cause a disruption to the higher-order chromatin fibre structure should sediment more slowly. These library-containing templates could be reconstituted as one sample and separated by sedimentation, and so the histone:DNA ratio and reconstitution efficiency could be controlled more accurately than has been possible in this study. Analysis of these sequences may allow the identification of sequences and features which promote or inhibit the complete folding of the higher-order chromatin fibre. Comparing these results with the higher-order chromatin structure *in vivo*, for example as determined by Micro-C or RICC-seq may allow the identification of loci where the DNA sequence plays an important role in higher-order chromatin structure and loci where other factors have a dominant role in determining chromatin structure.

The amount of linker histone added to *601* fibres may need to be further optimised to create a homogeneous population of fibre structures, as there are two distinct species of *601* fibres seen within the sucrose gradient. If the same phenomenon occurs using other DNA templates, and the peaks seen following sucrose gradient sedimentation are insufficiently homogeneous to be separated, this could also confound this analysis. At a linker histone:nucleosome ratio of 1.8:1, the two peaks seem to be close to converging and becoming one species, but this is at a surprisingly high ratio of linker histones to nucleosomes.

A nucleosomal library of mononucleosomes was generated by MNase digestion of RPE1 cells at various time points. The average size of these mononucleosomes was 177 bp, and these could be cloned into the centre of a "601" template alongside a 20 bp linker sequence (the repeat length of RPE1 cells is 197 bp, however, suggesting that the exonuclease activity of the MNase has trimmed the mononucleosomes down to this length). While the size of the nucleosomal library is sufficienttly complex to provide ample coverage of the genome, it would need to be cloned into a "601" sequence with maximum efficiency. Therefore, this has not yet been trialled.

6.4 Is the "601" a Suitable Model to Elucidate the Structure of Chromatin Within Cells?

The strongest biological nucleosome positioning sequences examined have been found to have at least a six-fold lower affinity for the histone octamer than artificial nucleosome positioning sequences including the "601" (Thåström et al., 1999). It has been suggested that the genome has not evolved to highly position nucleosomes over precise sites, as the "601" sequence does *in vitro*, but allows for flexibility in nucleosome positioning to enable DNA sequence to act in equilibrium with other cellular factors to contributes to gene regulation.

The "601" sequence may be a useful model for investigating the structure of repetitive regions of the genome such as centromeres and telomeres that contain satellite repeats. Indeed, prior to the discovery of the "601" by Lowary and Widom (1998), Luger et al. (1997) used α-satellite DNA sequences to reconstitute regular nucleosomes *in vitro*, though the DNA sequence was made palindromic to accommodate the symmetry of the nucleosome. DNA templates such as the *BLG* may provide a better dynamic model for euchromatin, as these have a variable sequence, with different repeat lengths coded into the histone affinity of their DNA sequence, creating a fibre which is more easily perturbed and will allow access to chromatin remodellers, histone modifiers and other proteins.

6.5 Limitations

An advantage of the "601" chromatin fibre model is that it is relatively easy to optimise reconstitution by salt dialysis *in vitro* with the use of a competitor DNA molecule. However, I have found that the comparatively low affinity of biologically-derived DNA sequences for the histone octamer means a competitor DNA molecule will bind histone octamers at a similar rate to the template sequences. A constant challenge throughout the experiments described in this thesis has been to identify a suitable histone:DNA ratio to use in each type of reconstitution. It is difficult to compare the higher-order folding of the three different fibres, when there is no internal control for fibre saturation. The different ratios and the different reconstitution strategies employed (in the presence vs. absence of competitor) also mean that it is difficult to directly compare different samples used between experiments in this thesis. Batch-to-batch variation was also seen, even when using the same samples of DNA and core histones to prepare chromatin within a short space of time, for example, in magnetic tweezer experiments, where two different samples prepared at the same ratio appear to have different average numbers of nucleosomes per fibre (Figure 4.12).

Some variation in the histone:DNA ratio is unlikely to affect the reconstitution of the 601 sequence in the presence of competitor, as the competitor will sequester any excess histones histones (Huynh et al., 2005). In this thesis, the *601* sequence has been reconstituted in the presence of competitor DNA at ratios between 1.3:1 and 1.7:1, with little variation in the number of nucleosomes per fibre according to various assays. However, for the *BLG* and the *601/BLG* and for the *601* in the absence of a competitor, further assays may be required to ensure that fibres are correctly saturated. I have altered the histone:DNA ratio throughout my PhD (eg between SAXS experiment 1 and experiment 2, section 4.5) based on new results, but as the saturation may even vary slightly between batches reconstituted at the same histone:DNA ratio, fibre saturation may need to be confirmed on each individual sample. Analytical ultracentrifugation is a technique which may provide a convincing measurement of the average number of nucleosomes found on each chromatin fibre that could be undertaken in future. When performing experiments such as SAXS this should ieally be performed on the same batch of chromatin to ensure that the differences in higher-order folding and dynamics of chromatin fibres can be attributed solely to differences in higher-order folding, and are not the result of over or undersaturation of chromatin fibres.

The results of single-molecule force spectroscopy, presented in section 4.7, show that there is a clear difference in the unfolding dynamics of chromatin fibres based on different DNA sequences where the number of nucleosomes formed on each fibre can be confirmed based on the fibre extension at high forces (6-50pN).

6.6 Does DNA Sequence Impact Chromatin Structure *in vivo*?

Multiple studies have found a high degree of similarity between histone occupancy *in vivo* and reconstituted nucleosome arrays, however, DNA sequence is only one of many factors that affects nucleosome positioning *in vivo* and histone variants and modifications, DNA modification, and other proteins such as chromatin remodellers or transcription machinery have also been found to influence nucleosome positioning. The role that other factors play in the formation of a higher-order chromatin structure is likely even more dominant, though the influence of the DNA sequence compared to other factors is likely to vary throughout the genome.

Where the “601” sequence has been cloned into cellular DNA and its structure analysed, its nucleosome positioning properties have varied between studies. When the “601” was cloned into the mouse *c-myc* gene and transfected into a human cell line, the “601” sequence (when placed in a forward orientation) was found to increase transcriptional pausing compared to other sequences at the +1 nucleosome position (Jimeno-González et al., 2015). However,

when the "601" sequence was cloned into the yeast genome, it was not found to strongly position nucleosomes and its histone occupancy was depleted relative to the surrounding regions when placed in an ORF. This suggests that nucleosome positioning is highly dependent on the context of the sequence and that different factors may govern histone occupancy and nucleosome positioning *in vivo* compared to *in vitro*.

In vivo, the wide-ranging mechanisms for regulating gene expression likely allow chromatin fibres to form that are more heterogeneous than in this controlled study. On a single DNA sequence, Hermans et al. (2017) found significant differences in the structure of native yeast chromatin using single-molecule force spectroscopy, though it is not known whether the higher-order structure is preserved well during the pull-down procedure.

6.7 Advantages of Sequencable Templates for *in vitro* Reconstitution

The "601" template cannot be sequenced due to its repetitive nature, which limits the types of analyses that can be used to study chromatin structure. While for physical structural assays, such as x-ray crystallography, the regularity of the "601" is advantageous, the fact that it cannot be sequenced makes it impossible to use many molecular biology techniques. For example, Micro-C (Hsieh et al., 2015) and RICC-seq (Risca et al., 2017) may be useful assays to study the conformation of higher-chromatin fibres, and whether they adopt a one or two-start helical structure, but both depend on a sequencable DNA template. To date, these assays have only been used to study higher-order nucleosome interactions *in vivo* on the tetra-nucleosome scale.

It may be possible to generate a sequencable "601" construct, which uses different sequences of linker DNA (of the same or different lengths). However, this would still not be sequencable in assays where the DNA is digested completely to the 147 base pair core particle, for example by micrococcal nuclease, necessary for Micro-C, as these particles would still have identical DNA sequences. Sequence mutations within the 147 bp "601" core may solve this, but also may alter the strong nucleosome positioning properties of the "601" site.

6.8 Conclusion

I have shown in this thesis that while DNA sequences containing strong nucleosome positioning sites fold into higher-order chromatin as compact as the "601" repeat sequence, these fibres have a more heterogeneous structure and are more easily perturbed under tension. Sequences containing a weak nucleosome positioning site however, have a disrupted higher-order chromatin structure which can be distinguished from a compacted

structure by sucrose gradient sedimentation. I suggest that thes fibres may be an appropriate model to test the ability of different DNA sequences to fold into a compacted higher-order chromatin fibre and hope that the results presented here will provide a useful foundation for such further work.

References

Allan, J., Hartman, P.G., Crane-Robinson, C., and Aviles, F.X. (1980a). The structure of histone H1 and its location in chromatin. Nature *288*, 675–679.

Allan, J., Staynov, D.Z., and Gould, H. (1980b). Reversible Dissociation of Linker Histone from Chromatin with Preservation of Internucleosomal Repeat. Proc. Natl. Acad. Sci. U. S. A. *77*, 885–889.

Allan, J., Mitchell, T., Harborne, N., Bohm, L., and Crane-Robinson, C. (1986). Roles of H1 domains in determining higher order chromatin structure and H1 location. J. Mol. Biol. *187*, 591–601.

Allan, J., Fraser, R.M., Owen-Hughes, T., and Keszenman-Pereyra, D. (2012). Micrococcal Nuclease Does Not Substantially Bias Nucleosome Mapping. J. Mol. Biol. *417*, 152–164.

Almouzni, G., and Wolffe, A.P. (1993). Nuclear Assembly, Structure, and Function: The Use of Xenopus in Vitro Systems. Exp. Cell Res. *205*, 1–15.

Anderson, J.D., and Widom, J. (2001). Poly(dA-dT) Promoter Elements Increase the Equilibrium Accessibility of Nucleosomal DNA Target Sites. Mol. Cell. Biol. *21*, 3830–3839.

Athey, B.D., Smith, M.F., Rankert, D.A., Williams, S.P., and Langmore, J.P. (1990). The diameters of frozen-hydrated chromatin fibers increase with DNA linker length: evidence in support of variable diameter models for chromatin. J. Cell Biol. *111*, 795–806.

Auerbach, R.K., Euskirchen, G., Rozowsky, J., Lamarre-Vincent, N., Moqtaderi, Z., Lefrançois, P., Struhl, K., Gerstein, M., and Snyder, M. (2009). Mapping accessible chromatin regions using Sono-Seq. Proc. Natl. Acad. Sci. *106*, 14926–14931.

Bak, A.L., Zeuthen, J., and Crick, F.H. (1977). Higher-order structure of human mitotic chromosomes. Proc. Natl. Acad. Sci. U. S. A. *74*, 1595–1599.

Baldi, S., Krebs, S., Blum, H., and Becker, P.B. (2018). Genome-wide measurement of local nucleosome array regularity and spacing by nanopore sequencing. BioRxiv 272526.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120–124.

Bao, Y., White, C.L., and Luger, K. (2006). Nucleosome core particles containing a poly(dA.dT) sequence element exhibit a locally distorted DNA structure. J. Mol. Biol. *361*, 617–624.

Becker, J.S., McCarthy, R.L., Sidoli, S., Donahue, G., Kaeding, K.E., He, Z., Lin, S., Garcia, B.A., and Zaret, K.S. (2017). Genomic and Proteomic Resolution of

Heterochromatin and Its Restriction of Alternate Fate Genes. Mol. Cell *68*, 1023-1037.e15.

Bell, J.C., Jukam, D., Teran, N.A., Risca, V.I., Smith, O.K., Johnson, W.L., Skotheim, J.M., Greenleaf, W.J., and Straight, A.F. (2018). Chromatin-associated RNA sequencing (ChAR-seq) maps genome-wide RNA-to-DNA contacts. ELife *7*, e27024.

Belmont, A.S., and Bruce, K. (1994). Visualization of G1 chromosomes: a folded, twisted, supercoiled chromonema model of interphase chromatid structure. J. Cell Biol. *127*, 287–302.

Belotserkovskaya, R., Oh, S., Bondarenko, V.A., Orphanides, G., Studitsky, V.M., and Reinberg, D. (2003). FACT Facilitates Transcription-Dependent Nucleosome Alteration. Science *301*, 1090–1093.

Blank, T.A., and Becker, P.B. (1995). Electrostatic Mechanism of Nucleosome Spacing - ScienceDirect. J. Mol. Biol. *252*, 305–313.

Boeger, H., Griesenbeck, J., Strattan, J.S., and Kornberg, R.D. (2003). Nucleosomes Unfold Completely at a Transcriptionally Active Promoter. Mol. Cell *11*, 1587–1598.

Bondarenko, V.A., Steele, L.M., Ujvári, A., Gaykalova, D.A., Kulaeva, O.I., Polikanov, Y.S., Luse, D.S., and Studitsky, V.M. (2006). Nucleosomes can form a polar barrier to transcript elongation by RNA polymerase II. Mol. Cell *24*, 469–479.

Bouazoune, K., Miranda, T.B., Jones, P.A., and Kingston, R.E. (2009). Analysis of individual remodeled nucleosomes reveals decreased histone–DNA contacts created by hSWI/SNF. Nucleic Acids Res. *37*, 5279–5294.

Boyle, A.P., Davis, S., Shulha, H.P., Meltzer, P., Margulies, E.H., Weng, Z., Furey, T.S., and Crawford, G.E. (2008). High-Resolution Mapping and Characterization of Open Chromatin across the Genome. Cell *132*, 311–322.

Brouwer, T., Kaczmarczyk, A., and van Noort, J. (in preparation).

Brown, S.A., Imbalzano, A.N., and Kingston, R.E. (1996). Activator-dependent regulation of transcriptional pausing on nucleosomal templates. Genes Dev. *10*, 1479–1490.

Caplan, A., Kimura, T., Gould, H., and Allan, J. (1987). Perturbation of chromatin structure in the region of the adult beta-globin gene in chicken erythrocyte chromatin. J. Mol. Biol. *193*, 57–70.

Carruthers, L.M., Bednar, J., Woodcock, C.L., and Hansen, J.C. (1998). Linker Histones Stabilize the Intrinsic Salt-Dependent Folding of Nucleosomal Arrays: Mechanistic Ramifications for Higher-Order Chromatin Folding. Biochemistry *37*, 14776–14787.

Chien, F.-T., and van Noort, J. (2009). 10 years of tension on chromatin: results from single molecule force spectroscopy. Curr. Pharm. Biotechnol. *10*, 474–485.

Choi, J.K., and Kim, Y.-J. (2008). Epigenetic regulation and the variability of gene expression. Nat. Genet. *40*, 141–147.

Chow, J., and Heard, E. (2009). X inactivation and the complexities of silencing a sex chromosome. Curr. Opin. Cell Biol. *21*, 359–366.

Cirillo, L.A., McPherson, C.E., Bossard, P., Stevens, K., Cherian, S., Shim, E.Y., Clark, K.L., Burley, S.K., and Zaret, K.S. (1998). Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. EMBO J. *17*, 244–254.

Cirillo, L.A., Lin, F.R., Cuesta, I., Friedman, D., Jarnik, M., and Zaret, K.S. (2002). Opening of Compacted Chromatin by Early Developmental Transcription Factors HNF3 (FoxA) and GATA-4. Mol. Cell *9*, 279–289.

Clark, D.J., and Thomas, J.O. (1986). Salt-dependent co-operative interaction of histone H1 with linear DNA. J. Mol. Biol. *187*, 569–580.

Compton, J.L., Bellard, M., and Chambon, P. (1976). Biochemical evidence of variability in the DNA repeat length in the chromatin of higher eukaryotes. Proc. Natl. Acad. Sci. *73*, 4382–4386.

Cremer, T., Kurz, A., Zirbel, R., Dietzel, S., Rinke, B., Schröck, E., Speicher, M.R., Mathieu, U., Jauch, A., Emmerich, P., et al. (1993). Role of Chromosome Territories in the Functional Compartmentalization of the Cell Nucleus. Cold Spring Harb. Symp. Quant. Biol. *58*, 777–792.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing Chromosome Conformation. Science *295*, 1306–1311.

Diagenode (2015). Chromatin Assembly Kit.

Diamond Light Source (2018). Equipment Description - Small Angle Scattering.

Dingwall, C., Lomonossoff, G.P., and Laskey, R.A. (1981). High sequence specificity of micrococcal nuclease. Nucleic Acids Res. *9*, 2659–2673.

Dong, F., and van Holde, K.E. (1991). Nucleosome positioning is determined by the (H3-H4)2 tetramer. Proc. Natl. Acad. Sci. U. S. A. *88*, 10596–10600.

Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T.J. (2003). Chromatin Fiber Folding: Requirement for the Histone H4 N-terminal Tail. J. Mol. Biol. *327*, 85–96.

Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R.R., and Richmond, T.J. (2004). Nucleosome Arrays Reveal the Two-Start Organization of the Chromatin Fiber. Science *306*, 1571–1573.

Drew, H.R., and Travers, A.A. (1985). DNA bending and its relation to nucleosome positioning. J. Mol. Biol. *186*, 773–790.

Fan, J.Y., Rangasamy, D., Luger, K., and Tremethick, D.J. (2004). H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding. Mol. Cell *16*, 655–661.

Fan, Y., Nikitina, T., Morin-Kensicki, E.M., Zhao, J., Magnuson, T.R., Woodcock, C.L., and Skoultchi, A.I. (2003). H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo. Mol. Cell. Biol. *23*, 4559–4572.

Fan, Y., Nikitina, T., Zhao, J., Fleury, T.J., Bhattacharyya, R., Bouhassira, E.E., Stein, A., Woodcock, C.L., and Skoultchi, A.I. (2005). Histone H1 Depletion in Mammals Alters Global Chromatin Structure but Causes Specific Changes in Gene Regulation. Cell *123*, 1199–1212.

Feinberg, A.P., and Tycko, B. (2004). The history of cancer epigenetics. Nat. Rev. Cancer *4*, 143–153.

Fenouil, R., Cauchy, P., Koch, F., Descostes, N., Cabeza, J.Z., Innocenti, C., Ferrier, P., Spicuglia, S., Gut, M., Gut, I., et al. (2012). CpG islands and GC content dictate nucleosome depletion in a transcription-independent manner at mammalian promoters. Genome Res. *22*, 2399–2408.

Field, Y., Kaplan, N., Fondufe-Mittendorf, Y., Moore, I.K., Sharon, E., Lubling, Y., Widom, J., and Segal, E. (2008). Distinct Modes of Regulation by Chromatin Encoded through Nucleosome Positioning Signals. PLoS Comput Biol *4*, e1000216.

Finch, J.T., and Klug, A. (1976). Solenoidal model for superstructure in chromatin. Proc. Natl. Acad. Sci. U. S. A. *73*, 1897–1901.

Finlan, L.E., Sproul, D., Thomson, I., Boyle, S., Kerr, E., Perry, P., Ylstra, B., Chubb, J.R., and Bickmore, W.A. (2008). Recruitment to the Nuclear Periphery Can Alter Expression of Genes in Human Cells. PLoS Genet. *4*.

Fisher, E.A., and Felsenfeld, G. (1986). Comparison of the folding of .beta.-globin and ovalbumin gene containing chromatin isolated from chicken oviduct and erythrocytes. Biochemistry *25*, 8010–8016.

Fraser, R.M., Allan, J., and Simmen, M.W. (2006). In Silico Approaches Reveal the Potential for DNA Sequence-dependent Histone Octamer Affinity to Influence Chromatin Structure in Vivo. J. Mol. Biol. *364*, 582–598.

Fraser, R.M., Keszenman-Pereyra, D., Simmen, M.W., and Allan, J. (2009). High-Resolution Mapping of Sequence-Directed Nucleosome Positioning on Genomic DNA. J. Mol. Biol. *390*, 292–305.

Fyodorov, D.V., and Kadonaga, J.T. (2003). Chromatin assembly in vitro with purified recombinant ACF and NAP-1. Methods Enzymol. *371*, 499–515.

Fyodorov, D.V., Zhou, B.-R., Skoultchi, A.I., and Bai, Y. (2018). Emerging roles of linker histones in regulating chromatin structure and function. Nat. Rev. Mol. Cell Biol. *19*, 192–206.

Gencheva, M., Boa, S., Fraser, R., Simmen, M.W., A. Whitelaw, C.B., and Allan, J. (2006). In Vitro and in Vivo Nucleosome Positioning on the Ovine β-Lactoglobulin Gene Are Related. J. Mol. Biol. *361*, 216–230.

Gilbert, N., and Allan, J. (2001). Distinctive higher-order chromatin structure at mammalian centromeres. Proc. Natl. Acad. Sci. *98*, 11949–11954.

Gilbert, N., Boyle, S., Fiegler, H., Woodfine, K., Carter, N.P., and Bickmore, W.A. (2004). Chromatin Architecture of the Human Genome: Gene-Rich Domains Are Enriched in Open Chromatin Fibers. Cell *118*, 555–566.

Grishaev, A. (2012). Sample preparation, data collection and preliminary data analysis in biomolecular solution X-ray scattering. Curr. Protoc. Protein Sci. Editor. Board John E Coligan Al *CHAPTER*, Unit17.14.

Gross, D.S., and Garrard, W.T. (1988). Nuclease Hypersensitive Sites in Chromatin. Annu. Rev. Biochem. *57*, 159–197.

Gurley, L.R., D'anna, J.A., Barham, S.S., Deaven, L.L., and Tobey, R.A. (1978). Histone Phosphorylation and Chromatin Structure during Mitosis in Chinese Hamster Cells. Eur. J. Biochem. *84*, 1–15.

Han, M., and Grunstein, M. (1988). Nucleosome loss activates yeast downstream promoters in vivo. Cell *55*, 1137–1145.

Hansen, J.C., Connolly, M., McDonald, C.J., Pan, A., Pryamkova, A., Ray, K., Seidel, E., Tamura, S., Rogge, R., and Maeshima, K. (2017). The 10-nm chromatin fiber and its relationship to interphase chromosome organization. Biochem. Soc. Trans. BST20170101.

Hassan, A.H., Prochasson, P., Neely, K.E., Galasinski, S.C., Chandy, M., Carrozza, M.J., and Workman, J.L. (2002). Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. Cell *111*, 369–379.

van der Heijden, T., van Vugt, J.J.F.A., Logie, C., and van Noort, J. (2012). Sequence-based prediction of single nucleosome positioning and genome-wide nucleosome occupancy. Proc. Natl. Acad. Sci. U. S. A. *109*, E2514–E2522.

Hermans, N., Huisman, J.J., Brouwer, T.B., Schächner, C., Heusden, G.P.H. van, Griesenbeck, J., and Noort, J. van (2017). Toehold-enhanced LNA probes for selective pull down and single-molecule analysis of native chromatin. Sci. Rep. *7*, 16721.

van Holde, K.E. (1989). Chromatin (New York: Springer-Verlag).

Hong, L., Schroth, G.P., Matthews, H.R., Yau, P., and Bradbury, E.M. (1993). Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. J. Biol. Chem. *268*, 305–314.

Hsieh, T.-H.S., Weiner, A., Lajoie, B., Dekker, J., Friedman, N., and Rando, O.J. (2015). Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. Cell *162*, 108–119.

Huynh, V.A.T., Robinson, P.J.J., and Rhodes, D. (2005). A Method for the In Vitro Reconstitution of a Defined "30 nm" Chromatin Fibre Containing Stoichiometric Amounts of the Linker Histone. J. Mol. Biol. *345*, 957–968.

Jimeno-González, S., Ceballos-Chávez, M., and Reyes, J.C. (2015). A positioned +1 nucleosome enhances promoter-proximal pausing. Nucleic Acids Res. *43*, 3068–3078.

Jin, C., and Felsenfeld, G. (2007). Nucleosome stability mediated by histone variants H3.3 and H2A.Z. Genes Dev. *21*, 1519–1529.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions in the human genome. Nat. Genet. *41*, 941–945.

Joti, Y., Hikima, T., Nishino, Y., Kamada, F., Hihara, S., Takata, H., Ishikawa, T., and Maeshima, K. (2012). Chromosomes without a 30-nm chromatin fiber. Nucleus *3*, 404–410.

Kaczmarczyk, A. (2019). Nucleosome Stacking in Chromatin Fibres Probed withe Single Molecule Force- and Torque-spectroscopy. Leiden University.

Kaczmarczyk, A., Allahverdi, A., Brouwer, T.B., Nordenskiöld, L., Dekker, N.H., and Noort, J. van (2017). Single-molecule force spectroscopy on histone H4 tail cross-linked chromatin reveals fiber folding. J. Biol. Chem. jbc.M117.791830.

Kadam, S., McAlpine, G.S., Phelan, M.L., Kingston, R.E., Jones, K.A., and Emerson, B.M. (2000). Functional selectivity of recombinant mammalian SWI/SNF subunits. Genes Dev. *14*, 2441–2451.

Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J., et al. (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. Nature *458*, 362–366.

Kim, J., Wei, S., Lee, J., Yue, H., and Lee, T.-H. (2016). Single-Molecule Observation Reveals Spontaneous Protein Dynamics in the Nucleosome. J. Phys. Chem. B *120*, 8925–8931.

Kimura, T., Mills, F.C., Allan, J., and Gould, H. (1983). Selective unfolding of erythroid chromatin in the region of the active β-globin gene. Nature *306*, 709–712.

Kornberg, R.D. (2007). The molecular basis of eukaryotic transcription. Proc. Natl. Acad. Sci. *104*, 12955–12961.

Kouzine, F., Gupta, A., Baranello, L., Wojtowicz, D., Ben-Aissa, K., Liu, J., Przytycka, T.M., and Levens, D. (2013). Transcription-dependent dynamic supercoiling is a short-range genomic force. Nat. Struct. Mol. Biol. *20*, 396–403.

Kowalski, A., and Pałyga, J. (2011). Chromatin compaction in terminally differentiated avian blood cells: the role of linker histone H5 and non-histone protein MENT. Chromosome Res. *19*, 579–590.

Kruithof, M., Chien, F.-T., Routh, A., Logie, C., Rhodes, D., and van Noort, J. (2009). Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30-nm chromatin fiber. Nat. Struct. Mol. Biol. *16*, 534–540.

Kubik, S., Bruzzone, M.J., and Shore, D. (2017). TFIID or not TFIID, a continuing transcriptional SAGA. EMBO J. *36*, 248–249.

Kujirai, T., Ehara, H., Fujino, Y., Shirouzu, M., Sekine, S.-I., and Kurumizaka, H. (2018). Structural basis of the nucleosome transition during RNA polymerase II passage. Science *362*, 595–598.

Langmore, J., and Paulson (1983). Low angle x-ray diffraction studies of chromatin structure in vivo and in isolated nuclei and metaphase chromosomes. J. Cell Biol. *96*, 1120–1131.

Langmore, J.P., and Schutt, C. (1980). The higher order structure of chicken erythrocyte chromosomes in vivo. Nature *288*, 620–622.

Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. Nature *547*, 236–240.

Laybourn, P.J., and Kadonaga, J.T. (1991). Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II. Science *254*, 238–245.

Li, G., and Widom, J. (2004). Nucleosomes facilitate their own invasion. Nat. Struct. Mol. Biol. *11*, 763–769.

Lichter, P., Cremer, T., Borden, J., Manuelidis, L., and Ward, D.C. (1988). Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. Hum. Genet. *80*, 224–234.

Lorch, Y., LaPointe, J.W., and Kornberg, R.D. (1987). Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell *49*, 203–210.

Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J. Mol. Biol. *276*, 19–42.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature *389*, 251–260.

Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Preparation of nucleosome core particle from recombinant histones. Methods Enzymol. *304*, 3–19.

Maeshima, K., Rogge, R., Tamura, S., Joti, Y., Hikima, T., Szerlong, H., Krause, C., Herman, J., Seidel, E., DeLuca, J., et al. (2016). Nucleosomal arrays self-assemble into supramolecular globular structures lacking 30-nm fibers. EMBO J. *35*, 1115–1132.

Maier, V.K., Chioda, M., Rhodes, D., and Becker, P.B. (2008). ACF catalyses chromatosome movements in chromatin fibres. EMBO J. *27*, 817–826.

Makde, R.D., England, J.R., Yennawar, H.P., and Tan, S. (2010). Structure of RCC1 chromatin factor bound to the nucleosome core particle. Nature *467*, 562–566.

Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P., Sandstrom, R., Qu, H., Brody, J., et al. (2012). Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. Science *337*, 1190–1195.

Mavrich, T.N., Ioshikhes, I.P., Venters, B.J., Jiang, C., Tomsho, L.P., Qi, J., Schuster, S.C., Albert, I., and Pugh, B.F. (2008). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. Genome Res. *18*, 1073–1083.

McDowall, A.W., Smith, J.M., and Dubochet, J. (1986). Cryo-electron microscopy of vitrified chromosomes in situ. EMBO J. *5*, 1395–1402.

McGhee, J.D., Wood, W.I., Dolan, M., Engel, J.D., and Felsenfeld, G. (1981). A 200 base pair region at the 5′ end of the chicken adult β-globin gene is accessible to nuclease digestion. Cell *27*, 45–55.

Meng, H., Andresen, K., and van Noort, J. (2015). Quantitative analysis of single-molecule force spectroscopy on folded chromatin fibers. Nucleic Acids Res. *43*, 3578–3590.

Misteli, T., Gunjan, A., Hock, R., Bustin, M., and Brown, D.T. (2000). Dynamic binding of histone H1 to chromatin in living cells. Nature *408*, 877–881.

Miyazawa, Y., and Thomas, C.A. (1965). Nucleotide composition of short segments of DNA molecules. J. Mol. Biol. *11*, 223–237.

Musladin, S., Krietenstein, N., Korber, P., and Barbaric, S. (2014). The RSC chromatin remodeling complex has a crucial role in the complete remodeler set for yeast PHO5 promoter opening. Nucleic Acids Res. gkt1395.

Nagai, S., Davis, R.E., Mattei, P.J., Eagen, K.P., and Kornberg, R.D. (2017). Chromatin potentiates transcription. Proc. Natl. Acad. Sci. U. S. A. *114*, 1536–1541.

Naughton, C., Sproul, D., Hamilton, C., and Gilbert, N. (2010). Analysis of Active and Inactive X Chromosome Architecture Reveals the Independent Organization of 30 nm and Large-Scale Chromatin Structures. Mol. Cell *40*, 397–409.

Naughton, C., Avlonitis, N., Corless, S., Prendergast, J.G., Mati, I.K., Eijk, P.P., Cockroft, S.L., Bradley, M., Ylstra, B., and Gilbert, N. (2013). Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. Nat. Struct. Mol. Biol. *20*, 387–395.

Neumann, H., Hancock, S.M., Buning, R., Routh, A., Chapman, L., Somers, J., Owen-Hughes, T., van Noort, J., Rhodes, D., and Chin, J.W. (2009). A Method for Genetically Installing Site-Specific Acetylation in Recombinant Histones Defines the Effects of H3 K56 Acetylation. Mol. Cell *36*, 153–163.

Ngo, T.T.M., Yoo, J., Dai, Q., Zhang, Q., He, C., Aksimentiev, A., and Ha, T. (2016). Effects of cytosine modifications on DNA flexibility and nucleosome mechanical stability. Nat. Commun. *7*, 10813.

Nishino, Y., Eltsov, M., Joti, Y., Ito, K., Takata, H., Takahashi, Y., Hihara, S., Frangakis, A.S., Imamoto, N., Ishikawa, T., et al. (2012). Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. EMBO J. *31*, 1644–1653.

Noll, H., and Noll, M. (1989). Sucrose gradient techniques and applications to nucleosome structure. In Methods in Enzymology, P.M. Wassarman, and R.D. Kornberg, eds. (Academic Press), pp. 55–116.

Noll, M., and Kornberg, R.D. (1977). Action of micrococcal nuclease on chromatin and the location of histone H1. J. Mol. Biol. *109*, 393–404.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation center. Nature *485*, 381–385.

North, J.A., Shimko, J.C., Javaid, S., Mooney, A.M., Shoffner, M.A., Rose, S.D., Bundschuh, R., Fishel, R., Ottesen, J.J., and Poirier, M.G. (2012). Regulation of the nucleosome unwrapping rate controls DNA accessibility. Nucleic Acids Res. *40*, 10215–10227.

Nozawa, R.-S., Boteva, L., Soares, D.C., Naughton, C., Dun, A.R., Buckle, A., Ramsahoye, B., Bruton, P.C., Saleeb, R.S., Arnedo, M., et al. (2017). SAF-A Regulates Interphase Chromosome Structure through Oligomerization with Chromatin-Associated RNAs. Cell *169*, 1214-1227.e18.

Ohno, M., Ando, T., Priest, D.G., Kumar, V., Yoshida, Y., and Taniguchi, Y. (2019). Sub-nucleosomal Genome Structure Reveals Distinct Nucleosome Folding Motifs. Cell *176*, 520-534.e25.

Olins, A.L., and Olins, D.E. (1974). Spheroid Chromatin Units (ν Bodies). Science *183*, 330–332.

Ou, H.D., Phan, S., Deerinck, T.J., Thor, A., Ellisman, M.H., and O'Shea, C.C. (2017). ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. Science *357*, eaag0025.

Owen-Hughes, T., Utley, R.T., Côté, J., Peterson, C.L., and Workman, J.L. (1996). Persistent site-specific remodeling of a nucleosome array by transient action of the SWI/SNF complex. Science *273*, 513–516.

Paoletti, J., Magee, B.B., and Magee, P.T. (1977). The structure of chromatin: interaction of ethidium bromide with native and denatured chromatin. Biochemistry *16*, 351–357.

Pennings, S., Meersseman, G., and Bradbury, E.M. (1994). Linker histones H1 and H5 prevent the mobility of positioned nucleosomes | PNAS. Proc. Natl. Acad. Sci. *91*, 10275–10279.

Perales, R., Zhang, L., and Bentley, D. (2011). Histone Occupancy In Vivo at the 601 Nucleosome Binding Element Is Determined by Transcriptional History ▿. Mol. Cell. Biol. *31*, 3485–3496.

Peterson, C.L., and Hansen, J.C. (2008). Chicken Erythrocyte Histone Octamer Preparation. Cold Spring Harb. Protoc. *2008*, pdb.prot5112.

Phansalkar, N., More, S., Sabale, A., and Joshi, M. (2011). Adaptive local thresholding for detection of nuclei in diversity stained cytology images. In 2011 International Conference on Communications and Signal Processing, pp. 218–220.

Poirier, M.G., Bussiek, M., Langowski, J., and Widom, J. (2008). Spontaneous access to DNA target sites in folded chromatin fibers. J. Mol. Biol. *379*, 772–786.

Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., Canfield, T.K., et al. (2014). Topologically associating domains are stable units of replication-timing regulation. Nature *515*, 402–405.

Radman-Livaja, M., and Rando, O.J. (2010). Nucleosome positioning: How is it established, and why does it matter? Dev. Biol. *339*, 258–266.

Ramirez-Carrozzi, V.R., Braas, D., Bhatt, D.M., Cheng, C.S., Hong, C., Doty, K.R., Black, J.C., Hoffmann, A., Carey, M., and Smale, S.T. (2009). A Unifying Model for the Selective Regulation of Inducible Transcription by CpG Islands and Nucleosome Remodeling. Cell *138*, 114–128.

Reinke, H., and Hörz, W. (2004). Anatomy of a hypersensitive site. Biochim. Biophys. Acta BBA - Gene Struct. Expr. *1677*, 24–29.

Ricci, M.A., Manzo, C., García-Parajo, M.F., Lakadamyali, M., and Cosma, M.P. (2015). Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes In Vivo. Cell *160*, 1145–1158.

Risca, V.I., Denny, S.K., Straight, A.F., and Greenleaf, W.J. (2017). Variable chromatin structure revealed by in situ spatially correlated DNA cleavage mapping. Nature *541*, 237–241.

Robinson, P.J., and Rhodes, D. (2006a). Structure of the '30 nm' chromatin fibre: A key role for the linker histone. Curr. Opin. Struct. Biol. *16*, 336–343.

Robinson, P.J.J., Fairall, L., Huynh, V.A.T., and Rhodes, D. (2006b). EM measurements define the dimensions of the "30-nm" chromatin fiber: Evidence for a compact, interdigitated structure. Proc. Natl. Acad. Sci. U. S. A. *103*, 6506–6511.

Robinson, P.J.J., An, W., Routh, A., Martino, F., Chapman, L., Roeder, R.G., and Rhodes, D. (2008). 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. J. Mol. Biol. *381*, 816–825.

Rodríguez-Campos, A., and Azorín, F. (2007). RNA Is an Integral Component of Chromatin that Contributes to Its Structural Organization. PLOS ONE *2*, e1182.

Rogge, R.A., Kalashnikova, A.A., Muthurajan, U.M., Porter-Goff, M.E., Luger, K., and Hansen, J.C. (2013). Assembly of Nucleosomal Arrays from Recombinant Core Histones and Nucleosome Positioning DNA. J. Vis. Exp. JoVE.

Routh, A., Sandin, S., and Rhodes, D. (2008). Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. Proc. Natl. Acad. Sci. U. S. A. *105*, 8872–8877.

Rubin, R.L., and Moudrianakis, E.N. (1972). Co-operative binding of histones to DNA. J. Mol. Biol. *67*, 361–374.

Sambrook, J., and Russell, D.W. (2001). Molecular Cloning: A Laboratory Manual (CSHL Press).

Samejima, K., and Earnshaw, W.C. (2005). Trashing the genome: the role of nucleases during apoptosis. Nat. Rev. Mol. Cell Biol. *6*, 677–688.

Scacchetti, A., Brueckner, L., Jain, D., Schauer, T., Zhang, X., Schnorrer, F., Steensel, B. van, Straub, T., and Becker, P.B. (2018). CHRAC/ACF contribute to the repressive ground state of chromatin. Life Sci. Alliance *1*, e201800024.

Schalch, T., Duda, S., Sargent, D.F., and Richmond, T.J. (2005). X-ray structure of a tetranucleosome and its implications for the chromatin fibre. Nature *436*, 138–141.

Schwabish, M.A., and Struhl, K. (2004). Evidence for Eviction and Rapid Deposition of Histones upon Transcriptional Elongation by RNA Polymerase II. Mol. Cell. Biol. *24*, 10111–10117.

Schwarz, P.M., and Hansen, J.C. (1994). Formation and stability of higher order chromatin structures. Contributions of the histone octamer. J. Biol. Chem. *269*, 16284–16289.

Segal, E., and Widom, J. (2009a). What controls nucleosome positions? Trends Genet. *25*, 335–343.

Segal, E., and Widom, J. (2009b). Poly(dA:dT) Tracts: Major Determinants of Nucleosome Organization. Curr. Opin. Struct. Biol. *19*, 65–71.

Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thåström, A., Field, Y., Moore, I.K., Wang, J.-P.Z., and Widom, J. (2006). A genomic code for nucleosome positioning. Nature *442*, 772–778.

Sekinger, E.A., Moqtaderi, Z., and Struhl, K. (2005). Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. Mol. Cell *18*, 735–748.

Shimamoto, Y., Tamura, S., Masumoto, H., Maeshima, K., and Misteli, T. (2017). Nucleosome–nucleosome interactions via histone tails and linker DNA regulate nuclear rigidity. Mol. Biol. Cell *28*, 1580–1589.

Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions. Science *311*, 844–847.

Simpson, R.T., Thoma, F., and Brubaker, J.M. (1985). Chromatin reconstituted from tandemly repeated cloned DNA fragments and core histones: a model system for study of higher order structure. Cell *42*, 799–808.

Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.-M., Zhu, P., and Li, G. (2014). Cryo-EM Study of the Chromatin Fiber Reveals a Double Helix Twisted by Tetranucleosomal Units. Science *344*, 376–380.

Spadafora, C., Bellard, M., Lee Compton, J., and Chambon, P. (1976). The DNA repeat lengths in chromatins from sea urchin sperm and gastrula cells are markedly different. FEBS Lett. *69*, 281–285.

Sridhar, B., Rivas-Astroza, M., Nguyen, T.C., Chen, W., Yan, Z., Cao, X., Hebert, L., and Zhong, S. (2017). Systematic Mapping of RNA-Chromatin Interactions In Vivo. Curr. Biol. *27*, 602–609.

Staynov, D.Z. (2000). DNase I digestion reveals alternating asymmetrical protection of the nucleosome by the higher order chromatin structure. Nucleic Acids Res. *28*, 3092–3099.

Stevens, T.J., Lando, D., Basu, S., Atkinson, L.P., Cao, Y., Lee, S.F., Leeb, M., Wohlfahrt, K.J., Boucher, W., O'Shaughnessy-Kirwan, A., et al. (2017). 3D structures of individual mammalian genomes studied by single-cell Hi-C. Nature *544*, 59–64.

Strom, A.R., Emelyanov, A.V., Mir, M., Fyodorov, D.V., Darzacq, X., and Karpen, G.H. (2017). Phase separation drives heterochromatin domain formation. Nature *547*, 241–245.

Struhl, K. (1985). Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. Proc. Natl. Acad. Sci. *82*, 8419–8423.

Struhl, K., and Segal, E. (2013). Determinants of nucleosome positioning. Nat. Struct. Mol. Biol. *20*, 267–273.

Suter, B., Schnappauf, G., and Thoma, F. (2000). Poly(dA·dT) sequences exist as rigid DNA structures in nucleosome-free yeast promoters in vivo. Nucleic Acids Res. *28*, 4083–4089.

Szerlong, H.J., and Hansen, J.C. (2011). Nucleosome distribution and linker DNA: connecting nuclear function to dynamic chromatin structure. Biochem. Cell Biol. Biochim. Biol. Cell. *89*, 24–34.

Takata, H., Hanafusa, T., Mori, T., Shimura, M., Iida, Y., Ishikawa, K., Yoshikawa, K., Yoshikawa, Y., and Maeshima, K. (2013). Chromatin compaction protects genomic DNA from radiation damage. PloS One *8*, e75622.

Taniguchi, T., and Takayama, S. (1986). High-order structure of metaphase chromosomes: evidence for a multiple coiling model. Chromosoma *93*, 511–514.

Tate, P.H., and Bird, A.P. (1993). Effects of DNA methylation on DNA-binding proteins and gene expression. Curr. Opin. Genet. Dev. *3*, 226–231.

Tate, V.E., and Philipson, L. (1979). Parental adenovirus DNA accumulates in nucleosome-like structures in infected cells. Nucleic Acids Res. *6*, 2769–2785.

Thåström, A., Lowary, P.T., Widlund, H.R., Cao, H., Kubista, M., and Widom, J. (1999). Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences1. J. Mol. Biol. *288*, 213–229.

Therizols, P., Illingworth, R.S., Courilleau, C., Boyle, S., Wood, A.J., and Bickmore, W.A. (2014). Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. Science *346*, 1238–1242.

Thoma, F., Koller, T., and Klug, A. (1979). Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. J. Cell Biol. *83*, 403–427.

Thomas, J.O., Rees, C., and Finch, J.T. (1992). Cooperative binding of the globular domains of histones H1 and H5 to DNA. Nucleic Acids Res. *20*, 187–194.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. Nature *489*, 75–82.

Tirosh, I., and Barkai, N. (2008). Two strategies for gene regulation by promoter nucleosomes. Genome Res. *18*, 1084–1091.

Truong, D.M., and Boeke, J.D. (2017). Resetting the Yeast Epigenome with Human Nucleosomes. Cell *171*, 1508-1519.e13.

Turner, A.L., Watson, M., Wilkins, O.G., Cato, L., Travers, A., Thomas, J.O., and Stott, K. (2018). Highly disordered histone H1−DNA model complexes and their condensates. Proc. Natl. Acad. Sci. 201805943.

Valouev, A., Johnson, S.M., Boyd, S.D., Smith, C.L., Fire, A.Z., and Sidow, A. (2011). Determinants of nucleosome organization in primary human cells. Nature *474*, 516–520.

Villeponteau, B., and Martinson, H.G. (1987). Gamma rays and bleomycin nick DNA and reverse the DNase I sensitivity of beta-globin gene chromatin in vivo. Mol. Cell. Biol. *7*, 1917–1924.

Villeponteau, B., Lundell, M., and Martinson, H. (1984). Torsional stress promotes the DNAase I sensitivity of active genes. Cell *39*, 469–478.

Wang, M.D., Schnitzer, M.J., Yin, H., Landick, R., Gelles, J., and Block, S.M. (1998). Force and Velocity Measured for Single Molecules of RNA Polymerase. Science *282*, 902–907.

Weintraub, H., and Groudine, M. (1976). Chromosomal subunits in active genes have an altered conformation. Science *193*, 848–856.

Widlak, P., and Garrard, W.T. (2006). Unique features of the apoptotic endonuclease DFF40/CAD relative to micrococcal nuclease as a structural probe for chromatin. Biochem. Cell Biol. Biochim. Biol. Cell. *84*, 405–410.

Widlak, P., Li, P., Wang, X., and Garrard, W.T. (2000). Cleavage Preferences of the Apoptotic Endonuclease DFF40 (Caspase-activated DNase or Nuclease) on Naked DNA and Chromatin Substrates. J. Biol. Chem. *275*, 8226–8232.

Widom, J. (2001). Role of DNA sequence in nucleosome stability and dynamics. Q. Rev. Biophys. *34*, 269–324.

Williams, S.P., and Langmore, J.P. (1991). Small angle x-ray scattering of chromatin. Radius and mass per unit length depend on linker length. Biophys. J. *59*, 606–618.

Williamson, I., Berlivet, S., Eskeland, R., Boyle, S., Illingworth, R.S., Paquette, D., Dostie, J., and Bickmore, W.A. (2014). Spatial genome organization: contrasting

views from chromosome conformation capture and fluorescence in situ hybridization. Genes Dev. *28*, 2778–2791.

Williamson, I., Lettice, L.A., Hill, R.E., and Bickmore, W.A. (2016). Shh and ZRS enhancer colocalisation is specific to the zone of polarising activity. Development *143*, 2994–3001.

Wolffe, A. (1998). Chromatin: Structure and Function (Academic Press).

Woodcock, C.L. (1994). Chromatin fibers observed in situ in frozen hydrated sections. Native fiber diameter is not correlated with nucleosome repeat length. J. Cell Biol. *125*, 11–19.

Woodcock, C.L., Frado, L.L., and Rattner, J.B. (1984). The higher-order structure of chromatin: evidence for a helical ribbon arrangement. J. Cell Biol. *99*, 42–52.

Woodcock, C.L., Skoultchi, A.I., and Fan, Y. (2006). Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. Chromosome Res. *14*, 17–25.

Wu, C., and Travers, A. (2019). Modelling and DNA topology of compact 2-start and 1-start chromatin fibres. Nucleic Acids Res.

Wu, C., McGeehan, J.E., and Travers, A. (2016). A metastable structure for the compact 30-nm chromatin fibre. FEBS Lett. *590*, 935–942.

Xiao, F., Widlak, P., and Garrard, W.T. (2007). Engineered apoptotic nucleases for chromatin research. Nucleic Acids Res. *35*, e93.

Yin, H., Wang, M.D., Svoboda, K., Landick, R., Block, S.M., and Gelles, J. (1995). Transcription Against an Applied Force. Science *270*, 1653–1657.

Yuan, G.-C., Liu, Y.-J., Dion, M.F., Slack, M.D., Wu, L.F., Altschuler, S.J., and Rando, O.J. (2005). Genome-Scale Identification of Nucleosome Positions in S. cerevisiae. Science *309*, 626–630.

Zhang, Y., Moqtaderi, Z., Rattner, B.P., Euskirchen, G., Snyder, M., Kadonaga, J.T., Liu, X.S., and Struhl, K. (2009). Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. Nat. Struct. Mol. Biol. *16*, 847–852.

Zhou, B., Li, X., Luo, D., Lim, D.-H., Zhou, Y., and Fu, X.-D. (2019a). GRID-seq for comprehensive analysis of global RNA–chromatin interactions. Nat. Protoc. 1.

Zhou, B.-R., Jiang, J., Feng, H., Ghirlando, R., Xiao, T.S., and Bai, Y. (2015). Structural Mechanisms of Nucleosome Recognition by Linker Histones. Mol. Cell *59*, 628–638.

Zhou, J., Fan, J.Y., Rangasamy, D., and Tremethick, D.J. (2007). The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. Nat. Struct. Mol. Biol. *14*, 1070–1076.

Zhou, K., Gaullier, G., and Luger, K. (2019b). Nucleosome structure and dynamics are coming of age. Nat. Struct. Mol. Biol. *26*, 3.

Zhuo, B., Yu, J., Chang, L., Lei, J., Wen, Z., Liu, C., Mao, G., Wang, K., Shen, J., and Xu, X. (2017). Quantitative analysis of chromatin accessibility in mouse embryonic fibroblasts. Biochem. Biophys. Res. Commun. *493*, 814–820.

Appendix 1

Vector Maps of all DNA Sequence Template Constructs

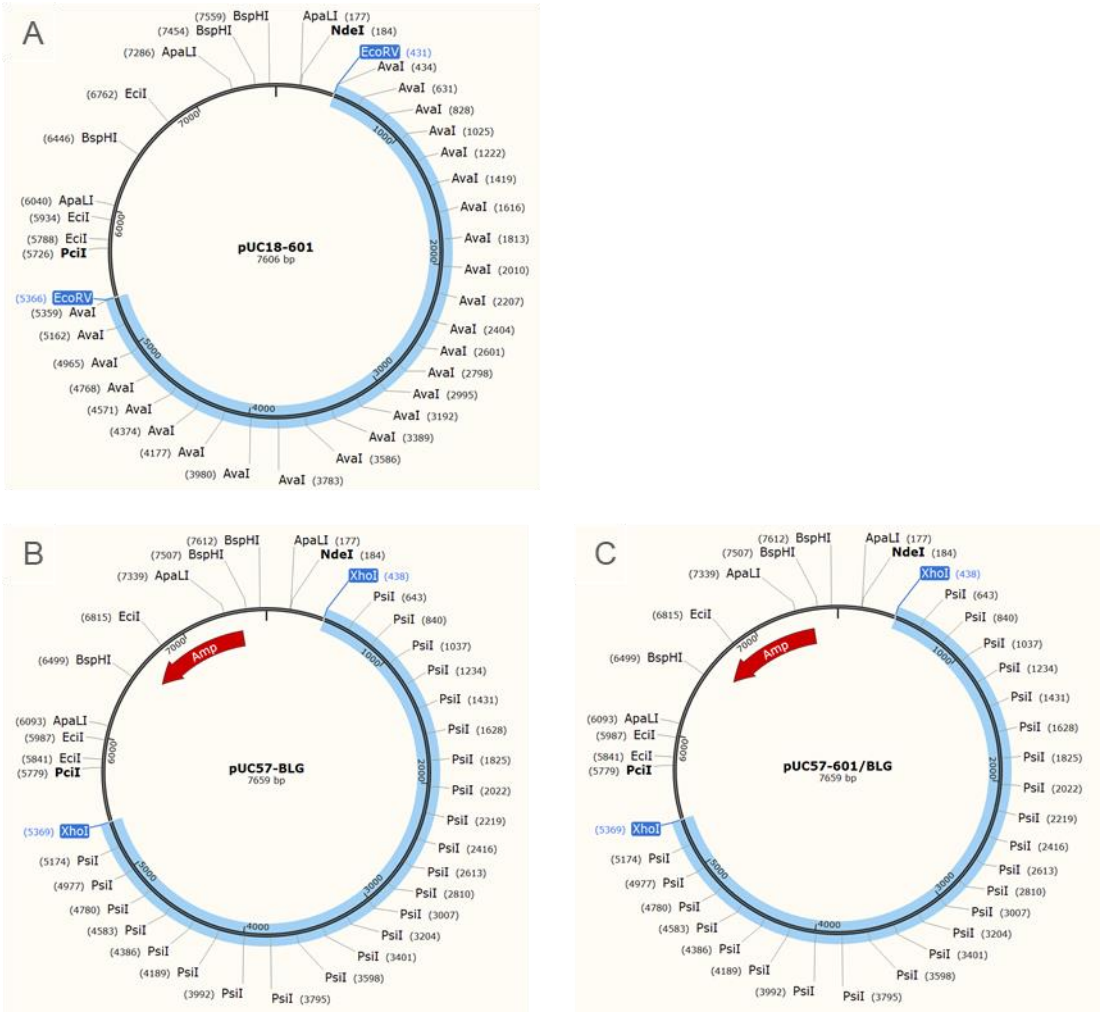

Figure S1. Vector maps of plasmid constructs containing 601, BLG and 601/BLG templates.
Sequence inserts are highlighted in blue. A) *601* in a pUC18 vector. B) *BLG* in a pUC57 vector. C) *601/BLG* in a pUC57 vector.

The vector maps of each of the DNA sequence template constructs, with restriction enzyme sites pertinent to this thesis, are shown in Figure S1. The template sequences are highlighted in blue. The 601/LA construct is identical to the 601, with only the central nucleosome positioning site altered. The 25 nucleosome positioning sites within each fibre are located between each of the AvaI or PsiI restriction sites. The sequences of eah of these sites are shown in Table S1. In the 601/BLG construct, alternating BLG nucleosome positioning sequences (BLG1, BLG3, BLG5 etc.) are replaced with the 601 sequence, as described in Figure 3.1A.

601	GAGAAGGTCGCTGTTCAATACATGCACAGGATGTATATATCTGACA CGTGCCTGGAGACTAGGGAGTAATCCCCTTGGCGGTTAAAACGCGG GGGACAGCGCGTACGTGCGTTTAAGCGGTGCTAGAGCTTGCTACGA CCAATTGAGCGGCCTCGGCACCGGGATTCTCCAGGGCGGCCGCGTA TAGGGTCCATCCC
BLG1	GATCCCTGGTCAGGGAACCATTAATAAGATCCCACATGCTGCAGGG CAACAAGCCCCCAAGCTGCAACCACTGAGCTGCAACCGCTGCAGTG CCCACAGGCCACGACCAGAGAAAGCCCACATACAGCAGGGAAGAC CCAGCACAACCGGAAAAAGGAGTTTGGTGGAATACAGCTGTGAAG CCGTCTGGTCCTGGA
BLG2	TTTCTATTTCTTCCGGGTTCAGTCTTGGGAGATTGTACATGCCTAGG AATGTGTCCGTTTCTTCTAGGTTGTCCATTTTATTGGACATGCATGG GAGCACACAGCACCGACCAGCGAGACTCATGCTGGCTTCCTGGGGC CAGGGCTGGGGCCCCAAGCAGCATGGCATCCTAGAGTGTGTGAAA GCCCACTGACCC
BLG3	GATTGGTGGCACCCAGATTTCCTAAGCTCGCTGGGGAACAGGGCGC TTGTTTCTCCCTGGCTGACCTCCCTCCTCCCTGCATCACCCAGTTCTG AAAGCAGAGCGGTGCTGGGGTCACAGCCTCTCGCATCTAACGCCGG TGTCCAAACCACCCGTGCTGGTGTTCGGGGGGCTACCTATGGGGAA GGGCTTCTCAC
BLG4	CCTCCAGAGGCTCCAGGGAGGGATCCTTGCCCCCCCGCTGCTGCCT CCAGCTCCTGGTGCCGCACCCTTGAGCCTGATCTTGTAGACGCCTCA GTCTAGTCTCTGCCTCCGTGTTCACACGCCTTCTCCCCATGTCCCCT CCGTGTCCCCGTTTTCTCTCACAAGGACACCGGACATTAGATTAGCC CCTGTTCCAG
BLG5	CTGCTACAGTCACCAACAGTCTCTCTGGGAAGGAAACCAGAGGCCA GAGAGCAAGCCGGAGCTAGTTTAGGAGACCCCTGAACCTCCACCCA AGATGCTGACCAGGCCAGCGGGCCCCCTGGAAAGACCCTACAGTTC AGGGGGGAAGAGGGGCTGACCCGCCAGGTCCCTGCTATCAGGAGA CATCCCCGCTATCA
BLG6	AGGCACAAGGCACCCACAGCCTGCTGGGTACCGACGCCCATGTGG ATTCAGCCAGGAGGCCTGTCCTGCACCCTCCCTGCTCGGGCCCCCTC TGTGCTCAGCAACACACCCAGCACCAGCATTCCCGCTGCTCCTGAG

	GTCTGCAGGCAGCTCGCTGTAGCCTGAGCGGTGTGGAGGGAAGTGTCCTGGGAGATTTA
BLG7	CGTCCTGGGGTTATTATGACTCTTGTCATTGCCATTGCCATTTTTGCTACCCTAACTGGGCAGCAGGTGCTTGCAGAGCCCTCGATACCGACCAGGTCCTCCCTCGGAGCTCGACCTGAACCCCATGTCACCCTTGCCCCAGCCTGCAGAGGGTGGGTGACTGCAGAGATCCCTTCACCCAAGGCCACGGTCACATG
BLG8	GGCTCTGACCTGTCCTTGTCTAAGAGGCTGACCCCGGAAGTGTTCCTGGCACTGGCAGCCAGCCTGGACCCAGAGTCCAGACACCCACCTGTGCCCCCGCTTCTGGGGTCTACCAGGAACCGTCTAGGCCCAGAGGGGGACTTCCTGCTTGGCCTTGGATGGAAGAAGGCCTCCTATTGTCCTCGTAGAGGAAGCCA
BLG9	GCCTGAGGATGAGCCAAGTGGGATTCCGGGAACCGCGTGGCTGGGGGCCCAGCCCGGGCTGGCTGGCCTGCATGCCTCCTGTATAAGGCCCCAAGCCTGCTGTCTCAGCCCTCCACTCCCTGCAGAGCTCAGAAGCACGACCCCAGCTGCAGCCATGAAGTGCCTCCTGCTTGCCCTGGGCCTGGCCCTCGCCTGTG
BLG10	GCCTGGCCCTCGCCTGTGGCGTCCAGGCCATCATCGTCACCCAGACCATGAAAGGCCTGGACATCCAGAAGGTTCGAGGGTGGCCGGGTGGGTGGTGAGTTGCAGGGCGGGCAGGGGAGCTGGGCCTCAGAGACCAAGAGAGGCTGTGACGTTGGGTTCCCATCAGTCAGCTAGGGCCACCTGACAAATCCCCCGCT
BLG11	GCATTCTGGAGGCTGGAAGCCCAAGATCCAGGTGTTGGCAGGGCTGGCTTCTCCTGCGGCCGCTCTCTGGGGAGCAGACGGCCGTCTTCTCCATGTCCTCTGCGCGCCCTGATTTCCTCTTCCTGTGAGGCCACCAGGCCTGCTGGAAACACGCCTGCCTGCGCAGCTTCACACGACCTTTGTCATCTCTTTAAAGG
BLG12	GGATGCCCAGAGTGCCCCCCTGAGAGTGTACGTGGAGGAGCTGAAGCCCACCCCCGAGGGCAACCTGGAGATCCTGCTGCAGAAATGGTGGGCGTCTCTCCCCAACATGGAACCCCCACTCCCCAGGGCTGTGGACCCCCCGGGGGGTGGGGTGCAGGAGGGACCAGGGCCCCAGGGCTGGGGAAGAGGGCTCAGAG

BLG13	TGGAAACGAAGCAGTGTGGGGATAGGCCCGTGTGAAGGCTGCTGG GAGGCAGCAGACCTGGGTCTTCGGGGCTCAAGCAGTTCCCGCTACC AGCCCTGTCCACCCTCAGACGGGGGTCAGGGTGCAGGAGAGAGCT GGATGGGTGTGGGGGCAGAGATGGGGACCTGAACCCCAGGGCTGC CTTTTGGGGGTGCCTG
BLG14	GATCGATGGTGAGTGCCGGGTCCCTGGGGGACACCCACCACCCCCG CCCCCGGGGACTGTGGACAGGTTCAGGGGGCTGGCGTCGGGCCCTG GGATGCTAAGGGACTGGTGGTGATGAAGACACTGCCTTGACACCTG CTTCACTTGCCTCCCTGCCACCTGCCCGGGGCCTTGGGGCCGGTGGC CATGGGCAGGTC
BLG15	CGGCTGGGCGGGGCTAACCCACCAGGGTGACACCCGAGCTCTCTTT GCTGGGGGGCGGGCGGTGCTCTGGGCCCTCAGGCTGAGCTCAGGA GGTACCTGTGCCCTCCCAGGGGTAACCGAGAGCCGTTGCCCACTCC AGGGGCCCAGGTGCCCCACGACCCCAGCCCGCTCCACAGCTCCTTC ATCTCCTGGAGACA
BLG16	CTGACTGGAGGCCCTGCACTGACTGACGCCAGGGTGCCCAGCCCAG GGTCTCTGGCGCCATCCAGCTGCACTGGGTTTGGGTGCTGGTCCTGC CCCCAAGCTGCCCGGACACCACAGGGCAGCCGGGGCTGCCCACTG GCCTCGGTCAGGGTGAGCCCCAGCTGCCCCCGCTCAGGGCTTGCCC CCGACAATGACCC
BLG17	TCCAGAGTTGACAGTGAGGGCTTCCTGGGCCCCATGCGCCTGGCAG TGGCAGCAGGGAAGAGGAAGCACCATTTCAGGGGTGGGGGATGCC AGAGGCGCTCCCCACCCCGTCTTCGCCGGGTGGTGACCCCGGGGGA GCCCCGCTGGTCGTGGAGGGTGCTGGGGGCTGACTAGCAACCCCTC CCCCCCCGTTGGAA
BLG18	CGCGTCCAGCCTTGAATGAGAACAAAGTCCTTGTGCTGGACACCGA CTACAAAAAGTACCTGCTCTTCTGCATGGAAAACAGTGCTGAGCCC GAGCAAAGCCTGGCCTGCCAGTGCCTGGGTGGGTGCCAACCCTGGC TGCCCAGGGAGACCAGCTGCGTGGTCCTTGCTGCAACAGGGGGTGG GGGGTGGGAGCTT
BLG19	GTCCCTGAGTCCCGCCAGGAGAGAGTGGTCGCATACCGGGAGCCA GTCTGCTGTGGGCCTGTGGGTGGCTGGGGACGGGGGCCAGACACAC AGGCCGGGAGACGGGTGGGCTGCAGAACTGTGACTGGTGTGACCG

	TCGCGATGGGGCCGGTGGTCACTGAATCTAACAGCCTTTGTTACCG GGGAGTTTCAATTAT
BLG20	CCCAAAATAAGAACTCAGGTACAAAGCCATCTTTCAACTATCACAT CCTGAAAACAAATGGCAGGTGACATTTTCTGTGCCGTAGCAGTCCC ACTGGGCATTTTCAGGGCCCCTGTGCCAGGGGGGCGCGGGCATCGG CGAGTGGAGGCTCCTGGCTGTGTCAGCCGGCCCAGGGGGAGGAAG GGACCCGGACAGCC
BLG21	ACCTGCAGACCCACTGCACTGCCCTGGGAGGAAGGGAGGGGAACT AGGCCAAGGGGGAAGGGCAGGTGCTCTGGAGGGCAAGGGCAGACC TGCAGACCACCCTGGGGAGCAGGGACTGACCCCCGTCCCTGCCCCA TAGTCAGGACCCCGGAGGTGGACAACGAGGCCCTGGAGAAATTCG ACAAAGCCCTCAAGGC
BLG22	ACGTCCTGGGCACACACATGGGGTAGGGGGTCTTGGTGGGGCCTGG GACCCCACATCAGGCCCTGGGGTCCCCCCTGTGAGAATGGCTGGAA GCTGGGGTCCCTCCTGGCGACTGCAGAGCTGGCTGGCCGCGTGCCC ACTCTTGTGGGGTGACCTGTGTCCTGGCCTCACACACTGACCTCCTC CAGCTCCTTCCA
BLG23	TCTCCTTCACCAATAAAGGCATAAACCTGTGCTCTCCCTTCTGAGTC TTTGCTGGACGACGGGCAGGGGGTGGAGAAGTGGTGGGGAGGGAG TCTGGCTCAGAGGATGACAGCGGGGCTGGGATCCAGGGCGTCTGCA TCACAGTCTTGTGACAACTGGGGGCCCACACACATCACTGCGGCTC TTTGAAACTTTCA
BLG24	GAGAAGGGGACGACAGAGGATGAGATGGTTGGATGGCATCACCAA CACAATGGACATGGGTTTGGGTGGACTCCAGGAGTTGGTGATGGAC AGGGAGGCCTGGCGTGCTGCGGTTTATGGGGTCACAAAGACTGAGT GACTGAACTGAGCTGAACTGAATGGAAATGAGGTATACAGCAAAG TGGGGATTTTTTAGA
BLG25	CACTTAATTACCAAAGCTGCTCCAAGAAAAAGCCCCTGTGCCTCTG AGCTTCCCGGCCTGCAGAGGGTGGTGGGGGTAGACTGTGACCTGGG AACACCCTCCCGCTTCAGGACTCCCGGGCCACGTGACCCACAGTCC TGCAGACAGCCGGGTAGCTCTGCTCTTCAAGGCTCATTATCTTTAAA AAAAACTGAGGT

LA	GATGCTTCAGAACATCATCAAACAAATGAACATAAAACATCATTTT TGTTTACTTGGAAGGGGAGATAAAATCCACTGAAGTGGAAATGCAT AGCAAAGATACATACAATGAGGCAGGTATTCTGAATTCCCTGTTAG TCTGAGGATTACAAGTGTATTTGAGCAACAGAGAGACATTTTCATC ATTTCTAGTCTGA

Table S1.
DNA sequences of the 197 bp 601 repeat, each 197 bp site derived from the β-lactoglobulin gene used within the BLG template, and the 197 bp low affinity nucleosome positioning site derived from the β-lactoglobulin gene.

Appendix 2

Methods of Reconstitution by Salt Dialysis

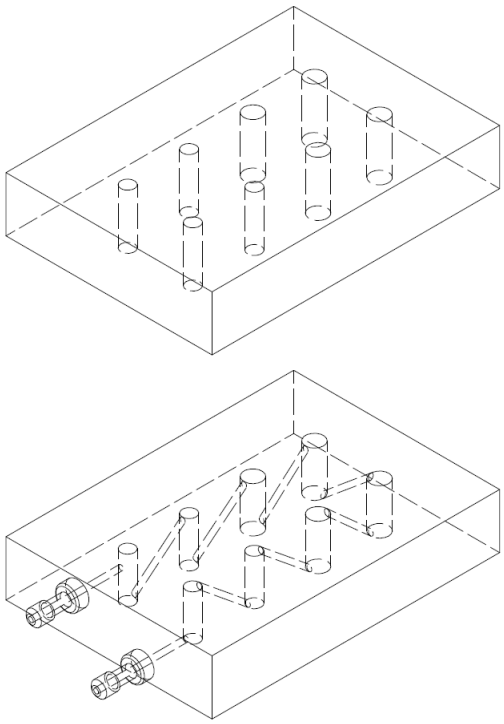

Figure S2. 3D Modelled Dialysis Reconstitution Unit.

During the two-step reconstitution dialysis used throughout this thesis (and the one-step reconstitution dialysis used to prepare chromatin for single-molecule force spectroscopy), chromatin was placed within a Slide-A-Lyzer MINI Dialysis Device with 10,000 MWCO (Thermo Fisher). These caps hold up to 100µl. Larger devices are available, but fit into 15 or 50 ml tubes and are therefore unsuitable for gradient dialysis.

To dialyse volumes greater that 100µl, a dialysis reconstitution unit (Figure S2) was 3D printed. A dialysis membrane with 10,000 MWCO was placed between the upper and lower units which were then clamped together, creating 8 adjoining chambers in the lower unit and 8 reservoirs in the upper unit. Chromatin samples were placed in the reservoirs in the upper units and a salt gradient solution was pumper through the lower unit. Chambers of different diameters hold 100µl or 200µl of chromatin. This unit was printed in VeroClear (Stratasys) which is inert and transparent, allowing the user to see if bubbles were accumulating inside the chambers in the lower unit.

While this effectively dialysed chromatin, the rate of salt dialysis was found to be slower than when using the Slide-A-Lyzer devices, most likely due to differences in the type of

membrane used. For consistency, larger amounts of chromatin were therefore reconstituted in multiple Slide-A-Lyzer dialysis devices and combined following dialysis.

Appendix 3

Comparison of Chicken and *Xenopus* Core Histone Titrations

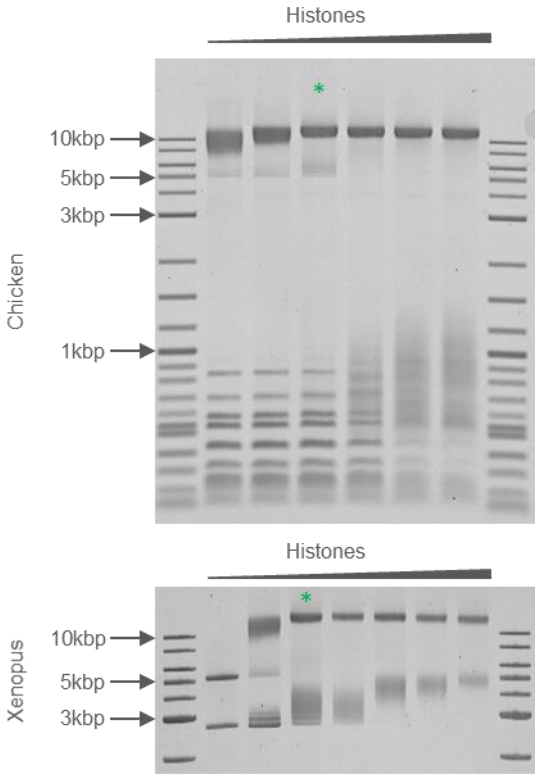

Figure S3. Titrations of *Xenopus* and Chicken Core Histone Octamers.
Gel showing shift of *601* template reconstituted with chicken octamers at ratios of 1, 1.2, 1.4, 1.6, 1.8 and 2:1 in the presence of a digested vector backbone or *Xenopus* octamers reconstituted at 0, 1.20, 1.40, 1.65, 1.90, 2.10 and 2.35:1 ratios in the presence of the undigested vector backbone.

In this thesis, both recombinant *Xenopus* and purified chicken erythrocyte core histones were used to reconstitute chromatin for different experiments. As described in section 3.5, a 1.4:1 ratio of *Xenous* octamers:DNA was found to be required to saturate the *601* template in the presence of a competitor, and a similar ratio was found to be required to saturate a purified template with chicken histones. Here, both chicken and *Xenopus* core histones are used to reconstitute the *601* template in the presence of a competitor (vector backbone competitor has been digested to smaller fragments when reconstituted with chicken histones, which would not be expected to impact these results). As indicated by the asterisk, both of these templates appear to saturate at a 1.4:1 ratio, showing that there is no difference in the ability of each type of histone to reconstitute these templates.