EPFL

Thèse n° 8932

Dynamic Chromatin Invasion and Remodeling by the Transcription Factor Rap1

Présentée le 3 décembre 2021

Faculté des sciences de base
Laboratoire de chimie biophysique des macromolécules
Programme doctoral en chimie et génie chimique

pour l'obtention du grade de Docteur ès Sciences

par

Maxime Nicolas MIVELAZ

Acceptée sur proposition du jury

Prof. K. Sivula, président du jury
Prof. B. Fierz, directeur de thèse
Prof. F. Mattiroli, rapporteuse
Prof. G. L. Hager, rapporteur
Prof. P. Rivera Fuentes, rapporteur

École polytechnique fédérale de Lausanne

2021

Acknowledgements

I would first like to thank Prof. F. Mattiroli, Prof. G. L. Hager and Prof. P. Rivera Fuentes for their contributions to this thesis and the insightful discussion during the defence.

Prof. Fierz, thank you for your mentorship during these past years, culminating in this thesis.

I extend my warmest thanks to past and present members of LCBM. Without your friendship, all these years would have seemed longer.

To my friends and family, thank you for your unconditional and infinite support.

Finally, to my wife, Lucile, tout simplement merci.

Abstract

Pioneer transcription factors (pTFs) bind to target sites within compact chromatin, initiating chromatin remodelling and control the recruitment of downstream factors important in gene regulation. Using single-molecule fluorescence approaches, we reconstituted a chromatin system replicating the yeast promoter architecture. We demonstrate that the yeast general transcription factor Rap1 can bind DNA within the nucleosome without altering nucleosome structure. Rap1 binding kinetics are modulated by DNA motif sequence, position within the nucleosome, and the tested local chromatin landscapes (free DNA, mono-nucleosome and chromatin fibres). Using single-molecule FRET, we observed that Rap1 binding decompacts chromatin fibres without inducing nucleosome loss. Rap1 must therefore collaborate with chromatin remodelers such as RSC to clear its binding sites from nucleosomes, generating the nucleosome free region observed in active promotor regions observed *in vivo*. We show using MNase-seq experiments that Rap1 biases the remodelling direction of RSC when present. Together, our results provide mechanistic insight into how Rap1 establishes an active promoter architecture and controls the initial phases of gene expression.

We further explore the collaboration of transcription factors and chromatin remodellers by establishing a modular single-molecule FRET-based assay capable of following mono-nucleosome remodelling in real-time. For our proof of concept, we used the TF-CR pair of Rap1 and yeast Chd1.

Key words: Rap1, Transcription Factors, Chromatin Remodeling, single-molecule TIRF, Chromatin Fibers

Résumé

Les facteurs de transcription pionniers sont capables de lire et se lier à l'ADN présente dans de la chromatine compactée. Une fois liés, ces facteurs peuvent recruter d'autres co-facteurs impliqués dans la régulation de l'expression génétique. Nous avons recréé l'architecture d'un promoteur type de levure que nous avons utilisé pour étudier la dynamique d'interaction d'un facteur de transcription pionnier en s'appuyant sur des méthodes de microscopie à base de fluorescence capable de suivre des molécules individuelles. Nous avons démontré que la protéine issue de la levure Rap1, est capable de se lier à ses motifs de reconnaissance d'ADN contenue dans un nucléosome.

De plus, nous avons constaté que la séquence, la position à l'intérieur du nucléosome ainsi que la complexité de l'environnement où se trouve le site de liaison (ADN libre, mono-nucléosome ou chromatine) influencent la cinétique de liaison. Une fois lié à un site de liaison dans un mono-nucléosome, Rap1 ne perturbe pas la structure du nucléosome en soit. Si le site de liaison se trouve à l'intérieur de la chromatine, nous observons une décompaction de la chromatine induite par Rap1. *In vivo*, les régions d'ADN contenues dans un promoteur de gène actif sont dépourvues de nucléosome.

Nos expériences conduites avec des complexes de remaniement de la chromatine tels que RSC (membre de la famille SWI/SNF) ont démontrées que ces derniers collaborent avec le facteur de transcription Rap1. En effet, la présence de Rap1 influence la direction du remaniement de la chromatine par RSC. Ensemble Rap1 et RSC sont capables de créer des régions dépourvues de nucléosome telles que celles observées *in vivo*.

Afin de mieux comprendre les mécanismes d'interactions entre les facteurs de transcription et les complexes de remaniement de la chromatine, nous avons mis en place une méthodologie capable de suivre le remaniement d'un mono-nucléosome en temps réel. Cette méthode s'appuie sur le FRET entre deux fluorophores, qui nous indique tous changements de distances à l'échelle moléculaire.

Mots clef : Rap1, Facteur de Transcription, Complexes de Remaniement de la Chromatine, Chromatine

Abbreviations

ATAQ-seq - Assay for Transposase-Accessible Chromatin with high throughput sequencing

bHLH – basic helix-loop-helix

BRCT - BRCA1 C-terminal

C2H2 ZF - Cys2-His2 Zinc finger

CBP - CREB-binding protein

CGI – CpG islands

CHD – Chromodomain-helicase DNA binding

ChIP – Chromatin immunoprecipitation

ChIP-Seq - Chromatin immunoprecipitation followed by DNA sequencing

Co-IP – Co-immunoprecipitation

CR – Chromatin Remodeller

CryoEM – Cryogenic Electron microscopy

CTCF - CCCTC-binding factor

CUT&RUN - Cleavage under targets and release using nuclease

DBD – DNA binding domain

dCAS – Dead cas9

DNA – Deoxyribose nucleic acid

DPE - Downstream promotor element

EMSA - Electrophoretic mobility shift assays

eRNA - enhancer RNA

FISH – Fluorescence in situ hybridization

FKBP - FK506 binding protein

FN – Fragile nucleosome

FRAP – Fluorescence recovery after photobleaching

FRB - FKBP12- rapamycin binding domain

FRET - Föster Resonance Energy Transfer

GFR - General transcription factors

GpC - Guanosine and cytosine sequences

HD – Homeodomain

HDAC - Histone deacetylases

HP1 - Heterochromatin protein 1

INO80 – Inositol requiring 80

Inr - Initiator motif

ISWI - Imitation switch

K_d – Affinity of binding

kDa – Kilo Dalton

k_{off} – Dissociation rate

k_{on} – Association rate

MD – Molecular dynamics

MLL - Mixed-lineage leukaemia

MNase – Monoccocal nuclease

NDR – Nucleosome depleted region

NFR – Nucleosome free region

NGS – Next generation sequencing

PDB – Protein Data Bank

PHD – Plant homeodomain

PIC - Preinitiation complex

PRC2 - Polycomb repressive complex 2

pTF – Pioneer Transcription Factor

PTM - Post-translational modification

PWM – Position weight matrix

Rap1 – Repressor Activator protein 1

RSC – Remodeling the structure of chromatin

SANT-SLIDE – (Swi3, ADA2, N-CoR, TFIIB – SANT-like ISWI domain)

Selex - Systematic evolution of ligands by exponential enrichment

smFRET – Single-molecule FRET

SPT – Single-particle tracking

SWI/SNF – Switch defective/sucrose nonfermenting

TAD - Topologically associated domains

TALES - Transcription activator-like factors

TBP – TATA-binding protein

TF – Transcription Factor

TIR – Total internal reflection

TSS – Transcription start site

ZF – Zinc finger

Contents

Acknowledgements iii

Abstract iv

Résumé v

Abbreviations vi

Contents ix

1.Introduction 1

1.1 DNA, Nucleosomes and Chromatin 2

1.2 The dynamics of DNA, Nucleosomes and Chromatin 6

1.3 Chromatin Effectors 10

1.3.1 Transcription Factors 10

1.3.2 Pioneer Transcription Factors 13

1.3.3 Chromatin Remodellers 16

1.4 Techniques for studying chromatin effectors 18

1.4.1 *In Vivo* Biological roles of chromatin and chromatin effectors 18

1.4.2 *In Vitro* studies of chromatin effectors 25

1.4.3 *In Vitro* Structural Biology 26

1.5 Hallmarks of DNA architecture in vivo 29

1.5.1 Heterochromatin and Euchromatin distinct nuclear regions 29

1.5.2 Architectural features of genetic Loci 32

2 Aims 37

3 Pioneer transcription factor chromatin-binding dynamics 39

3.1 The refractory effect of chromatin on transcription factor binding 40

3.1.1 The yeast general regulatory factor Repressor Activator Protein 1 (Rap1) 41

3.1.2 Cloning and Expression of Rap1 in sf9 insect cells 43

3.1.3 *In vitro* Rap1 characterization 46

3.2 Single-molecule characterization of Rap1 DNA binding 47
3.2.1 Single-molecule total internal reflection microscopy..... 47
3.2.2 smTIRF of Rap1 on free RPL30 promotor DNA..... 50
3.2.3 smTIRF of Rap1 on nucleosomes with RPL30 promotor DNA..... 52
3.2.4 The Rpl30 promotor region nucleosome set 52
3.2.5 Characterizing Rap1 binding to nucleosomes using EMSA..... 54
3.2.6 smTIRF characterization of Rap1 binding to nucleosomes..... 54
3.2.7 Induced nucleosome dynamics by Rap1 binding probed by ensemble FRET... 57
3.3 Single-molecule Rap1 binding to chromatin fibres 59
3.3.1 Cloning and reconstitution of chromatin arrays 59
3.3.2 Single-molecule colocalization TIRF on chromatin fibres..... 63
3.3.3 Single-molecule FRET on chromatin fibres 64
3.3.4 Rap1 binding to the WT RPL30 Promotor region 67
3.4 Discussion of Rap1 binding studies..... 68
4 Chromatin Remodelers as Chromatin Effector proteins..... 74
4.1 Chromatin remodelers as chromatin effector proteins..... 75
4.2 Expression and purification of RSC 76
4.3 RSC remodelling assays 77
4.4 Next-Generation Sequencing MNase-Seq RSC remodelling assays 80
4.5 Discussion of RSC remodelling studies..... 81
5 Single-molecule FRET-based real-time remodelling assays 85
5.1 Chd1 as a model Chromatin Remodeler 85
5.2 Expression and purification of yeast Chd1 86
5.3 Chd1 remodels nucleosomes to central positions 88
5.4 Generating asymmetric FRET nucleosomes..... 88

5.5 Ensemble FRET measurements using +70 Cy3B DNA and H2AN110C A647 asymmetric nucleosomes 90
5.6 smFRET measurements using +70 Cy3B DNA and H2AN110C A647 asymmetric nucleosomes 90
5.7 smFRET measurements using +70 Cy3B DNA and H2AN110C A647 symmetric nucleosomes 91
5.8 Discussion on real-time FRET-based TF-CR nucleosome remodelling 93
6 Thesis Summary and overall discussion 95
7 References 99
8 Materials and Methods 118
8.1 Expression and purification of Rap1-Halo 118
8.2 Expression and purification of recombinant histones 119
8.3 Large scale generation of recombinant plasmids 119
8.4 Large scale restriction digest and purification of recombinant plasmids 120
8.5 Oligonucleotide labelling 121
8.6 Production of labelled DNA fragments 121
8.7 Ligation and purification of 1 x 601 DNA to biotin anchor 122
8.8 Mononucleosome (Nuc) formation 122
8.9 Electrophoretic mobility shift assays (EMSA) 123
8.10 Convergent 3-piece convergent DNA ligation for the synthesis of 12x601 DNA ... 123
8.11 Reconstitution of 12-mer chromatin fibres 124
8.12 Preparation of microfluidic chambers for sm-FRET/TIRF experiments 125
8.13 Single-molecule TIRF (sm-TIRF) co-localization microscopy measurements 125
8.14 Photobleaching test for JF-549 Rap1-Halo 126
8.15 Ensemble FRET measurements 127
8.16 Single-molecule FRET (smFRET) measurements for Chd1 remodelling 127

8.17 Image processing, single-molecule trace extraction and trace analysis 128
8.18 Single-molecule FRET (smFRET) conformation analysis 129
8.19 Statistical analysis 131
8.20 Nucleosome shift assays with RSC, Nap1 and Rap1 131
8.21 RSC sliding and MNase-seq 132
8.22 Expression and purification of Chd1 133
8.23 Chd1 sliding 134
9 Supplementary Figure 135
10 Supplementary Tables 137
11 CV 140

1.Introduction

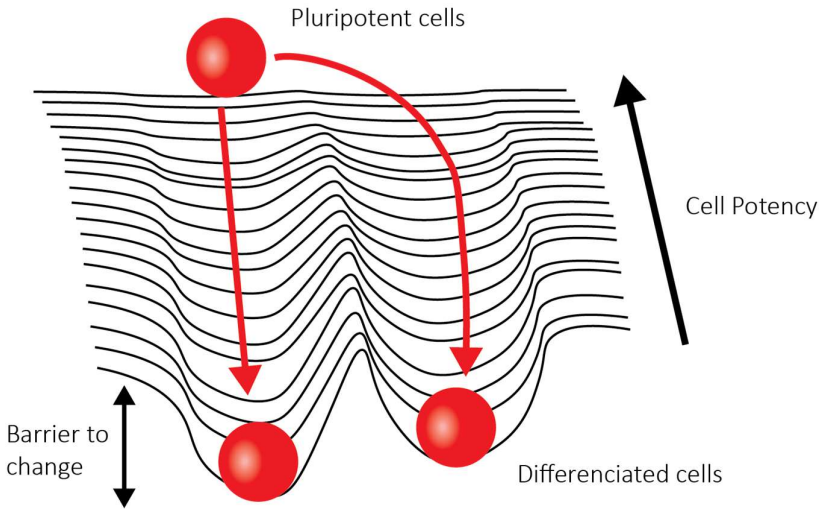


Figure 1 – Waddington landscape depicting cell differentiation.

Differentiation is a process by which cells transition from a pluripotent state (capable of self-renewal and able to form any cell type) to fully differentiated states (defined in shape and function). This process can be schematized using the Waddington landscape (**Fig 1**), which Conrad Waddington first proposed in 1957 (1). Cells, represented by a ball, start at the top of an inclined plane. During differentiation, the ball rolls down the plane to settle in different valleys representing cells in their fully differentiated state. The incline of the slope represents the change in cell potency. In contrast, the height of the valleys represents the barriers to change from one type of differentiated cell to another (transdifferentiating). One can appreciate that the heights of the valleys increase as we travel down the landscape, suggesting that the barriers to change increase as cells differentiate. As the DNA content of the cell during differentiation does not change, it was postulated that changes to factors beyond the genetic code (epigenetic) must be behind these barriers. Early experiments showed that fully differentiated cell nuclei injected into enucleated oocytes were capable of forming whole animals, demonstrating that in the right conditions, differentiation is a reversible process (2). Now, we understand that differentiation entails wholesale changes in the proteomes, epigenetics and transcriptional patterns of cells. Central to this process is gene expression regulation, where a myriad of proteins activate or repress specific genes. This fundamental process has been investigated during this thesis at the level of gene regulation, more precisely, the steps that initiate gene expression. This process requires DNA reading and effector proteins acting in concert, including transcription factors and chromatin remodellers.

Below I will provide an overview of the concepts and methodologies that this work is based on, as well as a state of the literature geared towards understanding the work highlighted in this thesis.

1.1 DNA, Nucleosomes and Chromatin

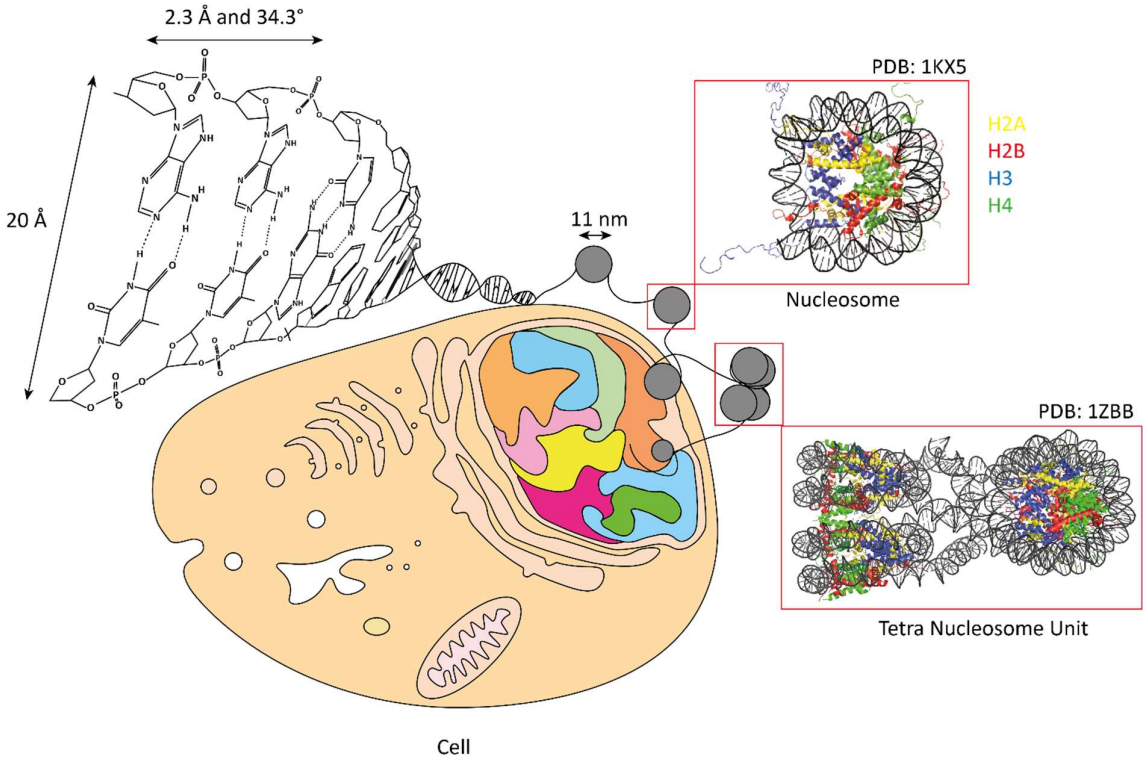


Figure 2- DNA forms a double-stranded helix that is wound 1.7 times around an octameric core consisting of histone proteins H2A, H2B, H3 and H4. These are called nucleosomes which can then interact with each other to form higher-order structures such as the tetrameric nucleosome unit. Together, the formation of nucleosomes allows for the compaction of 2 m of DNA into the cell's nucleus.

Deoxyribonucleic acid (DNA) molecules are biopolymers constituted of 4 nucleotides (Adenosine, Guanosine, Thymidine and Cytosine) together encoding all the information necessary for the function of all cell types. Nucleotides are monomeric units that include a nucleobase and deoxyribose. Monomers are covalently bonded together by a phosphodiester linkage forming a strand of DNA (**Fig. 2**) (3). Nucleobases contain hydrogen bond donors and acceptors; these allow the nucleobases to interact in a specific manner, with A – T and G -C forming interaction partners called base pairs (bp). This means that two anti-parallel DNA strands with complementary nucleobases can anneal to form the characteristic double-helix. This double helix can take different conformations, with different helix senses or overall structures. In cells, the B-DNA form is preferably found. B-DNA has a right-handed sense, a 20 Å diameter, with a distance of 2.3 Å and a rotation of 34.3° between each base (3). In this

conformation, ~10.4 bp are required for a whole twist, containing a major groove that is 22 Å wide and a minor groove that is 12 Å wide. Formation of this double helix from single strands to form duplex DNA is a spontaneous process involving hydrogen bonding between nucleobases and Van der Waals forces and the hydrophobic effect of nucleobase stacking. These favourable effects counteract the charge-charge repulsion between the phosphodiester backbones of each strand. The biophysical properties of DNA are governed by these forces making DNA a negatively charged rigid polymer with a persistence length of 40 and 75 nm at ionic strengths comparable to the cellular environment, which consists of 100 - 150 bp (4–6). This rigidity poses a problem for Eukaryotic cells, which possess a nucleus with an average diameter of 6 μm (7). In order to pack the genetic material within the nucleus, histone proteins associate with DNA to form the nucleosome (**Fig. 2**).

This DNA-protein complex is composed of an octameric core of histone proteins, including two copies of each H2A, H2B, H3 and H4, wrapped in 1.7 left-handed turns of DNA (8). This structure packs ~147 bp of DNA in an 11 nm diameter disk 5 nm thick, thus compacting DNA 4-fold compared to free DNA. Histones are among the most evolutionary conserved proteins. They are basic and positively charged, interacting with each other through hydrophobic packing (9–12). An estimated 80% of DNA in vivo is thought to be found in nucleosomes (13). Outside of the canonical core histones (H2A, H2B, H3 and H4), there exist other histones which confer different biophysical properties to the nucleosome. *In vivo* reconstitution of nucleosomes starts with the association of H3 – H4 and a single H2A-H2B dimer forming a hexasome. This is later followed by the final H2A- H2B dimer deposition via the chaperones (14,15). Nucleosomes exhibit 2-fold symmetry, and the position of this dyad symmetry axis passes through a centrally located nucleotide pair (16). This base pair (the 'dyad') is often used as a positional marker to describe super helical locations (SHL) of the DNA vis-à-vis the octameric core. This nomenclature starts from 0 at the dyad and assigns the minor grooves facing inward with half units (i.e. SHL 0.5, SHL 1.5 etc.) and outward-facing minor grooves as whole units (i.e. 1, 2, 3 etc.) (**Fig 3A-B**). DNA – Histone interactions are mainly mediated through electrostatic interactions from the positive histone surface and the negative DNA phosphodiester backbone. Within a nucleosome, several contacts exist between the DNA and octameric core; these occur at sites where the minor groove of the DNA faces towards the protein surface. These areas bring the phosphodiester backbone in close proximity to the core

histones, also allowing the formation of water-mediated hydrogen bonds with specific arginine's (R) (12,17) (**Fig 3B**). *In vitro* studies suggest that histone-DNA interactions contribute differently to nucleosome stability. At the dyad, interactions contribute the most and get weaker as you move towards the entry-exit sites (**Fig 3C**) (13). Additional interactions are made via the unstructured histone tails, and these regions are rich in lysines (K) and arginines and site within both major and minor grooves of DNA. This being said, the contribution to the thermal stability of the nucleosome from the histone tail – DNA interactions have been shown to be marginal (18).

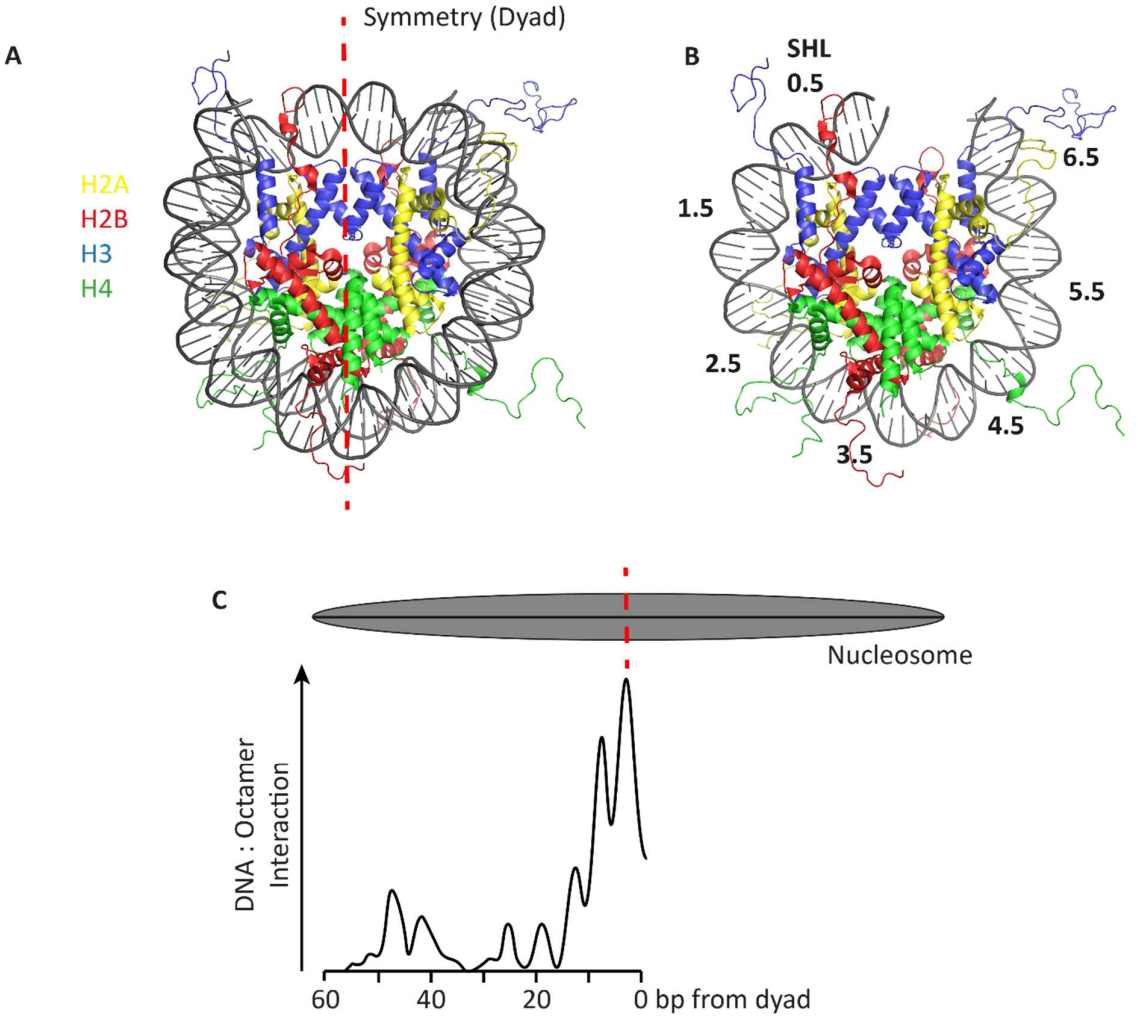


Figure 3 – A) The nucleosome has an axis of symmetry that passes through a centrally positioned nucleic acid at the dyad position. B) It is from the dyad that the superhelix locations (SHL) are defined, with the half numbers at minor grooves facing the octameric core. C) The interaction between DNA and the nucleosome is spread over the nucleosomes, with the strongest contact points near the dyad region.

Nucleosomes have the propensity to interact with each other, forming inter-nucleosomal contacts via a region called the acidic patch. This glutamic acid (E) and aspartic acid (D) rich region on histones H2A – H2B not only allows for the stacking of nucleosomes but is also a target for the recruitment of nucleosome targeting proteins.

Multiscale structures composed of DNA, proteins and RNA called chromatin have been observed *in vivo* and *in vitro.* Within cells, this structure is dynamic and changes during the cell cycle, becoming more compacted until cell division is completed. During interphase, some regions of the genome are compact (Heterochromatin) and others less compact (Euchromatin). These regions can be as simple as single nucleosomes scattered along DNA, much like beads on a string, or form condensed higher-order chromatin. *In vivo* observations have yielded various possible chromatin conformations with fibres ranging in diameter of 5 – 24 nm (19). Clusters of nucleosomes encompassing kilobases of DNA have been identified and named "clutches" (20). These clutches are interspersed between nucleosome depleted regions (20). Molecular dynamic simulations integrating data acquired *in vivo* on these nano-domains suggest that short linkers and small nucleosome-free regions favour more compact clutches (21,22). At the same time, longer linkers favour a more flexible overall assembly resembling a "sea of nucleosomes"(21,22). In contrast, *in vitro* reconstituted chromatin fibres have a highly defined 3D structure due to the packing of nucleosomes. Models of these fibres include the most energetically stable conformation, where each nucleosome in series is flanked by its neighbour (one-start of solenoid model) **(Fig4A)** (23,24). An alternative conformation proposed is the two-start model (**Fig 4A**). In this structure, nucleosomes are stacked between N -2 and N+2 nucleosomes with the DNA between each nucleosome (linker DNA) zig-zagging in the centre (25). The tetranucleosome presented above (**Fig 2**) is found in this two-start conformation and is the basic repeat unit. The centrally positioned linker DNA has been shown to be 20 – 100 bp in length and varies between species and in different tissues of the same organism (26). These structures have diameters of ~30 nm; hence, they are called 30 nm chromatin fibres. Although discrepancies exist between *in vivo* and *in vitro* structures at larger scales, with less regular chromatin *in vivo* being more globular and less well organized, the *in vitro* models remain a good approximation of small DNA regions (20).

Here I have presented a static view of DNA, nucleosomes and chromatin. In reality, these are all dynamic macromolecules that can interchange between free DNA, nucleosomes and

chromatin. These dynamics play a crucial role in gene regulation as they determine the access of DNA to essential DNA binding proteins.

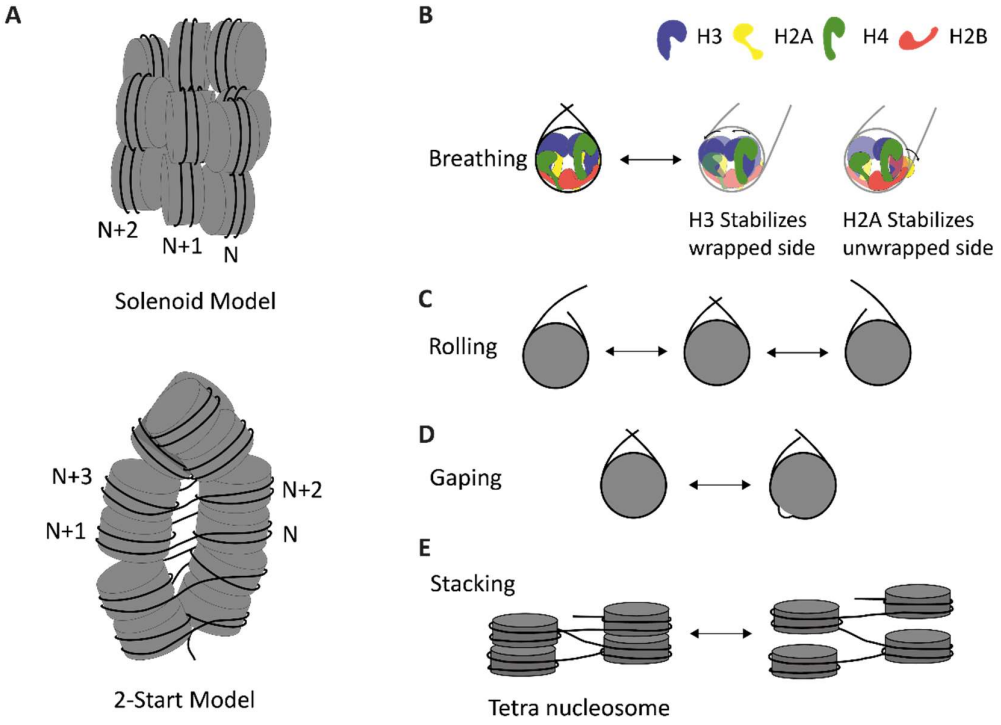


Figure 4 – A) Higher-order nucleosome stacking forms chromatin. The exact in vivo structure of chromatin is still unknown, with two models, the Solenoid model and the 2-Start model is proposed. B) Nucleosomes and chromatin are dynamic structures. This scheme depicts the different movements that have been observed.

1.2 The dynamics of DNA, Nucleosomes and Chromatin

Classic worm-like chain models of DNA predict stiffness over short distances (close to the persistence length of ~40-70 nm) but flexibility over large distances (27). This is mainly due to base-pair stacking and repulsion from the phosphodiester backbone, but the contribution of these two effects to stiffness remains contested. What has been observed is that local stiffness depends on sequence, ionic strength and the presence of DNA 'bending' proteins (indeed, DNA appears more flexible *in vivo*) (27). DNA's propensity to deform (twist, bend, stretch etc.) has proven very difficult to study. The fact that these properties are often linked, the size of these deformations and the timescales on which they occur render their study challenging (28,29). This being said, the contribution of DNA dynamics to protein – DNA interactions has been probed using molecular dynamics (MD) simulations and have indeed been found to be significant (30). Structural deformations along the DNA, known as Twist

defects, can occur spontaneously and be propagated along DNA or be induced by DNA interacting proteins, such as the nucleosome (31,32). These include under twisting where the turn per base pair is less than 25° and over-twisting with torsion angles per base pair of over 25° (28,33).

We can observe various modes of DNA movements on the nucleosome level, including breathing, sliding, and gaping (**Fig. 4B-D**) (34). DNA – histone interactions at the entry-exit site are weak, allowing the DNA to unwrap and wrap in a motion called DNA 'breathing' (35) **(Fig.4B)**. These millisecond thermal fluctuations spontaneously expose DNA that was shielded (both electrostatically and sterically) within the nucleosome, and this mechanism has been shown to be important in binding nucleosomal DNA by DNA binding proteins (36,37) further discussed later. Less favourable spontaneous sliding of DNA around a nucleosome can also occur. For this, it has been suggested that twist defects (where DNA structures with more / fewer number of base pairs are present per DNA turn) reposition DNA by 1 – 2 bp in a rotation coupled manner (38) **(Fig.4C)**. These movements have a significant energy barrier to overcome, and these movements become more frequent at higher temperatures. *In vivo,* these movements are also coupled to ATP utilizing molecular machines such as chromatin remodellers. The combination of breathing and sliding may contribute to gaping, a phenomenon where the nucleosome accommodates more than the canonical ~147 bp (39) **(Fig.4D)**. These movements at the DNA – histone interface are often accompanied by conformational changes within the octamer. Cryo-electron microscopy (Cryo-EM) images of nucleosomes have shown that DNA breathing coincided with stabilizing changes within the octameric core. Changes include a rearrangement of H3, where on the unwrapped side H3 is shifted away; this shift propagates to the H3 on the wrapped side, moving it closer to the dyad (40). This movement may stabilize the DNA on the unwrapped side. Additionally, H2A-H2B flexibility is required to accommodate breathing. Octamers containing fixed H2A-H2B dimers cannot stabilize unwrapped DNA; for this, H2A-H2B dimers must be able to rearrange towards the unwrapped DNA. This is only possible in a dynamic octamer system (40).

Furthermore, DNA translocation can induce rearrangements of the H4 tail which usually interacts with SHL 2.5, but during translocation, it interacts with SHL 2 (41). This makes the tail more flexible, maybe acting as a signal of nucleosome distortion (41). Although these movements are small, their existence shows the plasticity of the nucleosome, and these

dynamics are required for stability and recognition by DNA/histone binding proteins. Larger wholesale changes are also possible within the nucleosome, none more dramatic than the loss of the H2A-H2B dimer. The spontaneous formation of a hexasome (nucleosome lacking an H2A-H2B dimer) is energetically unfavourable (34). The mechanism of histone loss is most likely achieved by ATP-hydrolysing molecular machines. Losing an H2A-H2B histone dimer decreases the DNA – histone interaction at the entry-exit site. Thus the first 30 - 40 bp remain completely unwrapped (42). Less stable nucleosomes (more likely to form hexasomes) such as those containing the histone variant H2A.Z are hallmarks of certain DNA regions such as promotors. This highlights the importance of nucleosome dynamics in gene regulation.

Inter-nucleosome contacts also exhibit dynamics within chromatin; this is done through the interaction of the acidic patches of neighbouring individual histones (43,44). A multi-register model has been suggested to describe 30 nm fibre dynamics. This model is based on tetranucleosome units that interchange in different registers (**Fig 4E**). These interconvert at the 100-millisecond timescale going through an open conformation (44). Interestingly, it has been shown *in vitro* that torsional stress stabilizes higher-order chromatin structures, although the *in vivo* role of this remains elusive (45). Compaction is further aided by the insertion of histone H1 within the internucleosomal linker DNA regions. Although H1 interactions are dynamic (remaining bound in the min timescales compared to the hours for the canonical histones), its presence is thought to be structurally important in silenced gene regions (46). Spontaneous thermal fluctuations of chromatin are thought to allow access to DNA within chromatin fibres to DNA binding proteins. The *In vivo* chromatin landscape is very diverse, with the presence of post-translational modifications (PTM) of histone proteins (47) and chromatin remodelling proteins (48–50), both of which act on local chromatin structure (10).

An active mechanism to control nucleosome dynamics and chromatin compaction is the post-translational modification of histone proteins. This chemical toolbox of modifications ranges from the enzymatic addition of methyl groups, phosphate groups and even the ligation of proteins or lipids on histones (51,52). PTMs can directly influence nucleosomes and inter-nucleosome contacts by changing the local electrostatic environment (i.e. acetylation/phosphorylation of lysine) and adding steric bulk (i.e., adding the 8.6 kDa protein ubiquitin), or they can act as signalling posts to recruit proteins that alter chromatin

conformation. These modifications can occur on the flexible histone tails or within the histone core. One of the more studied PTMs is the acetylation of histone H3 at lysine 56 (H3K56ac). This PTM is found in the entry/exit region of the nucleosome, where K56 can interact with the phosphodiester backbone of DNA. When acetylated, K56 loses its ability to interact electrostatically with DNA, resulting in a 7-fold increase in DNA unwrapping (53). This is one example amongst many found in the entry/exit region that change DNA – histone contacts. Within the nucleosome core, PTMs have been shown to destabilize and even hinder nucleosome formation. PTMs also directly affect chromatin dynamics by disrupting inter-nucleosome contacts. Two examples of this are the acetylation of histone tail H4 (H4K16ac) and ubiquitination. This region is rich in lysine and arginines, which can form salt bridges with the acidic patch of neighbouring nucleosomes. Acetylation impedes the formation of these interactions by stabilizing the formation of an α-helix in the H4 tail, thus decompacts chromatin structure (54,55). Furthermore, histone PTMs are often found in combinations, leading to the postulation of the histone code (47).

Until now, I have omitted to discuss the role that DNA sequence plays in dynamics, as it turns out DNA sequence is a major governing factor when it comes to the biophysical properties of DNA, nucleosomes and chromatin. It has been observed that certain sequences have a propensity of winding around the octameric core to form nucleosomes, whereas others are refractory (but not necessarily inhibitory). *In vitro,* the laboratory of J. Widom screened random sequences to find those with the highest propensity of forming nucleosomes (56). His lab found a sequence that forms highly defined nucleosomes (exhibit little sliding and gaping, thus remaining in a defined position along the DNA). With this sequence, the 601 Widom sequence, the crystal structure of the nucleosome was not only elucidated at 2.8 Å but a set of 'rules' were postulated for DNA sequences that would easily accommodate around a histone core (8,56). It was seen that A – T base pairs interspersed at 10 bp intervals facilitated DNA wrapping around the nucleosome. This coincides with the minor grooves with DNA – histone contacts that occur at ~10.4 bp intervals. A-T base pairs are easier to deform compared to those with G – C, but poly A-T or G-C sequences are both particularly rigid (16,26,57,58). Nucleosome positioning is not the only aspect that is influenced. DNA breathing and hexasome formation have also been shown to depend on sequence (42,59). These features are difficult to investigate *in vivo,* where DNA and nucleosomes are subject to

binding and remodelling by transcription factors, chromatin remodellers and histone-modifying enzymes. What is clear is that the distribution of nucleosomes *in vivo* is not homogenous and is in some regions highly dynamic due to the effect of chromatin effector proteins.

1.3 Chromatin Effectors

1.3.1 Transcription Factors

Chromatin effectors can specifically bind to DNA sequences and/or PTMS of histone tails and/or recognize octameric core regions directly, i.e. the acidic patch. A large class of chromatin effectors are transcription factors (TF). These DNA reading proteins bind specific motifs that are 6 – 30 bp long via DNA binding domains (DBD), interestingly, the number of TFs encoded, their size (but not the size of the DBD directly) and the number of unique domains per TF, scale with the size of the genome of the organism (60). Concentrations of potential DNA binding motifs within the nucleus are in the order of mM. Hence a TF must have a high enough affinity to bind specifically. TFs have large stretches of intrinsically disordered regions that are thought to contribute contacts between DNA and TFs that come from Van der Waals interactions, followed by hydrogen bonding to nucleobases and water-mediated contacts (61,62). Therefore, the free energy of binding is mediated not only via the DBD but across the TF. This has recently been highlighted by the sensing of DNA shape as well as or in lieu of DNA sequence (discussed later in this chapter), making TF site predictions difficult. In prokaryotes, most TFs are constituted of a DBD and an effector domain that directly recruits the RNA polymerase (60,63). Due to the smaller size of the prokaryotic genome, the DNA binding motifs are unique enough to confer sufficient specificity for the prokaryotic TFs (64). In Eukaryotes, however, DNA binding motifs are often degenerate, with a majority of binding events *in vivo* in regions with no discernible TF binding motif (60). This observation has led to the postulation that various strategies for specific recognition of Eukaryotic TFs to their motifs are required.

Sequence recognition occurs in both B-DNA's minor and major grooves via Van der Waals interactions with the nucleobases (65). Major groove binding offers more possible contact points and is more common, often via insertion of α-helices from TFs DBDs. Efforts have been made to see whether there exists a preference of certain amino acids in TF DBDs to interact with specific DNA nucleobases, this has shown some trends, but the multiplicity of interactions makes direct readout alone unspecific.

TFs can be categorized into family lineages. One method sorts TFs into structural families' as a function of their DBDs. In general, the average length of a DBD is around 60 amino acids, with 60 – 70 different DBD architectures per organism (66). In humans, the largest family is the Cys2-His2 Zinc finger (C2H2 ZF), with ~700 members. This small domain containing two cysteines and two histidines that coordinate an ion of zinc can recognise small stretches of DNA (3 -4 bp) (**Fig 5A)** (67,68). DNA recognition is achieved through the ββα fold induced by zinc-binding, which arranges around the DNA strand, positioning the α-helix to make direct contact with the 3 – 4 bp. For added specificity, multiple C2H2 ZF domains can be found within a single TF, CCCTC-binding factor (CTCF), for example, contains 11 such domains (69).

The Homeodomain (HD) DBD family is the next most abundant. Its members contain a helix-turn-helix motif, allowing them to recognize short AT-rich sequences (**Fig 5B)** (70). Homeodomain containing TFs are linked to tissue-specific expression, reflecting their role in cell-fate determination (67). The DNA motifs that HD TFs bind have been found to be very similar, in some cases even identical, but *in vivo,* they regulate different sets of genes. It seems that low-affinity binding sites are required for coordinated regulation of HD TF controlled gene expression. Experiments in which a substitution of low-affinity binding sites for high-affinity binding sites was performed showed ectopic gene activation and altered spatial expression patterns, indicating that binding site affinity regulates expression (71,72). In this family, homo/hetero-dimerization helps to confer specificity. Added protein-protein interaction domains allow the recognition of binding partners *in vivo*, providing additional interaction interfaces with DNA or nucleosomes (70,73). HD proteins have also been shown to contain PTM recognizing domains such as plant homeodomain (PHD) fingers that can read H3K4me3 (an activating mark), allowing multivalent binding (both to PTMS and DNA sequence) (70).

The last DBD family I would like to introduce are the basic helix-loop-helices (bHLH). This family intrinsically forms dimers and trimers that recognize 6 bp long sequences called the E-box (**Fig 5C)**. The basic N-terminal α-helix confers DNA binding specificity (binding to the major groove), whilst the HLH structure allows dimerization by hydrophobic contacts and Van der Waals forces (74,75). This HLH motif allows for to formation of hetero-dimers, further diversifying the recognition sequences of the bHLH family.

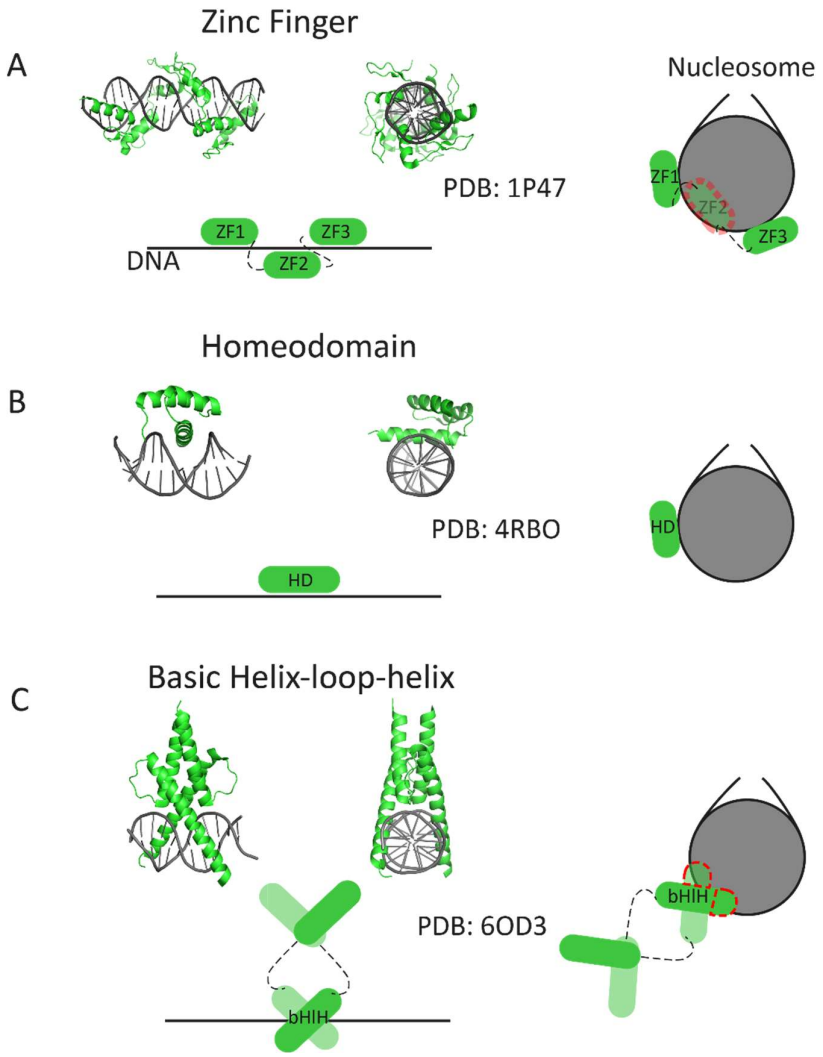


Figure 5 Transcription factors are categorized into structural subfamilies by their DNA binding domains (DBD). A) the zinc finger family often contain multiple DBDs able to recognize sort sequences. The crystal structure PDB:1P47 shows that the DBD subunits rotate around the DNA following the major grooves. This makes members of this subfamily unlikely to bind nucleosomal DNA (red lines indicated potential steric clashes). B) Homeodomain containing transcription factors can bind a single major groove. This potentially allows members to bind nucleosomal DNA as no steric clash is induced upon binding with the nucleosome. C) Basic Helix-loop-helix TFs saddle DNA when bound. This also renders nucleosome DNA binding difficult as t induces steric clashes.

These 3 super families may have different DNA binding domains, but they share common strategies to recognize and bind DNA motifs. These include multiple copies of DBDs,

combinatorial protein-protein association with other TFs, hetero-dimerization and recognition of PTMs. An additional strategy consists of TFs containing heterotypic DBDs, whereby a single TF can contain multiple types of DBD (67). This further diversifies DNA binding motifs recognized. As mentioned above, direct base readout has been shown to be only part of the TF binding mechanism; DNA shape has recently been shown to play a key modulating role. This can also participate in cooperative binding of TF via changing DNA physical properties such as major/minor groove width, propeller twist, to name a few. This is achieved by inducing DNA bending, the formation of kinks or stretching DNA upon binding of a TF. These physical changes can be induced by TF binding, DNA sequence and nucleosome formation. The classic example of this is the interferon-β enhanceosome, where 3 distinct TFs that do not physically interact bind cooperatively, although the mechanism is still not fully understood (76). On the DNA level, PolyA tracts weaken the stacking of bases allowing DNA to bend more freely. These stretches can induce kinking of DNA or compression of the minor groove (62). This phenomenon is thought to be able to regulate the capacity of TFs to stabilise the binding of another TF through DNA. This has been shown to have a periodicity consistent with the helicity of DNA, placing TFs on the same DNA face (77). An example of this is ComK, where the binding of ComK to a polyA tract influences the binding kinetics of another ComK TF 18 bp away through DNA mediated mechanical deformations (78). DNA mediated TF cooperation is still a nascent field due to the difficulty of studying such systems. The presence of nucleosomes acts as a barrier for the majority of TFs. Therefore nucleosome breathing, remodelling and, in extreme cases, eviction are required for stable binding. This is most evident in promotor regions of activated genes. TF cooperativity is the ability of one TF to target nucleosomal DNA, thus changing nucleosome properties and allowing the binding of another factor previously hindered by the nucleosome. This role has been observed for a subset of TFs called pioneer transcription factors.

1.3.2 Pioneer Transcription Factors

Induction of lineage-specific genes in embryonic cells leads to differentiation, but with a majority of the genome found wrapped around nucleosomes, how are genes in environments refractory to TF binding activated? Two main strategies have been postulated; first, the nucleosome containing genetic elements involved in regulating gene expression must be removed, allowing facile access of TFs to the underlying DNA. Secondly, the DNA held within the nucleosome is still accessible to proteins with distinct nucleosome binding features (79). Early studies of protein binders of lineage-specific DNA elements led to the discovery of TFs capable of nucleosome binding. Amongst these were the FoxA family of TFs. These factors are implicated in the lineage determination in the gut endoderm region that will later form the liver (80). The binding of FoxA to upstream DNA regulatory regions keeps cells in a poised multipotent state able to form hepatoblasts. *In vitro,* FoxA was shown to bind nucleosomal DNA, inducing nuclease sensitivity (81). Work from the Zaret group demonstrated that FoxA specifically bound chromatin fibres *in vitro*, but did not perturb its structure (82). Furthermore, fibres containing hitone H1 and a FoxA binding site became hypersensitive to DNaseI in the prescence of FoxA, induicating that FoxA acts directly on chromatin structure (82).

In vivo, FoxA demonstrates low mobility, which has been attributed to non-specific scanning of the genome by FoxA (83). Since the discovery of the FoxA family in mouse embryonic endoderm lineage cells, others have been found in Humans, Yeast and other organisms. pTFs share their ability to scan non-specific DNA regions and bind to nucleosomal targets. Predicting the ability of TFs to bind nucleosomes from their protein domains remains difficult. However, shared structural features have been highlighted in pTFs, chief among them is how the DBD interacts with DNA. Structural families with DBDs able to bind a single face of the DNA, such as the HTH and C2H2-ZF (single domain), make good DBD candidates for pTF activity. This stems from the α-helices within the DBDs that can easily access a solvent-exposed major groove within nucleosomal DNA (**Fig 5)** (84). Conversely, bHLH DBDs with the N-terminal α-helices interacting perpendicularly with the DNA axis can potentially clash sterically with the octameric core of the DNA. This family of TFs have shown a dependency on the length of their N-Terminal α-helices, with short α-helices seen to remain capable of binding to nucleosomes, which is not the case for longer helices (84,85). Further characterization of the interaction landscape of different DBD structural families have not

only shown differences in binding abilities but also in binding locations within the nucleosome (86). Certain TF families can bind periodically, following solvent-exposed DNA grooves (HD for example). Other families have demonstrated preferences for binding at the dyad or are forced to exploit nucleosome movements such as breathing to reveal DNA binding motifs. This is true for TFs that bind in the entry-exit region, such as those from the bHLH family, which are dependent on nucleosome breathing rates. TFs binding in the entry-exit region has been shown to be capable of compensating their reduced binding with longer dissociation rates (the site exposure model) (36,87). Once a pTF is bound, it can facilitate the recruitment of other previously occluded TFs, using mechanisms as discussed above (87–89).

Even though pTFs have been shown to bind and alter DNA-nucleosome interactions, nucleosome eviction and full recapitulation of the *in vivo* promotor landscape by simple incubation with pTFs and TFs *in vitro* remains elusive. This has led to the investigation of the role of other chromatin effector proteins, including the ATP utilizing chromatin remodeller machinery.

1.3.3 Chromatin Remodellers

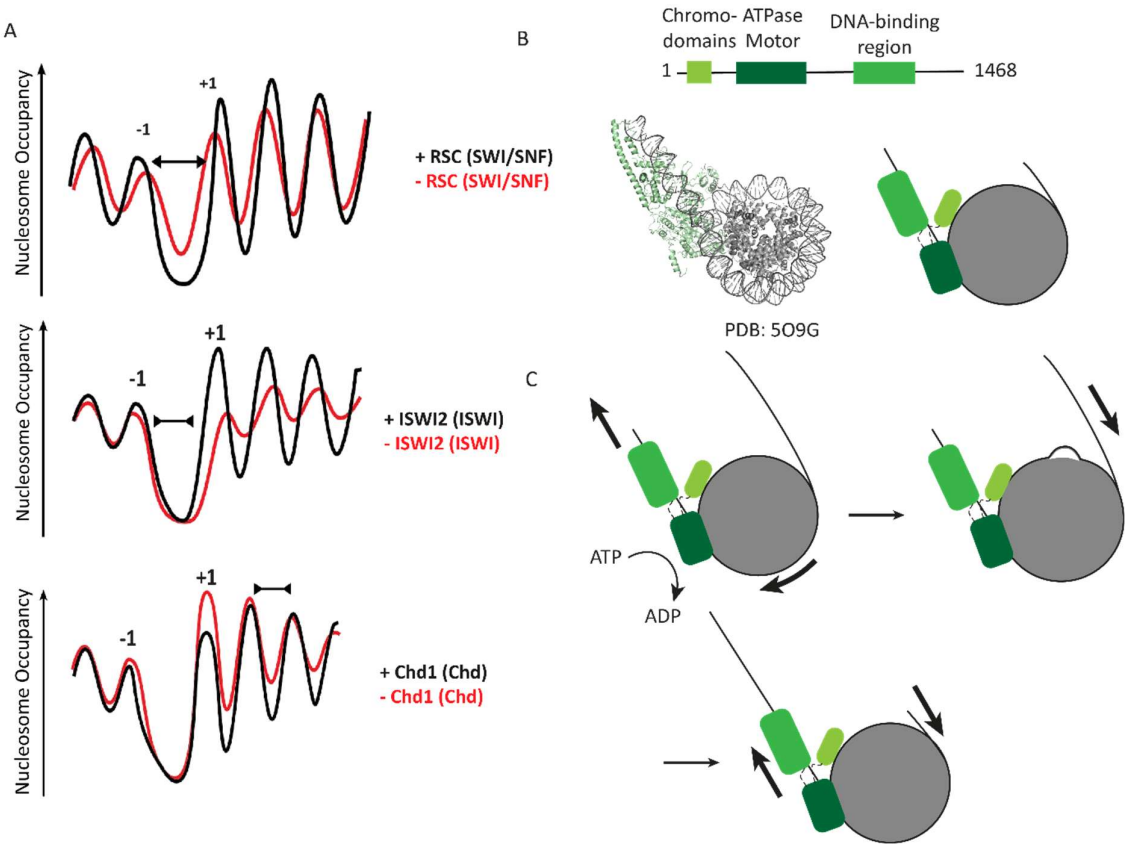


Figure 6 A) MNase-Seq experiments in vivo were done in both WT and remodeller depleted yeast strains (Kubik et al., NSMB; 2019). Traces shown are those are archetypical for the 3 major remodeller families SWI/SNF, ISWI and Chd. These have been shown to have activity in different regions in vivo. B) Protein domains and crystal structure of the yeast remodeller Chd1. C) Scheme of the proposed remodelling mechanism based on the crystal structure and biochemical assays. The exact mechanism remains to be determined.

Chromatin remodelers (CR) are large (up to a megadalton) biological machinery that utilize ATP hydrolysis to move/remove, assemble and edit individual nucleosomes, shaping the overall chromatin landscape. Along with an ATPase dependent DNA-translocating ability, chromatin remodelers share other features, including nucleosome binding affinity (as well as recognition of some histone PTMs). They have domains/protein-subunits controlling ATPase activity and the interaction with other proteins such as transcription factors (49,90,91). CRs are essential in regulating chromatin organization throughout cellular processes, including replication, DNA damage repair and gene regulation (49,92). Specialized functions and phylogenetic traits separate CRs into four subfamilies that are highly conserved from Yeast to

Humans. These are Switching defective/sucrose nonfermenting (SWI/SNF), imitation switch (ISWI), chromodomain-helicase-DNA binding (CHD) and inositol requiring 80 (INO80) sub families. Genetic studies in model organisms (mainly yeast) have revealed the scope of reactions catalyzed by each CR subfamily, SWI/SNF CRs are active at promotor regions regulating transcription (48,93–96) (**Fig 6A**), ISWI and CHD CRs evenly space intragenic nucleosomes (48,97–99) and INO80 regulates the turnover of histone variants (100,101). Remodelers contain domains or interact with subunits that can directly read out the PTM landscape. These domains include bromodomains that recognize acetylated lysine's (including H3K14ac), chromodomains that recognize H3K4me2 and plant homeodomains which also recognize methylated lysine's (49). Different genomic regions contain specific PTMs, the presence of PTM reading domains has been shown to recruit CRs directly. In addition, depletion of the CHD family member CHD1 *in vivo* led to a loss of H3K4me3 and H3K36me3 domains (102). The extent to which these PTM binding domains are involved in reading the PTM landscape and whether these regions are only involved in recruitment or influence activity is still an open question.

Recent advances in cryogenic-electron microscopy (Cryo-EM) have led to many structural studies showing CR-nucleosome contacts (103–105). In conjunction with *in vitro* single-molecule studies, DNA translocation mechanisms are being elucidated (91,106,107). For example, Chd1 has been shown to anchor itself on the nucleosome by binding the alpha helix α1 on histone H3 and the basic N-terminal tail of H4 via its ATPase domain (**Fig 6B**). Concurrently Chd1 binds the entry/exit DNA with its DNA binding domain and SHL 2 with its ATPase domain. Energy transfer from ATP hydrolysis to DNA translocation is thought to occur by inducing twisting of the DNA being pumped by Chd1 (106). Upon twist diffusion (relaxation), DNA is translocated along the nucleosome (**Fig 6C**). More precisely, during ATP hydrolysis, Chd1 feeds DNA towards the dyad. During this phase, the nucleosome accommodates an extra 1 – 3 bp (gaping). Eventually, these accumulated nucleobases near the dyad relax by pushing the DNA out of the entry/exit and diffuse the induced over twisting. It has been suggested that all CRs use a similar mechanism of DNA translocation. This has been postulated as the *hourglass model* (91).

Regulating this ATPase activity goes beyond just specific recruitment. The SANT-SLIDE motif found in many CRs makes direct contact with flanking DNA. Ablation or mutations in these

regions led to aberrant positioning of nucleosomes by CRs (108). The DNA sequence also seems to play a role in remodelling directionality. RSC, a member of the SWI/SNF family, has been shown to be influenced by paired motifs of poly(dA:dT) tracts in the proximity of regions with high GC content (93,109).

At the cellular level, chromatin remodelers co-exist and act in similar loci—the question of how CRs interact remains. Furthermore, the interplay between CRs and other nuclear proteins has been shown for selected pairs of CRs and TFs. Still, questions on both the biological relevance and mechanism of interactions remain. For transcription by the polymerase, the rate-limiting step to transcription is forming the Pre-initiation complex (PIC). Working hand-in-hand, TFs and CRs have been shown to facilitate the recognition and binding of TFs, facilitating PIC formation and transcription initiation (109,110).

1.4 Techniques for studying chromatin effectors

In the following section, we will discuss the techniques developed to study complex biological systems to understand not only the underlying biology but also the mechanisms underlying these processes.

1.4.1 *In Vivo* Biological roles of chromatin and chromatin effectors

Changes in cell fate or phenotype from cells that have deletions or insertions of coding and non-coding regions have proved invaluable in understanding the central role played by effector proteins. The most striking example is the Yamanaka factors, Sox2, Oct4, Klf4 and C-Myc. These four TFs can reprogram fully differentiated cells into pluripotent stem cells when ectopically expressed (111). This change in potency implies large-scale changes in the proteomes, epigenetics and transcriptional patterns of cells. The drive to further understand gene regulation led to the advent of methodologies to locate DNA-binding sites for proteins of interest, such as transcription factors. Chromatin immunoprecipitation (ChIP) uses specific antibodies to target a protein of interest (TF, Histones, CR, protein Tag etc.) from cell extracts of fixed cells, which can then be isolated and analysed (112). In the chromatin field, variants

of this methodology have been extensively used; these are Co-Immunoprecipitation (Co-IP), ChIP-ChIP and ChIP-Seq (113). Co-IP experiments detect protein-protein interactions by analysing the macromolecules bound to the precipitated protein of interest. With this, we can find protein-protein interacting partners. However, the mechanism of the interaction cannot be inferred. For this, *in vitro* techniques have been developed (discussed below). Precipitating a protein of interest, followed by a second precipitation using another antibody specific for another protein (ChIP-ChIP), retrieves longer distance interaction for proteins that are found in the proximity of each other *in vivo* **(Fig. 7)** (114). "Interacting" here may not mean direct contact, but two proteins in the same locus may act as co-factors or regulators of a specific molecular process. Once again, the role of each protein pair needs to be further studied either with genomic techniques (Anchor-Away, deletions and mutations *in vivo*) or biochemically *in vitro*. Using ChIP, we can also probe Protein-DNA interactions, sequence preferences can be found, and DNA binding motifs posited by sequencing the DNA that is co-precipitated with a specific protein (ChIP-seq) (112,113) (**Fig. 7**). Suppose a TF is pulled down (after sheering DNA via sonication). In that case, sequencing reveals DNA binding motifs, which can be averaged, and a position weight matrix (PWM) can be generated to display the overall sequence preferences. However, it has been shown that the mere presence of a matching sequence *in vivo* is a poor predictor of whether or not a TF regulates said gene (67). As TFs can bind partial motifs, it is more likely that a TF is scanning the DNA rather than specifically bound at any given time. This skewed distribution of sequences can lead to inaccurate determinations of DNA binding motifs that do not directly reflect binding affinity but rather abundance (67). Further limitations include that measurements are not done at equilibrium, the requirement

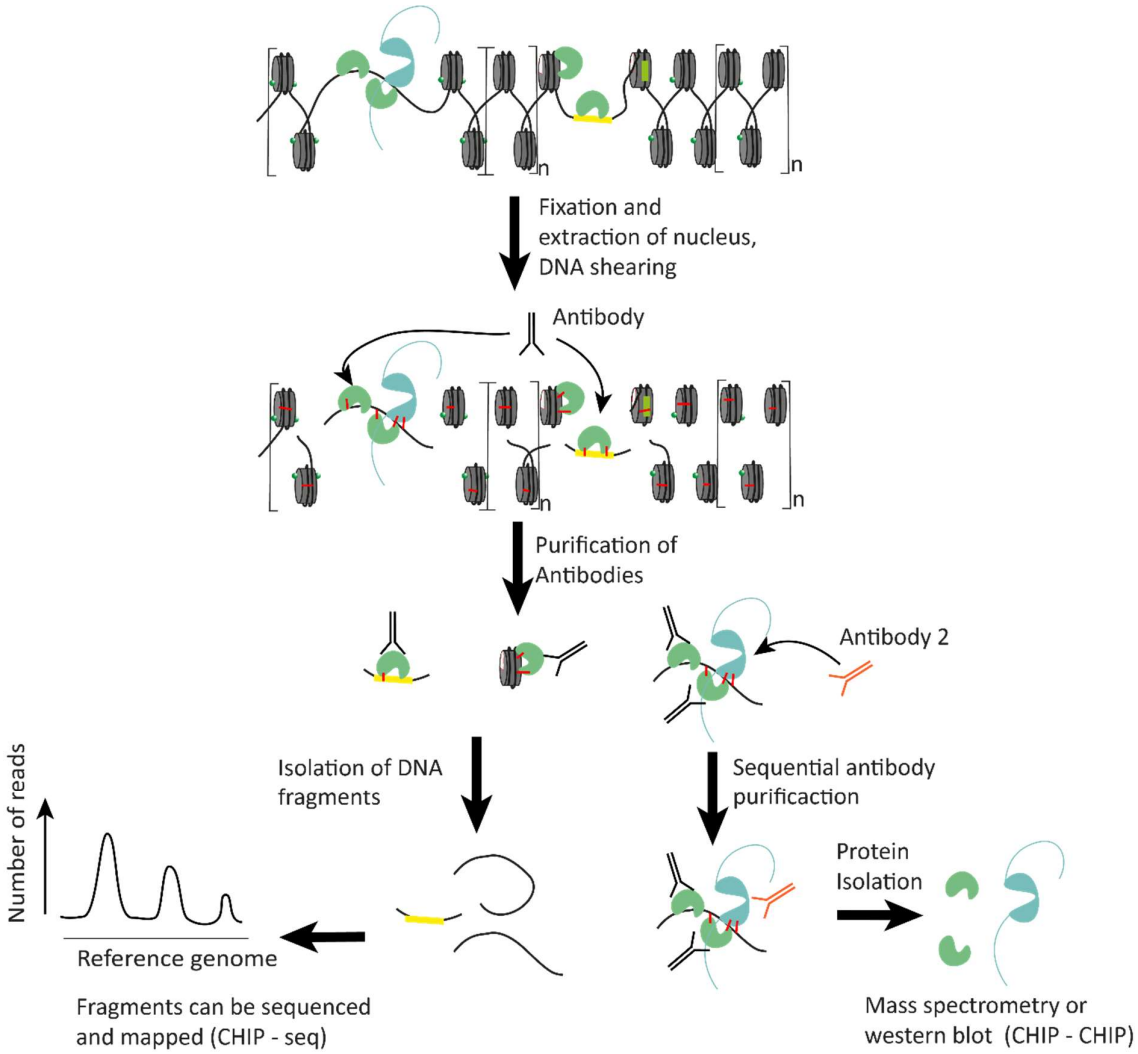


Figure 7 Common biochemical assays techniques used to study the interaction between proteins and DNA in vivo. These include chromatin immunoprecipitation (ChIP) which can also be coupled to sequencing (ChIP-Seq) to find the positions within a genome that a protein binds. Protein-Protein interactions can also be studied using sequential ChIP experiments (ChIP-ChIP).

for specific antibodies and the influence of the chromatin landscape. Nonetheless, ChIP-Seq data is often the starting point for planning experiments *in vivo* and *in vitro.* The encyclopaedia of DNA elements (ENCODE) consortium has mapped transcription factor association within the human genome and their association with PTMS and nucleosome positions (115). Mapping DNase hypersensitive regions using DNase-Seq reports on the accessibility of regions. Using this method, the role of effector proteins can be directly probed *in vivo* (**Fig 8**). Changes in accessible DNA regions due to deletions or depletions of TFs or CRs can map the influence of these effector proteins within the whole genome. From these studies a cell-to-cell heterogeneity was observed, suggesting differences in gene regulation between cells. More strikingly, a large proportion of changes were detected in intergenic and

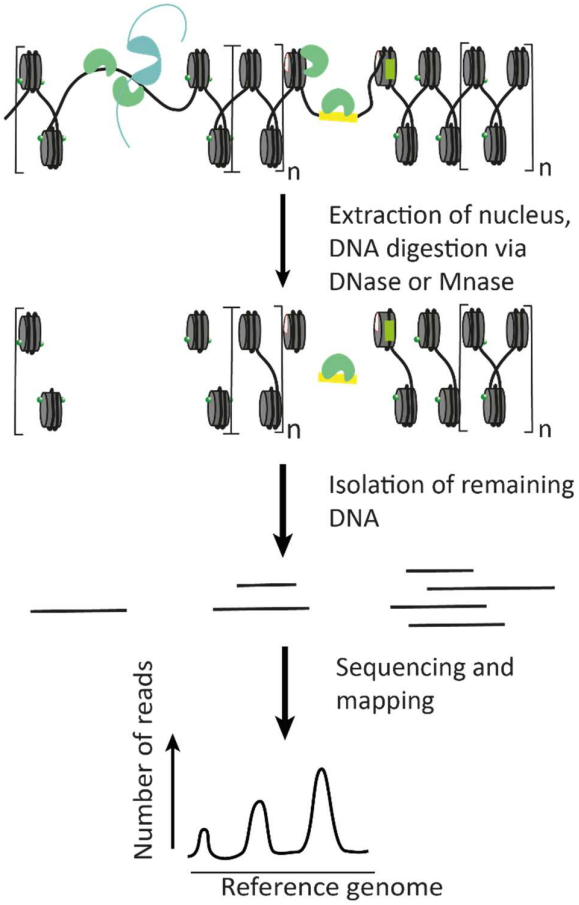


Figure 8 The principle of DNase Hypersensitivity or Mnase where a nucleus is extracted and exposed to purified DNA cutting enzymes (DnaseI or Mnase). Exposed DNA is digested preferentially, nucleosomes and DNA binding proteins can partially protect DNA from these cutters. After DNA purification and sequencing these protected DNA areas can be mapped to a reference genome to find areas containing nucleosomes or other DNA binding proteins.

intronic regions, long considered 'junk' DNA (116). DNase hypersensitive regions and high GC content sequences located close to the TSS were invariable between cell types, hinting at the role played by changes in chromatin structure during gene regulation. Monococcal nuclease (MNase) is an alternative endo/exo-nuclease more often used for experiments mapping nucleosomes precisely. MNase cleaves differently depending on the concentrations used, cleaving DNA closer to the nucleosome (4-5bp closer than DNaseI) as it uses a different cutting mechanism. Both DNaseI and Mnase have different sequence biases (117). MNase-seq has been extensively used to explore the role of TFs and CRs in regulatory regions. An alternative to measuring changes in DNA accessibility by DNaseI/MNase digestion is to use a transposase that can recognise accessible chromatin environments (118,119). Assay for Transposase-Accessible Chromatin with high throughput sequencing (ATAQ-Seq) uses the hyperactive Tn5 transposase to splice NGS compatible primers in regions where T5n can bind. This simultaneously fragments and labels accessible DNA regions, which can then be sequenced

and mapped. This method uses fewer cells (even single cells) to gain valuable insight into the chromatin landscape. DNaseI and Mnase are indirect measurements of the activity of chromatin effectors. These experiments cannot distinguish active roles from passive roles played by proteins. Therefore they are not suited to mechanistic studies. Furthermore, the methods mentioned above are not targeted. Thus they rely on the intrinsic interaction of the 'reader' (DnaseI, Mnase and Tn5), which can introduce certain biases. Cleavage under targets and release using nuclease (CUT&RUN) is a method to map specific contacts between proteins of interest and DNA (120). The main advantages of this method are that only a subset of DNA is cleaved and that sample preparation is devoid of fixation (removing certain antibody biases). Targeting is done by the WT protein of interest, which is then located by antibodies. Antibody binding to target proteins is accompanied by DNA cleavage via a protein-A-MNase fusion which recognizes the heavy chain of antibodies. Together, this system couples the endogenous targeting of proteins, contrary to DNaseI and MNase methodologies, whilst still using NGS as a read out in isolated nuclei. Data analysis of all of these methods remains a challenge, with millions of reads to map to a reference genome and interpretation remains indirect. These tools are widely used in the chromatin field to directly map changes in protein binding or chromatin environment following mutations/deletions of proteins of interest, application of drugs that change the PTM landscape, amongst other genetic manipulations. However, in some cases, the protein of interest is essential, and deletions are lethal to the organism. A rapid and controlled depletion method has been devised using fusion proteins to contour this problem.

For a long time, temperature-sensitive mutants were used to induce the depletion of controllable conditional mutant in yeast. However, this method was often leaky and limited by the amount of characterized temperature-sensitive targets. Recently a different approach has been established, whereby a nuclear protein of interest and either FK506 binding protein (FKBP) or FKBP12- rapamycin binding domain (FRB) is made as a fusion protein. The anchor, a cognate cytosolic protein fusion (usually a membrane protein) with the complement of FKBP/FRB, is also expressed as a fusion protein within the same cell. The FK506 and FRB protein fragments are used as conditional interacting partners, which depend on the presence of Rapamycin (121,122). Therefore, in the absence of Rapamycin, the nuclear protein of interest is free to move throughout the cell. However, the addition of Rapamycin forces the

interaction of the FK506/FRB and sequesters the protein of interest to the cytosol, depleting the nucleus (**Fig 9A**). For nuclear proteins, the shuttling of ribosomal proteins can be utilised by making a ribosomal protein anchor. Rapamycin incubation and interaction of the FK506/FRB, this complex is preferentially exported (**Fig 9B**). FRB Anchor-away, as it is called, has been applied to ChIP- and Mnase-Seq to great effect to study TFs and CRS as well as their interactions (48,110).

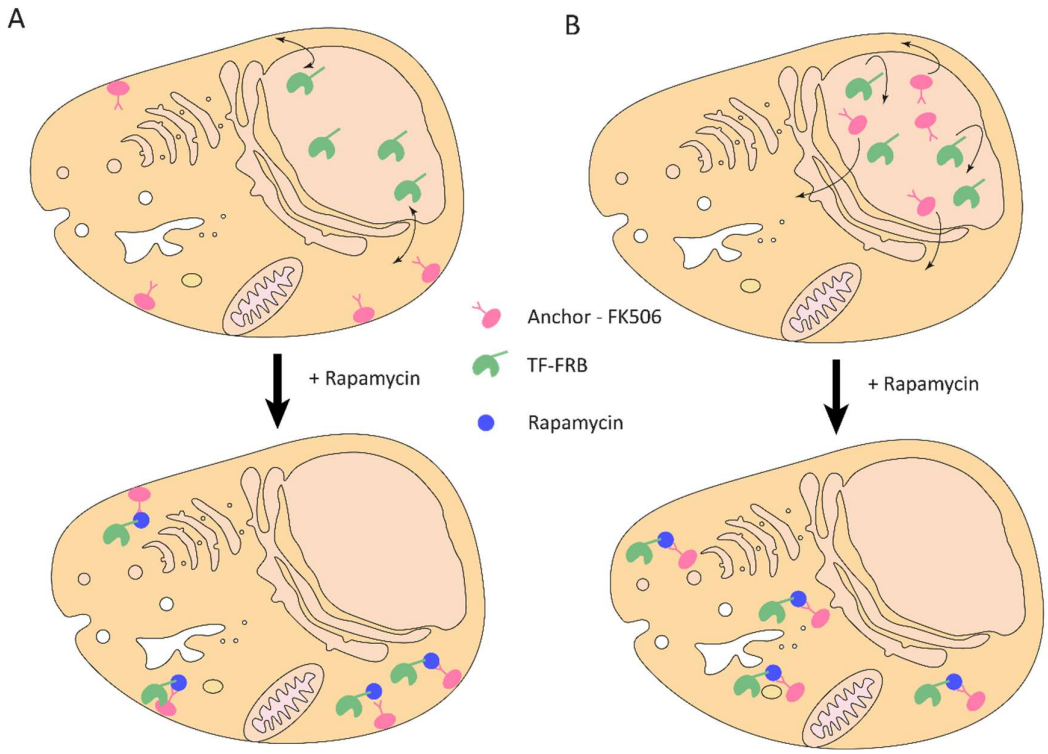


Figure 9 A) Anchor-away depletion of a TF that can shuttle between the nucleus and the cytoplasm. The addition of rapamycin forces the dimerization of the TF-FRB to the anchor bound to the cell membrane Anchor-FK506. This in effect depletes the nucleus of the TF. B) Anchor-away of a TF that is only found in the nucleus. The Anchor -FK506 can shuttle between the nucleus and cytoplasm but is preferentially found in the cytoplasm. Addition of rapamycin induces dimerization and depletes the nucleus of the TF.

On top of DNA-Protein interactions and local chromatin landscape dynamics, a key question is whether the 3-dimensional genome organisation, including conformations of active promotors and enhancers, plays a role in gene regulation. To investigate the nuclear 3D organisation, physically touching genomic loci can be crosslinked using chemical fixatives. Following a digestion and purification protocol, crosslinked DNA sequences are ligated to each other, allowing for specific PCR or high-throughput sequencing to map all genomic contact points (123,124). Chromosome capture technology can map active and inactive genomic TADs and finer structures such as point-like contacts of loci bound by CTCF (124,125). Participation of scaffolding proteins in processes like DNA damage response and replication is still debated,

but chromosome capture technology has been key in investigating these phenomena. Identifying interactions of regulatory sequences with surrounding partners has shown that most contacts are enriched 2 – 3-fold compared to random collision rates. However, the nature of this interaction cannot be inferred beyond simple proximity. These studies show that the spatial distribution of genes is not random in interphase and that establishment of this spacing is key to regulating cell fate. Chromosome capture technologies have identified regulatory sequences that warrant further biochemical characterization.

Beyond sequencing-based technologies, fluorescent-based measurements have also been established to study chromatin architecture and individual effector proteins. Before chromosome conformation capture technologies, image-based systems such as fluorescent in situ hybridization (FISH) was used to follow the spatial distance between genomic loci (125). Image-based techniques are limited by the diffraction limit ~ 200 nm, factoring in the size of the probes itself, makes resolving sequences below 100 kb difficult to obtain. Furthermore, only a limited number of probes can be used simultaneously, making this technique low throughput. Dynamic *in vivo* measurements have also been devised to follow DNA movements. Using tagged versions of an inactivated CRISPR-Cas system (Dead-Cas or dCas), specific sequences can be targeted and followed in live cells. Early versions of this methodology were limited to repetitive sequences because they used Zinc finger (ZF) and transcription activator-like factors (TALEs) (126,127). The dCas system can be tuned using different guide RNA sequences. However, they require the presence of a short regulatory DNA sequence upstream of the recognition site. Super-resolution advances, including the use of DNA paint, have increased the spatial resolution of imaged-based techniques, but samples for this imaging method remain static as they are fixed (128–130). Conversely, nuclear protein studies have also benefitted from image-based methodologies. Fluorescent recovery after photobleaching (FRAP) and Single-particle tracking (SPT) has been used to decipher the dynamics of protein-DNA interactions. Interactions spanning a broad spectrum of timescales have been revealed, spanning from hours for histone proteins to seconds for TFs (131).

The desire to bridge a biological outcome to a biophysical event remains. The complexity and heterogeneity of the cellular environment and the technical challenges of studying processes on multiple size and time scales means we are often left inferring mechanisms from multiple experiments. The coupling of genetic studies to identify key players followed by *in vitro*

characterization has been the gold standard in elucidating the mechanistic role of effector proteins.

1.4.2 *In Vitro* studies of chromatin effectors

Contrary to the *in vivo* environment, *In vitro* experiments allow for the use of chemically defined milieus and temporal control. The use of recombinantly expressed proteins and known DNA sequences facilitate kinetic and structural studies of the interaction of chromatin and effector proteins. Amongst the panoply of methodologies, gel-based electrophoresis has been the gold standard in biology for decades. Electrophoretic mobility shift assays (EMSA) have been used to determine the binding kinetics on proteins to DNA and nucleosome substrates (132). This simple technique involves titrating either the protein or the binding substrate to a fixed concentration of the other binding partner. Using native gels, a super shifting of bands within the gel is observed upon complex formation. Therefore, the proportion of sequestered substrate can be calculated for each reaction condition, and a binding affinity (Kd) can be calculated. This method requires an incubation step as the measurements (the gel) should be done at equilibrium as we observe an ensemble. DNA sequences containing DNA binding motifs can be selected using data from ChIP experiments, the advantage being that sequences can be individually tested. A caveat of this method is that buffer conditions can be optimized to increase binding affinities, meaning that Kds obtained may not be equivalent *in vivo*. *De novo* sequence motif discovery can also be done *in vitro* using systematic evolution of ligands by exponential enrichment (Selex) (133–135). Barcoded DNA sequences are pooled and incubated to compete for binding with tagged protein of interest. After purification of the protein and recovery of the bound DNA, PCR amplifies all bound sequences. A second round of competition incubation is done to enrich strong binding sequences further. After several rounds of binding, purification and amplification, enriched sequences can be identified, and high-affinity binding motifs discovered. This method has the advantage of being independent of abundance artefacts found *in vivo*. Multiplexed methodologies of this strategy using all known human TF DBDs and combinations of TFs (cooperative binding) have been used to discover and refine TF-DNA and TF-nucleosome

interactions (86,133,136). As these experiments are done on an ensemble, no kinetic data can be inferred. For this, image-based technologies have been efficiently used with known protein-DNA/nucleosome/chromatin pairs. One of these techniques is single-molecule colocalization total internal reflection microscopy (smTIRF).

TIRF microscopy uses high incident light (angle generated from a prism or through the objective) that is greater than the critical angle, therefore total internally reflected (TIR) (137). TIR occurs when light travels within a medium with a high refractive index and meets a lower refractive index media (i.e. Glass – Water interface). At this boundary, an infinitely extended plane is formed called the evanescent field. Its penetration depth into the aqueous medium is only ~100 nm, meaning only molecules close to the surface are illuminated, rendering low background images. Using surface-immobilized fluorophore labelled DNA/nucleosomes and orthogonally labelled recombinant proteins of interest, TIRF illumination can follow colocalization events, therefore inferring binding dynamics of the system. This method has been used for a host of proteins to follow kinetics in different DNA contexts to study search and binding dynamics (87,138). Binding is only one half of the DNA-Protein interaction puzzle. The other half is what the consequences of binding are? Are there structural changes, for instance? The methods mentioned in this section are not tuned to tackle these questions. For this, more structural biology approaches are needed.

1.4.3 *In Vitro* Structural Biology

Purified recombinant histone proteins have been shown to self-associate and fold into an octameric core when dialysed from denaturing conditions to high salt (139). The addition of DNA, usually a nucleosome positioning sequence such as Widom 601, and dialysis from high slat to native conditions spontaneously forms nucleosomes (139). This method has extensively been used to reconstruct mononucleosomes and chromatin fibres. With this, defined homogenous nucleosome populations can easily be generated, facilitating the interpretation of *in vitro* experiments. Several methodologies have been exploited to tease out chromatin's structural and/or dynamic properties with/without chromatin effectors. X-ray crystallography was the first methodology explored, allowing to resolve the static

structure of whole proteins (140). These high resolution (1 – 5 Å) snapshots include the nucleosome, chromatin and bound effector proteins (141,142). However, due to the size of the structures interrogated, the crystallization of samples proved to be a real bottleneck for this methodology (143). Vitrification in ice of samples instead of crystallization has sped up sample preparation and generated a new avenue of experimentation, Cryo-Electron Microscopy (Cryo-EM) (144). With this, the structures of large multi-subunit complexes have been successfully defined with sufficient resolution to extract conformational data. These include complexes of the CRs Chd1, RSC and FACT with the nucleosome (104,145,146). Some caveats of these experiments include the use of fixatives and ATP analogues which may trap CRs in inactive conformations. Indeed, many conformations exist, including functional and inhibitory states (not counting inherent flexibility). CryoEM is uniquely suited to disentangle these states by grouping classes of molecules with different conformations within a single sample. However, functional interpretation of these structures can require the coupling of information gained with dynamic *in vivo* methods such as Nuclear Magnetic Resonance (NMR) or optical methods (Circular dichroism and Föster Resonance Energy Transfer (FRET)), allowing us to develop further our mechanistic understanding of large complexes (140,147).

When discussing dynamic processes, the notion of timescales is central. Structural changes and biological processes occur over a large spectrum of timescales, from nano-seconds to minutes. Methodologies are usually fine-tuned for a range of these timescales. Here we will discuss FRET as an example, which is the non-radiative transfer of energy between two suitable chromophores (148). This interaction of the transition moments of the donor and acceptor chromophores is proportional to the distance between them, making FRET a "molecular ruler" that can read distances from 2.5 – 10 nm. Labelled molecules/complexes containing a donor and acceptor chromophore pair are free to adopt all possible conformations in an aqueous environment (and at steady-state). The resulting FRET efficiencies from these changes can be monitored over time. This can be done for a whole population as an ensemble or a the single-molecule level. Single-molecule FRET (sm-FRET) has the advantage of following heterogenous structural populations and changes that would be masked when observed as an ensemble (149). FRET measures the distances between 2 labelled regions. Therefore monitoring of multiple changes in different regions requires multiple constructs with labels in these different positions. Currently, in a single FRET

experiment, a maximum of 3 dyes can be used (3-color FRET) at a time. Therefore, monitoring structural changes across a whole protein entails cloning, expression, purification and specific labelling for each change in position. Although this approach is low-throughput, it has proven invaluable in studying the dynamics of the nucleosome and chromatin (34).

1.5 Hallmarks of DNA architecture in vivo

With the plethora of *in vivo* and *in vitro* techniques detailed above, the regulation of gene expression and changes in chromatin structure involved in this process have been extensively studied. This section provides an overview of the state of the current literature in this field.

1.5.1 Heterochromatin and Euchromatin distinct nuclear regions

A sub-organelle organisation within the nucleus of eukaryotic cells was first observed in the 1930s using light microscopy, where light and dark regions within the nucleus appeared after DNA staining (7). These regions, heterochromatin (darker) and euchromatin (lighter) have since been extensively studied and major differences in their genetic content, nuclei localization, and proteomic content have been highlighted **(Fig 10A)**. However, this seemingly binomial distribution of DNA regions is overly simplistic. We now know that Heterochromatin can further be split into two categories. These are facultative heterochromatin (able to decondense and contains silenced genes) and constitutive chromatin (remains condensed throughout the cell cycle). The establishment and maintenance of Hetero- / Eu-chromatin is an active process involving different proteins and mechanisms (150). Throughout differentiation, the distribution and presence of these different chromatic regions depend on the cell types. Stem cells, for example, contain fewer PTMs associated with heterochromatin and have more dispersed heterochromatic architectures than differentiated cells (151). Whereas in differentiated cells in interphase, Heterochromatin is typically found at the nucleus periphery and contains few genes. Heterochromatin mainly consists of pericentromeric, telomeric regions of chromosomes and repetitive DNA sequences (satellite DNA, transposons etc.) (152). A small level of transcription remains in heterochromatin, although this may be regulated by the cell cycle, notably replication (153,154). Euchromatin is enriched in gene containing DNA sequences and is found scattered around the nucleus. The nucleolus is a prime example of euchromatin. This subregion of euchromatin is rich in ribosomal genes from 10 different chromosomes and is active in all cell types (155).

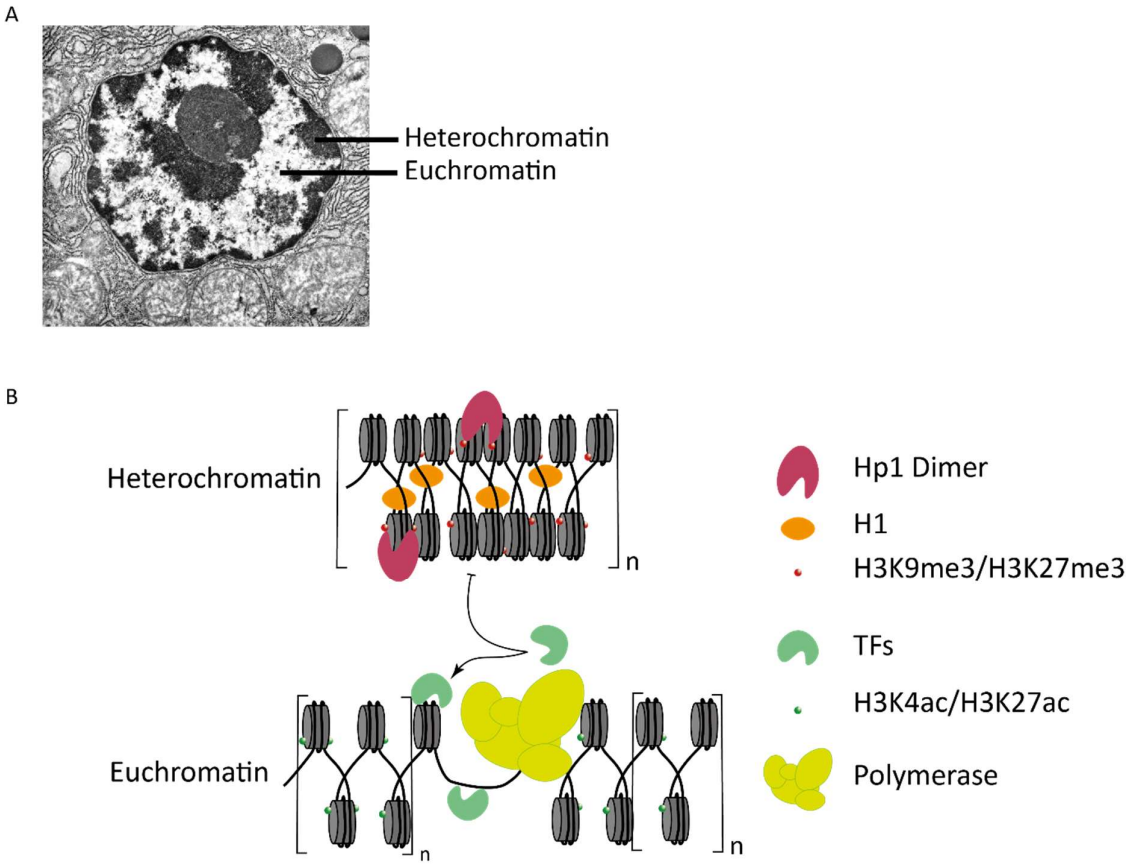


Figure 10 A) Electron micrograph of a cell with its DNA stained to see the different genomic regions. Euchromatin is less dense therefore is less stained compared to Heterochromatin which appears darker. B) Scheme of the different chromatin regions and their associated proteins such as heterochromatin protein 1 (Hp1) and Histone H1 (H1) for Heterochromatin. Different posttranslational modifications are also shown to highlight the differences between Euchromatin and Heterochromatin.

These DNA states are established and maintained by dynamic features. Thus, the passage through the cell cycle disrupts the boundaries between the different chromatin regions, requiring the reestablishment of PTMs and topologically associated domains (TADS) to take place (152).

Within heterochromatin, these features include hypoacetylation, histone PTMs such as histone H3 lysine (K) 9 methylation (H3K9me2/3) and H3K27me, as well as recruitment of silencing proteins including histone H1 **(Fig 10B)** (152). Removal of acetylation by histone deacetylases (HDACs) and methylation of histones by methyltransferases such as polycomb repressive complex 2 (PRC2), which installs H3K27me3, allow for the recruitment of proteins that maintain chromatin condensation. The enigmatic silencing protein, heterochromatin protein 1 (HP1), binds neighbouring nucleosomes containing H3K9me2/3. This multivalent binding not only increases HP1 affinity it also maintains chromatin condensed (44,156). HP1 has been observed to have two populations within heterochromatin, both a dynamic and a

stable HP1 pool. The differences between these populations remain unclear. A recent hypothesis revolves around the ability of HP1 to form liquid-like droplets via phase-separation (157–159), enriching local concentrations of heterochromatin associated proteins. These PTMs also allow for the interaction with the nuclear lamina, spatially sequestering silenced regions to the nuclear periphery (160). Although located preferentially on the nuclear periphery, heterochromatin still shows largescale movements within the nucleus (161). Meaning that the heterochromatic regions may collide or be in close proximity of euchromatin regions and compete for the deposition of hallmark PTMs, which can hinder cell activity (162). Therefore the cell uses countermeasures to avoid excessive spreading. Such barriers can be nucleosome depleted regions, high nucleosome turnover, specific recruitment of anti silencing proteins via PTMs, amongst others (150). It is only through cycles of deacetylation, methylation, HP1 binding and recruitment of other heterochromatin associated proteins that heterochromatin is established and maintained.

Euchromatin is generally characterized by the accessibility of the underlying DNA by DNA binding proteins and enzymes **(Fig 10B)**. DNaseI hypersensitivity is an example of this, whereby nuclear extracts are incubated with the DNase cleaving enzyme DNaseI and sequences are analysed through sequencing (163). The broad specificity of DnaseI means that any exposed DNA regions are potentially cleaved, but densely compacted chromatin regions are refractory to DNaseI cleavage. Euchromatin shows hypersensitivity to DnaseI, meaning it has a globally open chromatin architecture. These studies also highlighted the presence of RNA polymerase II and transcription factors, seen by UV-crosslinking experiments and “cold spots” (sequences lacking hypersensitivity) within certain hypersensitive regions (163). Active gene regions are rich in histone acetylation (at positions including H3K9, H3K3, H3K27 and H4K16) but also contain specific methylation marks (H3K4me2/3 and H3K27me2) (164–166). Acetylation of the core histones H3/H4 has been shown to directly perturb chromatin compaction by masking the positive charges on the histone tails, reducing their interaction with DNA and perturbing inter-nucleosome stacking. Together this results in more loosely packed chromatin regions (52,164,167,168).

Whether active or silenced, single genomic loci require specific genetic elements such as promotors, enhancers and core promotor elements in diverse conformations and flavours. These underlying genetic sequences and dynamic regulatory elements (PTMS, histone

variants and associated proteins) allow for the complete spatial and temporal control of gene expression.

1.5.2 Architectural features of genetic Loci

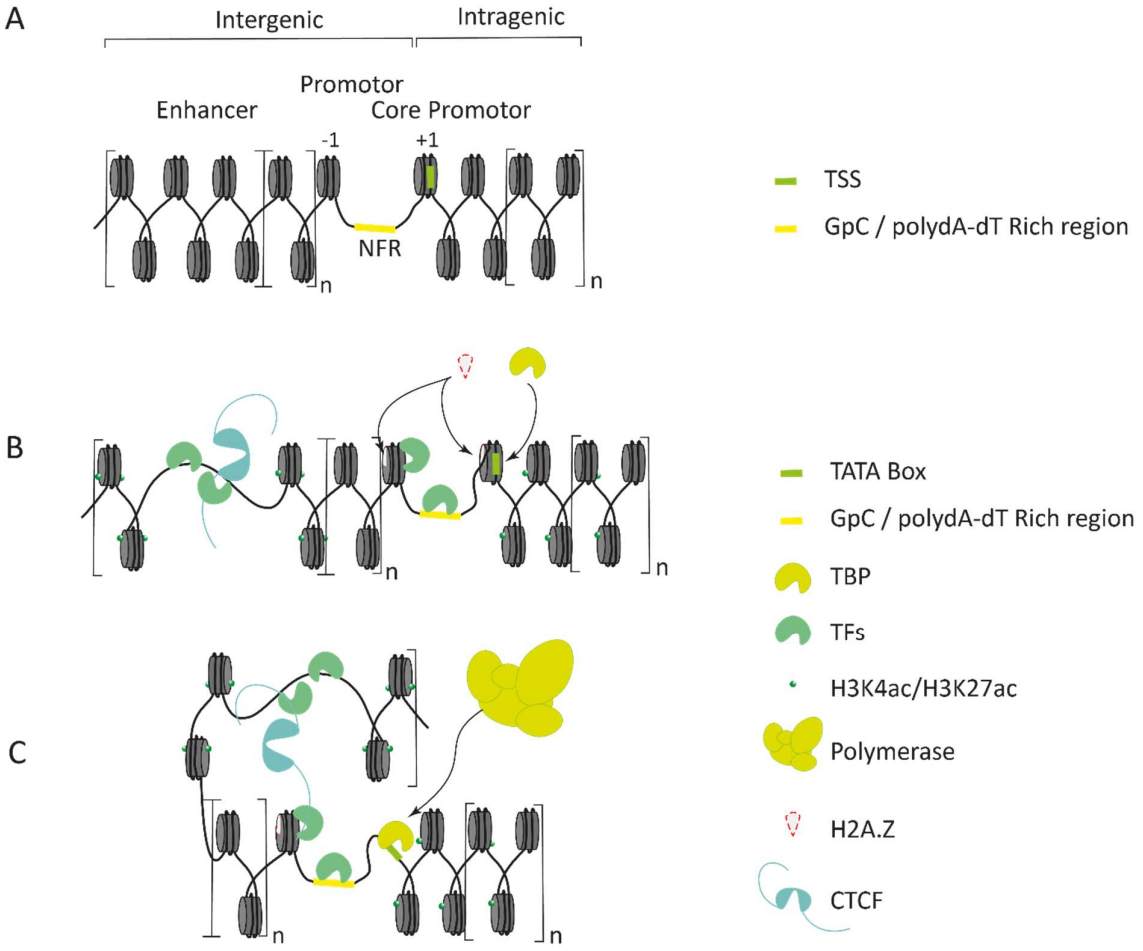


Figure 11 A) Drawn is the typical architecture of a promotor region. Highlighted are the nucleosome free region (NFR) composed of islands rich in Guanosine and Cytosine (GpC) as well as poly-Adenosine-Thymidine (PolydA-dT) sequences. Finally, the transcription start site (TSS) is also depicted. B) Requirement of chromatin effectors such as transcription factors (TF) and Chromatin remodellers change the landscape around the promotor region and the upstream enhancers. These changes include loss of nucleosomes and histone exchange. C) Upstream enhancers are brought into close proximity to active promotor regions by the loop formation protein CTCF. Finally the polymerase is recruited with initiates transcription.

Baked into the genetic sequence of all genes, specific signals exist for RNA polymerases to start transcription (transcription start sites, TSS) and terminator signals. The sequence of events leading to RNA polymerase recognition and binding of the TSS required for transcription is highly orchestrated and differs in complexity between organisms. In bacteria (the simplest), the RNA polymerase holoenzyme (consisting of σ-factor and core enzymes) weakly binds DNA until finding its target sequence; promotor regions upstream of the TSS (7).

Stable binding of σ-factor to the promotor allows for the melting of the double helix forming the transcription bubble and transcription initiation by the core enzymes. Bacterial promotor regions are found close to TSSs (10 – 35 bps) and can be more or less strong depending on their sequences (169,170). Different σ-factors recognize different promotor sequences, much like the transcription factors in eukaryotes (7,170).

Transcription initiation is the rate-limiting step in gene expression. In eukaryotes, this step has evolved by adding layers of control, including histones, enhancers and post-translational modifications. **Figure 11A** shows a typical gene locus. The core promotor region containing the TSS retains the basic structure seen in bacteria and is sufficient for basal expression (171). The highly conserved TATA box motif is the better-known eukaryotic core-promotor motif (usually found 30bp upstream of the TSS). It is only found in 5% of fly promotor regions and is thought to be sparsely used in other organisms (171,172). Other motifs such as the initiator motif (found overlapping the TSS and differ in sequence between organisms, Inr), as well as the downstream promotor element (DPE) initially found in flies, constitute the most widespread promotor elements (171). Much like the σ-factor in bacteria, eukaryotes rely on a host of general transcription factors (GTF) to position the RNA polymerase complex. These consist of multiple sub units (TFIIA, TFIIB, TFIID, TFIIF, TFIIS, TFIIE and TFIIH) that form the preinitiation complex (PIC)(173,174). TFIID, which contains the TATA-box binding protein (TBP), binds first even at loci devoid of TATA box motifs, followed by TFIIA, TFIIB and the Pol II – TFIIF complex (171,175). The binding of GTFs occurs at promotor elements which are often devoid of nucleosomes (even in transcriptionally silent genes but not in inactivated loci (heterochromatin)) (176). These regions have been defined as nucleosome-free regions (NFR). However, it is still unknown if these regions are genuinely devoid of nucleosomes or if they contain very labile nucleosomes containing H2A.Z or H3.3, dubbed 'fragile' nucleosomes (171,177). These NFRs often contain specific DNA sequences, including GC rich sequences, often referred to as CpG islands (CGI**,** mainly found in vertebrates), or homopolymeric dA:dT tracts, both of which are unamenable to nucleosome formation (176,178). Mapping nucleosomes around promotor regions have shown 3 distinct promotor types differing in the number of TSS's, PTMs and 'shape'.

Enhancers are sequences upstream of the TSS (up to megabases in mammals) that can activate gene expression (**Fig 11B**). Unlike promoters, enhancers are position and orientation

independent vis-à-vis the TSS. Enhancers themselves are found to be 100 bp to 100 kbp long, with the latter called super-enhancers. These regions contain an abundance of transcription factor binding motifs and are depleted in nucleosomes (DNase I Hypersensitive regions) as well as hyperacetylation (H3K27ac) and hypermethylation (with more H3K4me than H3K4me3, contrary to promotors) (179). These PTMs are deposited by p300 and CREB-binding protein (CBP) for acetylation and the mixed-lineage leukaemia (MLL) complexes for methylation, the presence of these factors are clear indications of active enhancers (180). Due to the large distance between enhancers and promoters, a looping mechanism has been suggested, whereby the active enhancer is brought into proximity of promotors in 3D space to activate transcription (**Fig 11B-C)**. This is supported by chromatin capture techniques which revealed contacts between distal enhancers with promotor regions.

Loop inducing and maintenance factors have been identified. These include Mediator, CCCTC-binding factor (CTCF) and Cohesin (181). CTCF and Cohesin colocalize within genomes. They can help partition and insulate active genes from inactive genes, forming topologically associated domains (TAD) (181,182). With its ring-like structure, Cohesin forms around DNA and feeds DNA through its lumen until encountering CTCF. This is known as the loop extrusion model. This model, however, cannot explain how only certain CTCF sites play key roles, whereas the majority of sites do not play a role whatsoever. The model also cannot explain why the majority of active enhancers have not been shown to be physical loops by high-resolution chromatin capture techniques (181). Interestingly, enhancer regions show transcriptional activity and generate small enhancer RNA (eRNA) that are quickly degraded by the exosome (179). The role of transcription at enhancers on gene regulation is not well understood, but one mechanism proposed includes a scaffolding role where eRNA recruits specific proteins, trapping TFs in a certain area (183). This trapping by eRNAs could induce phase separation, increasing the concentration of activating factors and excluding silencing proteins. Together, the idea of forming a membraneless sub-region within the nucleus containing specific genomic regions (enhancers, promotors and TSSs'), enriched by certain factors associated with transcription (Polymerases, TFs etc.), is the basis of the idea of liquid-liquid phase-separated droplets that drive transcription.

What has been observed in cells is that promotors containing the PTMs H3K4me3 and H3K27ac with a TATA-box as well as an Inr site found close to the TSS are found in terminally

differentiated cells. Transcription initiation occurs in a defined (sharp) locus (**Fig12A**) (171). Housekeeping genes tend to have a broader transcription initiation signal within a well-defined NFR. Phased nucleosomes containing H3K4me3 and H3K27ac flank this region (**Fig12B**) (171). These nucleosomes are commonly referred to as +1 and -1. In vertebrates, this promotor type is associated with CGIs. Finally, developmental genes (e.g. transcription factors involved in morphogenesis) can be found in stem cells containing activating marks and repressive PTMs. This state is called bivalent and thought to poise the gene locus for either repression or activation depending on the cell's developmental fate. These bivalent marks are H3K4me3 or H3K36me3 and H3K27me3 which can be found both within a single nucleosome (**Fig12C**). In vertebrates, CGIs are present and even found in multiple copies. Histone variants such as the labile H2A.Z are installed by chromatin remodellers in the promotor region, increasing DNA accessibility to effector proteins (184). Additionally, highly conserved long non-coding sequences separate this promotor type from elements such as enhancers (171). Together this chromatin landscape gives rise to dispersed transcription initiation signals sandwiched between two positioned nucleosomes.

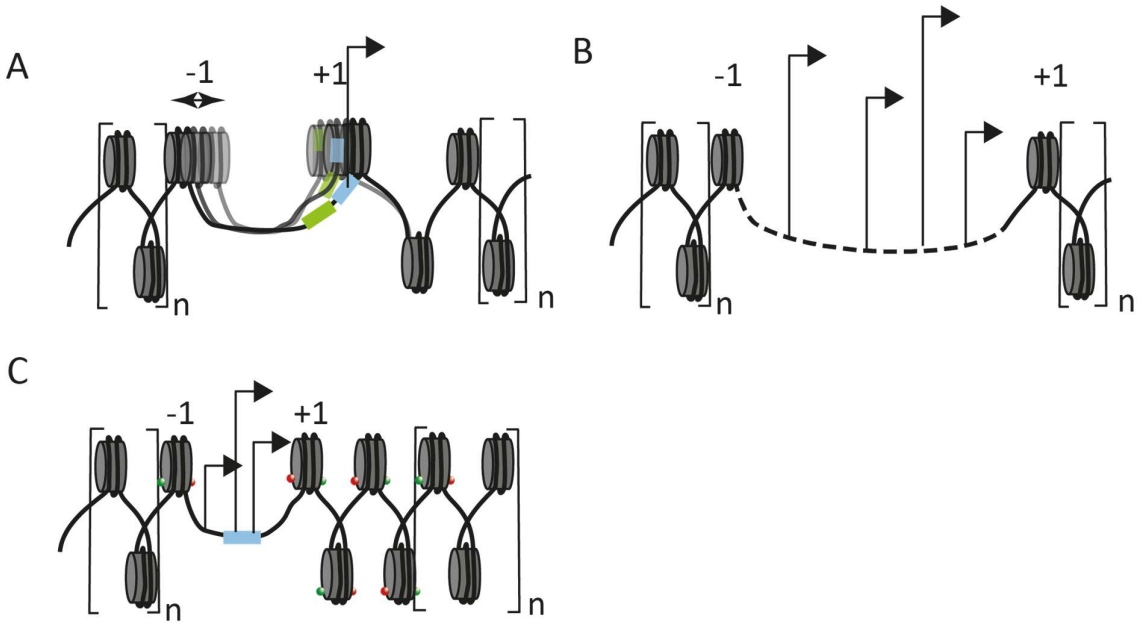


Figure 12 A) Scheme of promotor regions in differentiated cells showing a defined transcription peak (arrows) around the core promoter region. B) Housekeeping genes tend to have broad nucleosomes free regions with transcription initiation occurring within the whole area. C) Genes involved in development have bivalent histone marks, poising the loci for inhibition or transcription.

Evidently, the establishment and maintenance of these essential gene-regulating promotors and enhancers require more than just differences in DNA sequence to recruit the PIC. Wholesale chromatin landscape changes are an active process; for this, chromatin reading and effector proteins must be involved.

2 Aims

Previous studies highlighted the existence of transcription factors (TF) capable of nucleosomal DNA binding, pioneer factors (pTF). pTFs have been found to have distinct DBDs able to bind a single DNA face, which reduces steric clashes with the histone octameric core of nucleosomes. pTFs are thought to participate in transcription initiation by binding upstream regulatory sequences inducing changes in the local chromatin environment, and potentially recruiting co-factors. Open questions remain regarding the mechanistic role of pTFs within this process. These include whether the nucleosome modulates pTF binding and if pTF-nucleosome binding affects nucleosome stability.

In this work, I used colocalization single-molecule total internal reflection microscopy (smTIRF) to follow the changes in binding dynamics of the yeast general transcription factor Rap1 in different DNA substrates (**Chapter 3**). These include free DNA, mononucleosomes and chromatin arrays. We further explore the contribution to binding dynamics vis-a-vis the position within the nucleosomes and sequence composition of the Rap1 DNA binding sites. To better understand the effect of Rap1 binding on chromatin structure, we employed ensemble and single-molecule FRET to track changes in the structures of mononucleosomes and chromatin fibres in the presence and absence of Rap1. This work showed that Rap1 acts like a pTF, able to bind DNA sequences within chromatin fibres and decompact chromatin.

In vivo, nucleosome loss and formation of a nucleosome free region has been observed in the presence of Rap1. To recapitulate the formation of these nucleosome regions in promotor loci upon gene activation, we investigated the interplay between Rap1 and the chromatin remodeler RSC *in vitro* (**Chapter 4**). These experiments, including gel and MNase-Seq assays, led us to conclude that Rap1 biases the direction of RSC remodelling. This effect was more significant when using native DNA substrates compared to the Widom 601 nucleosome positioning sequence.

In order to broaden the investigation into the mechanisms governing the interplay between TFs and CRs, a smFRET based approach is discussed as a means to follow the real-time remodelling of a nucleosome substrate (**Chapter 5**). Initial data on the development of this modular platform using the model TF-CR pair composed on Rap1 and Chd1 shows that

remodelling can be tracked. However, the positions of the FRET dyes need to be further optimized as they currently display high rates of photobleaching.

Together, this thesis aims to better understand the dynamics of transcription factor binding in different chromatin contexts, using Rap1 as a model protein.

3 Pioneer transcription factor chromatin-binding dynamics

This chapter outlines work that was published in:

Mivelaz, M., Cao, A.M., Kubik, S., Zencir, S., Hovius, R., Boichenko, I., Stachowicz, A.M., Kurat, C.F., Shore, D. & Fierz, B. Chromatin Fiber Invasion and Nucleosome Displacement by the Rap1 Transcription Factor. *Mol Cell* **77**, 488-500 e9 (2020).

and

Mivelaz, M. & Fierz, B. Observing protein interaction dynamics to chemically defined chromatin fibres by colocalization single-molecule fluorescence microscopy. *Methods* (2020).

This work includes contributions from Anna-Maria Stachowicz for Rap1 constructs, smFRET data in chromatin fibres from Anne Marinette Cao.

Summary of chapter

The yeast general transcription factor Rap1 was expressed as a Halo fusion protein. After labelling with JF549, the binding affinity of tagged Rap1-Halo was tested against a set of PCR generated DNA substrates using single-molecule colocalization TIRF microscopy. It was found that the Rap1 dwell times on a free DNA substrate containing a Rap1 site is in the mins – hours time range. This same microscopy technique was again employed with mono-nucleosome and chromatin arrays containing 12 nucleosomes to test Rap1's pioneering abilities. We observed that Rap1 dwell times reduced 10 and 20 fold respectively compared to free DNA, but Rap1 could still access and bind nucleosomal DNA. Furthermore, by tiling the Rap1 sites in 3 bp increments, we saw that Rap1 binding affinity demonstrates some periodicity. Although Rap1 is always able to bind a partial motif, its binding affinity is modulated. Ensemble FRET using mono-nucleosomes labelled with FRET compatible dyes in the linker DNA were used to follow potential structural changes of the nucleosome upon Rap1 binding. Titration of Rap1 did not reveal a decrease in FRET, suggesting that Rap1 binding does not destabilize the overall nucleosome structure. Finally, single-molecule FRET studies were performed using reconstituted chromatin fibres containing FRET pairs reporting on the overall compaction of the fibres. These studies demonstrated that the presence of Rap1 disrupted overall chromatin fibre compaction, with Rap1 decompacting chromatin fibres.

3.1 The refractory effect of chromatin on transcription factor binding

The chromatin landscape at active gene loci is strictly defined (**Fig 11A**), with core promoter elements sequestered in a nucleosome upstream of the TSS (the +1 nucleosome). This sequestration restricts the formation of the PIC, thus down-regulating gene expression. Further downstream after the nucleosome free region (NFR), a second well-defined nucleosome can be found (the -1 nucleosome). This entire region comprising the -1 and +1 nucleosome as well as the NFR contains DNA binding motifs for various TFs. However, access to these motifs is thought to be restricted by the presence of nucleosomes. As mentioned above, a subset of TFs, called pioneer transcription factors, are thought to access DNA motifs even if they are sequestered within a nucleosome. These pTFs are thought to have unique features such as a DNA binding domain capable of binding a single face of DNA when solvent-exposed, even when only partially exposed. *In vivo* experiments have highlighted the key role played by pTFs in transcription initiation and the changes in nucleosome occupancy induced by the presence of a pTF. However, the mechanism by which these pTFs alter the chromatin landscape remains unclear. In this chapter, the questions of how pTFs bind their target DNA within chromatin and their direct effect upon binding is explored. The aim is to better understand the role played by pTFs, whether pTFs directly evict nucleosomes or change chromatin dynamics to ultimately generate the archetypical active promotor landscape. For this, the yeast general transcription factor Rap1 is used as a model protein for *in vitro* studies with reconstituted chromatin fibres.

3.1.1 The yeast general regulatory factor Repressor Activator Protein 1 (Rap1)

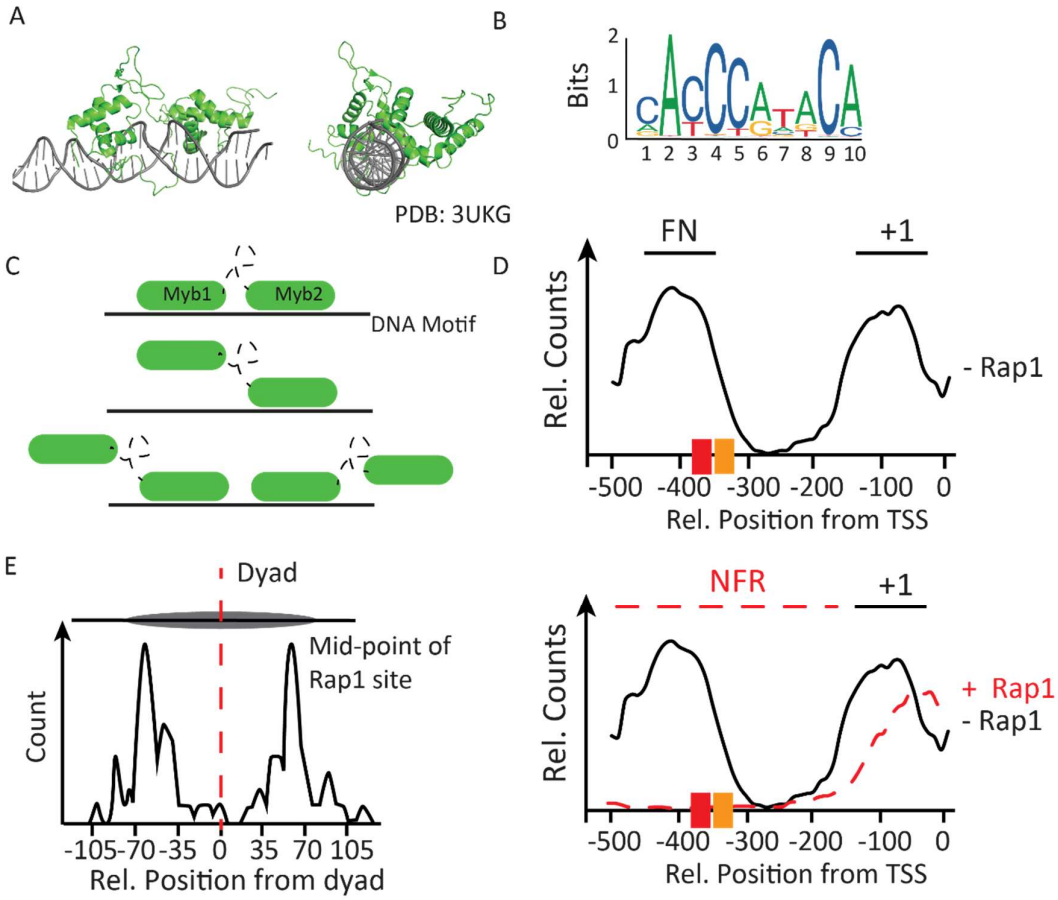


Figure 13 Repressor Activator Protein 1 (Rap1) A) Chrystal structure of the Rap1 DBD which consists of 2 tandem Myb-like helix-turn-helix domains capable of binding a single face of DNA. B) Rap1's DNA binding DNA motif. C) Rap1 can bind DNA in multiple conformations either engaging both Myb like DBDs or s single domain. This means that a single 10-12 bp binding domain can accommodate 2 Rap1 proteins. D) ChIP-Seq data of the Rpl30 promotor region in the presence or absence of Rap1. The stereotypical +1 and -1 nucleosomes are present in the absence of Rap1, however when Rap1 is present we see a depletion of the -1 nucleosome and an increase in the NFR. E) Mapping Rap1 binding sequences within nucleosomes show that the entry-exit regions of nucleosomes are enriched with Rap1 seqeunces.

Rap1 is an essential protein in budding yeast. It is involved in telomeric integrity, repression of noncoding divergent transcription and controls roughly 5 % of yeast genes (109,110,185–187). Rap1 is involved in the expression of nearly all ribosomal proteins, representing ~50 % of all RNA polymerase II initiation events (185). *In vivo* mapping of Rap1 showed high nucleosomal occupancy, which may be due to the distribution of strong Rap1 binding sites in DNA loci in promotor regions (188). This was further explored *in vitro,* and Rap1 was shown to be able to bind its DNA motif within nucleosomal DNA (189). The capacity of Rap1 to bind

nucleosomal DNA comes from its tandem Myb-like 3-helix bundles, which can bind DNA binding motifs consisting of 12 – 13 bp (**Fig13A-B)** (190–192). These DNA binding domains (DBD) confer the ability to bind a single face of DNA, therefore reducing possible steric clashes when binding nucleosomal DNA (192). Rap1 is thought to bind its motif in different conformations with its tandem domains, either engaging a single Myb-like domain per 12 - 13 bp DNA binding motif or both tandem Myb-like domains simultaneously (**Fig13C**) (191). Mapping the distribution of Rap1 DNA motifs revealed the existence of motif clusters, with multiple DNA binding motifs separated by 20 – 100 bp (185,188).

As Rap1 is an essential protein, deletion experiments to follow changes in chromatin landscape or transcription rates in the absence of Rap1 are not feasible. To circumvent this problem, depletion methodologies have been utilized, including anchor-away and auxin degron experiments. Mapping changes in nucleosome distribution in these Rap1 depleted cells demonstrated the key role of Rap1 in tagging nucleosomes for removal/remodelling (48,109,185). ChiP-Seq and other data show that Rap1 motif-containing nucleosomes are removed in the presence of Rap1 (**Fig 13D**). In sequences with dual Rap1 motifs, such as in the locus of the promotor region of 60S ribosomal protein L30 (RPL30), mutating individual sites contributes differently to nucleosome removal (185). These sites not only have different sequences (therefore different Rap1 binding affinities) but they are spaced out differently within the nucleosome, potentially modulating Rap1 recruitment. The distribution of Rap1 sequences within nucleosomes is enriched of DNA binding motifs within the first 2 solvent-exposed major grooves (**Fig 13E**) (193).

Sequence-specific binding of Rap1 also recruits other TFs, including Fhl1, ifh1, sfp1 and Hmo1, which all contribute to RP gene regulation (185,194). The differing presence of these factors at different RPG promotor regions can be categorized into two distinct architectures. Category I (CatI) promotors have a HmoI binding site directly downstream from Fhl1/ihf1, whereas, in category II (CatII) promoters, Hmo1 binding sites are not present. Additionally, CatI promotors have Rap1 sites further away from the TSS compared to CatII (185). These distances are roughly -400 bp for CatI and -200 bp for CatII (185).

Rap1 has reported low nano-molar binding affinity, and it's well-established *in vivo* effect on the chromatin landscape makes it an ideal candidate for elucidating the binding mechanism of TFs and their effect on the chromatin landscape. As a DNA binding target, the Rpl30

promotor region, which contains two Rap1 DNA binding motifs, allows to test the influence of DNA binding sequence affinity and position within the nucleosome. For these reasons, both Rap1 and the Rpl30 promotor region were chosen for the experiments detailed below.

Recent literature pointed out the contribution to DNA binding of regions outside of the DBD. Therefore purification of full-length Rap1 was desired (65). Additionally, recombinantly expressed histones and PCR generated template DNA were produced to reconstitute nucleosomes containing different Rap1 binding motifs. Together, these allow for the *in vitro* study of Rap1 DNA binding using single-molecule colocalization microscopy. For these experiments, both Rap1 and the target DNA need to be labelled with dyes of non-overlapping excitation and emission spectra to localise them on a glass slide.

3.1.2 Cloning and Expression of Rap1 in sf9 insect cells

Rap1 is a large 827 amino acid (aa) protein with a molecular weight of 92 kDa (pI = 4.83) composed of 5 characterized domains (**Fig 14A**). The BRCA1 C-terminal (BRCT) domain has been shown to be important in regulating genes involved in glycolysis (195). However, it plays an unknown mechanistic role in Rap1. Studies on other proteins with this domain have shown that BRCT is capable of binding phosphopeptides, participating in DNA binding and forming BRCT-BRCT dimers (196). Tandem Myb-like helix-turn-helix DNA binding domains constitute the DBD, the most characterized region of Rap1. It has been suggested that the first Myb like domain is capable of binding H3/H4 dimers, thus acting as a SANT domain and contributing to Rap1 nucleosome binding (197). Overexpression of Rap1 is cytotoxic, but deletion of a small region rescues this phenotype. The Tox domain is found at the C-terminus of the DBD (Fig 14B) (198). Further downstream, an acidic activation domain (Act) has been suggested to interact with transcription cofactors such as TFIID (199). Finally, at the C-terminus the RCT domain was shown to recruit gene silencing proteins Sir3p and Sir4p (200).

Previous unpublished work from the laboratory of Prof. David Shore (University of Geneva) suggested that expression of full-length Rap1 required the use of a baculovirus-insect cell expression system. Therefore, a Rap1 construct with an N-terminal streptavidin (Strep) tag followed by a maltose-binding protein (MBP) solubility tag fused to Rap1 via a cleavable linker with a TEV-protease site and a self-labelling C-terminal HALO tag was cloned (**Fig 14B**).

Expression tests showed an optimal expression time of 3 days at 27°C, after which cells were harvested by centrifugation and flash-frozen (**for complete methods, see Materials and Methods section**). After purification using a Strep column (**Fig14C**), the MBP-Rap1-Halo construct was labelled using a Halo-JF549 fluorophore and was cleaved using TEV-protease to remove the MBP tag (**Fig14D**). Final purification was done using size exclusion chromatography to remove free excess dye and separate the Rap1-Halo from the free MBP and TEV-protease (**Fig 14E**).

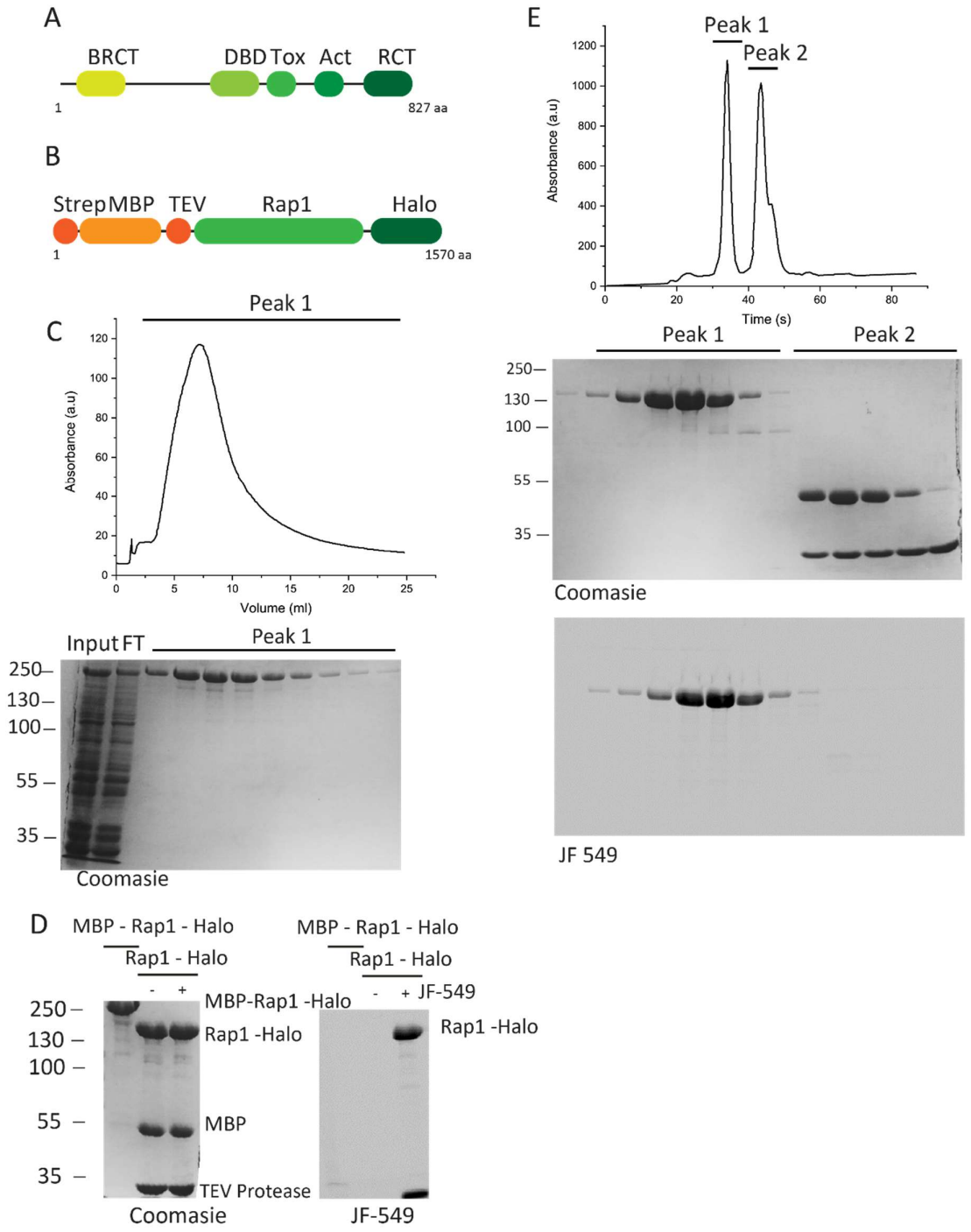


Figure 14 A) Known Rap1 protein domains including the BRCA1 C-terminal (BRCT) domain, tandem Myb-like helix-turn-helix DNA binding domains (DBD), cytotoxicity inducing domain (Tox), acidic activation domain (Act) and the Rap1 C-terminal domain (RCT). B) Scheme of the fusion protein expressed using the baculovirus expression system. With a streptavidin tag (Strep) fused to a maltose binding protein (MBP) for solubility. A cleavable TEV site allows the removal of MBP.Finally, a C-terminal self-labelling Halo tag allows for conjugation with a fluorophore. C) UV chromatogram of Strep tag purification which shows a single large peak. Below is the SDS-PAGE gel of each fraction to identify those containing Rap1. D) SDS-PAGE gel of the overnight TEV protease cleavage reaction of MBP-Rap1-Halo. This was done on two separate fractions of Rap1, one labelled with JF-549 and one un labelled. E) Size exclusion chromatography FPLC purification to separate Rap1-Halo from the cleaved MBP tag and TEV protease. Below is the SDS-PAGE gels of each fractions from the corresponding peaks.

3.1.3 *In vitro* Rap1 characterization

With labelled and purified recombinant Rap1, it is now possible to test the binding characteristics of our construct. This also acts as a quality control to confirm that the full-length Rap1 construct is functional. A typical methodology used is electromobility shift assays (EMSA), which has a gel-based readout. The protein of interest is titrated against a fixed concentration of target DNA, and a complex is allowed to form. After 10 mins, the reaction mixture is added to a native polyacrylamide gel and run. Free target DNA will run to the correct size through the gel. In contrast, bound DNA will form a DNA:Protein complex that will migrate differently (usually a super shifting from the original size to a higher molecular weight).

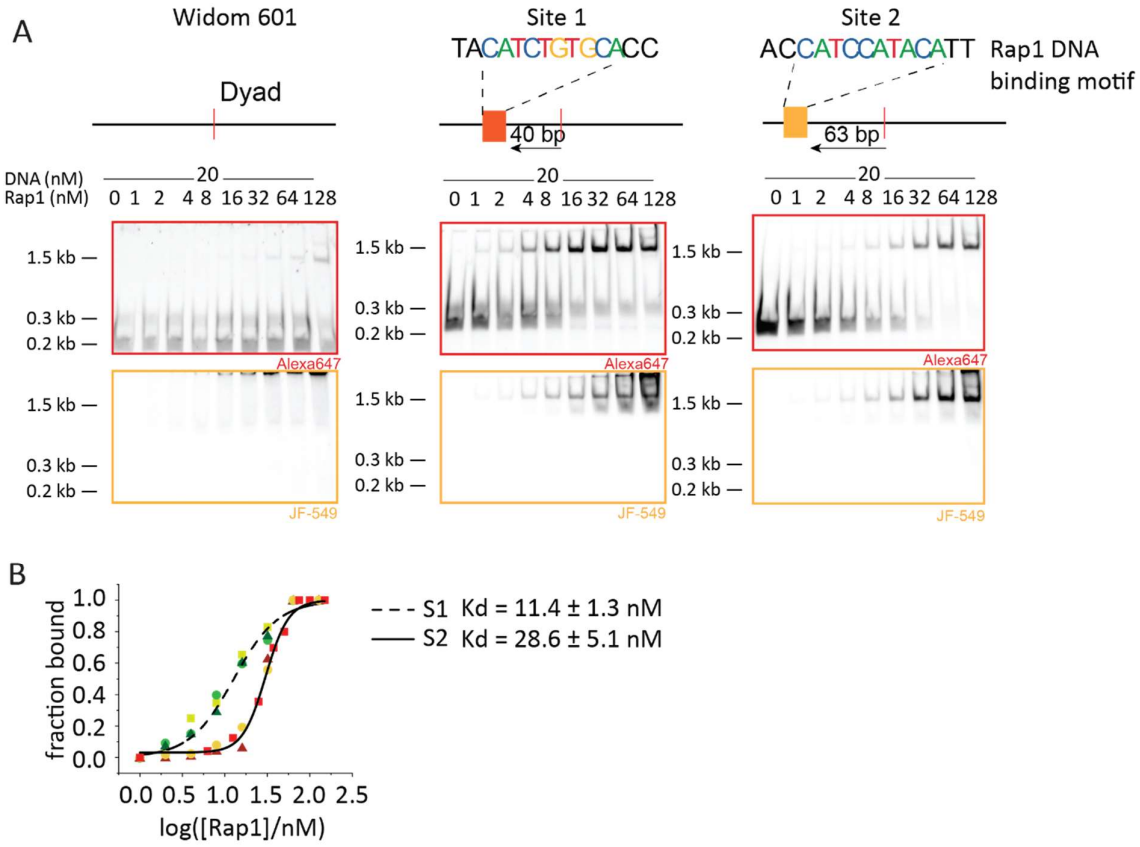


Figure 15 A) Electromobility Shift Assays (EMSA) of labelled Rap1-Halo with DNA constructs either devoid of Rap1 binding sites (Widom 601) or based on the RPL30 promoter region with Site 1 (S1) and Site 2 (S2) tested separately. B) Densitometry quantification of Rap1 DNA binding for both S1 and S2.

EMSAs were performed with purified labelled Rap1, and a target ~ 200 bp labelled DNA sequence consisting of a nucleosome positioning sequence, the Widom 601, with either of

the 2 Rap1 binding motifs from the RPL30 promoter region at one time. These binding sites differ in sequence and position from the dyad as mapped from ChIP-Seq data *in vivo* (185). As a control, just the Widom 601 was used. These experiments show that Rap1 is functional and that it preferentially binds to DNA sequences containing its DNA binding motif (**Fig 15A**). Binding affinities were calculated using densitometry to quantify the proportion of free DNA and DNA:Protein complex using the fluorescence readout. Data were fit using a non-linear sigmoidal model to find the binding affinities of Rap1 to both Rpl30 site 1 (S1) and Rpl30 site 2 (S2). In agreement with previously published literature S1 was found to have a higher affinity than S2, with K_Ds of 11.4 nM and 28.6 nM, respectively (**Fig 15B**).

These experiments show that we have purified, labelled, functional Rap1 able to bind its DNA motifs with similar binding affinities as previously described in the literature (185). However, EMSA is a technique that looks at an ensemble. Therefore binding kinetic characteristics such as association (k on) and dissociation (k off) rates cannot be inferred. For this, a single-molecule approach is required. As K_D is derived from $K_D = \frac{k_{off}}{k_{on}}$, teasing out the contribution of association and dissociation rates gives us better insight into the dynamic process of Rap1 search and binding dynamics. For example, it has been observed that some TFs can compensate their reduced association kinetics to nucleosomes by increasing dwell times, thus compensating a potential decrease in their K_D (87).

3.2 Single-molecule characterization of Rap1 DNA binding

3.2.1 Single-molecule total internal reflection microscopy

Single-molecule imaging using total internal reflection microscopy (smTIRF) has extensively been used to follow protein dynamics (156,201,202). TIRF illumination is generated when angled light meets an interface with a lower refractive index. From this, a uniform evanescent field is generated, penetrating the low refractive index medium by ~200 nm. Glass has a refractive index of n = 1.517. In contrast, water or the cell's cytosol has a refractive index of n = 1.33 and n = 1.38, respectively. The low penetration depth allows for high signal to noise,

as only a small fraction of molecules which are at the surface are illuminated. However, this illumination depth is a limitation for cell experiments, as you are constrained to imaging the cell surface. Therefore, using a glass slide as support and either an aqueous medium or a cell surface as samples would be amenable to the use of TIRF illumination (Fig 16A).

Microchannels are assembled to reduce the volume of reagents required (20 – 30 µl). They also allow us to exchange buffers or flow in proteins at defined time points (**Fig 16B**) (202). Furthermore, these airtight microchannels also reduce the gaseous exchange, reducing the concentration of O_2 in solution (in collaboration with oxygen scavengers). O_2 is a well-known contributor to the irreversible photochemical modification of fluorophores by excited triplet state O_2 molecules, also known as photobleaching. Therefore, reducing O_2 or reducing excited triplet state O_2 are two strategies to increase the lifetime of fluorophores, ultimately increasing the number of photons emitted per fluorophore before photobleaching. We employ glucose oxidase – catalase (GODCAT) as a glucose-dependent oxygen depletion system and a combination of 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox), cyclopolyenes (cyclooctatetraene or COT) and nitrobenzoic alcohol (NBA) as additives to quench excited triplet O_2. This being said, photobleaching remains a concern for all experiments and bleaching rates for fluorophores should be monitored for different buffers, dyes and illumination conditions.

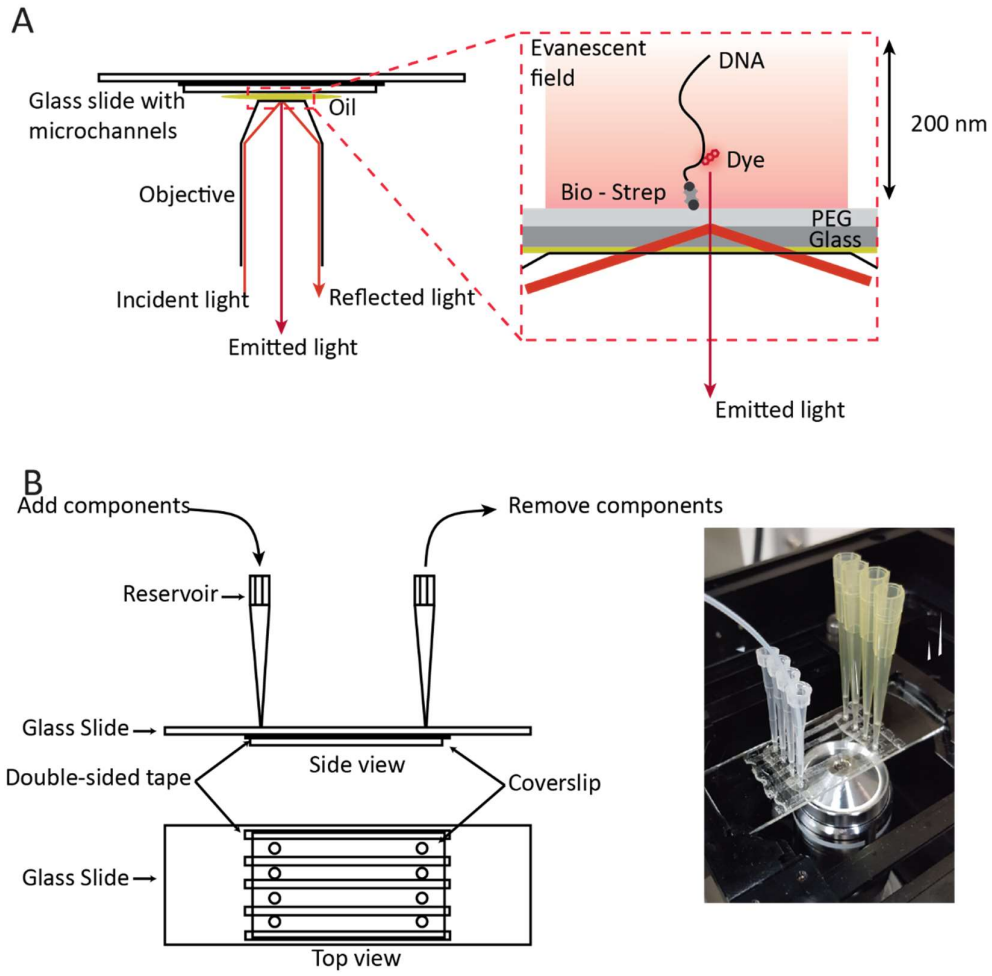


Figure 16 Total Internal Reflection Microscopy (TIRF) A) scheme of the evanescent wave generated by the high angle incident laser generating TIRF: B) Microchannels used for single-molecule TIRF experiments.

During a smTIRF colocalization experiment, labelled DNA is immobilized on the surface of a microchannel through the biotin-streptavidin interaction. Subsequently, protein labelled with a dye having an orthogonal excitation and emission spectrum is flowed into the same microchannel. Illumination in the wavelength of the dye on the DNA is used to find surface-immobilized DNA molecules. In contrast, illumination of the dye on the protein allows us to follow the position of the free moving protein molecules (**Fig 17A**). Illumination can be continuous for both channels or sequential (e.g. one Red pulse (DNA), three green pulses (protein) and so forth); this depends on the optical setup (**Fig 17B**). Furthermore, the times between pulses or dark times can be programmed, thereby potentially extending the time of an experiment by limiting photobleaching. When both dyes (DNA and protein) are localized in the same area (defined by the experimenter), an interaction can be inferred between both

molecules. Colocalization times, including the time a single protein remains bound to its target and the interval between subsequent binding events, inform us of the target protein pair dynamics (**Fig 17B**).

As these are defined *in vitro* experiments, the DNA sequences, protein constructs and buffer components can be defined and controlled. They are allowing for the real-time interrogation of protein association and dissociation kinetics in different chromatin contexts.

3.2.2 smTIRF of Rap1 on free RPL30 promotor DNA

The binding kinetics of Rap1-Halo labelled with JF 549 to the identical DNA sequences characterized using EMSA showed a similar overall tendency. S1 binding by Rap1 was the longest, too long to quantify with our experimental setup as we were limited by photobleaching with bound times that were over 40 mins long (**Fig 18A**). However, cumulative lifetime histograms of S2 binding times by Rap1 were compiled, and the data fit a bi-exponential function (**Fig 18B**). This suggests that Rap1 has been observed to exist in 3 different states, free Rap1, Rap1 bound with a short residence time and finally Rap1 tightly bound (**Fig 18C**). The residence times for the short events was τ_{off1} = 12.4 ± 4.5 s and τ_{off2} = 452 ± 115 s for the tightly bound Rap1 events. The length of these events meant that the illumination sequence had to be optimized to limit photobleaching. For this, dark times of 600 ms were inserted between each 100 ms illumination. Our time resolution is, therefore, 700 ms (100 ms t_{on} + 600 ms t_{off}). This means that a molecule that is bound for 452 s is only excited for 63 s in total. Photobleaching experiments using Rap1 JF-549 displayed a photobleaching time constant of 168 s (**Fig S1**). We can correct for photobleaching by dividing our histogram by the mono-exponential photobleaching fit with the 168 s time constant. With this, we see that the tightly bound Rap1 events are actually 721 ± 183 s long (Table S2). In total, 35 % of all observed events were short, whereas the remaining 65 % belonged to the long, tightly bound state (Table S2).

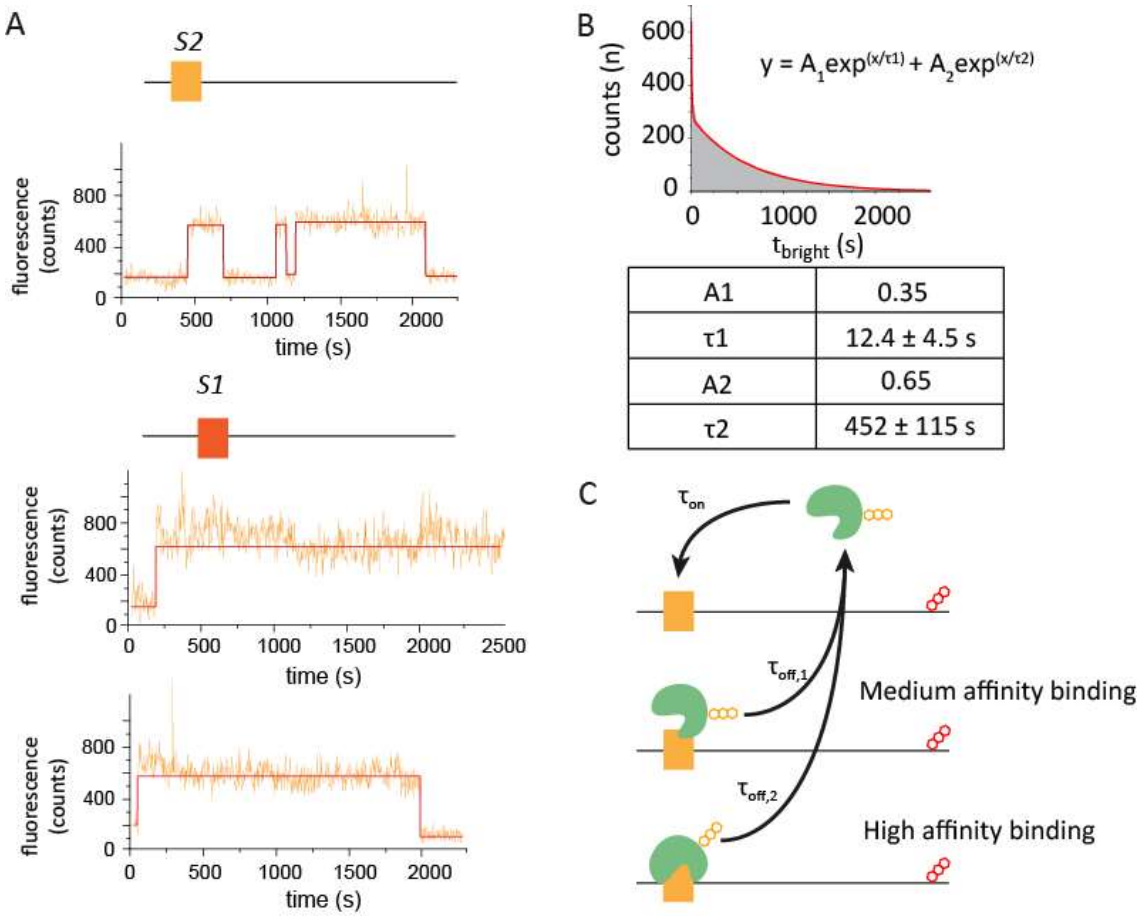


Figure 17 Single-molecule TIRF colocalisation experiments of a RPL30 promotor region DNA mimic using the Widom 601 seqeunce. A) Site 1 (S1) and site 2 (S2) Rap1 binding motifs were independetnly tested to measure the difference in affinity due to their different sequences. S2 showed binding events in the seconds to minutes, whereas S1 binding was too long to measure using our experimental setup. B) Quantification of S2 binding using a cumulative histogram of the duration of binding events. These data are fit with a bi-exponential function. C) Model of Rap1 DNA binding using the information gathered from our smTIRF colocalization experiemtns. We describe 3 states, free unbound Rap1 which can bind with a τ_{on} either in a medium affinity conformation. This could be Rap1 engaging the Rap1 binding site with a single Myb-like domain, and a high affinity binding bmode with both Myb-like domains specifically bound.

This experiment demonstrates that we can use smTIRF colocalization as a method to decipher Rap1 binding kinetics *in vitro*. We can also see that Rap1 binds to S1 with longer dwell times compared to S2. As mentioned above, *in vivo,* Rap1 binding sites are found within nucleosomes, including in the Rpl30 promotor region. Therefore, free DNA only recapitulates the remodelled nucleosome landscape. To probe the effect on Rap1 binding when its DNA binding motif is found within nucleosomal DNA, our smTIRF colocalization assay can be extended by using nucleosomes in lieu of free DNA.

3.2.3 smTIRF of Rap1 on nucleosomes with RPL30 promotor DNA

The presence of dual Rap1 DNA binding motifs within a single promoter region is found in many Rap1 target sites (185,188). In the Rpl30 promotor region, both S1 and S2 are found in the -1 nucleosome. S1 is positioned 40 bp from the dyad and S2 63 bp away. This places S1 at SHL 4.5 and S2 at SHL 6.5 within a nucleosome (**Fig 19A**). *In vivo* transcription assays using YFP fluorescence as a reporter, indicated that both S1 and S2 contributed to overall transcription. However, mutations of S1 had a more dramatic effect in reducing YFP transcription (185). This study also performed Rap1 ChIP; Rap1 was less enriched when S1 was mutated, suggesting a link between DNA binding motif strength and transcriptional output, as well as cooperative binding between S1 and S2.

Other studies demonstrated that TFs were excluded from nucleosome binding (including the bHlH DBD family). High throughput *in vitro* experiments built on this finding, adding that some DBD families had a periodic binding preference. Requiring the DNA binding motif to be solvent-exposed for the DBD to be able to bind (e.g. HD family members) (86). DNA within the nucleosome is highly bent, and some TF DBDs require this opening of the gyre to bind. Examples of this are members of the T-box family or the High mobility group, which preferentially bind highly bent DNA near the nucleosome dyad.

Rap1 sites are found preferentially at the periphery of nucleosomes, although some have been shown to exist close to the dyad. Insertion of 1 of the 3-α-helices of the helix-turn-helix DBD would suggest that Rap1 may exhibit a periodical binding preference. However, the size of the DNA binding motif means that a partial DNA binding motif is always solvent-exposed. Both EMSA and smTIRF colocalization were used on a set of Rpl30 promotor DNA constructs to determine the effect on Rap1 binding of DNA sequence and nucleosomal position (**Fig 19B**).

3.2.4 The Rpl30 promotor region nucleosome set

All 4 human core histone proteins H2A, H2B, H3 (C110A) and H4, were expressed and purified independently. Equimolar amounts of H3 and H4 are mixed with an excess of H2A and H2B in

denaturing buffer. The solution is titrated into a high salt buffer to form octamers and purified by gel filtration (**Fig 19C**) (203).

For nucleosome reconstitution, purified octamers are titrated against DNA (the same DNA as previously described; A Widom 601 NPS template with/without a Rap1 binding site, biotin and Alexa 647). Super-shifting of the DNA band indicates correct nucleosome assembly (**Fig 20A**). A defined band also indicates that nucleosomes are all phased (**Fig 20B**). This ability of

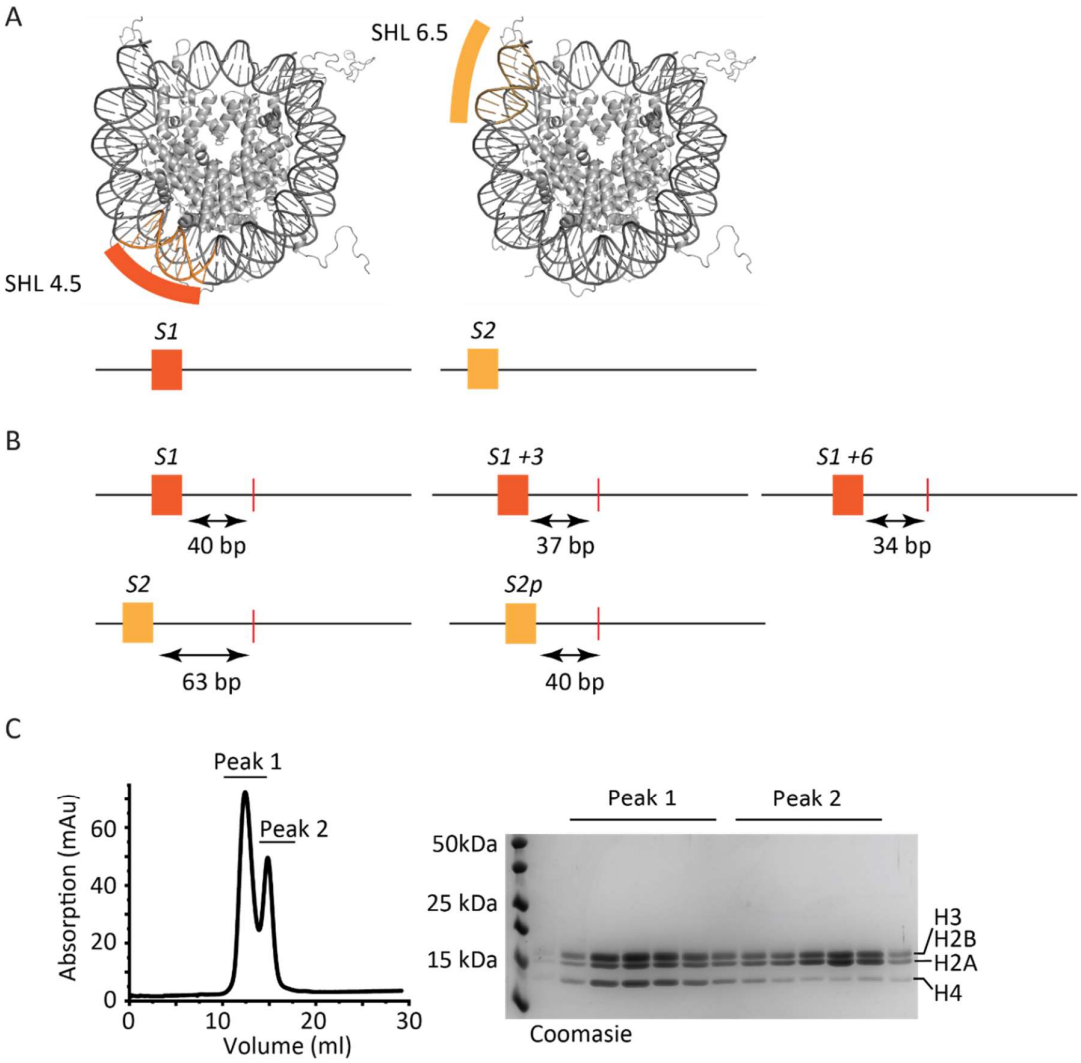


Figure 18 A) Scheme of the relative positions of Site 1 and Site 2 within the nucleosomes, based on ChIP-seq information of the Rpl30 promotor. B) DNA set of Rap1 binding sites with their relative positions from the dyad.C) FPLC purification profile and SDS-PAGE gel of purified octamer assemblies used to reconstitute nucleosomes.

nucleosome positioning sequences to repeatedly generate homogenous samples allows us to guarantee the correct positioning of the inserted Rap1 site. This being said, some genomic sequences can also be reconstituted into nucleosomes, albeit less defined.

3.2.5 Characterizing Rap1 binding to nucleosomes using EMSA

Focusing on S1, reconstituted nucleosomes with S1 in its natural position (+ 40 bp from the dyad) and S1 shifted by +3 and +6 bp were tested to see whether Rap1 exhibits periodicity in binding (**Fig 20C**). Calculated affinities for these constructs show some variation, with the +3 shifted S1 position having the highest affinity (48.5 ± 7.4 nM), but Rap1 is not excluded from binding the nucleosome (**Fig 20C**). This ability could stem from the RPL30 Rap1 binding motif itself, as nucleotides crucial to binding (as seen in the binding motif position weight matrix representing single letters A in position 2 and the C's in positions 4 and 9) are spread across the DNA binding motif, always leaving a partial solvent-exposed motif. Compared to DNA binding alone, we see a 4-fold decrease in binding affinity in the presence of the nucleosome. Whether this decrease is due to highly bent DNA within the nucleosome or steric hindrance remains to be tested.

3.2.6 smTIRF characterization of Rap1 binding to nucleosomes

Previous studies have suggested that binding affinities can be maintained for certain TFs binding to nucleosomes by compensating the decrease of association kinetics by increased dwell times (87). Our S1 EMSA studies point to a positional redundancy of Rap1 binding motifs due to the long and bi-partite Rap1 binding sequence. To further test this, we used the S2 binding sequence, which showed long dwell times on free DNA and a 2-fold decrease in affinity compared to S1 using EMSA. Nucleosomes were immobilized within microchannels, and 50 – 100 pM labelled Rap1-Halo were flowed in and imaged. Analysis of the cumulative histograms revealed a high proportion of very short-lived events (τ_{off0} = 0.2 – 0.7 msec), around 50 % of all detected interactions were short-lived in constructs with a Rap1 motif present.

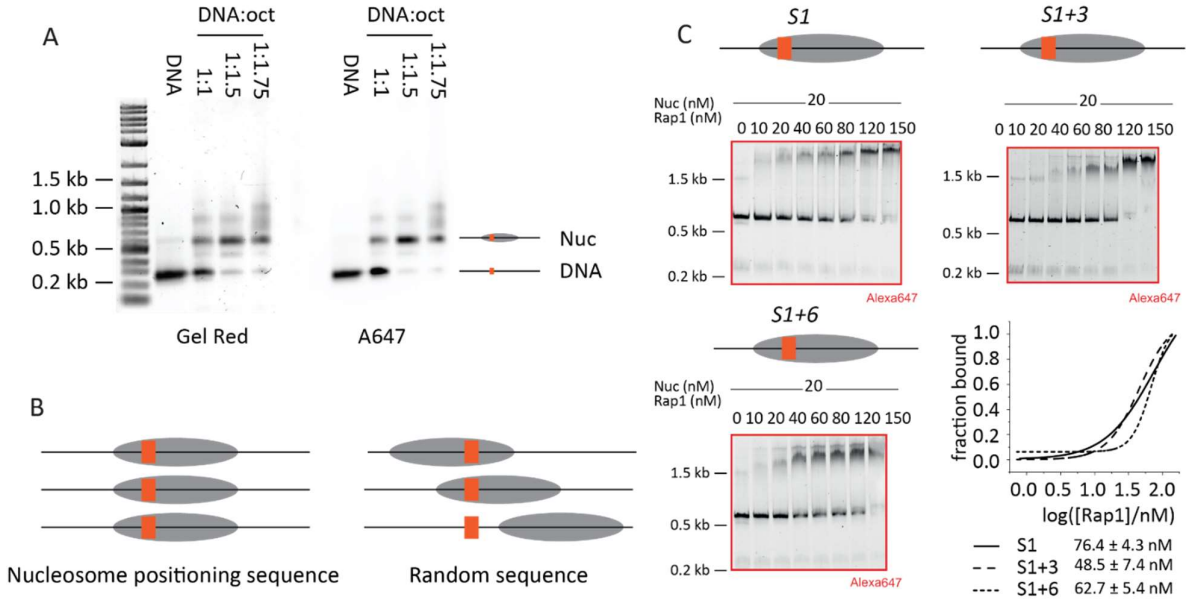


Figure 19 A) Nucleosome reconstitution by titration of octamers to a fixed amount of DNA, Nucleosome formation can be seen as a super-shifting of the free DNA band. B) Nucleosome phasing using nucleosome positioning sequences (NPS). An NPS such as the Widom 601 allow for homogenous phased nucleosomes populations, compared to randomly distributed populations as can be seen using random genomic DNA. C) Electrophoretic mobility shift assays (EMSA) using nucleosomes with tilled site 1 (S1) positions.

Interestingly, nucleosomes devoid of Rap1 show an increased proportion of these short events (83 %) and the remaining proportion of events with dwell times of $\tau_{off,1}$ = 3.5 ± 3 s (**Fig 21A-B**). This suggests that these short events are unspecific Rap1-DNA binding events, most likely involved in the search dynamics of Rap1 for its target DNA motif. These short unspecific searching interactions were similar for all constructs (**Fig21C**). Specific Rap1 binding times were 20 – 500 fold higher than those measured on Widom 601 only nucleosomes. S1, although further buried in the nucleosome, showed the highest specific dwell times, $\tau_{off,1}$ = 18 ± 11 s and $\tau_{off,2}$ > 100s. With S2 having dwell times of $\tau_{off,1}$ = 8.4 ± 1.4 s and $\tau_{off,2}$ = 46 ± 3 s. As mentioned above, S1 and S2 are not identical in sequence and display a 2-fold difference in binding affinity in free DNA. This 2-fold change seems to be conserved in nucleosome binding despite S1's position further within the nucleosome. To further explore this effect, S2 was permutated to sit at the same position as S1 (40 bp from the dyad (SHL 4.5) instead of 63 bp away (SHL 6.5)). This permutation, S2p, was associated with a 6-fold decrease in dwell times compared to the wt position, $\tau_{off,1}$ = 2.4 ± 0.4 s and $\tau_{off,2}$ = 7.7 ± 1.9 s. This drastic reduction argues against a universal evolution of Rap1 sites to present a rotation independent solvent-accessible Rap1 site for all sequences. It would seem that the S1 DNA sequence is more amenable to Rap1 binding within the nucleosome compared to S2. Tiling experiments using the S2 and other Rap1 binding sequences could be performed within the nucleosome to

explore this hypothesis further. The shorter dwell times observed compared to DNA meant that the imaging conditions had to be modified to include a dark time of 0.3 msec (compared to 600msec for free DNA). Cumulative dwell time histograms were shown to fit best with a tri-exponential model (except for Widom 601 only). To avoid overfitting, we plotted residuals of the fit (**Fig 21D**). The scatter of points should be random above and below 0 for an accurate fit from this analysis. If the dispersion is overfitted, then the scatter will be biased +/- 0.

Rap1 nucleosome binding presents a 10-fold reduction in dwell times compared to free DNA, and nucleosome position acts as a rheostat of Rap1 binding affinities. Additionally, we do not see nucleosome eviction during our microscopy experiments. If this would be the case, adding Rap1 into the microchannel would result in dwell times similar to DNA. This is not the case, even for experiments where Rap1 and nucleosomes have been incubated for over 10-mins. This was somewhat expected as no nucleosome eviction was seen during our EMSAs. This points to the fact that Rap1 by itself is not able to recapitulate the change in nucleosome landscape seen *in vivo*. We, therefore, wondered what effect did Rap1 have on the structure of the nucleosome, whether it induced rolling or breathing upon nucleosome binding.

Conversely, we were interested to see if breathing was a prerequisite for stable Rap1 binding within the nucleosome.

3.2.7 Induced nucleosome dynamics by Rap1 binding probed by ensemble FRET

Forster resonance energy transfer (FRET) is a non-radiative dipole interaction between a dye pair with overlapping emission/absorption spectra (**Fig 22B**). This photophysical interaction occurs through space and is dependent on the distance and orientation of the dye pairs (amongst other factors). This relationship has turned the use of FRET into a molecular ruler in biophysical studies able to report changes in distance in the range of 1 – 10 nm. Previous FRET

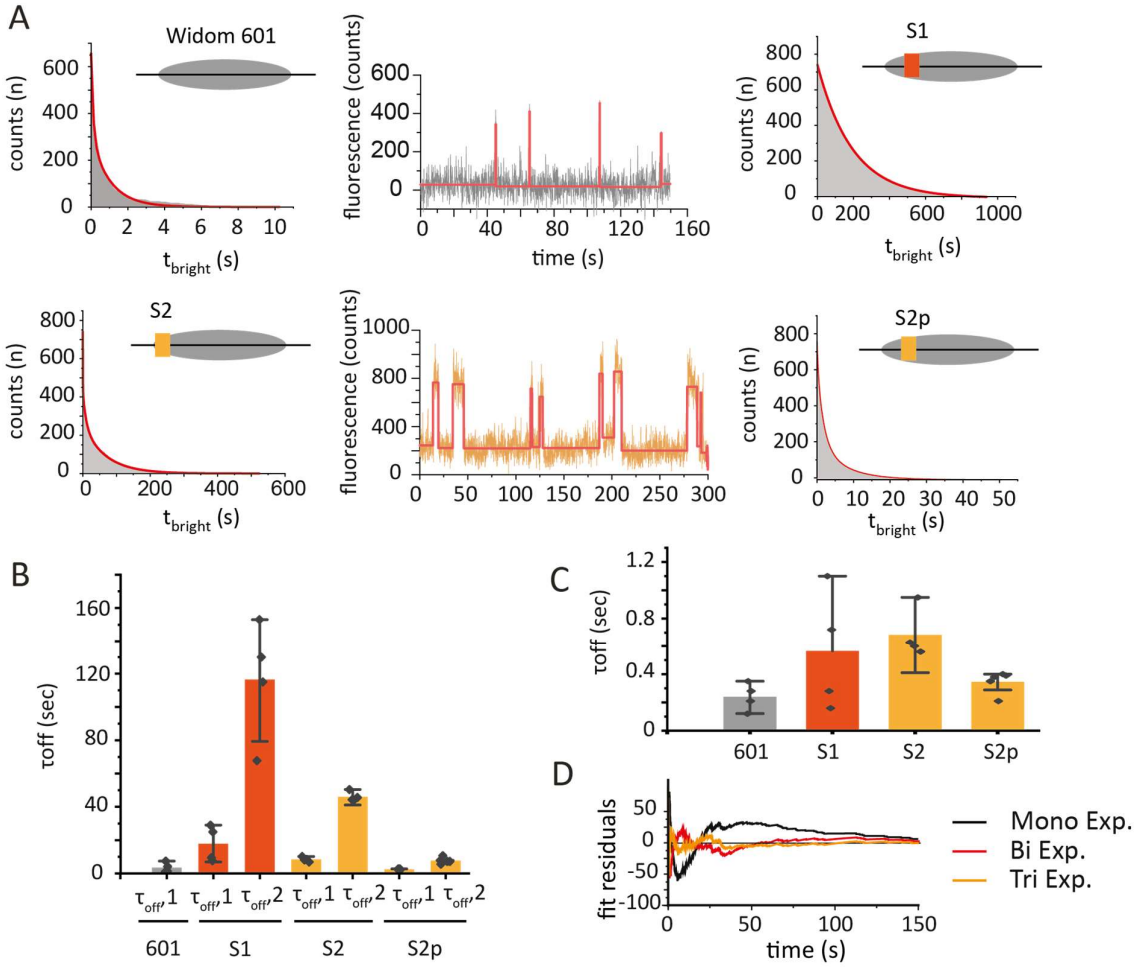


Figure 20 A) Representative cumulative binding time traces for all mononucleosomes tested Widom 601, Site 1 (s1), Site 2 (S2) and permutated Site 2 (S2p). Additionally, individual traces show probing short events for the Widom 601 and longer events for S2. B) Specific dwell times obtained from exponential decay fits for all nucleosomes. C) Unspecific dwell times obtained from exponential decay fits. D) Residual fit analysis of different exponential fits too limit under and over sampling of data.

studies demonstrated that dye pairs placed within the linker DNA of nucleosomes could report on the conformation of these regions (44). Using the known FRET pair Cy3b – Alexa647

with a reported R_0 of 62 Å, we reconstituted the FRET pair within the linkers, which exhibits a high FRET efficiency of E_{FRET} ~ 0.8 (Theoretical E_{FRET} = 0.88 based on PDB: 1ZBB and using Kalinin et al., software) (**Fig 22B-C**) (204,205). These ensemble measurements also showed an absence of FRET when only the doubly labelled DNA was present (**Fig 22C-D**). Rap1 was titrated against S1 / S2 containing nucleosomes using 1, 5 and 10 molar equivalences. Ensemble FRET measurements were recorded after 10 mins incubation at room temperature. To our surprise, Rap1 binding did not alter nucleosome structure. Furthermore, Rap1 does not seem to be dependent on nucleosome breathing, as no loss or reduction in FRET is seen upon binding (**Fig 22C-D**). As a control, 0.8 M NaCl was added to monitor the FRET change upon nucleosome dissociation. FRET values measured in 0.8 M NaCl were close to those of free DNA. Together, this demonstrates that Rap1 does not induce a change in structure upon nucleosome binding or that DNA breathing is required for Rap1 binding. Curiously, the S2 mononucleosome consistently showed lower FRET values compared to the S1 nucleosome. This may be attributed to the position of the S2 motif within the linker, potentially changing the angle of the exit/entry DNA. As previously mentioned some TFs can compensate a decrease in binding kinetics through increased dwell times. This is thought to occur through a trapping mechanism, blocking the nucleosome in an open state upon binding (87). As Rap1 does not alter the nucleosome structure, Rap1 does not seem to utilize this mechanism during nucleosome binding.

The *In vivo* chromatin landscape is not composed of isolated nucleosomes but rather multiple scattered nucleosomes forming chromatin fibres. To better reconstruct this *in vivo* architecture, chromatin fibres were reconstituted *in vitro* composed of a 12-nucleosome array, with a modular central nucleosome (N6) (**Fig 23A**). Using this array, the capacity of Rap1 to search and bind its motif was probed, as well as the real-time change in chromatin compaction induced by this process.

3.3 Single-molecule Rap1 binding to chromatin fibres

3.3.1 Cloning and reconstitution of chromatin arrays

Facile and modular access to DNA arrays was previously established in the lab (156). The strategy revolves around a combination of large recombinantly generated pieces with smaller centrally placed PCR generated ones. The repetitive nature of the tandem Widom 601 sequences means that PCR cannot produce the sequence. Conversely, the addition of modifications such as dyes cannot be done recombinantly. Therefore a mixed approach is required; an example of this is also proposed by the Sczepanski lab (206). For this, they recombinantly express the 12 x nucleosome array with recognition sites for a nickase (single-strand phosphodiester backbone cutter). Once the backbone is cleaved, the short piece of DNA can dissociate and be replaced by a PCR generated fragment of identical sequence before being ligated enzymatically.

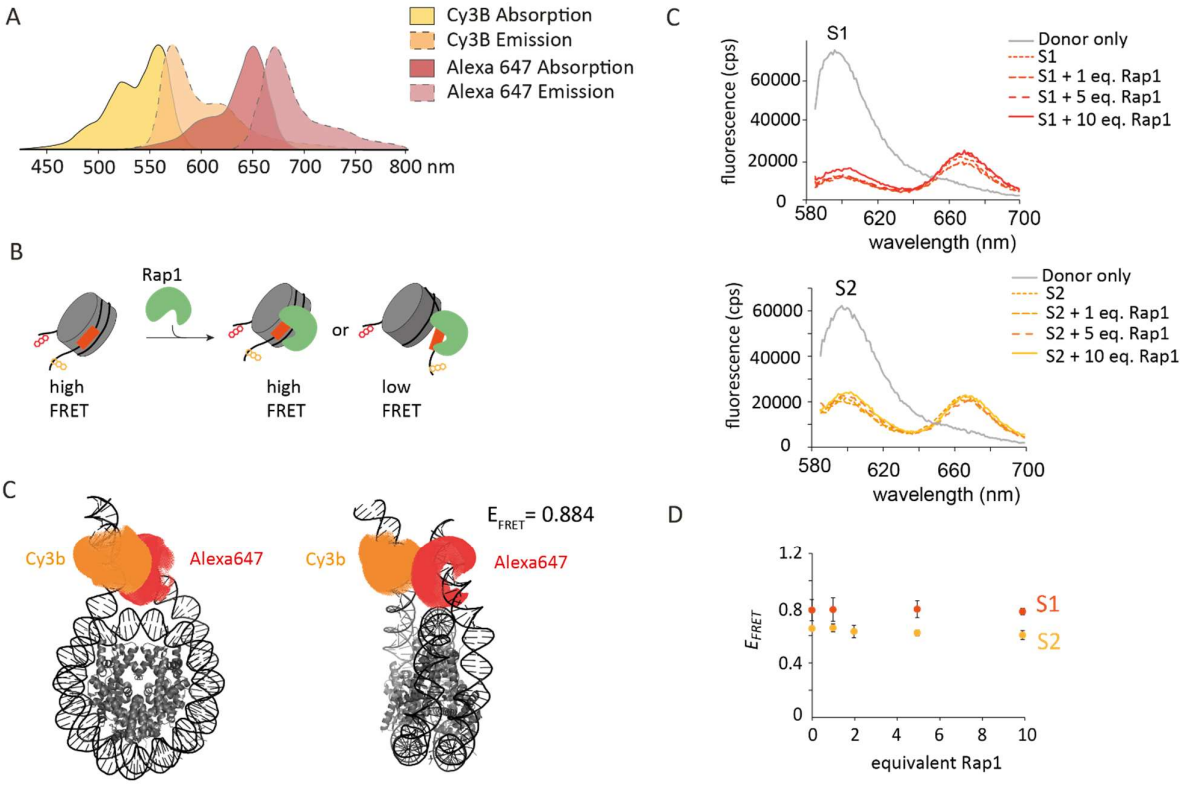


Figure 21 A) Spectral overlap of Cy3B and Alexa647 allowing FRET B) Scheme of the ensemble FRET experiment highlighting the two possible mechanisms of Rap1 nucleosome binding. Either Rap1 binds DNA on one face without disrupting nucleosome shape. Or Rap1 is dependent on DNA breathing and unwinds DNA upon binding, which would result in a loss in FRET. C) Scheme using PDB:1ZBB and software from the Seidel group to calculate theoretical FRET efficiencies. We find a potential E_{FRET} of 0.884. C) Ensemble FRET measurements of both Site 1 (S1) and Site 2 (S2) nucleosomes and donor only in the presence of Rap1 as indicated. D) Quantification of E_{FRET} shows no change even in the presence of excess Rap1.

Two different routes were taken to generate arrays for this work, a 3-piece (**shown in Fig 23B**) and a 5-piece ligation. Both routes have 2 large recombinantly expressed pieces, and either 1 or 3 PCR generated pieces. The requirement of 2-dyes for single-molecule FRET experiments means that 2 PCR generated sequences within two nucleosomes need to be inserted. Whereas in single-molecule colocalization, only a single modification is inserted, requiring fewer PCR generated pieces. An example of 3-piece purification is presented in **figure 23D**. Plasmid DNA containing 5 and 6 Widom 601 repeats, respectively (recP12 and recP45), are digested with nonpalindromic restriction enzymes DraIII and BsaI (**Fig23C**). After complete digestion, the plasmid backbone is further digested with EcoRV to facilitate the isolation of the fragment of interest by Polyethylene glycol (PEG) purification (**Fig2C-D**). The central fragment (P3) is produced through PCR and digested with DraII and BsaI (**Fig23E**). Initially, the recP12 and P3 are ligated with an excess P3 (**Fig 23F**) before being PEG purified (**Fig23G**). This P123 fragment is then ligated to recP45 (**Fig23H-I**) and a biotin anchor. Yielding a final 12 – nucleosome array with a dye, a modified central nucleosome and a biotin anchor (**Fig23J**).

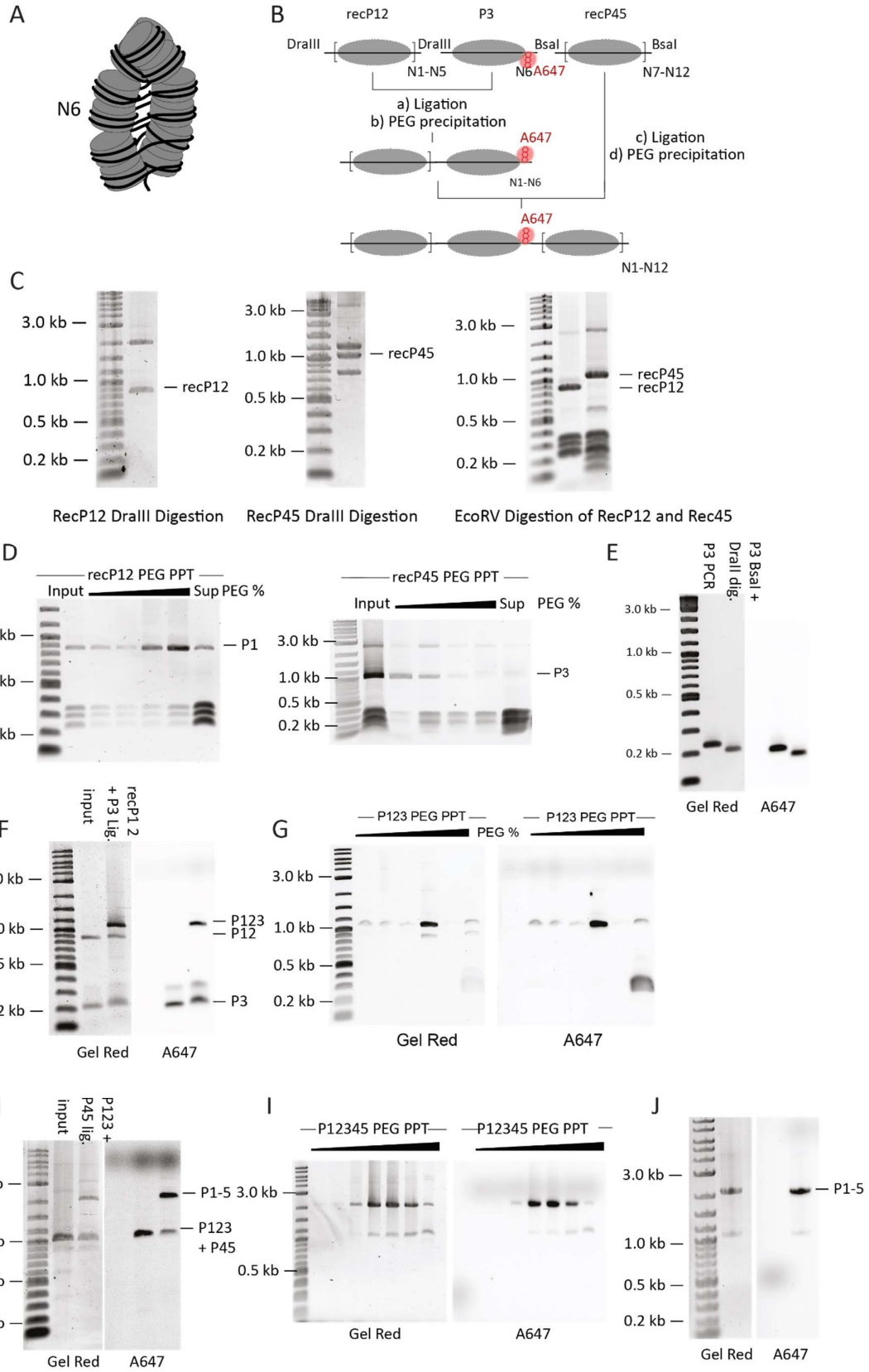


Figure 22 A) Scheme of 12mer array DNA highlighting the N6 central nucleosome B) 3-piece strategy to generate 12-mer DNA C-J) Gels showing each step of the process until final assembly.

For the 5-piece ligation used in the FRET experiments, 2 extra steps are required. The recP1 (containing 4 Widom 601 sequences) is first ligated to the PCR P2 fragment, then purified and followed by the ligation of P12 to P3 (another PCR fragment). At the same time, recP5 (containing 5 Widom 601 sequences) is ligated to P4 and purified. Finally, both P123 and P45 have ligated to form the whole 12x fragment. This work was done by Dr Anne Marinette-Cao (EPFL, LCBM).

Once the full-length DNA fragments are purified, chromatin reconstitution is done as explained above by dialysis from high salt (2M KCl/NaCl) to low salt (10 – 100 mM KCl/NaCl) in the presence of octamers. The addition of a small competitor DNA fragment from the mouse mammary tumour virus (MMTV) is used as quality control. This small fragment can form nucleosomes, albeit with a lower affinity than Widom 601. Therefore, after saturation of Widom 601 sequences, MMTV nucleosome will start to form (**Fig 24**).

Once again, octamers are titrated to find the optimal ratio for the formation of chromatin. To ensure that arrays are fully saturated, arrays are incubated with the restriction enzyme ScaI.

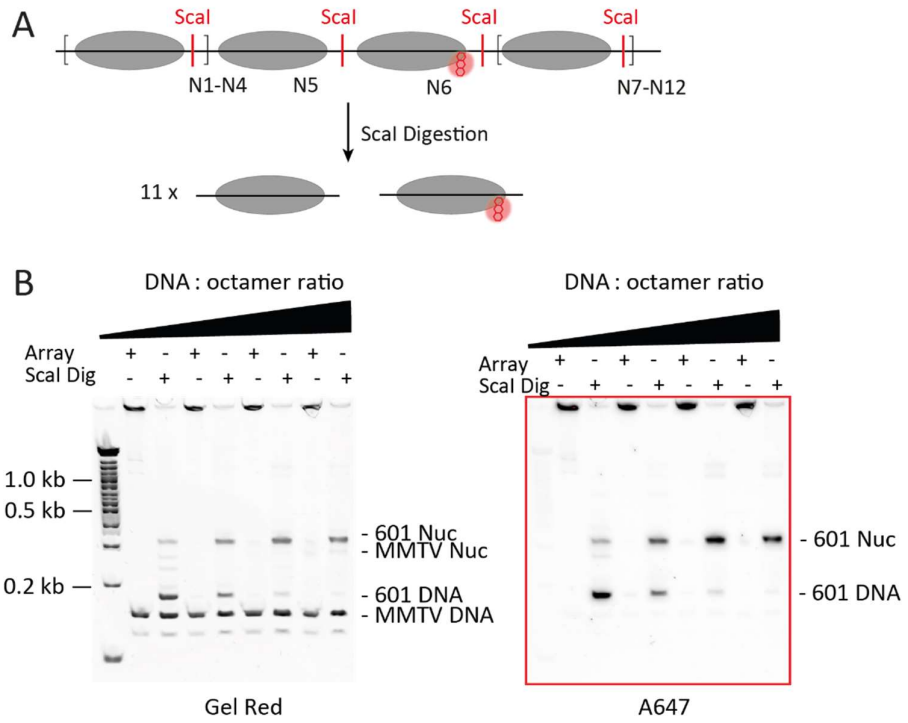


Figure 23 A) Scheme showing ScaI restriction sites between each 1 x Widom 601 nucleosome B) ScaI digestion and native PAGE gel after reconstitution of 12 mer chromatin arrays using increasing ratios of Octamer: DNA. We see that increasing ratios deplete the free 601 DNA and form more nucleosomes. At the highest ratio, MMTV nucleosomes start to form, indicating the saturation of 601 DNA.

ScaI restriction sites are found in the linkers between each NPS. Digestion by ScaI will free individual nucleosomes, which can be visualized on native acrylamide gels to check for the presence of free array DNA and MMTV nucleosomes, both of with are used as indicators of chromatin array quality (**Fig 24**).

3.3.2 Single-molecule colocalization TIRF on chromatin fibres

Similarly to the smTIRF colocalization experiments on mononucleosomes, chromatin arrays are immobilized on the surface of microchannels in the presence of 130 mM of KCl. In this solution, chromatin arrays are found in a compact state as the monovalent cations shield the negative charges from the phosphodiester backbone reduction their repulsion. This, in turn, allows the stacking of nucleosomes, forming 30 nm chromatin fibres (**Fig 4**). 50 – 100 pM labelled Rap1-Halo is added into the microchannels, and the location of both arrays and Rap1 are monitored over time.

Chromatin arrays devoid of Rap1 sites only showed fast events, τ_{off0} = 0.2 ± 0.1 s and τ_{off1} = 1.4 ± 0.7 s. These probing, unspecific binding events were similar to those seen in mononucleosomes. Specific binding of Rap1 to S1 and S2 containing arrays was still observed, although dwell times were reduced 3-fold compared to mononucleosomes (**Fig 25A**). This means that Rap1 can still bind its DNA motif within compact chromatin, although the proportion of specific longer-lived binding events reduces ~10-fold. Long dwell times of τ_{off2} = 25.6± 4.0 s and τ_{off2} = 16.8 ± 2.9 s for S1 and S2 respectively represented only 4 % of all measured events (compared to 30 % and 11 % in mononucleosomes) (**Fig 25B**). These fast events were of similar magnitude across all chromatin fibres tested (**Fig 25C**). This can be explained by the relative increase in the amount of unspecific DNA in the array compared to mononucleosomes, meaning the probability of random probing events is higher than the probability of Rap1 finding its specific target. If we only look at the rate of specific Rap1 association (spec. kon) for all experiments from DNA to chromatin arrays, we see that they are all very similar (**Fig 25D**). This suggests that Rap1 can hop and slide across the entire chromatin array, therefore increasing the chances of Rap1 finding its target.

Although Rap1 Is still able to bind chromatin fibres, dwell times are reduced 3-fold compared to nucleosomes and over 20-fold compared to free DNA. Rap1 seems to use random

unspecific probing events to search for its binding site, potentially utilizing the spontaneous thermal movements of chromatin to aid its search (44). So far, we have looked at the chromatin landscape's effects on Rap1 binding kinetics, but what happens to the chromatin landscape upon Rap1 binding?

3.3.3 Single-molecule FRET on chromatin fibres

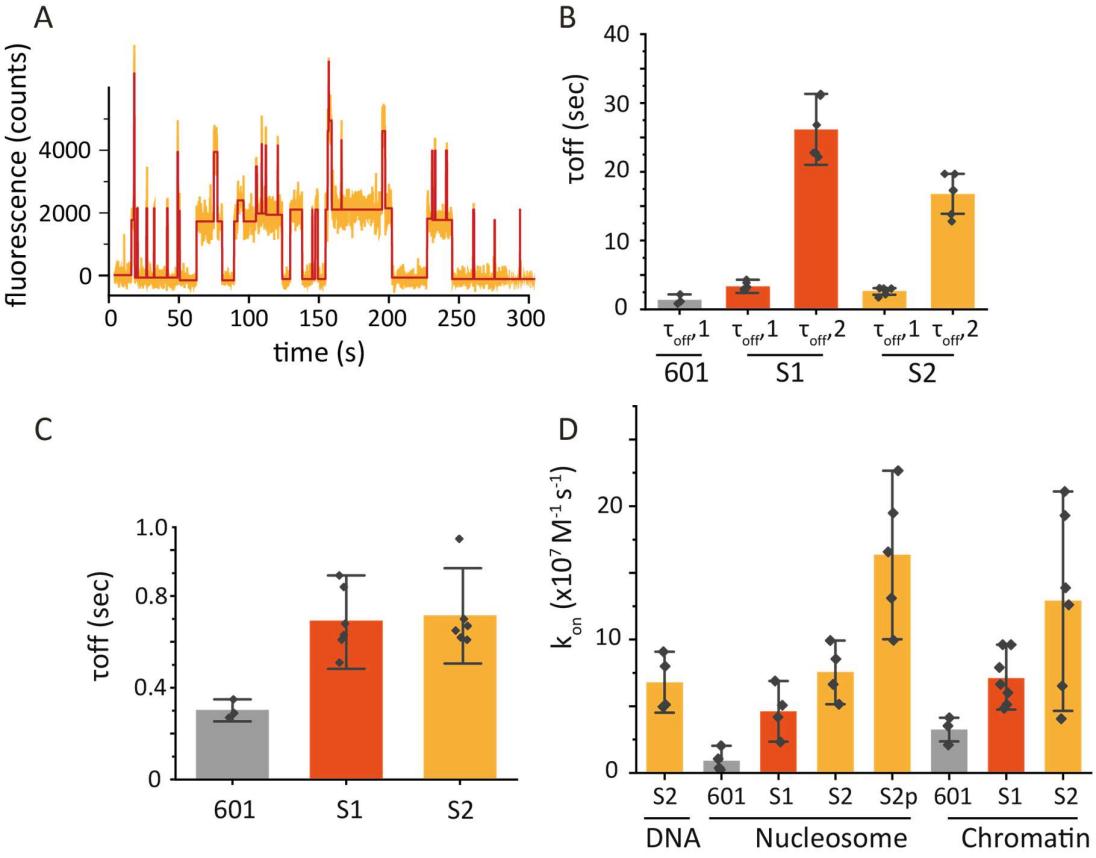


Figure 24 A) Example trace showing a mix of short unspecific binding events and longer specific events. B) Extracted dwell times from cumulative histograms for specific events. A tri-exponential function was used to fit the S1 and S2 data. C) Extracted dwell times from cumulative histograms of unspecific events. D) Comparison off all on rate binding kinetics for all constructs tested.

This work was done by Dr Anne Marinette-Cao (EPFL, LCBM). However, I find it important for this thesis to highlight her findings. Dr Marinette-Cao immobilized the identical S2 chromatin arrays used above for colocalization experiments with the addition of FRET compatible dyes in nucleosomes N5 and N7, surrounding the N6 nucleosome containing the Rap1 DNA motif. This FRET pair reports on the compaction of the central tetranucleosome unit. Chromatin fibre compaction can be modulated by varying salt concentrations (**Fig 26A**). In a solution

containing 40 mM KCl, a high FRET efficiency state of E_{FRET} ~0.35 was observed. Whereas in the presence of 150 mM KCl or 4 mM of the well-known compaction agent $MgCl_2$ E_{FRET} was ~0.5 and ~0.6 respectively, additional populations at E_{FRET} ~0.3 and E_{FRET} < 0.1 are also present. We, therefore, infer that relatively open chromatin has an E_{FRET} value of 0.3, and compact chromatin displays an E_{FRET} of 0.5 for this construct and FRET pair. The E_{FRET} < 0.1 population may come from array defects such as octamer undersaturation or shifted nucleosomes.

Titration of Rap1 from 50 to 500 pM in 150 mM KCl displayed a concentration-dependent overall reduction of the E_{FRET} ~0.5 population and a concomitant increase in the E_{FRET} ~0.3 population. This phenomenon was only observed when a specific Rap1 site was present (**Fig 26B**). From this data, we can conclude that specific Rap1 binding traps chromatin fibres in a more open state. On a structural level, S2 is found in the linker region of the chromatin array. Therefore Rap1 binding would induce a steric clash impeding nucleosome stacking. Interestingly, the presence of Rap1 also induced measurable movements within the chromatin arrays. These transitory anti-correlated dye fluctuations may be due to Rap1 scanning the chromatin array and forcing the array to remain momentarily decompacted.

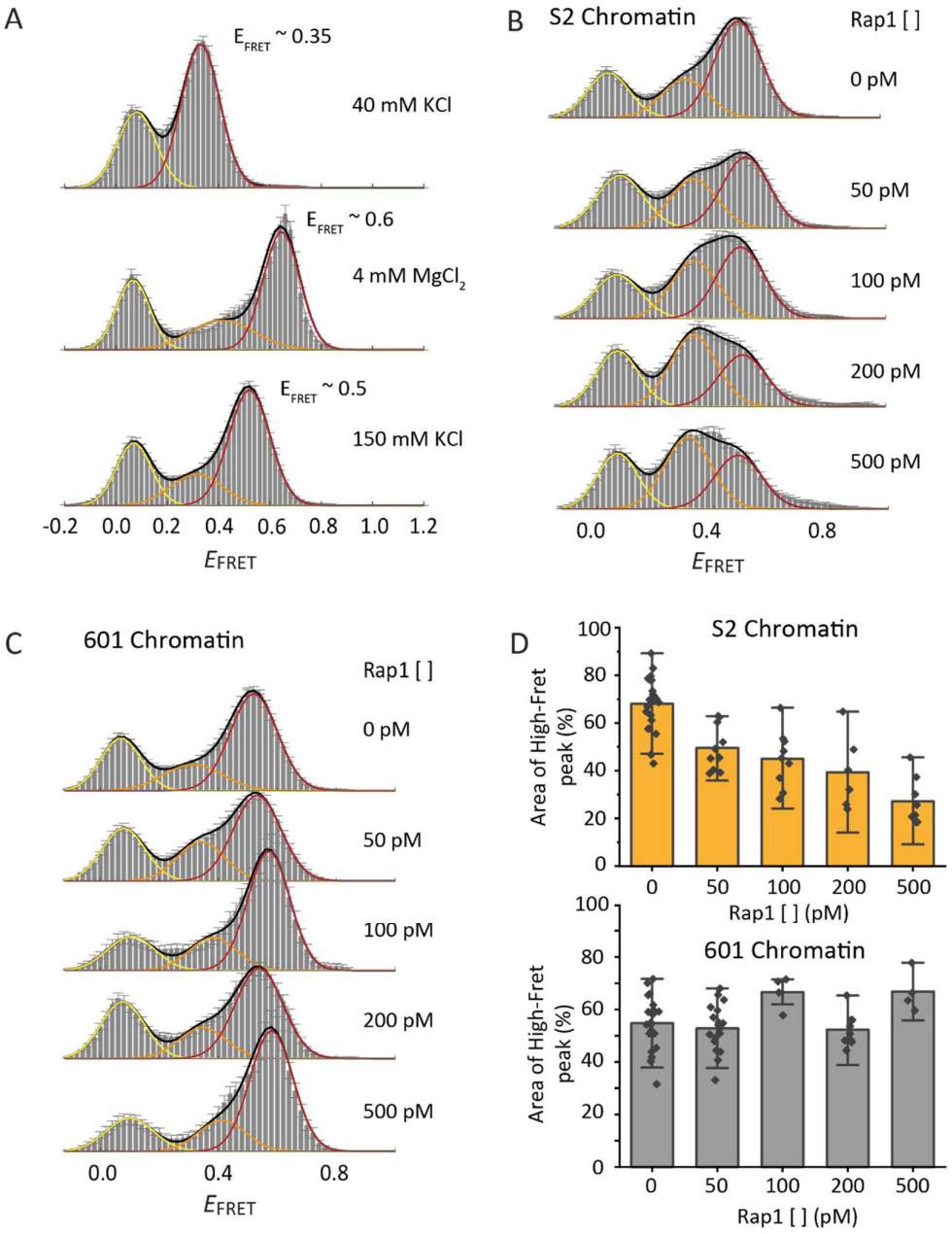


Figure 25 This work was done by Anne MArinette Cao A) Cumulative histograms of smFRET on chromatin fibres fitted with multiple Gaussians. We can see the change in FRET when known compaction agents are added as the high-FRET population (red Gaussian) shifts from 0.35 to 0.6 when $MgCl_2$ is added. A population showing no FRET is always present (yellow Gaussian). B) Cumulative histograms of smFRET on S2 containing chromatin fibres during Rap1 titration. We see a decrease in high-FRET population and an increase in the mid-FRET (Orange population). C) Cumulative histograms of smFRET on 12x601 chromatin fibres during Rap1 titration. D) Quantification of the relative area under the high-FRET gaussian fit. We see that for S2 Chromatin (top) that the population decrease with increasing Rap1 concentrations. For 12 x 601 Chromatin the population remains unchanged even at high Rap1 concentrations.

3.3.4 Rap1 binding to the WT RPL30 Promotor region

As mentioned above, nucleosome positioning sequences are a powerful tool for *in vitro* studies, with the caveat of limiting nucleosome dynamics. For this reason, we used the Widom 601 for our smTIRF localization and smFRET studies as they required predictable and homogenous nucleosome and chromatin arrays. However, we were interested to see the effects when using the natural Rpl30 promotor region. Nucleosomes that were reconstituted with the Rpl30 sequence migrated to lower molecular weights in polyacrylamide gels compared to nucleosomes with DNA of the same length of Widom 601 (**Fig 27A**). This is an indication that the nucleosomes may not be phased between both respective DNA constructs. Focusing on S1 in the Rpl30 promotor nucleosome, ensemble FRET experiments using the same FRET pairs and FRET positions as reported in **Chapter 3.2.6** were performed. Measured FRET values were lower than those using the Widom 601, E_{FRET} ~ 0.35 (compared to E_{FRET} ~ 0.8) (**Fig 27B-C)**. Much like our previous data, the titration of Rap1 did not change the ratio

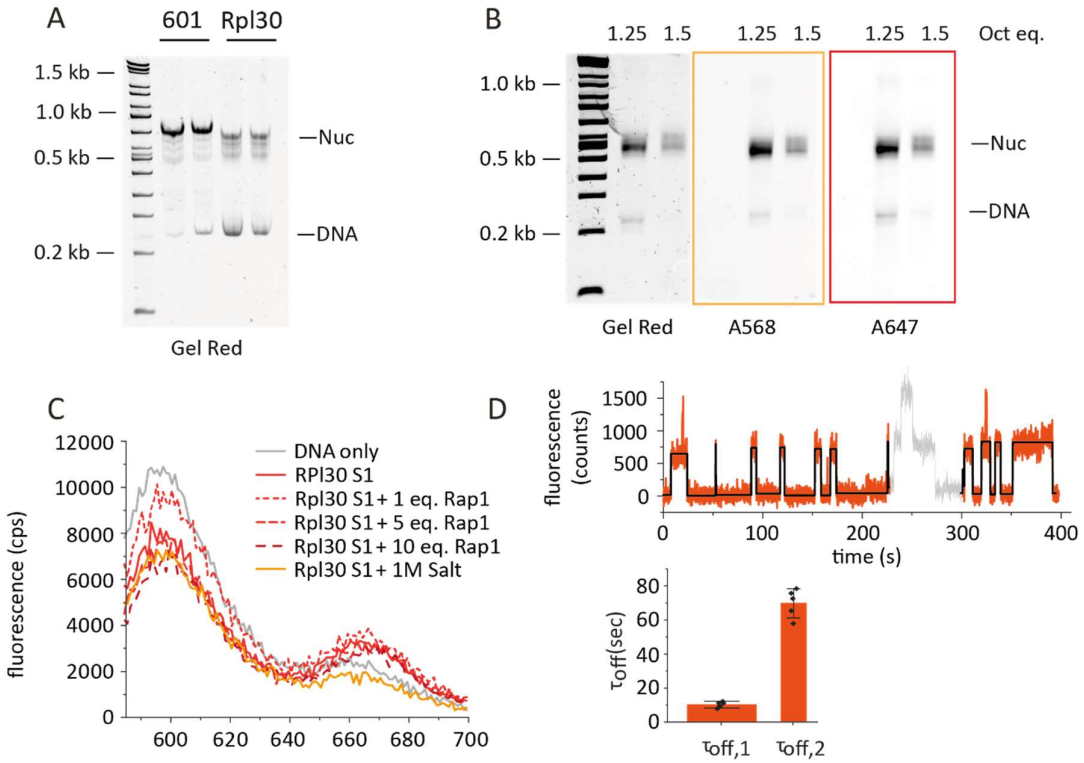


Figure 26 A) Native-PAGE gel to compare 1 x 601 Widom nucleosomes with Rpl30 native DNA, we see that they migrate differently indicating a difference in nucleosome phasing. B) Native Page gel of Rpl30 nucleosomes used for ensemble FRET studies. C) Ensemble FRET measurements showing no change when Rap1 is titrated. The ensemble FRET efficiency is lower compared to 601 nucleosomes. D) smTIRF colocalization measurements. Typical individual trace (top) and dwell time quantification (below).

between donor and acceptor fluorescent counts. We next performed smTIRF localization experiments with S1 containing wt Rpl30 mononucleosomes. Surprisingly these yielded dwell times that were shorter than those measured in Widom 601 mononucleosomes (τ_{off2} = 69.6± 8.3 s compared to τ_{off2} = 116.4± 36.0 s for the S1 in Widom 601 (**Fig 27C**)).

The use of the natural sequence proved challenging. Firstly, our standard quality control using polyacrylamide gels indicated well-formed nucleosomes. Ensemble FRET measurements suggest that these nucleosomes are not centrally positioned. This renders interpretation of smTIRF colocalization experiments complex, as we cannot speculate on the position of S1 within the nucleosome. Promotor region DNA sequences often contain sequences that are not amenable to nucleosome formation. These include polyA tracts or CpG islands. Rpl30 indeed contains a large polyA tract close to the centre of the -1 promotor nucleosome sequence used. Mnase-sequencing experiments described further showed a large distribution of positions within these nucleosomes. We can say that Rap1 seems to behave similarly with both Widom 601 and within the natural -1 promotor nucleosomes.

3.4 Discussion of Rap1 binding studies

These results confirm those from previous studies showing that Rap1 can bind nucleosomes, also showing that Rap1 forms a stable complex when binding in free DNA (dwell times increase 10-fold) (**Fig 28**). Rap1's nucleosome binding capacity probably stems from both its DNA target motif split in partial motifs and Rap1's DNA binding domain. Firstly, the base pairs with a unique possible nucleic acid (2,4 and 9 shown in Fig 13) are positioned in a way that at least one will always be found in a solvent-exposed major groove. This presents a partial DNA binding motif to Rap1 at all times. By tiling S1, we tested this and saw that by moving S1 by +3 base pairs, we could increase Rap1's binding affinity. If we look at the crystal structure and the locations of the 3 invariable nucleic acids, we can see that in the +3 position, all 3 nucleic acids are more favourably solvent-exposed. In the WT and +6 positions, a subset of these positions face the octameric core or are found in solvent-exposed minor grooves, reducing the possibilities of Rap1 binding. Since Rap1 has two Myb-like HTH domains within its DBD, it would be interesting to further study its multivalent binding capacity. This is also reflected in

the DNA binding sequences themselves, with two half-sites within each DNA binding motif. To better understand the importance of these bipartite partial binding sites, we could proceed to study mutants of S1 to test whether one partial binding motif is sufficient or dominant. The single-molecule colocalization TIRF experiments using Rap1 showed that it could interact with DNA in different binding modes, yielding different dwell times. This is also in line with previous studies showing this effect, whereby Rap1 can engage its binding motif with either a single or both of its tandem Myb-like helix-turn-helix domains (191). Rap1 with a single functional Myb-like domain could help us elucidate the structure of these modes, as we would expect that the high-dwell time population would disappear, leaving only events with dwell times of mins – secs.

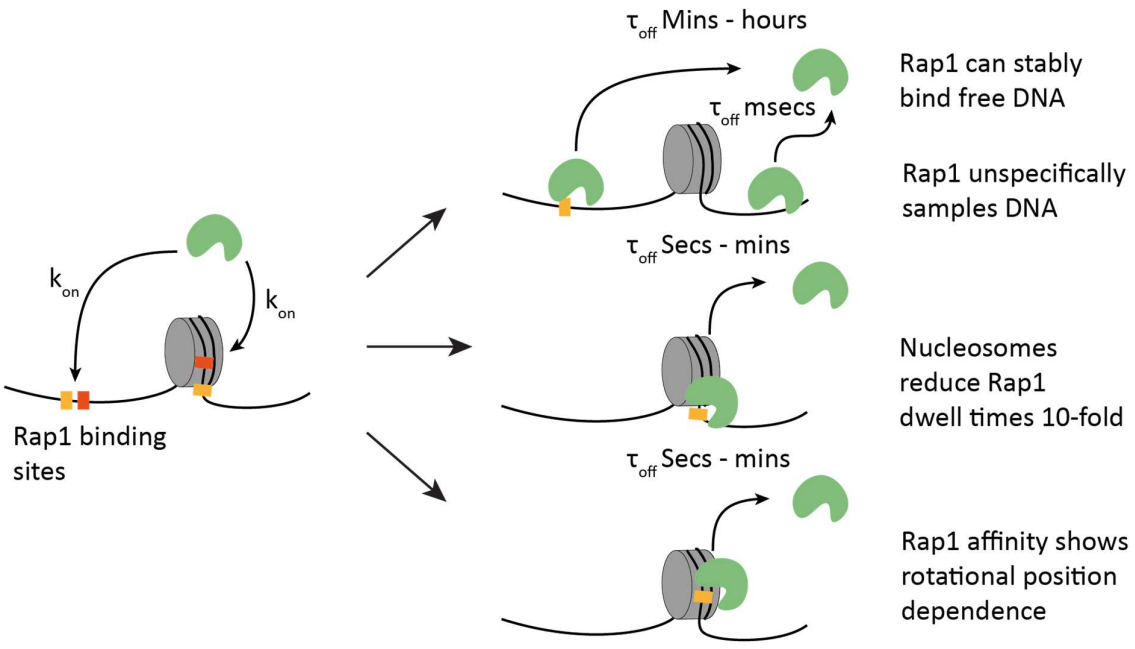


Figure 27 Rap1 binding model showing Rap1's pioneering abilities. Rap1 can freely sample free DNA and nucleosomal DNA with the same association kinetics. Once Rap1 binds its DNA motif, it stays stably bound for minutes to hours. However, the presence of nucleosomes reduces the dwell times 10-fold.

Investigating the search dynamics involved in Rap1 binding, a large proportion of binding events were short-lived (milliseconds - seconds). Experiments using sequences devoid of Rap1 motifs suggested that these short events occur when Rap1 is unspecifically sampling the DNA substrate. The overall on-rate of Rap1 slightly increases for the chromatin substrate. However, if we only look at the on-rates for specific binding events, we see that it is the same despite the presence of nucleosomes or chromatin. *In vivo* TFs have been observed to require

~100 binding events before finding their targets (207). Our experiments cannot elucidate the hoping and sliding mechanism used by Rap1, but we see that Rap1 can use these probing events to find its target DNA. It has been suggested in the literature that a mix of 1D (sliding) and 3D (hoping) diffusion search mechanisms are involved (208–210). TF sliding is thought to be hindered by roadblocks such as nucleosomes and other DNA associated proteins, causing TF dissociation (211). Our small chromatin fragments do not replicate the proportion of unspecific DNA to specific DNA motif ratio as we are currently limited with our technological toolbox to perform such studies. Generating long (megabases) DNA fragments that can be precisely chromatinated to varying degrees remains challenging.

The binding experiments presented above were done in highly defined *in vitro* settings with known DNA sequences and imaging conditions tailored to reduce photobleaching. Therefore, direct comparison with *in vivo* measurements done by single-particle tracking (SPT) using labelled TFs reported in the literature is challenging. Compared to the Rap1 data presented above, which had dwell times in the hours – minutes, reported *in vivo* dwell times exhibit dwell times in the range of minutes - seconds (207,212–214). This being said, recent studies finding binding times of ~40 mins for the serum response factor (215). One explanation for differences in dwell times observed may be due to technical limitations in *In vivo* single-particle tracking. Firstly, unlike in *in vitro* experiments with defined buffers designed to reduce photobleaching, the environment *in vivo* is not ideal for prolonged imaging of fluorophores. For JF549 (the same dye used for the experiments in this thesis), the bleaching rate τ was found to be ~2 s (with an irradiation energy of 1 kW/cm^2), compared to the settings used for single-molecule TIRF *in vitro* (163 s) (214). SPT experiments for the serum response factor used dark times of up to 1 min to follow the longer dwell times. This may introduce 'gaps' in tracking, making it challenging to assign single-particle trajectories (215).

Another important factor to consider is the cell volume, which requires specific imaging conditions for *in vivo* SPT experiments. These include using illumination methods that can optically section the sample, such as HiLo or light-sheet illumination, instead of illuminating the cell volume in its entirety (216). This has the advantage of only illuminating a subset of labelled proteins within the nucleus. However, with 3D diffusion, labelled particles can move out of focus, and the connectivity (which allows high fidelity tracking) of a single particle between acquired frames can be lost. This can also artificiality reduce dwell times.

An additional source of artificial shortening of dwell times is that exponential fits used to describe the data in this work were corrected for photobleaching using bleaching curves measured using the same microscope settings, which is not the case for all *in vivo* data (217).

In this work, we show that dwell times are modulated by DNA sequence and chromatin context. We can infer this as we have strict control of the sample injected into the microarrays. Therefore, we can attribute changes in Rap1 binding to differences in the DNA sequences or chromatin landscape between experiments. In cells, one would expect that there exist a plethora of specific and partial binding sites within a large palette of DNA landscapes, even within a single cell. This would then mean that the 2 – 3 defined states per DNA motif as observed *in vitro* would be observed as a continuum of dwell times constituted of distinct individual states. Therefore, the use of bi/tri exponential functions that imply the existence of 3 -4 defined states would be an average of these different binding states. The use of exponential decays can be seen in many *in vivo* SPT experiments with TFs, where a slow fraction, often characterized as specific binding, and a fast fraction thought to be unspecific sampling are extracted using exponential decay fits. However, with no knowledge of the underlying DNA sequences or chromatin context onto which TFs bind to during an *in vivo* SPT experiment, it is difficult to infer the biological nature of these states. Therefore, alternative models to describe SPT data have been proposed, including the use of a power-law (217). This statistical model describes a relationship where a relative change in one state implicates a proportional change in another state. This function depicts a heavily skewed tail, perhaps better accustomed to accommodating rare events such as longer dwell times (217). Evidence in favour of using a power-law distribution to fit data stems from the fact that cumulative histogram data of the number of counts plotted against time on a log scale can display a linear relationship instead of the exponential decay. A power law does not assign discrete states but instead implies a continuum of possible states, thus better reflecting the environment within a cell (217).

On top of the numerous possible DNA motifs and chromatin landscapes, TFs *in vivo* may also face competition with other factors for binding DNA motifs, or they may be evicted by other accessory proteins such as chromatin remodelers. This would also reduce the observed dwell times for TFs. Conversly, some factors can aid in loading TFs onto chromatin *in vivo*. This is thought to be the case for FoxA in some enhancer loci, whereby the prescence of IL6/STAT3

allows access to FoxA (218). Taken together, *in vitro* measurements are done in an idealised context. Thus measurements reflect the best possible scenario for TF binding. This does not mean that insights gained by *in vitro* measurements cannot be translated to the *in vivo* setting. However, the actual binding dynamics of Rap1 to the RPl30 are most likely shorter than those measured in this work but longer than the seconds to mins tine scales measured for other factors using SPT.

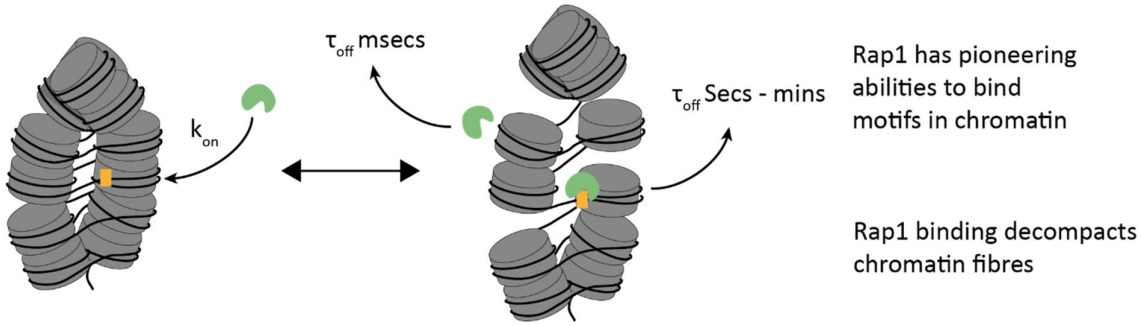


Figure 28 Rap1 binding to chromatin decreases the compaction of the chromatin fibre. This is only the case for fibres containing Rap1 target motifs.

FRET experiments using chromatin fibres done by Anne Marinette-Cao to follow the overall compaction of the fibres revealed that only in the presence of Rap1 and a specific Rap1 target motif did overall compaction change. One would expect that Rap1 could probe the entirety of the DNA within the chromatin fibre by exploiting the thermal fluctuations of chromatin fibres that occur in the µs – ms time range. These short non-specific events, which had dwell times of 0.2 s – 1.7 s overall, may induce a FRET change that we cannot observe with our setup—using integration times of 100 ms, the FRET changes induced by unspecific probing may be too quick to measure. These measurements showed an increase in the proportion of dynamic (changes in FRET efficiencies within a single chromatin fibre) traces scaled with Rap1 concentrations. However, at 500 pM Rap1, we would expect chromatin fibres to be bound by multiple Rap1 molecules at a given time.

Nevertheless, the number of dynamic traces was still higher in fibres containing a Rap1 site than those devoid of Rap1 sites. *In vivo,* the presence of pTFs has been shown to induce DNaseI or MNase sensitivity (219–221). However, the reason for this change may be multifactorial and not directly due to the binding of pTFs but more significant changes induced by the recruitment of co-factors or chromatin remodelling machinery. Structural changes induced by Rap1 binding itself in the experiments presented above may be small and local. To further study this process, a FRET pair in a tetranucleosome register that does not contain

a Rap1 site but within a fibre containing a Rap1 binding motif may show us if these structural changes are only local or if they can spread through a fibre.

The use of the Widom 601 nucleosome positioning sequence is invaluable for *in vitro* studies. It assures a homogenous population of nucleosomes, reducing the possibility of sliding. Thus the Rap1 binding motifs are always in the same relative position. The caveat is that the Widom 601 is a very strong nucleosome binder. This sequence has been shown to influence the recruitment of certain TFs (222,223). Breathing and sliding rates, as well as DNA shape, can facilitate TF nucleosome binding; therefore, DNA sequence is non-negligible when exploring this process (59,224). We proceeded to do some of these experiments using the native Rpl30 nucleosome, discussed in **Chapter 3.3.4**. Overall these experiments showed Rap1 behaved similarly in both DNA contexts. However, the inherent heterogeneity of the native system rendered the interpretation of our data challenging. Another key aspect of these promotor regions that affect nucleosome dynamics is the presence of histone variants such as H2AZ and PTMS such as H3K27ac. What effects do these have on Rap1 remain to be explored.

In vivo experiments also showed that Rap1 played a key role in regulating gene activation of ribosomal protein genes (RPGs) (109,110,185–187). The presence of Rap1 induced changes in the local chromatin landscape in the promoter regions of these genes (48,109,185). Our work suggests that Rap1 by itself does not induce target nucleosome loss. However, Rap1 does disrupt tetra-nucleosome stacking within chromatin. Rap1 depletion *in vivo* has also been associated with the recruitment of co-factors such as HmoI and Fhl1 (185). HmoI physically interacts with the polymerase II (Pol II) pre-initiation complex component TFIID and TBP (225). Therefore, it is conceivable that Rap1 acts as an anchor for the formation of HmoI induced preinitiation complex formation, which would, in turn, evict the nucleosome in the promotor region. More recent work has shown the importance that chromatin remodelers have in shaping promotor and also intragenic regions.

4 Chromatin Remodelers as Chromatin Effector proteins

This chapter outlines work that was published in:

Mivelaz, M., Cao, A.M., Kubik, S., Zencir, S., Hovius, R., Boichenko, I., Stachowicz, A.M., Kurat, C.F., Shore, D. & Fierz, B. Chromatin Fiber Invasion and Nucleosome Displacement by the Rap1 Transcription Factor. *Mol Cell* **77**, 488-500 e9 (2020).

This work includes contributions from Christoph Kurat for the RSC complex purifications, Ruud Hovius for RSC assays, Sevil Zencir and Slawomir Kubik for MNase-Seq.

Summary of chapter

The RSC complex was expressed and purified by Christoph Kurat by tandem affinity tag purification and kindly provided for these assays. In-gel experiments incubating nucleosomes and RSC complex showed that RSC moved centrally positioned nucleosomes towards the periphery. This was seen as a change in the electrophoretic mobility of the nucleosomes before and after remodelling. When Rap1 was present, nucleosomes would be supershifted due to Rap1 binding, forming a Rap1-nucleosome complex that also reduced the movement within gels. Therefore, to recover free nucleosomes and see nucleosome remodelling, we added an excess of plasmid DNA to outcompete Rap1 nucleosome binding and liberate free nucleosomes. Only incubating Rap1 did not change the electrophoretic mobility of free nucleosomes, indicating that Rap1 by itself does not remodel nucleosomes. This is in accordance with our observations in **Chapter 3**. RSC remodelling assays generated a higher running species that excess plasmid DNA was not able to outcompete. Further assays using TMR labelled nucleosomes suggested that this species included intact octamers. Reconstituted asymmetric nucleosomes used to mimic RSC remodelled nucleosomes towards one end, with Rap1 sites in a free DNA portion migrated to the same higher running location, further indicating that this species was a Rap1-Nucleosome complex with stably bound Rap1. As the proportion of this species seemed to increase over time, we used MNase-seq to follow if the binding of Rap1 biased RSC remodelling away from its binding sites, thus generating this

stably bound Rap1 complex. MNase-Seq was performed by Sevil Zencir and Slawomir Kubik, and they saw that indeed Rap1 was able to bias RSC remodelling direction away from its own binding sites.

4.1 Chromatin remodelers as chromatin effector proteins

Chromatin remodelers can be categorized according to their phylogenetic traits with four highly conserved subfamilies. These are Switching defective/sucrose nonfermenting (SWI/SNF), **i**mitation switch (ISWI), chromodomain-helicase-DNA binding (CHD) and inositol requiring 80 (INO80) subfamilies. *In vivo* studies in yeast have shown that these subfamilies act on different parts of the genome. SWI/SNF CRs are active at promotor regions regulating transcription (48,93–96) (**Fig 6A**), ISWI and CHD CRs evenly space intragenic nucleosomes (48,97–99) and INO80 regulates the turnover of histone variants (100,101). In addition to the presence of an ATP-dependent domain, CRs contain domains that can recognize PTMs such as bromo-, chromo- and PHD-finger domains (226,227). Apart from these reader domains, very little is known about CRs and how they specifically target genomic regions. Some sequences have been found to influence the direction of remodelling, RSC (a member of the SWI/SNF family) has been shown to bias movement towards the 5' side of paired motifs of poly(dA:dT) tracts in the proximity of regions with high GC content (93,109). Depleting CRs such as RSC directly affects promotor region architecture, including narrowing the NDR; consequently, TBP sites relocate into +1 nucleosomes (109). This, in turn, reduces the recruitment and assembly of the PIC in genomic regions associated with RSC activity (109). As CRs are the drivers of largescale nucleosome landscape changes, regulation of their function needs to be controlled in time and space. TFs are ideal coregulators, as they can be expressed in response to stimuli, are expressed differentially across tissues and cell types, but most importantly, they recognize specific DNA sequences. Functional interactions between TFs and CRs have been previously demonstrated (93,109,110,228). Amongst these pairs, Rap1 has been shown to modulate RSC activity *in vivo* (109,110,185). These papers demonstrated that Rap1's and RSC's contribution to nucleosome remodelling is not ablated when either protein is depleted, indicating that both entities can act independently. The strength of this effect

varies depending on the genomic loci, with some loci showing additive or synergistic effects when both Rap1 and RSC are present. Although this functional pair has been identified, the mechanism and kinetics of this process remain largely unknown.

The approach we have taken to further study TF-CR interactions aimed to first study the Rap1-RSC pair to see whether we could recapitulate the NFR formed *in vivo.* Building on this, we would like to implement an *in vitro* pipeline to understand TF-CR relationships in general. For this, we used an *in vitro* next-generation sequencing MNase-seq technique to get base-pair resolution of nucleosome positions after remodelling by RSC with/without Rap1. In order to follow remodelling kinetics in real-time, we have started to develop a FRET-based system using Chd1 as our model remodeller.

4.2 Expression and purification of RSC

RSC is abundantly expressed in yeast (*S. cerevisiae*), and deletion of RSC is lethal. RSC consists of multiple subunits with a total weight over 1 MDa (49,229). A core of 7 subunits, including the ATPase containing domain Sth1, are conserved within the SWI/SNF remodelling family members. Whilst many subunits contribute to complex stability and formation, their roles remain elusive. However, the presence of 6 bromodomains (2 in each of Rsc1/2 and Rsc4, and one each in Rsc58 and Sth1) have all been shown to bind acetylated H3 tails (230,231). Rsc1 and Rsc2 are paralogs that also contain a DNA binding motif, HMG box (231). Other DNA binding motifs are found in Rsc3 and Rsc30, which are zinc cluster DNA-binding proteins capable of binding GC-rich regions (228). The combination of polyA tracts and GC-rich motifs are found at many sites associated with RSC activity, and it is postulated that this 'motif' drives RSC binding and remodelling (109).

RSC remodelling activity experiments have shown that it can move and evict nucleosomes (232,233). Recent structural studies have determined RSC's nucleosome binding conformation, with the ATPase motor bound at SHL 2 (104,234,235). Contacts with both the octamer's acidic patch and nucleosomal DNA anchor RSC to the nucleosome. Rsc3-Rsc30 were mapped and revealed that they could bind DNA in the NFR (104). The size and combinatorial subunit compositions of RSC make it challenging to purify and study *in vitro*.

Tandem-Affinity-Purifications (TAP) has been a successful methodology to purify RSC. The Kurat Lab (LMU, Munich) tagged Rsc2 with calmodulin-binding peptide and the IgG-binding portion of Protein A as reported in Wittmeyer et al., (236). Endogenous expression levels in *S. cerevisiae* of all RSC subunits (including Tap tagged Rsc2) are used. Purified RSC complex was provided by the Kurat lab (237).

4.3 RSC remodelling assays

Mononucleosomes containing both RPl30 Rap1 motifs S1 and S2 in a Widom 601 template were reconstituted and incubated with a 5x molar excess of Rap1 (**Fig 30A**). As before, a supershifting of the nucleosome band is observed. Incubation with excess plasmid DNA was sufficient to recover free nucleosomes (**Fig 30A**). Stable Rap1 binding, through the shifting of Rap1 sites into nucleosome free regions, would result in longer dwell times, thus inhibit nucleosome recovery by plasmid DNA incubation. Similar recovery rates over 90 mins showed that nucleosomes are stable and not shifted even in the presence of excess Rap1 (**Fig 30B**). This reinforces our findings described above, where Rap1 was not observed to alter nucleosome structure. The same mononucleosomes were incubated this time with RSC and the yeast histone chaperone yNap1 (**Fig 30C**). Nap1 was included as it was previously reported to increase RSC remodelling and nucleosome disassembly (238,239). In our hands, we see that yNap1 does not increase RSC kinetics but pushes the nucleosome further to the periphery (**Fig 30C**). This can be followed as a change in electromobility shift, as a new species, the remodelled nucleosome (rNuc), is now visible. In the presence of Rap1, RSC was still able to remodel nucleosomes at a similar rate. However, even in the presence of excess plasmid DNA, the free nucleosome could not be fully recovered (**Fig 30D**). We observed the presence of a new species (red box), which we thought may be Rap1 bound to free DNA or the nucleosome. This suggested that Rap1 was able to bind the remodelled species stably. To better follow the remodelling reaction in the presence of Rap1, we reconstituted nucleosomes with TMR labelled H2A. The choice to use these nucleosomes not only allows us to better follow species that super-shift within the thick plasmid DNA band, but we can also see whether the majority of nucleosomes still contain the H2A dimer. As mentioned above, RSC has previously been

reported to evict nucleosomes. This process is thought to go through a loss of consecutive H2A-H2B dimers to form tetramers (239).

With our labelled H2A-H2B dimers, we did not observe hexasome formation (**Fig 30E**). To confirm that the super-shifted species were indeed Rap1-bound nucleosomes, we reconstituted fully shifted nucleosomes with free Rap1 motifs. Incubation of these shifted nucleosomes recapitulated the super-shifted species observed after RSC remodelling (**Fig 30F**). From these experiments, we replicate previous experiments showing that RSC moves nucleosomes to the periphery. Unlike previous experiments, we do not observe hexasome formation. We also show that when Rap1 sites are in the nucleosome, incubation with excess plasmid DNA can outcompete Rap1 binding, recovering free nucleosomes. However, incubation of nucleosomes with outcompete RSC and Rap1 containing higher-running species. However, some species are not easily recovered by excess plasmid DNA; these were identified as end shifted nucleosomes bound by Rap1. Since the Rap1 sites are found only on one side of the nucleosome, stable Rap1 binding requires the nucleosomes to be shifted in the opposite direction. Over time, we see an increase of this stably bound species, indicating that nucleosomes may be preferentially shifted away from the Rap1 sites. Knowing this, we hypothesized that remodelling without Rap1 would yield a mix of end-shifted nucleosomes, whilst remodelling in the presence of Rap1 would have a directional preference. Therefore, if we add Rap1 after an RSC remodelling experiment, Rap1 would only stably bind the subfraction of nucleosomes remodelled away from the Rap1 sites. Experiments conducted in parallel with and without Rap1, with the addition of Rap1 to the reaction mixture initially without Rap1, showed a difference in the amount of stably bound nucleosomes (**Fig 30G**). Using Rap1 binding stability as an indicator of nucleosome position, we concluded that Rap1 must influence RSC remodelling direction. In collaboration with the laboratory of Prof. David Shore (University of Geneva), we went further with these experiments by incorporating high throughput sequencing as our readout to get base-pair precision information on the location of the remodelled nucleosome.

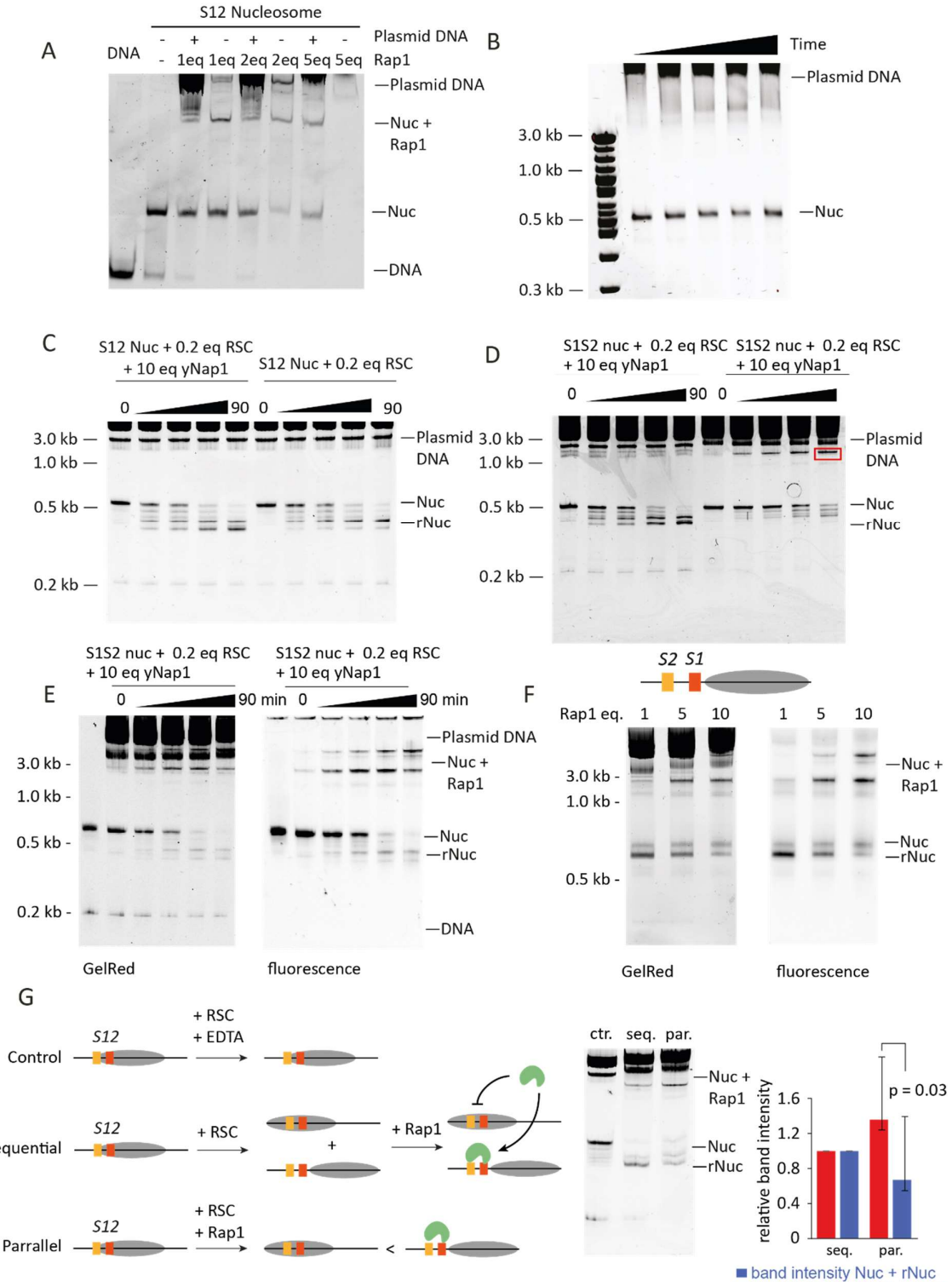


Figure 29 A) Recovery of Rpl30 S12 nucleosomes using excess plasmid DNA is tested by titrating Rap1. B) of Rpl30 S12 nucleosomes are incubated with 10 eq. Rap1 for 0 – 90 mins. At time points excess plasmid DNA is added to see whether nucleosomes remain stable over time. C) RSC remodelling assay with and without yNap1. D)RSC remodelling assay with and without Rap1. A new species can be observed. E) RSC remodelling reaction repeated with TMR labelled H2A Rpl30 S12 nucleosomes. F) Asymmetric S12 TMR labelled H2A nucleosomes are incubated with Rap1 to identify higher running species. G) RSC remodelling assays are done to test the influence of Rap1. For this Rap1 was either included after 90 mins remodelling (Sequential) or at the beginning (Parallel). Quantification of remodelling (left) shows that addition of Rap1 generates more stable Rap1 binding. This hints that Rap1 pushes RSC away from its binding sites.

4.4 Next-Generation Sequencing MNase-Seq RSC remodelling assays

RSC remodelling assays were performed using the same conditions as those presented above, in the presence and absence of Rap1. For these experiments, both a Widom 601 scaffold and the WT RPl30 promotor region were used, both containing S1 and S2 in their native positions. Following RSC remodelling, samples were split into 3 aliquots and treated with MNase (6 units, 3 units and 1 unit respectively) for 5 mins at 37 °C (**Fig 31A-B**). Different MNase concentrations have been shown to give different information. For example, TFs can protect underlying DNA from mild MNase digestion but not from high MNase. Reactions were stopped using EDTA and SDS before Proteinase K digestion at 60 °C for 1 h. Digested DNA was isolated using silica-based PCR purification columns, removing all proteins (**Fig 31B**). Our collaborators in the Shore Laboratory then prepared the sequencing libraries and proceeded with sequencing. Mapping the positions of Widom 601 nucleosomes showed that in the absence of remodelling, ~70 bp on each side of the dyad are protected from MNase digestion (**Fig 31C**). Following RSC remodelling, flanking sequences are better protected, with a clear directionality, shifting nucleosomes away from Rap1 sites. This confirms our previous observations that increased Rap1 dwell times must originate from the Rap1 binding to its DNA binding motifs found in free DNA regions. In the presence of Rap1, this directionality is conserved, with a slight increase in the directionality bias away from the Rap1 sites (**Fig 31C**). Nucleosomes containing the Rpl30 promotor DNA showed a less defined digestion pattern. This was expected when considering our previous experiments. However, with this data, it seems that nucleosomes are found more often off centre (slightly towards the Rap1 sites). However, it could be suggested that entry and exit regions are more mobile therefore less protected. The addition of RSC preferentially pushes the nucleosome towards the Rap1 sites (**Fig 30D**). This reflects *in vivo* data where the Rap1 sites in the RPl30 promotor region are encroached on by nucleosomes in the absence of Rap1. In the presence of Rap1, the distribution of the Rpl30 nucleosome changes significantly, with less strong protection from MNase digestion at the Rap1 binding sites. Together, these experiments suggest that the presence of Rap1 alone is sufficient to influence the direction of RSC remodelling.

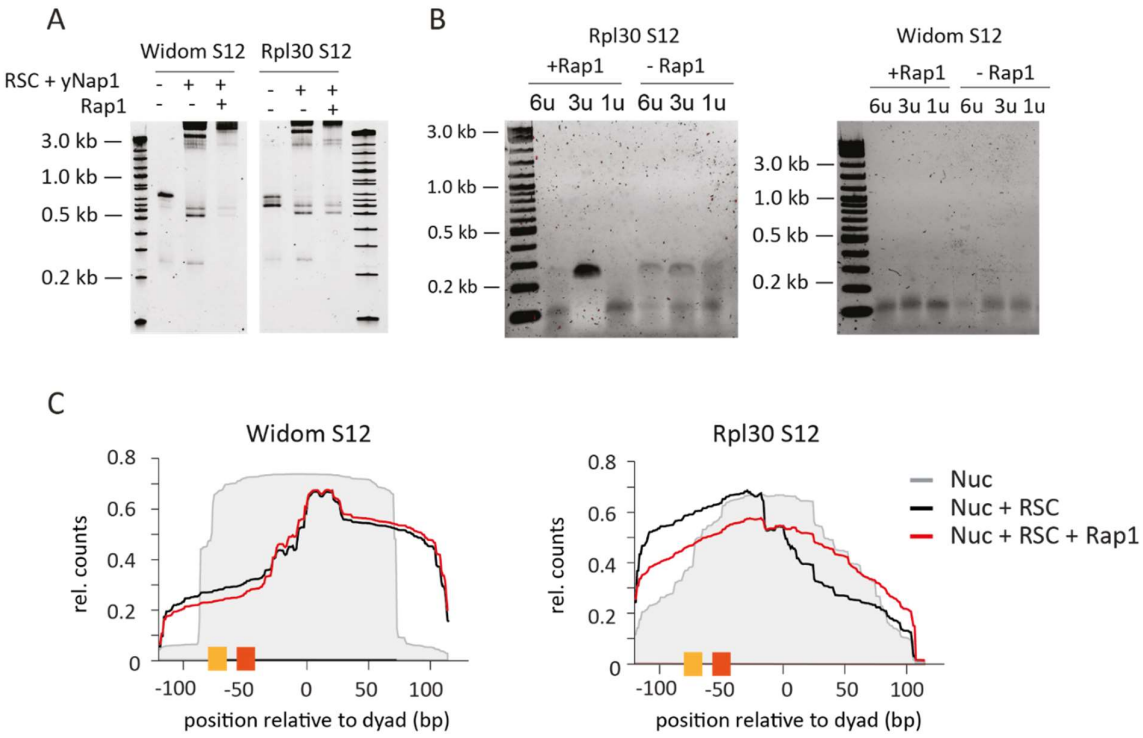


Figure 30 A) Native PAGE gel used as quality control of RSC assay used for sequencing for both Widom S12 and Rpl30 S12. B) Agarose gel of PCR purified RSC samples digested with indicated units Micrococcal nuclease (Mnase). C) Graphic output of high-through put sequencing showing the relative number of reads.

4.5 Discussion of RSC remodelling studies

In vivo ChIP-Seq, experiments identified RSC as a chromatin remodeler that shapes the chromatin landscape in promotor regions (109,185). Additionally, the directionality of RSC remodelling has been seen to be sensitive to the DNA sequence. With a propensity to bias movement towards the 5' side of paired motifs of poly(dA:dT) tracts in the proximity of regions with high GC content (93,109). Genomic regions associated with Rap1 binding show differential changes depending on the presence or absence of Rap1. The role of Rap1 in these observed changes remained unknown. Using our in gel and MNase-seq experiments, we demonstrate that Rap1 alone can influence RSC remodelling without the need for cofactors. This being said, the presence of co-factors may increase RSC directionality bias; however, this remains to be explored.

When mapped *in vivo*, Rap1 binding sites often sit within the first 2 solvent-exposed major grooves. When looking at maps showing a change in nucleosome occupancy after RSC

remodelling *in vivo,* we can see that the position of Rap1 sites vis-à-vis the closest nucleosome dyad dictates the direction of remodelling (**Fig 32A**). This is evident as the Rap1 sites sit adjacent to the most significant fold nucleosome occupancy change and not in the middle, which would have suggested a bilateral movement (109). Rap1 sites are mapped by the black line, whereas the transcription start site is in red. This comparison is made in the presence and absence of Rap1. Therefore we assume that RSC remodelling without Rap1 is purely a result of DNA sequence directionality (109). This results in the generation of a larger NFR in the promotor region of genes controlled by Rap1. Together with our *in vitro* data, we can now say that Rap1 is directly involved in this process. However, it remains to be seen if Rap1 co-factors increase the strength of this effect. It is noteworthy that the yeast general transcription factors Abf1 and Reb1 also showed some cooperativity with RSC, although no directionality was observed (109). The architecture of nucleosomes containing DNA binding motifs for these TFs showed a stereotypical arrangement with flanking poly(dA:dT) tracts and high GC content (109). Finally, for all 3 TFs tested, there was variability in the effect size within genomic loci associated with TF and CR binding. Suggesting that perhaps other factors such as binding strength could also play an important role.

Previous TF-CR interaction studies *in vitro* demonstrated that some CR could read through or remove bound TFs, whereas others could not (240). In a paper from the Michelle D. Wang laboratory, they tested the yeast CRs ISW1a and ySWI/SNF in combination with the DBD of the yeast transcription factor Gal4 (a member of the Zinc-finger DBD family, with a reported of Kd ~ 3.4 nM) and the Lac repressor (a member of the Helix-turn-helix motif DBD family, Kd not reported for the construct used but reported elsewhere to be in the range of ~ 0.1 nM) (240). Gal4 DBD was found to push nucleosomes away in the presence of ISWIa as it presents a barrier to remodelling. Whereas ySWI/SNF remodels nucleosomes towards Gal4 DBD and can evict Gal4 DBD. This was replicated with the Lac repressor TF to avoid potential Gal4 DBD specific interactions. When comparing these findings to our work, we can point to **two** main differences. Firstly, the incubation duration in these experiments was 1 – 10 minutes, which may not be sufficient to reach an equilibrium. Our endpoint measurements may thus provide a picture at equilibrium.

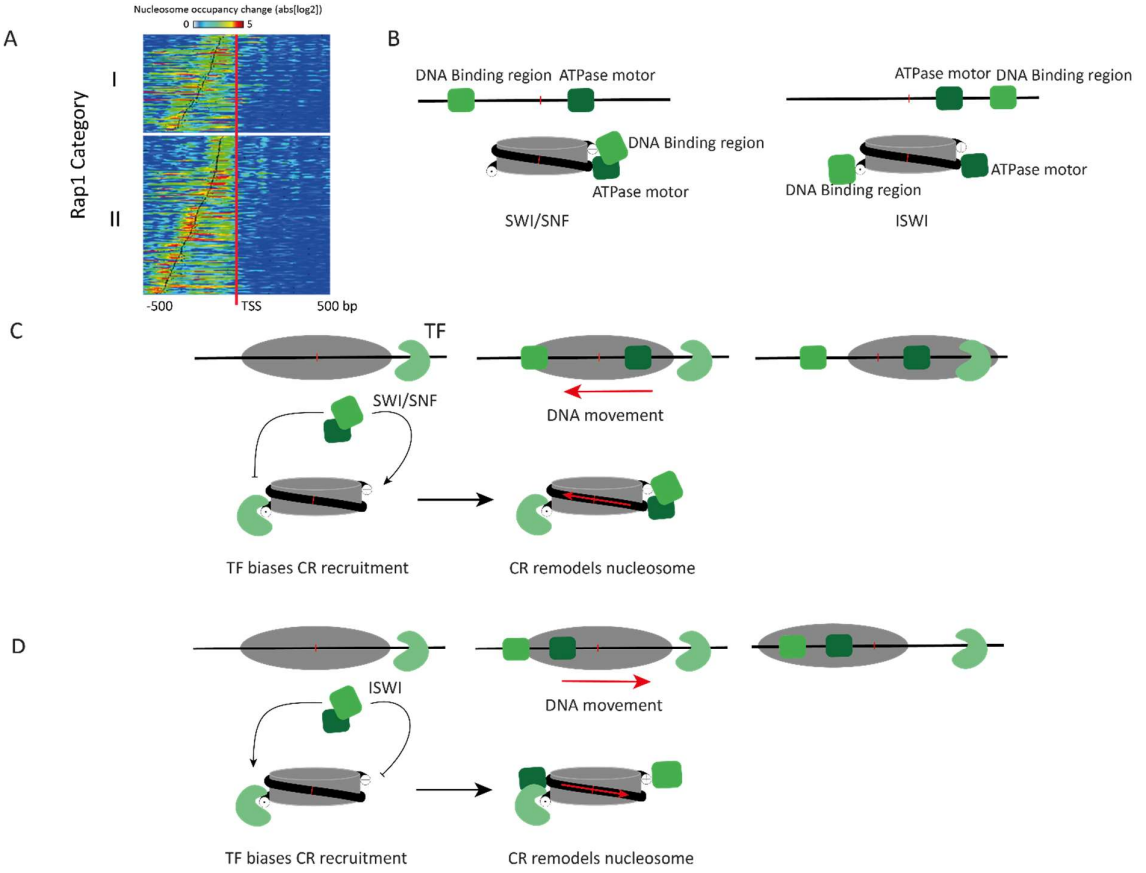


Figure 31 A) Change in ChIP-Seq signal from studies in the presence of Rap1 or during depletion. The change in nucleosome occupancy is mapped and shows that Rap1 is often found at the edge of the largest fold change indicating a directionality. B) Scheme of Swi/Snf and ISWI emodelling regulation based on previous literature. C) Proposed mechanism of transcription factor induced bias during remodelling of Swi/Snf family members as observed in Li et al, as well as our work. D) Proposed mechanism of transcription factor induced bias during remodelling of ISWI family members as observed in Li et al.

Secondly, the Gal4 DBD binding sites are initially found outside the Widom 601 nucleosome, and Gal4 has been shown to require DNA breathing to bind nucleosomal DNA. As opposed to our assay with the initial position of the Rap1 sites within the nucleosome, reducing but not abolishing Rap1's ability to bind. These points mean that we may be observing snapshots of the same process at different time points.

One explanation is that the presence of Gal4 DBD binding sites on one side of the nucleosome may bias the recruitment of SWI/SNF to the opposite side. Structural studies (including hydroxyl radical footprinting and cryoEM) have shown that the SWI/SNF family members RSC and BAF have DNA binding domains that bind on the same side of the nucleosome as the ATPase motor (**Fig32B**) (104,241,242). As CRs feed DNA into the dyad through the same side as they bind, direct competition by restricting access on one side may explain the observation that SWI/SNF remodels towards Gal4 DBD (**Fig32C**). On the other hand, ISWI family CRs have

been shown to have their DNA binding SANT-Slide domains bind on the opposite side of the nucleosome to the ATPase motor (91,243). This would mean CR binding on the same side as a TF could be unhindered by the presence of the TF (**Fig31D)**. This model is based on crosslinking studies with peptide mapping and is contradicted by recent structural models (146). For ISWI, it is postulated that a regulatory mechanism works through an inhibitory state of ISWI when the DBD is not bound on the opposite side, bringing the inhibitory NegC domain close to the ATPase motor limiting ATP access (90,91). This may be the state captured in the recent structures with both the ATPase motor and the DNA binding domain on the same side of the nucleosome.

Taken together, the data allows for the following overall model for Rap1 and RSC chromatin remodelling (**Fig33**) to be posited. Whereby Rap1 only weakly binds nucleosomes but is able to passively interfere with nucleosome-nucleosome stacking. However, Rap1 does not evict nucleosomes when bound. RSC (SWI/SNF family) may then initially bind more or less equally on both sides as Rap1 binds and unbinds rapidly. ATP hydrolysis and remodelling then either further masks or unveils the Rap1 binding sites depending on the remodelling direction. In the case of unveiling the Rap1 DNA binding sites, this allows high-affinity Rap1 binding to free DNA (Kd ~ 10 nM). This now partially recapitulates the Gal4 DBD experiments, as Rap1 is, in fact, tightly bound to one side of the remodelled nucleosome, potentially blocking RSC recruitment. As the free side of the nucleosome is devoid of free linker DNA as it is end positioned, RSC does not remodel the nucleosome towards the Rap1 sites.

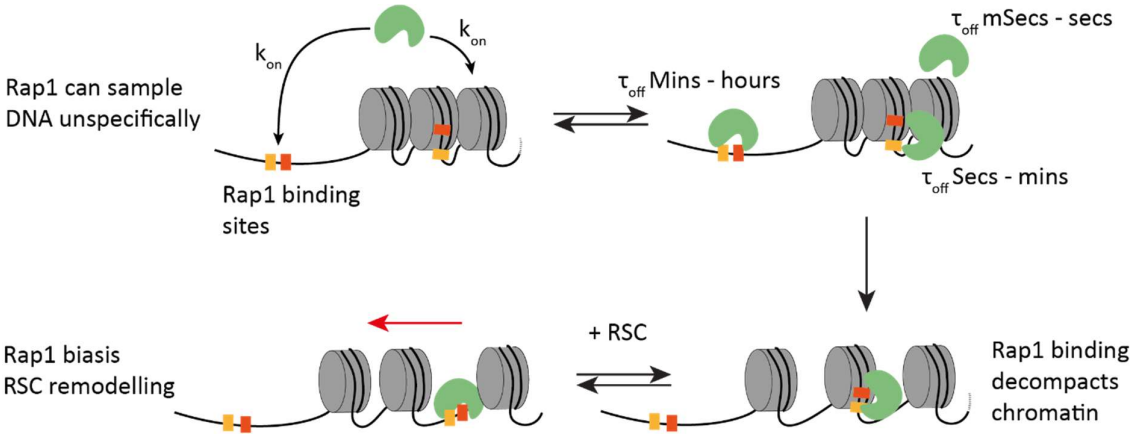


Figure 32 Overall Rap1 binding and interaction with RSC model. Rap1 can unspecifically scan DNA. When bound, Rap1 stably binds free DNA and transiently binds sites within nucleosomes. Rap1 binding decompacts chromatin structure. In collaboration with RSC, Rap1 pushes RSC away from its binding sites.

Disparate and conflicting structural data on CR regulation has led us to pursue the creation of a platform able to study CR remodelling and the influence of TFs in real-time.

5 Single-molecule FRET-based real-time remodelling assays

This chapter outlines unpublished work.

Chapter summary

The chromatin remodeler Chromodomain Helicase DNA binding protein 1 (Chd1) was expressed and purified using the baculovirus expression system. Chd1 was shown to be functional using gel assays and remodel asymmetric nucleosomes to a more central position. Reconstituted mono-nucleosomes were prepared containing a FRET pair, with a donor (Cy3B) at the entry-exit site on the DNA 70 bp from the dyad and an acceptor (A647) on histone H2A at position 110 (H2AN110C-A647). Using this FRET system, remodelling of Chd1 on mono-nucleosomes with both asymmetric and symmetric linker DNA was followed by single-molecule FRET. Experiments in the presence of Rap1 did not show changes in the remodelling dynamics, although the data is challenging to interpret. Together this chapter outlines preliminary steps in establishing a modular platform to follow chromatin remodelling in real-time.

5.1 Chd1 as a model Chromatin Remodeler

As mentioned above, CRs often require large multi-subunit complexes to function that are difficult to obtain in purified form with sufficient yields. Chromodomain Helicase DNA binding protein 1 (CHD1) is an exception to this as it is expressed as a single polypeptide chain containing the ATPase motor, chromodomains and a DNA-binding region (**Fig34A**). This allows for expression using the baculovirus-insect cell expression system, making it more easily obtainable. Chd1 has been shown to act intra-genically, evenly spacing nucleosomes (48,244). With this CR, we want to establish a modular FRET-based nucleosome system able to screen

the effect on CR remodelling with TFs binding at different superhelix locations, PTMS and nucleosome variants. To begin with, we have also chosen Rap1 as our model TF, as we have a good understanding of how Rap1 interacts with DNA in different contexts. Although few biological interactions have been reported between Rap1 and Chd1, as their function has been localized in different genomic regions, this TF-CR pair may still interact due to Rap1's ability to bind partial motifs *in vivo* (245). However, our primary motivation is to set up our assay using this pair and to be able to screen more relevant TF-CR in the future.

For this, we have to express and purify Chd1 and explore different modular nucleosome FRET constructs to best explore the changes in recruitment and remodelling dynamics of CRs.

5.2 Expression and purification of yeast Chd1

A vector encoding full-length *S. cerevisiae* Chd1 was purchased from the MRC PPU Reagents and Services facility (University of Dundee). This is the same construct used for the cryoEM structure of Chd1 with a nucleosome (105). The sequence of interest was cloned out by PCR with suitable overhangs to insert into a recipient PaceBac plasmid by the Gibson cloning protocol. Two constructs were made, MBP-Chd1 and MBP-Chd1-Halo (**Fig 34A**); the latter was not expressed but could be used to study Chd1 binding dynamics to different DNA substrates.

Expression tests showed an optimal expression time of 3 days at 27°C, after which cells were harvested by centrifugation and flash frozen. After purification using a Strep column (**Fig 34B**), the MBP-Chd1 construct was cleaved using TEV-protease to remove the MBP tag (**Fig 34C**), and final purification was done using size exclusion chromatography (**Fig 34D**). Chd1 was concentrated to 1.5 μM, aliquoted and flash frozen.

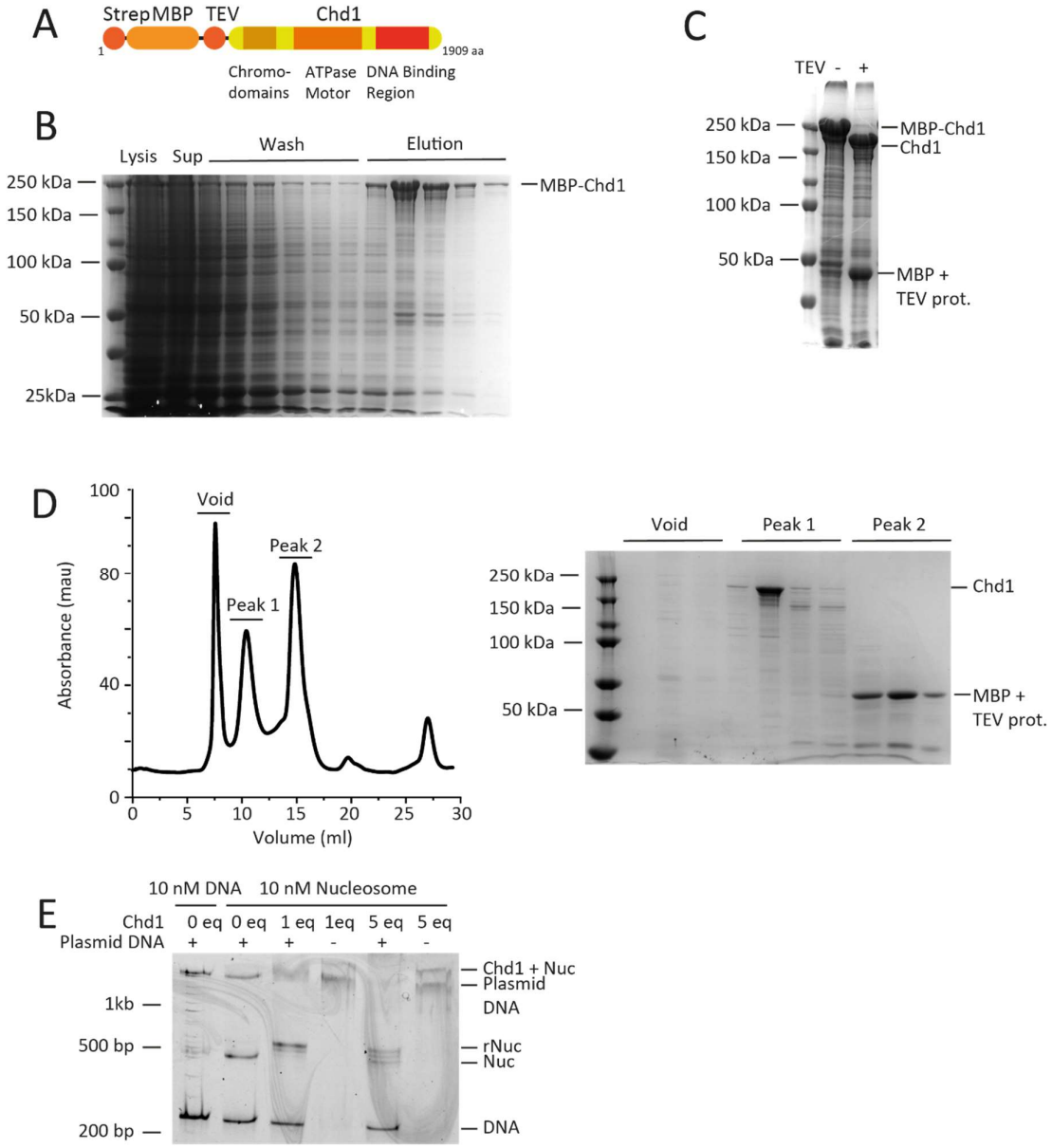


Figure 33 A) Scheme of expressed Chd1 construct. N terminus streptavidin tag (Strep) and maltose binding protein (MBP) are used for purification. TEV can be used to cleave Chd1 from strep-MBP. B) SDS-PAGE gel of streptavidin purification. C) SDS-PAGE gel of TEV protease cleavage overnight. D) FPLC size exclusion chromatogram of cleaved Chd1 to separate Chd1 from MBP and TEV protease. SDS-PAGe gel of FPLC purification fractions. E) Asymmetric nucleosomes are reconstructed and incubated with Chd1 in the prescence of absence of plasmid DNA.

5.3 Chd1 remodels nucleosomes to central positions

Knowing that Chd1 binds asymmetric nucleosomes to position them centrally, we cloned a 228 bp asymmetric nucleosome using a Widom 601 sequence and 78 bp linker (**Fig 34E, Table S1**). To first test the amount of Chd1 required for a standard remodelling experiment, 1 molar equivalent (eq) and 5 eq Chd1 was added to 10 nM asymmetric nucleosomes in the presence of 1 mM ATP. We see that 1 eq is sufficient to remodel asymmetric nucleosomes fully. Newly formed symmetrical nucleosomes are super-shifted compared to the asymmetric nucleosomes (as seen previously) (**Fig 33E**). Knowing that our purified Chd1 is functional, we made an asymmetric nucleosome containing a FRET compatible dye pair.

5.4 Generating asymmetric FRET nucleosomes

We first chose to recapitulate FRET positions in previously described smFRET experiments with Chd1 from the Deindl laboratory. Their construct has a donor at position +70 inserted into the DNA and an acceptor in histone H2A at position 120 (H2AT120C) (246,247). For our FRET system, we used H2AN110C, as it was available in our lab (**Fig 35A**). This change did not dramatically reduce the E_{FRET} at the starting position as calculated *in silico* to be 0.99 at the start position (**Fig 35A**). Following expression, inclusion body protein purification and ion-exchange chromatography, H2AN110C was purified by semipreparative HPLC and lyophilized (**Fig 35B**). Alexa Fluor 647 (Alexa647) maleimide was used to conjugate the cysteine cloned at position 110 in H2AN110C. H2AN110C was dissolved in PBS pH 7.4 with 1 eq TCEP for 20 mins before 3 eq of A647-maleimide was added. The reaction was followed by HPLC and was found to have been completed after 30 mins (**Fig 35B**). Labelled H2AN110C_A647 was analysed by LC-MS before being purified by preparatory HPLC and lyophilized (**Fig 35C**). Octamers were reconstituted as described above with a 50 : 50 mix of labelled H2AN110C_A647 and wt H2A (**Fig 35C)**. Statistically, with this mix, we would get 25% unlabelled octamers, 25% double labelled and 2 x 25% labelled on each side. The octamer is symmetrical, however, when we reconstitute nucleosomes, the DNA contains a single dye, meaning we will have 50 % of our singly labelled population on the same side as the dye and 50 % on the far side of the nucleosome. Therefore, only 25 % of our total sample will display high E_{FRET} and a single

acceptor bleaching step. Although this is not ideal, we can still get sufficient statistics from our assays with ~500 individual spots per field of view, meaning 125 traces per acquired movie. Purified octamers were then used to reconstitute nucleosomes with a Cy3B dye at position +70 and tested to see if Chd1 could remodel even in the presence of the dyes (**Fig 35D)**.

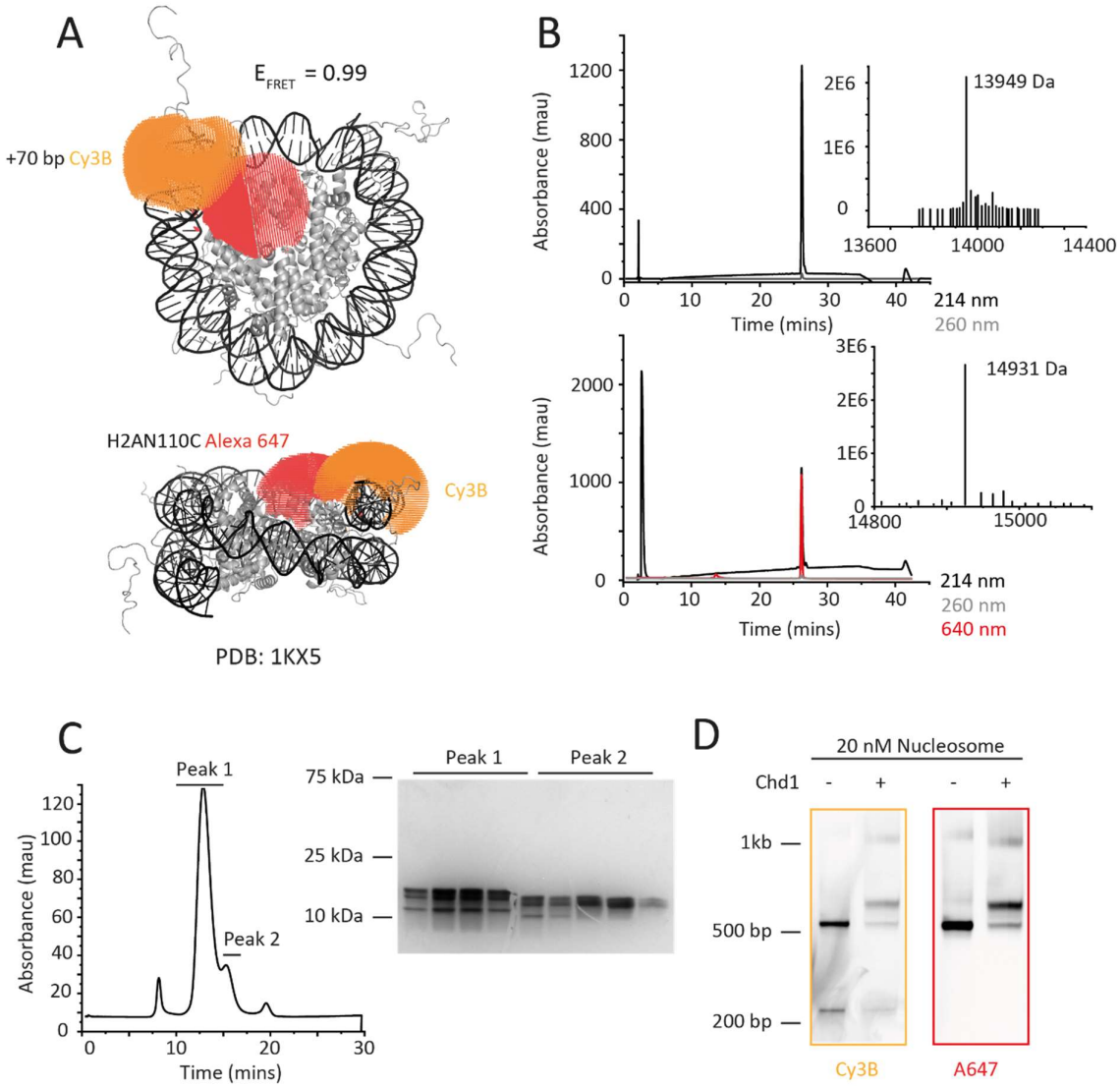


Figure 34A) FRET efficiency calculation using PDB:1KX5 and software from the Seidel group for Cy3B +70 and A647 H2AN110C. B) Analytical HPLC trace of purified H2AN110C and mass spectrometry deconvolution data. Analytical HPLC trace of purified H2AN110C labelled with Alexa 647 maleimide and mass spectrometry deconvolution data. C) FPLC size exclusion chromatogram of H2AN110C A647 incorporated octamers and SDS-PAGE gel of corresponding peaks. D) Nucleosomes containing both H2AN110C Alexa 647 and +70 Cy3b.

5.5 Ensemble FRET measurements using +70 Cy3B DNA and H2AN110C A647 asymmetric nucleosomes

To see if we can follow remodelling using our FRET nucleosomes, ensemble FRET measurement experiments were performed. With only a subset of nucleosomes in the correct high-FRET configuration, we cannot use this experiment to calculate the E_{FRET} of our dye pair. Nevertheless, we can see whether, upon remodelling, we can observe a change in E_{FRET}, indicating whether or not we can use this configuration for smFRET experiments. Measured spectra showed FRET only in the presence of the nucleosome, with an ensemble E_{FRET} ~ 0.25. After remodelling with Chd1 for 5 mins, E_{FRET} was ~ 0.1. This indicates that at least a subset of nucleosomes showing FRET has been remodelled to lower FRET states. This change in FRET is small as it concerns a subset of nucleosomes, with the majority of nucleosomes showing no change as they either lack an acceptor on H2A or the acceptor is on the far side. Kinetic experiments could be done by measuring E_{FRET} at different time points. However, we have found these types of ensemble FRET experiments to be difficult to interpret due to photobleaching. To better follow the remodelling kinetics, we proceeded to smFRET experiments.

5.6 smFRET measurements using +70 Cy3B DNA and H2AN110C A647 asymmetric nucleosomes

Asymmetric FRET nucleosomes were immobilized in a microchannel and imaged using alternate excitation smTIRF (ALEX smTIRF). This allows us to directly excite both the donor and the acceptor in alternation to monitor their presence, allowing us to exclude traces with bleaching or those presenting photophysical anomalies. Traces are gathered and curated in a semi-automated manner using a custom built MATLAB program suite to exclude traces missing either a donor or an acceptor. This leaves traces containing both dyes. Finally, upon acceptor bleaching, a concomitant increase in donor intensity should be observed. This selection criteria for traces was similar to that used previously (44). The asymmetric nucleosome displayed two FRET populations as predicted, and these had an E_{FRET} = 0.93 ± 0.05 and 0.63 ± 0.11, respectively (**Fig 36A**). These values agree with the *in silico* calculated values

of E_{FRET} = 0.99 and **E_{FRET} = 0.58.** The high-FRET value (E_{FRET} 0.93) corresponds to the dye pair on the same side of the nucleosome, and the mid-FRET value is when the dyes are on opposite sides. Previous studies suggested that even in the absence of ATP, Chd1 binding could perturb nucleosome conformation with the unwanted effect of changing the measured E_{FRET} (106). In our hands, in-gel experiments suggested that ATP was required for Chd1 to super-shift nucleosomes. Therefore ATP is required for stable binding. In our smFRET microarrays, Chd1 was flowed in without ATP, measured FRET values remained similar (E_{FRET} = 0.91 ± 0.02 and E_{FRET} = 0.63 ± 0.04) suggesting Chd1 transient binding does not influence the FRET pair (**Fig 36B**).

Experiments flowing in 10 nM Chd1 at 10 s in the presence of 1 mM ATP in imagining buffer containing 5 mM $MgCl_2$ showed clear remodelling (**Fig 36C**). We did not quantify the rates of remodelling as at this stage, we wanted to explore other constructs that could be used for all chromatin remodeller types. As these asymmetric nucleosomes force the CR to one side, remodelling will occur in a predictable direction but may not be unanimously accepted as a substrate for different CRs (248). We, therefore, hypothesized that a symmetrical nucleosome with long linkers could be better used as it more faithfully reflects what TFs and CRs could encounter *in vivo*.

5.7 smFRET measurements using +70 Cy3B DNA and H2AN110C A647 symmetric nucleosomes

Using the same dye pair and positions, however, now in a symmetrical nucleosome with 50 bp linkers on both sides, E_{FRET} was measured using ALEX TIRF smFRET as above (Table S1). With linkers on both sides, we hoped to capture an oscillating movement of Chd1 remodelling in both directions. The nucleosome alone showed similar E_{FRET} values as seen in the asymmetric construct, with measured values of E_{FRET} = 0.94 ± 0.03 and E_{FRET} = 0.56 ± 0.11 (**Fig 36D**).

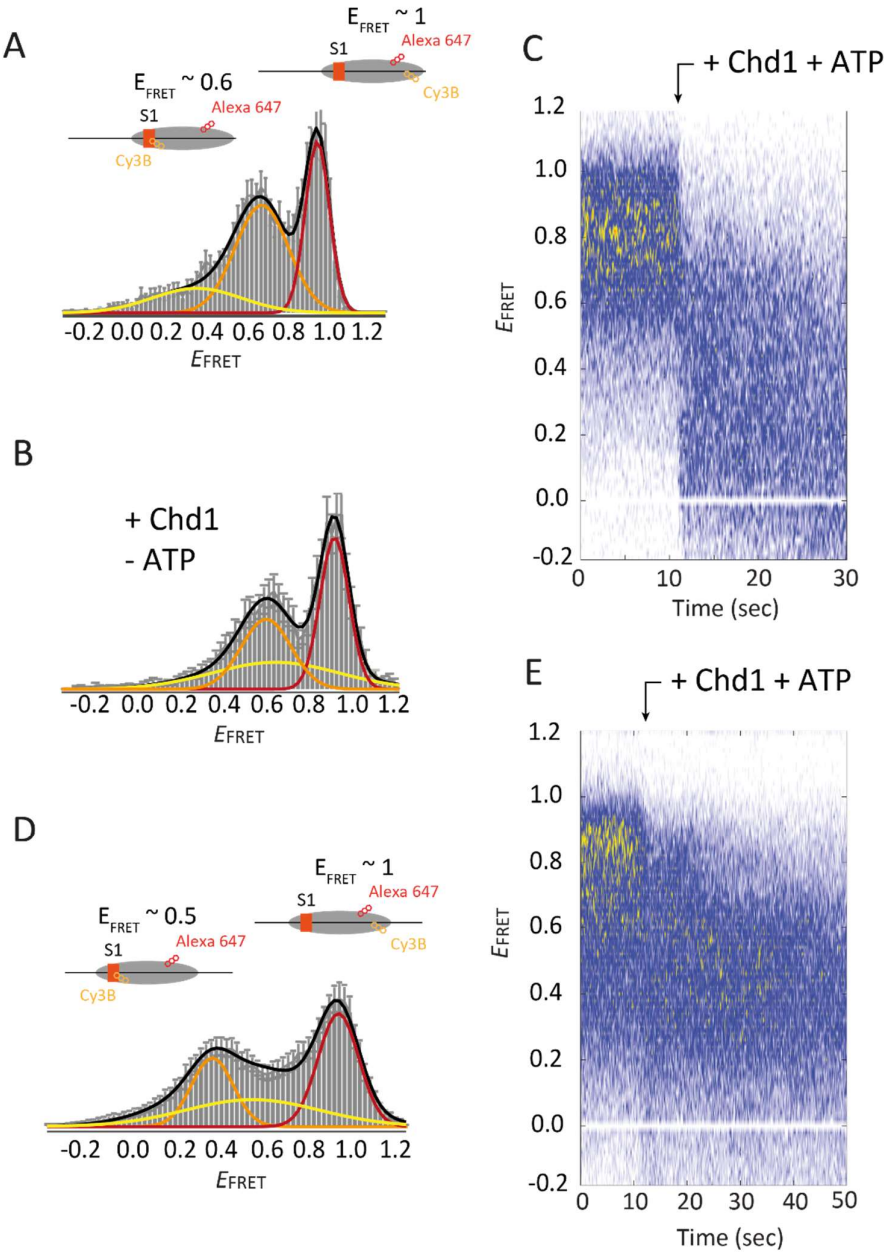


Figure 35 A) Cumulative histograms of smFRET on asymmetric nucleosomes. Two distinct FRET populations can be seen which are thought to originate from the position of the acceptor on the octamer Alexa647 vis-à-vis the donor on the DNA Cy3B. When they are on the same side we expect E_{FRET} = 1 and E_{FRET} = 0.6 when on opposite sides. B) Cumulative histograms of smFRET on asymmetric nucleosomes in the presence of Chd1, no change in FRET is observed compared to no Chd1. C) Cumulative time FRET trace of Chd1 injection experiment using asymmetric nucleosomes. After injection at 10s we see a rapid decrease in FRET. D) Cumulative histograms of symmetric nucleosomes, the high FRET (red) population is the same, whereas there is a decrease in the mid-FRET population (orange). D) Cumulative time FRET trace of Chd1 injection experiment. After injection at 10s we see a rapid decrease in FRET using symmetric nucleosomes.

The addition of Chd1 did show remodelling but no total loss of signal, such as for the asymmetric nucleosomes (**Fig 36E**). In literature, few examples of Chd1 remodelling on already centred nucleosomes have been reported, although some have shown little to no effect (248). We thought that the addition of preincubated Rap1 might show a different remodelling profile as Rap1 would block Chd1 on one side, replicating the RSC/ISWI data discussed previously (240). The Rap1 binding motif was placed at +75 and +85 bp away from the dyad and on the opposite side as the FRET pair. The addition of 1 nM Rap1, which has previously been shown to be sufficient to saturate Rap1 binding sites, did not show a significant change in the remodelling profile (**Fig 37A**).

5.8 Discussion on real-time FRET-based TF-CR nucleosome remodelling

Using different FRET pairs containing nucleosomes, we show that we can follow Chd1 remodelling in real-time. These experiments have shown us that Chd1 fully remodels asymmetric nucleosomes but does not fully remodel symmetrical nucleosomes. This is not unexpected as Chd1 is known to centre and equally space nucleosomes. However, by using 50 bp linkers, we thought that the Chd1 bound nucleosome would display an oscillatory behaviour as it would first remodel in one direction before moving it back. Some examples of this are seen in our data (**Fig 37B**). The exact regulatory mechanism of remodelling largely remains unknown, but as discussed above, ISWI and Chd1 are thought to have an inhibitory regulation of transcription (91,243). Chd1 remodelling seems to be confined in our symmetrical nucleosomes, although the DNA binding arm should have sufficient DNA to bind. Furthermore, the presence of Rap1 did not seem to induce a change in the remodelling profile. Our FRET positions do not allow us to decipher the direction of remodelling as we start at a high-FRET value, and any movement will induce a reduction in FRET. There may be a directionality change in the presence of Rap1 that we are unable to detect. Furthermore, it is evident that photobleaching increases in the presence of the CR. This is not due to the presence of ATP in the buffer, as experiments flowing in the buffer only did not show any change. This increased photobleaching rate may come from the fact that when remodelled, the dye on the DNA passes “through” the nucleosome, which means that rotation of DNA wedges the dye between the DNA and the surface of the octameric core. This change in

environment may reduce access to triplet quenchers and other dye stabilizing additives that increase photostability.

Furthermore, the presence of potential nucleophiles on the surface of the octamer may help increase the photobleaching/blinking rate, which is observed in single-molecule localization microscopy (249,250). This would also explain why the asymmetric nucleosome did not suffer from this phenomenon as the end labelled dye was pushed away from the octamer and not "through". To circumvent this issue, dyes could be placed further out of the nucleosome, whereby we start at low FRET and would observe high FRET upon remodelling (**Fig 37C**). This would also allow us to follow the direction of remodelling, as only remodelling in one particular direction would increase FRET.

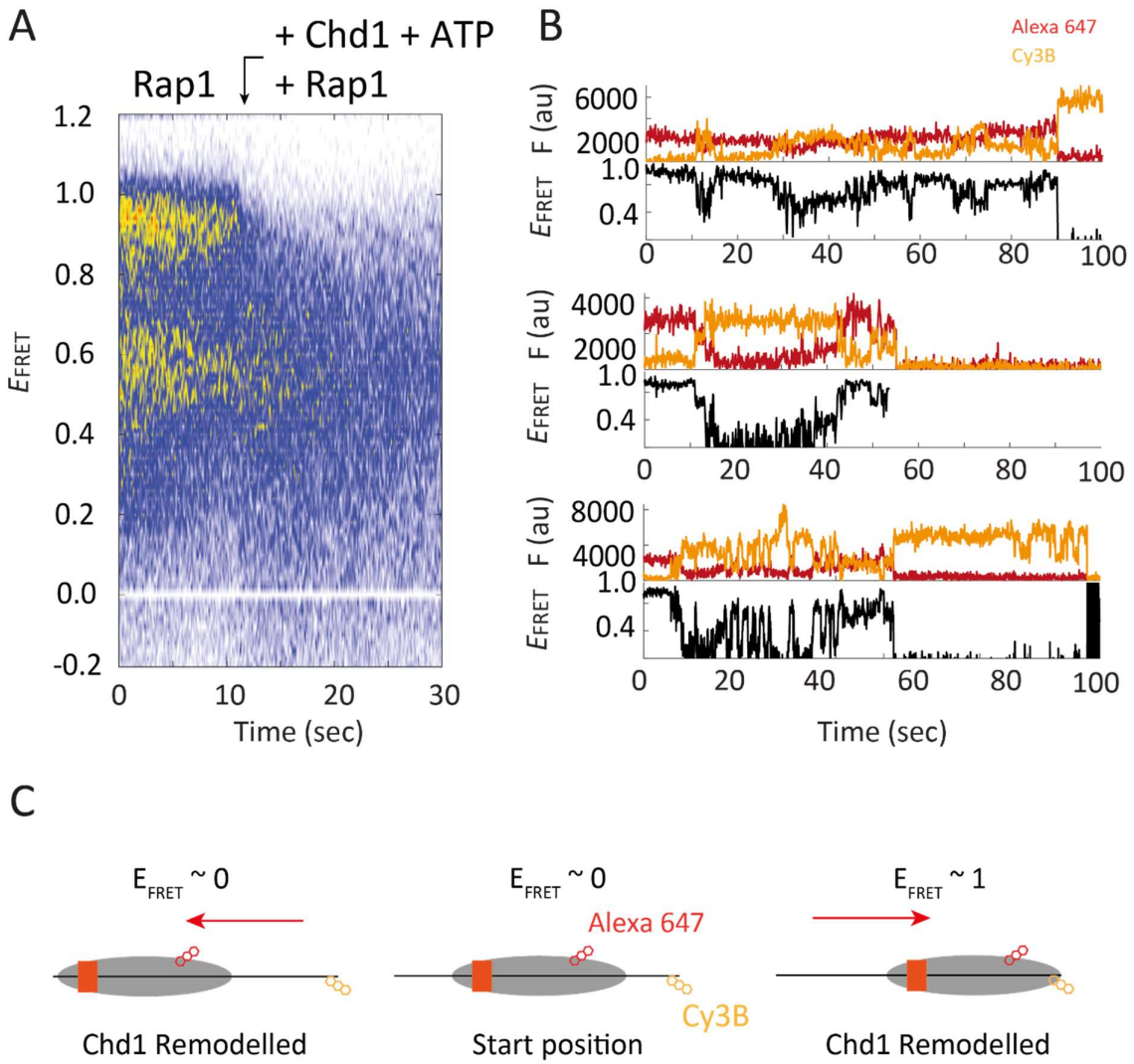


Figure 36 A) Symmetrical nucleosomes remodelled in the presence of Rap1. No difference is observed from remodelling without Rap1. B) Examples of traces where Chd1 remodels back and forth. C) Scheme of potential FRET pair location to minimise photobleaching and inform on the directionality of remodelling.

6 Thesis Summary and overall discussion

This thesis has explored the changes in the chromatin landscape in the promoter regions of genes in the initial phases of gene activation. This region compromises a nucleosome free region (NFR) flanked by a +1 nucleosome upstream and a -1 nucleosome downstream. The +1 nucleosomes contain the core promotor, including the TATA-box or Inr sequence important for preinitiation complex recruitment. The -1 nucleosome often contains binding motifs of TFs and pTFs important in regulating gene expression. We used the -1 promoter architecture of the RPL30 gene promotor loci to reconstitute defined nucleosomes containing DNA binding motifs for the yeast general transcription factor Rap1. Using *in vitro* gel-based and single-molecule TIRF colocalization microscopy assays, we show that Rap1 binding affinity is modulated by the sequence of its target motif (**Chapter 3**) and the DNA landscape in which its motifs are found. DNA binding was observed to be in the minutes to hour timescales (**Chapter 3**), whilst the presence of the nucleosome or chromatin fibres decreased dwell times 10- fold and 20-fold, respectively (**Chapter 3, Table S2**). These dwell times are longer than those observed with TFs using single-particle tracking *in vivo*. Rotational positioning of Rap1 sites can mask nucleic acids essential for Rap1 motif recognition, also modulating Rap1 binding affinity (**Chapter 3**). Although potentially buried within chromatin fibres, Rap1 was observed to have similar specific binding rates within all chromatin contexts assayed (**Chapter 3**). Upon Rap1 binding, overall chromatin compaction is reduced, although Rap1 does not affect nucleosome integrity (**Chapter 3**).

Taken together, we conclude that Rap1 binding alone is not sufficient to evict or significantly displace nucleosomes as observed *in vivo*. To further explore NFR formation by chromatin effectors, we used in-gel and Mnase-Seq experiments on mononucleosomes with the yeast SWI/SNF family chromatin remodeler RSC in the presence and absence of Rap1. We find that Rap1 influences RSC remodelling direction, pushing RSC away from Rap1 sites (**Chapter 4**).

Most experimental values reported in the literature are binding affinities, which reflect both association and dissociation rates. These binding affinities can also be thought of as binding site occupancy at a certain concentration of protein. For Rap1, the *in vivo* concentration is not known. Additionally, the abundance of telomeric Rap1 sites can sequester Rap1 away from

promotor regions, rendering the quantification of free Rap1 more difficult. Our experiments were conducted at concentrations between 50 – 100 pM, and this optimized range allowed us to observe statistically significant non-overlapping Rap1 events. We also used a single Rap1 motif at a time, but many Rap1 promotor sites have 2 distinct Rap1 sites. Knowing this, Rap1 could display cooperative binding, increasing its affinity *in vivo*. This effect could have been tested using binding assays containing both S1 and S2 Sites. Although we tried smTIRF colocalization experiments using both sites, they could not be analysed by our pipeline, as we cannot distinguish photobleaching from unbinding or the order of binding and unbinding for 2-proteins with the same dye on the same DNA fragment. EMSA or isothermal titration calorimetry experiments would be more suited for this. The mechanism of TF cooperativity remains an open question, especially for pTFs. The hierarchy of binding between TFs or TF-CR pairs has been explored for only a small proportion of pairs, such as Oct4 and Sox2(207). It was shown for this pair that Sox2 binding was required for Oct4 facilitated loading. The order of binding could determine whether a TF binds stably or not to its motif, potentially affecting the TFs downstream function *in vivo*. In our experiemnts we show that Rap1 does not require other chromatin effectors to bind chromatin, but this may not be the case for all TFs and pTF.

The question remains of how does this TF-CR relationship affect gene expression. I view the binding of gene-regulating regions (i.e. enhancers, promotors) as highly dynamic, with TFs and CRs continuously binding and unbinding. The binding rates for these chromatin effectors display a continuum based on DNA sequence, chromatin context (in particular chromatin condensation state and the presence of PTMs) and effector concentration (vis-à-vis the number of binding sites and binding affinity, including partial sites). Aided by unspecific binding, 1D sliding and hopping mechanisms, TFs can quickly scan the genome to find their specific motif sequences. CRs continuously act on nucleosomes, increasing and decreasing the size of the nucleosome free region in promotors and thus the rotational position within the -1 nucleosome containing the binding motifs for TFs. pTFs can bind their sites within this -1 nucleosome, acting as a feed-forward loop by influencing the recruitment and direction of CR action, thus modulating their binding affinities. I would posit that stable binding times of TFs (in nucleosome-free conditions) are found in the minute's timescales, increasing the chance of co-factor recruitment.

In my view, promotor architecture is hardcoded into the genome, thereby predetermining the ease of transcription initiation by having DNA sequences refractory to nucleosome formation generating large NFRs, strong TF binding sites or CR directionality motifs to maintain an environment that can recruit the PIC. These differences have been observed and depend on the gene type (homeostasis or developmental genes).

This view also implies that if TFs are present in sufficient concentrations, transcription will occur. This is supported by the concept of induced pluripotent stem cells, whereby fully differentiated cells regain pluripotency by ectopic expression of TFs. The chromatin landscape through cell division evicts most TFs and reduces the concentration of TFs stochastically. Here the asymmetric cell division of stem cells gives us a clue of how to maintain stemness, in other words, a more permissive transcriptional landscape. In asymmetric cell division, one cell keeps its pluripotent abilities, whereas the daughter cell starts to differentiate. In the case of the daughter cell, it has been observed that differentiation may be induced by the lack of specific external cues (i.e. metabolites, physical stresses from its microenvironment) that hinder the re-establishment of the fully competent state by restricting TF activity. An example of this type of regulation of TFs can be seen in the steroid receptor pathways, whereby some TFs are regulated by sequestration in the cytosol until certain metabolic cues trigger their release.

We can now start to recreate the Waddington landscape. Whereby repressive and activating hallmarks are constantly turned over and compete continuously. pTFs and TFs, with their ability to read the genome, act as flag posts for transcription. Through dilution of TFs and the eviction of most TFs during cell division, this process restarts constantly. However, through several cycles, genomic regions partially maintain certain PTMs favouring one state over the other. This is exasperated by the presence/absence of metabolic cues. Finally, topological features such as gene clusters or topologically associated domains tip the balance into a transcriptionally silent or active state by increasing the local concentrations of factors.

However, once bound, the direct association between TF binding dynamics and gene expression levels remains unclear. Transcription has been observed in cells to occur in bursts and not at a constant rate (131,251,252). In this process, several RNAs are transcribed before a silent period with no transcriptional activity. Overall, the number of transcripts from a given gene is relatively low, with 80% of yeast genes having only around 8 copies of RNA molecules

per gene found at a given time (252). Studies looking at the transcription rates and correlating this to transcriptional bursts have found a relationship between longer TF dwell times, TF concentration and overall transcription output (253–255). A further study using steroid receptors with modular binding dynamics showed reduced off times and not increased on times that generated the largest overall transcriptional bursts (256). Taken together, these works paint a nuanced picture beyond the relationship between a simple relationship between TF binding and transcriptional output. The existence of gene clusters and phase-separated droplets may also contribute to this bursting activity, increasing local TF and polymerase concentration, although this remains to be explored further.

To further explore the interaction between TFs and CRs, we show initial experiments for a FRET-based platform that follows nucleosome remodelling in real-time (**Chapter 5**). We plan to optimize this platform further to enable us to screen different TF-CR pairs as well as PTMs and histone variants.

7 References

1. Waddington CH. The strategy of the genes: a discussion of some aspects of theoretical biology [Internet]. Allen & Unwin; 1957. Available from: https://books.google.ch/books?id=anEJtgEACAAJ

2. Gurdon JB. The Developmental Capacity of Nuclei taken from Intestinal Epithelium Cells of Feeding Tadpoles. Development. 1962 Dec 1;10(4):622–40.

3. Berg JM, Tymoczko JL, Gatto GJ, Stryer L. Biochemistry. 2019.

4. Hagerman PJ. Flexibility of DNA. Annu Rev Biophys Biophys Chem. 1988;17(1):265–86.

5. Rittman M, Gilroy E, Koohy H, Rodger A, Richards A. Is DNA a worm-like chain in Couette flow?: In search of persistence length, a critical review. Sci Prog. 2009 Jul 1;92(2):163–204.

6. Brunet A, Tardin C, Salomé L, Rousseau P, Destainville N, Manghi M. Dependence of DNA Persistence Length on Ionic Strength of Solutions with Monovalent and Divalent Salts: A Joint Theory–Experiment Study. Macromolecules. 2015 Jun 9;48(11):3641–52.

7. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Molecular Biology of the Cell. 4th ed. Garland Science; 2002.

8. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997 Sep 18;389(6648):251–60.

9. Campos EI, Reinberg D. Histones: Annotating Chromatin. Annu Rev Genet. 2009 Nov 17;43(1):559–99.

10. Lai WKM, Pugh BF. Understanding nucleosome dynamics and their links to gene expression and DNA replication. Nat Rev Mol Cell Biol. 2017 Sep;18(9):548–62.

11. Zhou K, Gaullier G, Luger K. Nucleosome structure and dynamics are coming of age. Nat Struct Mol Biol. 2019 Jan;26(1):3–13.

12. McGinty RK, Tan S. Nucleosome Structure and Function. Chem Rev. 2015 Mar 25;115(6):2255–73.

13. Hall MA, Shundrovsky A, Bai L, Fulbright RM, Lis JT, Wang MD. High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. Nat Struct Mol Biol. 2009 Feb;16(2):124–9.

14. Aguilar-Gurrieri C, Larabi A, Vinayachandran V, Patel NA, Yen K, Reja R, et al. Structural evidence for Nap1-dependent H2A–H2B deposition and nucleosome assembly. EMBO J. 2016 Jul 1;35(13):1465–82.

15. Smith S, Stillman B. Stepwise assembly of chromatin during DNA replication in vitro. EMBO J. 1991 Apr 1;10(4):971–80.

16. Flaus A. Principles and practice of nucleosome positioning *in vitro*. Front Life Sci. 2011 Jun;5(1–2):5–27.

17. Cutter A, Hayes JJ. A Brief Review of Nucleosome Structure. FEBS Lett. 2015 Oct 7;589(20 0 0):2914–22.

18. Ausio J, Dong F, van Holde KE. Use of selectively trypsinized nucleosome core particles to analyze the role of the histone "tails" in the stabilization of the nucleosome. J Mol Biol. 1989 Apr 5;206(3):451–63.

19. Ou HD, Phan S, Deerinck TJ, Thor A, Ellisman MH, O'Shea CC. ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. Science. 2017 28;357(6349).

20. Ricci MA, Manzo C, García-Parajo MF, Lakadamyali M, Cosma MP. Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes In Vivo. Cell. 2015 Mar 12;160(6):1145–58.

21. Portillo-Ledesma S, Tsao LH, Wagley M, Lakadamyali M, Cosma MP, Schlick T. Nucleosome Clutches are Regulated by Chromatin Internal Parameters. J Mol Biol. 2021 Mar 19;433(6):166701.

22. Alvarado W, Moller J, Ferguson AL, de Pablo JJ. Tetranucleosome Interactions Drive Chromatin Folding. ACS Cent Sci. 2021 Jun 23;7(6):1019–27.

23. McGhee JD, Nickol JM, Felsenfeld G, Rau DC. Higher order structure of chromatin: Orientation of nucleosomes within the 30 nm chromatin solenoid is independent of species and spacer length. Cell. 1983 Jul 1;33(3):831–41.

24. Kruithof M, Chien F-T, Routh A, Logie C, Rhodes D, van Noort J. Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30-nm chromatin fiber. Nat Struct Mol Biol. 2009 May;16(5):534–40.

25. Dorigo B, Schalch T, Bystricky K, Richmond TJ. Chromatin fiber folding: requirement for the histone H4 N-terminal tail. J Mol Biol. 2003 Mar 14;327(1):85–96.

26. Szerlong HJ, Hansen JC. Nucleosome distribution and linker DNA: connecting nuclear function to dynamic chromatin structure. Biochem Cell Biol Biochim Biol Cell. 2011 Feb;89(1):24–34.

27. Peters JP, Maher LJ. DNA curvature and flexibility in vitro and in vivo. Q Rev Biophys. 2010 Feb;43(1):23–63.

28. Gore J, Bryant Z, Nöllmann M, Le MU, Cozzarelli NR, Bustamante C. DNA overwinds when stretched. Nature. 2006 Aug 17;442(7104):836–9.

29. Gross P, Laurens N, Oddershede LB, Bockelmann U, Peterman EJG, Wuite GJL. Quantifying how DNA stretches, melts and changes twist under tension. Nat Phys. 2011 Sep;7(9):731–6.

30. Etheve L, Martin J, Lavery R. Protein–DNA interfaces: a molecular dynamics analysis of time-dependent recognition processes for three transcription factors. Nucleic Acids Res. 2016 Nov 16;44(20):9990–10002.

31. Mukherjee S, Sarkar SN, Raychaudhuri P, Mazumdar SK. Propagation of elastic waves in DNA [Internet]. Vol. 6, International Journal of Mathematics and Mathematical Sciences. Hindawi; 6 [cited 2020 Apr 23]. p. eS0161171283000654. Available from: https://www.hindawi.com/journals/ijmms/1983/701593/

32. Nelson P. Transport of torsional stress in DNA. Proc Natl Acad Sci. 1999 Dec 7;96(25):14342–7.

33. Kannan S, Kohlhoff K, Zacharias M. B-DNA Under Stress: Over- and Untwisting of DNA during Molecular Dynamics Simulations. Biophys J. 2006 Oct 15;91(8):2956–65.

34. Fierz B, Poirier MG. Biophysics of Chromatin Dynamics. Annu Rev Biophys [Internet]. 2019 May 6 [cited 2019 Apr 5];48(1). Available from: https://www.annualreviews.org/doi/10.1146/annurev-biophys-070317-032847

35. Polach KJ, Widom J. Mechanism of Protein Access to Specific DNA Sequences in Chromatin: A Dynamic Equilibrium Model for Gene Regulation. J Mol Biol. 1995 Nov 24;254(2):130–49.

36. North JA, Shimko JC, Javaid S, Mooney AM, Shoffner MA, Rose SD, et al. Regulation of the nucleosome unwrapping rate controls DNA accessibility. Nucleic Acids Res. 2012 Nov;40(20):10215–27.

37. Li G, Levitus M, Bustamante C, Widom J. Rapid spontaneous accessibility of nucleosomal DNA. Nat Struct Mol Biol. 2005 Jan;12(1):46–53.

38. Brandani GB, Niina T, Tan C, Takada S. DNA sliding in nucleosomes via twist defect propagation revealed by molecular simulations. Nucleic Acids Res. 2018 Apr 6;46(6):2788–801.

39. Ngo TTM, Ha T. Nucleosomes undergo slow spontaneous gaping. Nucleic Acids Res. 2015 Apr 30;43(8):3964–71.

40. Bilokapic S, Strauss M, Halic M. Histone octamer rearranges to adapt to DNA unwrapping. Nat Struct Mol Biol. 2018 Jan;25(1):101–8.

41. Bilokapic S, Strauss M, Halic M. Structural rearrangements of the histone octamer translocate DNA. Nat Commun. 2018 Apr 6;9(1):1–11.

42. Brehove M, Shatoff E, Donovan BT, Jipa CM, Bundschuh R, Poirier MG. DNA sequence influences hexasome orientation to regulate DNA accessibility. Nucleic Acids Res. 2019 Jun 20;47(11):5617–33.

43. Schalch T, Duda S, Sargent DF, Richmond TJ. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. Nature. 2005 Jul 7;436(7047):138–41.

44. Kilic S, Felekyan S, Doroshenko O, Boichenko I, Dimura M, Vardanyan H, et al. Single-molecule FRET reveals multiscale chromatin dynamics modulated by HP1α. Nat Commun. 2018 Jan 16;9(1):235.

45. Kaczmarczyk A, Meng H, Ordu O, Noort J van, Dekker NH. Chromatin fibers stabilize nucleosomes under torsional stress. Nat Commun. 2020 Jan 8;11(1):1–12.

46. Hergeth SP, Schneider R. The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. EMBO Rep. 2015 Nov;16(11):1439–53.

47. Strahl BD, Allis CD. The language of covalent histone modifications. Nature. 2000 Jan;403(6765):41.

48. Kubik S, Bruzzone MJ, Challal D, Dreos R, Mattarocci S, Bucher P, et al. Opposing chromatin remodelers control transcription initiation frequency and start site selection. Nat Struct Mol Biol. 2019 Aug;26(8):744–54.

49. Clapier CR, Cairns BR. The Biology of Chromatin Remodeling Complexes. Annu Rev Biochem. 2009;78(1):273–304.

50. Hepp MI, Alarcon V, Dutta A, Workman JL, Gutiérrez JL. Nucleosome remodeling by the SWI/SNF complex is enhanced by yeast High Mobility Group Box (HMGB) proteins. Biochim Biophys Acta. 2014 Sep;1839(9):764–72.

51. Tessarz P, Kouzarides T. Histone core modifications regulating nucleosome structure and dynamics. Nat Rev Mol Cell Biol. 2014 Nov;15(11):703–8.

52. Bowman GD, Poirier MG. Post-Translational Modifications of Histones That Influence Nucleosome Dynamics. Chem Rev. 2015 Mar 25;115(6):2274–95.

53. Neumann H, Hancock SM, Buning R, Routh A, Chapman L, Somers J, et al. A Method for Genetically Installing Site-Specific Acetylation in Recombinant Histones Defines the Effects of H3 K56 Acetylation. Mol Cell. 2009 Oct 9;36(1):153–63.

54. Yang D, Arya G. Structure and binding of the H4 histone tail and the effects of lysine 16 acetylation. Phys Chem Chem Phys. 2011 Feb 1;13(7):2911–21.

55. Shogren-Knaak M, Ishii H, Sun J-M, Pazin MJ, Davie JR, Peterson CL. Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions. Science. 2006 Feb 10;311(5762):844–7.

56. Lowary PT, Widom J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J Mol Biol. 1998 Feb 13;276(1):19–42.

57. Brogaard K, Xi L, Wang J-P, Widom J. A map of nucleosome positions in yeast at base-pair resolution. Nature. 2012 Jun 28;486(7404):496–501.

58. Geggier S, Vologodskii A. Sequence dependence of DNA bending rigidity. Proc Natl Acad Sci. 2010 Aug 31;107(35):15421–6.

59. Culkin J, de Bruin L, Tompitak M, Phillips R, Schiessel H. The role of DNA sequence in nucleosome breathing. Eur Phys J E Soft Matter. 2017 Nov 30;40(11):106.

60. Kribelbauer JF, Rastogi C, Bussemaker HJ, Mann RS. Low-Affinity Binding Sites and the Transcription Factor Specificity Paradox in Eukaryotes. Annu Rev Cell Dev Biol. 2019;35(1):357–79.

61. Rohs R, West SM, Sosinsky A, Liu P, Mann RS, Honig B. The role of DNA shape in protein–DNA recognition. Nature. 2009 Oct 1;461(7268):1248–53.

62. Rohs R, Jin X, West SM, Joshi R, Honig B, Mann RS. Origins of Specificity in Protein-DNA Recognition. Annu Rev Biochem. 2010;79(1):233–69.

63. Perez-Rueda E, Hernandez-Guerrero R, Martinez-Nuñez MA, Armenta-Medina D, Sanchez I, Ibarra JA. Abundance, diversity and domain architecture variability in prokaryotic DNA-binding transcription factors. PloS One. 2018;13(4):e0195332.

64. Wunderlich Z, Mirny LA. Different gene regulation strategies revealed by analysis of binding motifs. Trends Genet. 2009 Oct 1;25(10):434–40.

65. Smith NC, Matthews JM. Mechanisms of DNA-binding specificity and functional gene regulation by transcription factors. Curr Opin Struct Biol. 2016 Jun 1;38:68–74.

66. Charoensawan V, Wilson D, Teichmann SA. Genomic repertoires of DNA-binding transcription factors across the tree of life. Nucleic Acids Res. 2010 Nov;38(21):7364–77.

67. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, et al. The Human Transcription Factors. Cell. 2018 Feb 8;172(4):650–65.

68. Mueller AL, Corbi-Verge C, Giganti DO, Ichikawa DM, Spencer JM, MacRae M, et al. The geometric influence on the Cys2His2 zinc finger domain and functional plasticity. Nucleic Acids Res. 2020 Jun 19;48(11):6382–402.

69. Filippova GN, Fagerlie S, Klenova EM, Myers C, Dehner Y, Goodwin G, et al. An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. Mol Cell Biol. 1996 Jun;16(6):2802–13.

70. Bobola N, Merabet S. Homeodomain proteins in action: similar DNA binding preferences, highly variable connectivity. Curr Opin Genet Dev. 2017 Apr 1;43:1–8.

71. Rister J, Razzaq A, Boodram P, Desai N, Tsanis C, Chen H, et al. Single–base pair differences in a shared motif determine differential Rhodopsin expression. Science. 2015 Dec 4;350(6265):1258–61.

72. Crocker J, Abe N, Rinaldi L, McGregor AP, Frankel N, Wang S, et al. Low Affinity Binding Site Clusters Confer Hox Specificity and Regulatory Robustness. Cell. 2015 Jan 15;160(1):191–203.

73. Bürglin TR, Affolter M. Homeodomain proteins: an update. Chromosoma. 2016;125:497–521.

74. Jones S. An overview of the basic helix-loop-helix proteins. Genome Biol. 2004 May 28;5(6):226.

75. Massari ME, Murre C. Helix-Loop-Helix Proteins: Regulators of Transcription in Eucaryotic Organisms. Mol Cell Biol. 2000 Jan;20(2):429–40.

76. Panne D. The enhanceosome. Curr Opin Struct Biol. 2008 Apr 1;18(2):236–42.

77. Kim S, Broströmer E, Xing D, Jin J, Chong S, Ge H, et al. Probing Allostery Through DNA. Science. 2013 Feb 15;339(6121):816–9.

78. Rosenblum G, Elad N, Rozenberg H, Wiggers F, Hofmann H. Allostery through DNA drives phenotype switching. bioRxiv. 2020 Jan 1;2020.07.04.187450.

79. Voss TC, Hager GL. Dynamic regulation of transcriptional states by chromatin and transcription factors. Nat Rev Genet. 2014 Feb;15(2):69–81.

80. Gualdi R, Bossard P, Zheng M, Hamada Y, Coleman JR, Zaret KS. Hepatic specification of the gut endoderm in vitro: cell signaling and transcriptional control. Genes Dev. 1996 Jan 7;10(13):1670–82.

81. Cirillo LA, McPherson CE, Bossard P, Stevens K, Cherian S, Shim EY, et al. Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. EMBO J. 1998 Jan 2;17(1):244–54.

82. Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. Mol Cell. 2002 Feb;9(2):279–89.

83. Sekiya T, Muthurajan UM, Luger K, Tulin AV, Zaret KS. Nucleosome-binding affinity as a primary determinant of the nuclear mobility of the pioneer transcription factor FoxA. Genes Dev. 2009 Apr 1;23(7):804–9.

84. Fernandez Garcia M, Moore CD, Schulz KN, Alberto O, Donague G, Harrison MM, et al. Structural Features of Transcription Factors Associating with Nucleosome Binding. Mol Cell. 2019 Sep 5;75(5):921-932.e6.

85. Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M, Zaret KS. Pioneer Transcription Factors Target Partial DNA Motifs on Nucleosomes to Initiate Reprogramming. Cell. 2015 Apr 23;161(3):555–68.

86. Zhu F, Farnung L, Kaasinen E, Sahu B, Yin Y, Wei B, et al. The interaction landscape between transcription factors and the nucleosome. Nature. 2018 Oct;562(7725):76–81.

87. Donovan BT, Chen H, Jipa C, Bai L, Poirier MG. Dissociation rate compensation mechanism for budding yeast pioneer transcription factors. Formosa T, Tyler JK, Formosa T, Spies M, editors. eLife. 2019 Mar 19;8:e43008.

88. Li S, Zheng EB, Zhao L, Liu S. Nonreciprocal and Conditional Cooperativity Directs the Pioneer Activity of Pluripotency Transcription Factors. Cell Rep. 2019 Sep 3;28(10):2689-2703.e4.

89. Michael AK, Grand RS, Isbel L, Cavadini S, Kozicka Z, Kempf G, et al. Mechanisms of OCT4-SOX2 motif readout on nucleosomes. Science. 2020 Jun 26;368(6498):1460–5.

90. Narlikar GJ, Sundaramoorthy R, Owen-Hughes T. Mechanisms and Functions of ATP-Dependent Chromatin-Remodeling Enzymes. Cell. 2013 Aug 1;154(3):490–503.

91. Clapier CR, Iwasa J, Cairns BR, Peterson CL. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. Nat Rev Mol Cell Biol. 2017 Jul;18(7):407–22.

92. Bartholomew B. Regulating the Chromatin Landscape: Structural and Mechanistic Perspectives. Annu Rev Biochem. 2014 Jun 2;83(1):671–96.

93. Krietenstein N, Wal M, Watanabe S, Park B, Peterson CL, Pugh BF, et al. Genomic nucleosome organization reconstituted with pure proteins. Cell. 2016 Oct 20;167(3):709-721.e12.

94. Yen K, Vinayachandran V, Batta K, Koerber RT, Pugh BF. Genome-wide Nucleosome Specificity and Directionality of Chromatin Remodelers. Cell. 2012 Jun 22;149(7):1461–73.

95. Boeger H, Griesenbeck J, Strattan JS, Kornberg RD. Nucleosomes Unfold Completely at a Transcriptionally Active Promoter. Mol Cell. 2003 Jun 1;11(6):1587–98.

96. Klein-Brill A, Joseph-Strauss D, Appleboim A, Friedman N. Dynamics of Chromatin and Transcription during Transient Depletion of the RSC Chromatin Remodeling Complex. Cell Rep. 2019 Jan 2;26(1):279-292.e5.

97. Gkikopoulos T, Schofield P, Singh V, Pinskaya M, Mellor J, Smolle M, et al. A Role for Snf2-Related Nucleosome-Spacing Enzymes in Genome-Wide Nucleosome Organization. Science. 2011 Sep 23;333(6050):1758–60.

98. Whitehouse I, Rando OJ, Delrow J, Tsukiyama T. Chromatin remodelling at promoters suppresses antisense transcription. Nature. 2007 Dec;450(7172):1031–5.

99. Zentner GE, Tsukiyama T, Henikoff S. ISWI and CHD Chromatin Remodelers Bind Promoters but Act in Gene Bodies. PLOS Genet. 2013 Feb 28;9(2):e1003317.

100. Mizuguchi G, Shen X, Landry J, Wu W-H, Sen S, Wu C. ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1 Chromatin Remodeling Complex. Science. 2004 Jan 16;303(5656):343–8.

101. Krogan NJ, Keogh M-C, Datta N, Sawa C, Ryan OW, Ding H, et al. A Snf2 Family ATPase Complex Required for Recruitment of the Histone H2A Variant Htz1. Mol Cell. 2003 Dec 1;12(6):1565–76.

102. Lee Y, Park D, Iyer VR. The ATP-dependent chromatin remodeler Chd1 is recruited by transcription elongation factors and maintains H3K4me3/H3K36me3 domains at actively transcribed and spliced genes. Nucleic Acids Res. 2017 Jul 7;45(12):7180–90.

103. Kobayashi W, Kurumizaka H. Structural transition of the nucleosome during chromatin remodeling and transcription. Curr Opin Struct Biol. 2019 Dec 1;59:107–14.

104. Ye Y, Wu H, Chen K, Clapier CR, Verma N, Zhang W, et al. Structure of the RSC complex bound to the nucleosome. Science [Internet]. 2019 Oct 31 [cited 2020 Sep 3]; Available from: https://science.sciencemag.org/content/early/2019/10/30/science.aay0033

105. Farnung L, Vos SM, Wigge C, Cramer P. Nucleosome–Chd1 structure and implications for chromatin remodelling. Nature. 2017 Oct;550(7677):539–42.

106. Sabantsev A, Levendosky RF, Zhuang X, Bowman GD, Deindl S. Direct observation of coordinated DNA movements on the nucleosome during chromatin remodelling. Nat Commun. 2019 Apr 12;10(1):1720.

107. Zhong Y, Paudel BP, Ryan DP, Low JKK, Franck C, Patel K, et al. CHD4 slides nucleosomes by decoupling entry- and exit-side DNA translocation. Nat Commun. 2020 Mar 23;11(1):1519.

108. Zhou CY, Johnson SL, Gamarra NI, Narlikar GJ. Mechanisms of ATP-Dependent Chromatin Remodeling Motors. Annu Rev Biophys. 2016 Jul 5;45(1):153–81.

109. Kubik S, O'Duibhir E, de Jonge WJ, Mattarocci S, Albert B, Falcone J-L, et al. Sequence-Directed Action of RSC Remodeler and General Regulatory Factors Modulates +1 Nucleosome Position to Facilitate Transcription. Mol Cell. 2018 Jul 5;71(1):89-102.e5.

110. Challal D, Barucco M, Kubik S, Feuerbach F, Candelli T, Geoffroy H, et al. General Regulatory Factors Control the Fidelity of Transcription by Restricting Non-coding and Ectopic Initiation. Mol Cell. 2018 Dec 20;72(6):955-969.e7.

111. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell. 2007 Nov 30;131(5):861–72.

112. Kuo M-H, Allis CD. In Vivo Cross-Linking and Immunoprecipitation for Studying Dynamic Protein:DNA Associations in a Chromatin Environment. Methods. 1999 Nov 1;19(3):425–33.

113. Furey TS. ChIP–seq and beyond: new and improved methodologies to detect and characterize protein–DNA interactions. Nat Rev Genet. 2012 Dec;13(12):840–52.

114. Furlan-Magaril M, Rincón-Arano H, Recillas-Targa F. Sequential Chromatin Immunoprecipitation Protocol: ChIP-reChIP. DNA-Protein Interact. 2009;253–66.

115. Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012 Sep;489(7414):57–74.

116. Qu H, Fang X. A Brief Review on the Human Encyclopedia of DNA Elements (ENCODE) Project. Genomics Proteomics Bioinformatics. 2013 Jun 1;11(3):135–41.

117. Allan J, Fraser RM, Owen-Hughes T, Keszenman-Pereyra D. Micrococcal Nuclease Does Not Substantially Bias Nucleosome Mapping. J Mol Biol. 2012 Mar 30;417–135(3):152–64.

118. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol. 2015;109(1):21.29.1-21.29.9.

119. Yan F, Powell DR, Curtis DJ, Wong NC. From reads to insight: a hitchhiker's guide to ATAC-seq data analysis. Genome Biol. 2020 Feb 3;21(1):22.

120. Skene PJ, Henikoff S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. eLife [Internet]. [cited 2021 Apr 27];6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5310842/

121. Bosch PS, Pepperl J, Basler K. Anchor Away – A Fast, Reliable and Reversible Technique To Inhibit Proteins in Drosophila melanogaster. G3 GenesGenomesGenetics. 2020 Mar 26;10(5):1745–52.

122. Haruki H, Nishikawa J, Laemmli UK. The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. Mol Cell. 2008 Sep 26;31(6):925–32.

123. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing Chromosome Conformation. Science. 2002 Feb 15;295(5558):1306–11.

124. McCord RP, Kaplan N, Giorgetti L. Chromosome Conformation Capture and Beyond: Toward an Integrative View of Chromosome Structure and Function. Mol Cell. 2020 Feb 20;77(4):688–708.

125. Yu M, Ren B. The Three-Dimensional Organization of Mammalian Genomes. Annu Rev Cell Dev Biol. 2017 Oct 6;33(1):265–89.

126. Wu X, Mao S, Ying Y, Krueger CJ, Chen AK. Progress and Challenges for Live-cell Imaging of Genomic Loci Using CRISPR-based Platforms. Genomics Proteomics Bioinformatics. 2019 Apr 1;17(2):119–28.

127. Birk UJ. Super-Resolution Microscopy of Chromatin. Genes [Internet]. 2019 Jun 28 [cited 2021 Apr 29];10(7). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6678334/

128. Xie L, Dong P, Chen X, Hsieh T-HS, Banala S, De Marzio M, et al. 3D ATAC-PALM: super-resolution imaging of the accessible genome. Nat Methods. 2020 Apr;17(4):430–6.

129. Beliveau BJ, Boettiger AN, Avendaño MS, Jungmann R, McCole RB, Joyce EF, et al. Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes. Nat Commun. 2015 May 12;6(1):7147.

130. Schnitzbauer J, Strauss MT, Schlichthaerle T, Schueder F, Jungmann R. Super-resolution microscopy with DNA-PAINT. Nat Protoc. 2017 Jun;12(6):1198–228.

131. Liu Z, Tjian R. Visualizing transcription factor dynamics in living cells. J Cell Biol. 2018 Jan 29;217(4):1181–91.

132. Hellman LM, Fried MG. Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. Nat Protoc. 2007 Aug;2(8):1849–61.

133. Jolma A, Yin Y, Nitta KR, Dave K, Popov A, Taipale M, et al. DNA-dependent formation of transcription factor pairs alters their binding specificity. Nature. 2015 Nov;527(7578):384–8.

134. Riley TR, Slattery M, Abe N, Rastogi C, Liu D, Mann RS, et al. SELEX-seq: A Method for Characterizing the Complete Repertoire of Binding Site Preferences for Transcription Factor Complexes. In: Graba Y, Rezsohazy R, editors. Hox Genes: Methods and Protocols [Internet]. New York, NY: Springer; 2014 [cited 2021 Apr 30]. p. 255–78. (Methods in Molecular Biology). Available from: https://doi.org/10.1007/978-1-4939-1242-1_16

135. Tuerk C, Gold L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science. 1990 Aug 3;249(4968):505.

136. Jolma A, Kivioja T, Toivonen J, Cheng L, Wei G, Enge M, et al. Multiplexed massively parallel SELEX for characterization of human transcription factor binding specificities. Genome Res. 2010 Jan 6;20(6):861–73.

137. Axelrod D. Total Internal Reflection Fluorescence Microscopy for Single-Molecule Studies. In: Roberts GCK, editor. Encyclopedia of Biophysics [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 2623–31. Available from: https://doi.org/10.1007/978-3-642-16712-6_479

138. Mivelaz M, Cao A-M, Kubik S, Zencir S, Hovius R, Boichenko I, et al. Chromatin Fiber Invasion and Nucleosome Displacement by the Rap1 Transcription Factor. Mol Cell. 2020 Feb 6;77(3):488-500.e9.

139. Luger K, Rechsteiner TJ, Richmond TJ. Preparation of nucleosome core particle from recombinant histones. In: Methods in Enzymology [Internet]. Academic Press; 1999 [cited 2019 Aug 22]. p. 3–19. (Chromatin; vol. 304). Available from: http://www.sciencedirect.com/science/article/pii/S0076687999040033

140. Dobson CM. Biophysical Techniques in Structural Biology. Annu Rev Biochem. 2019 Jun 20;88(1):25–33.

141. Luger K, Dechassa ML, Tremethick DJ. New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? Nat Rev Mol Cell Biol. 2012 Jul;13(7):436–47.

142. Tan S, Davey CA. Nucleosome structural studies. Curr Opin Struct Biol. 2011 Feb 1;21(1):128–36.

143. Shoemaker SC, Ando N. X-rays in the Cryo-Electron Microscopy Era: Structural Biology's Dynamic Future. Biochemistry. 2018 Jan 23;57(3):277–85.

144. Ognjenović J, Grisshammer R, Subramaniam S. Frontiers in Cryo Electron Microscopy of Complex Macromolecular Assemblies. Annu Rev Biomed Eng. 2019 Jun 4;21(1):395–415.

145. Liu Y, Zhou K, Zhang N, Wei H, Tan YZ, Zhang Z, et al. FACT caught in the act of manipulating the nucleosome. Nature. 2020 Jan;577(7790):426–31.

146. Sundaramoorthy R, Hughes AL, Singh V, Wiechens N, Ryan DP, El-Mkami H, et al. Structural reorganization of the chromatin remodeling enzyme Chd1 upon engagement with nucleosomes. Formosa T, editor. eLife. 2017 Mar 23;6:e22510.

147. Braitbard M, Schneidman-Duhovny D, Kalisman N. Integrative Structure Modeling: Overview and Assessment. Annu Rev Biochem. 2019;88(1):113–35.

148. Braslavsky SE, Fron E, Rodríguez HB, Román ES, Scholes GD, Schweitzer G, et al. Pitfalls and limitations in the practical use of Förster's theory of resonance energy transfer. Photochem Photobiol Sci. 2008;7(12):1444–8.

149. Lerner E, Barth A, Hendrix J, Ambrose B, Birkedal V, Blanchard SC, et al. FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. Boudker O, editor. eLife. 2021 Mar 29;10:e60416.

150. Allshire RC, Madhani HD. Ten principles of heterochromatin formation and function. Nat Rev Mol Cell Biol. 2018 Apr;19(4):229–44.

151. Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T. Hyperdynamic Plasticity of Chromatin Proteins in Pluripotent Embryonic Stem Cells. Dev Cell. 2006 Jan;10(1):105–16.

152. Janssen A, Colmenares SU, Karpen GH. Heterochromatin: Guardian of the Genome. Annu Rev Cell Dev Biol. 2018;34(1):265–88.

153. Chen ES, Zhang K, Nicolas E, Cam HP, Zofall M, Grewal SIS. Cell cycle control of centromeric repeat transcription and heterochromatin assembly. Nature. 2008 Feb;451(7179):734–7.

154. Lu J, Gilbert DM. Proliferation-dependent and cell cycle–regulated transcription of mouse pericentric heterochromatin. J Cell Biol. 2007 Nov 5;179(3):411–21.

155. Smith KP, Lawrence JB. Nuclear Compartmentalization. In: Lennarz WJ, Lane MD, editors. Encyclopedia of Biological Chemistry (Second Edition) [Internet]. Waltham: Academic

Press; 2013 [cited 2021 Mar 9]. p. 296–301. Available from: https://www.sciencedirect.com/science/article/pii/B978012378630200476X

156. Kilic S, Bachmann AL, Bryan LC, Fierz B. Multivalency governs HP1α association dynamics with the silent chromatin state. Nat Commun. 2015 Jun 18;6:7313.

157. Cheutin T, Gorski SA, May KM, Singh PB, Misteli T. In Vivo Dynamics of Swi6 in Yeast: Evidence for a Stochastic Model of Heterochromatin. Mol Cell Biol. 2004 Apr 15;24(8):3157–67.

158. Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, et al. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. Nature. 2017 Jul;547(7662):236–40.

159. Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH. Phase separation drives heterochromatin domain formation. Nature. 2017 Jul;547(7662):241–5.

160. van Steensel B, Belmont AS. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. Cell. 2017 May 18;169(5):780–91.

161. Akhtar A, Gasser SM. The nuclear envelope and transcriptional control. Nat Rev Genet. 2007 Jul;8(7):507–17.

162. Talbert PB, Henikoff S. A reexamination of spreading of position-effect variegation in the white-roughest region of Drosophila melanogaster. Genetics. 2000 Jan;154(1):259–72.

163. Gross DS, Garrard WT. Nuclease hypersensitive sites in chromatin. Annu Rev Biochem. 1988 Jun 1;57(1):159–97.

164. Wang Z, Zang C, Cui K, Schones DE, Barski A, Peng W, et al. Genome-wide Mapping of HATs and HDACs Reveals Distinct Functions in Active and Inactive Genes. Cell. 2009 Sep 4;138(5):1019–31.

165. Wang Z, Schones DE, Zhao K. Characterization of human epigenomes. Curr Opin Genet Dev. 2009 Apr 1;19(2):127–34.

166. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res. 2011 Mar;21(3):381–95.

167. Eskeland R, Freyer E, Leeb M, Wutz A, Bickmore WA. Histone acetylation and the maintenance of chromatin compaction by Polycomb repressive complexes. Cold Spring Harb Symp Quant Biol. 2010;75:71–8.

168. Zhang R, Erler J, Langowski J. Histone Acetylation Regulates Chromatin Accessibility: Role of H4K16 in Inter-nucleosome Interaction. Biophys J. 2017 Feb 7;112(3):450–9.

169. Shultzaberger RK, Chen Z, Lewis KA, Schneider TD. Anatomy of Escherichia coli sigma70 promoters. Nucleic Acids Res. 2007;35(3):771–88.

170. [edited by] Jocelyn E. Krebs ESG Stephen T Kilpatrick. Lewin's genes X [Internet]. Tenth edition. Sudbury, Mass. : Jones and Bartlett, [2011] ©2011; 2011. Available from: https://search.library.wisc.edu/catalog/9910124114202121

171. Haberle V, Stark A. Eukaryotic core promoters and the functional basis of transcription initiation. Nat Rev Mol Cell Biol. 2018 Oct;19(10):621–37.

172. FitzGerald PC, Sturgill D, Shyakhtenko A, Oliver B, Vinson C. Comparative genomics of Drosophila and human core promoters. Genome Biol. 2006 Jul 7;7(7):R53.

173. Buratowski S, Hahn S, Guarente L, Sharp PA. Five intermediate complexes in transcription initiation by RNA polymerase II. Cell. 1989 Feb 24;56(4):549–61.

174. Boeger H, Bushnell DA, Davis R, Griesenbeck J, Lorch Y, Strattan JS, et al. Structural basis of eukaryotic gene transcription. FEBS Lett. 2005 Feb 7;579(4):899–903.

175. Patel AB, Greber BJ, Nogales E. Recent insights into the structure of TFIID, its assembly, and its binding to core promoter. Curr Opin Struct Biol. 2020 Apr 1;61:17–24.

176. Jiang C, Pugh BF. Nucleosome positioning and gene regulation: advances through genomics. Nat Rev Genet. 2009 Mar;10(3):161–72.

177. Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, et al. H3.3/H2A.Z double variant–containing nucleosomes mark "nucleosome-free regions" of active promoters and other regulatory regions. Nat Genet. 2009 Aug;41(8):941–5.

178. Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, et al. Genome-wide analysis of mammalian promoter architecture and evolution. Nat Genet. 2006 Jun;38(6):626–35.

179. Field A, Adelman K. Evaluating Enhancer Function and Transcription. Annu Rev Biochem. 2020 Jun 20;89(1):213–34.

180. Shlyueva D, Stampfel G, Stark A. Transcriptional enhancers: from properties to genome-wide predictions. Nat Rev Genet. 2014 Apr;15(4):272–86.

181. Schoenfelder S, Fraser P. Long-range enhancer–promoter contacts in gene expression control. Nat Rev Genet. 2019 Aug;20(8):437–55.

182. Kim S, Yu N-K, Kaang B-K. CTCF as a multifunctional protein in genome regulation and gene expression. Exp Mol Med. 2015 Jun;47(6):e166–e166.

183. Arnold PR, Wells AD, Li XC. Diversity and Emerging Roles of Enhancer RNA in Regulation of Gene Expression and Cell Fate. Front Cell Dev Biol [Internet]. 2020 [cited 2021 Mar 30];7. Available from: https://www.frontiersin.org/articles/10.3389/fcell.2019.00377/full

184. Giaimo BD, Ferrante F, Herchenröther A, Hake SB, Borggrefe T. The histone variant H2A.Z in gene regulation. Epigenetics Chromatin. 2019 Jun 14;12(1):37.

185. Knight B, Kubik S, Ghosh B, Bruzzone MJ, Geertz M, Martin V, et al. Two distinct promoter architectures centered on dynamic nucleosomes control ribosomal protein gene transcription. Genes Dev. 2014 Aug 1;28(15):1695–709.

186. Lieb JD, Liu X, Botstein D, Brown PO. Promoter-specific binding of Rap1 revealed by genome-wide maps of protein–DNA association. Nat Genet. 2001 Aug;28(4):327–34.

187. Wu ACK, Patel H, Chia M, Moretto F, Frith D, Snijders AP, et al. Repression of Divergent Noncoding Transcription by a Sequence-Specific Transcription Factor. Mol Cell. 2018 Dec 20;72(6):942-954.e7.

188. Rhee HS, Pugh BF. Comprehensive Genome-wide Protein-DNA Interactions Detected at Single Nucleotide Resolution. Cell. 2011 Dec 9;147(6):1408–19.

189. Rossetti L, Cacchione S, De Menna A, Chapman L, Rhodes D, Savino M. Specific interactions of the telomeric protein rap1p with nucleosomal binding sites11Edited by A. Klug. J Mol Biol. 2001 Mar 9;306(5):903–13.

190. Feldmann EA, Galletto R. The DNA-binding domain of yeast Rap1 interacts with double-stranded DNA in multiple binding modes. Biochemistry. 2014 Dec 9;53(48):7471–83.

191. Feldmann EA, Koc KN, Galletto R. Alternative arrangements of telomeric recognition sites regulate the binding mode of the DNA-binding domain of yeast Rap1. Biophys Chem. 2015 Mar;198:1–8.

192. Matot B, Le Bihan Y-V, Lescasse R, Pérez J, Miron S, David G, et al. The orientation of the C-terminal domain of the Saccharomyces cerevisiae Rap1 protein is determined by its binding to DNA. Nucleic Acids Res. 2012 Apr;40(7):3197–207.

193. Koerber RT, Rhee HS, Jiang C, Pugh BF. Interaction of Transcriptional Regulators with Specific Nucleosomes across the Saccharomyces Genome. Mol Cell. 2009 Sep 24;35(6):889–902.

194. Reja R, Vinayachandran V, Ghosh S, Pugh BF. Molecular mechanisms of ribosomal protein gene coregulation. Genes Dev. 2015 Sep 15;29(18):1942–54.

195. Mizuno T, Kishimoto T, Shinzato T, Haw R, Chambers A, Wood J, et al. Role of the N-terminal region of Rap1p in the transcriptional activation of glycolytic genes in Saccharomyces cerevisiae. Yeast. 2004;21(10):851–66.

196. Leung CCY, Glover JM. BRCT domains. Cell Cycle. 2011 Aug 1;10(15):2461–70.

197. Song S, Perez JV, Svitko W, Ricketts MD, Dean E, Schultz D, et al. Rap1-mediated nucleosome displacement can regulate gene expression in senescent cells without impacting the pace of senescence. Aging Cell. 2020;19(1):e13061.

198. Freeman K, Gwadz M, Shore D. Molecular and Genetic Analysis of the Toxic Effect of Rap1 Overexpression in Yeast. Genetics. 1995 Dec;141(4):1253–62.

199. Johnson AN, Weil PA. Identification of a transcriptional activation domain in yeast repressor activator protein 1 (Rap1) using an altered DNA-binding specificity variant. J Biol Chem. 2017 Jul 4;292(14):5705–23.

200. Rai R, Chen Y, Lei M, Chang S. TRF2-RAP1 is required to protect telomeres from engaging in homologous recombination-mediated deletions and fusions. Nat Commun. 2016 Mar 4;7(1):10881.

201. Bryan LC, Weilandt DR, Bachmann AL, Kilic S, Lechner CC, Odermatt PD, et al. Single-molecule kinetic analysis of HP1-chromatin binding reveals a dynamic network of histone modification and DNA interactions. Nucleic Acids Res. 2017 Oct 13;45(18):10504–17.

202. Mivelaz M, Fierz B. Observing protein interaction dynamics to chemically defined chromatin fibers by colocalization single-molecule fluorescence microscopy. Methods. 2020 Dec 1;184:112–24.

203. Dyer PN, Edayathumangalam RS, White CL, Bao Y, Chakravarthy S, Muthurajan UM, et al. Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods Enzymol. 2004;375:23–44.

204. Ploetz E, Lerner E, Husada F, Roelfs M, Chung S, Hohlbein J, et al. Förster resonance energy transfer and protein-induced fluorescence enhancement as synergetic multi-scale molecular rulers. Sci Rep [Internet]. 2016 Sep 19 [cited 2021 Jun 12];6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5027553/

205. Kalinin S, Peulen T, Sindbert S, Rothwell PJ, Berger S, Restle T, et al. A toolkit and benchmark study for FRET-restrained high-precision structural modeling. Nat Methods. 2012 Dec 1;9(12):1218–25.

206. Banerjee DR, Deckard CE, Elinski MB, Buzbee ML, Wang WW, Batteas JD, et al. Plug-and-Play Approach for Preparing Chromatin Containing Site-Specific DNA Modifications: The Influence of Chromatin Structure on Base Excision Repair. J Am Chem Soc. 2018 Jul 5;140(26):8260–7.

207. Chen J, Zhang Z, Li L, Chen B-C, Revyakin A, Hajj B, et al. Single-Molecule Dynamics of Enhanceosome Assembly in Embryonic Stem Cells. Cell. 2014 Mar 13;156(6):1274–85.

208. Slutsky M, Mirny LA. Kinetics of Protein-DNA Interaction: Facilitated Target Location in Sequence-Dependent Potential. Biophys J. 2004 Dec 1;87(6):4021–35.

209. Swift J, Coruzzi G. A Matter of Time - How Transient Transcription Factor Interactions Create Dynamic Gene Regulatory Networks. Biochim Biophys Acta. 2017 Jan;1860(1):75–83.

210. Hettich J, Gebhardt JCM. Transcription factor target site search and gene regulation in a background of unspecific binding sites. J Theor Biol. 2018 Oct 7;454:91–101.

211. Murugan R. Theory of Site-Specific DNA-Protein Interactions in the Presence of Nucleosome Roadblocks. Biophys J. 2018 Jun 5;114(11):2516–29.

212. Goldstein I, Baek S, Presman DM, Paakinaho V, Swinstead EE, Hager GL. Transcription factor assisted loading and enhancer dynamics dictate the hepatic fasting response. Genome Res. 2017 Mar;27(3):427–39.

213. Swinstead EE, Miranda TB, Paakinaho V, Baek S, Goldstein I, Hawkins M, et al. Steroid Receptors Reprogram FoxA1 Occupancy through Dynamic Chromatin Transitions. Cell. 2016 Apr 21;165(3):593–605.

214. Callegari A, Sieben C, Benke A, Suter DM, Fierz B, Mazza D, et al. Single-molecule dynamics and genome-wide transcriptomics reveal that NF-kB (p65)-DNA binding times can be decoupled from transcriptional activation. PLOS Genet. 2019 Jan 17;15(1):e1007891.

215. Hipp L, Beer J, Kuchler O, Reisser M, Sinske D, Michaelis J, et al. Single-molecule imaging of the transcription factor SRF reveals prolonged chromatin-binding kinetics upon cell stimulation. Proc Natl Acad Sci. 2019 Jan 15;116(3):880–9.

216. Presman DM, Ball DA, Paakinaho V, Grimm JB, Lavis LD, Karpova TS, et al. Quantifying Transcription Factor Binding Dynamics at the Single-molecule Level in Live Cells. Methods San Diego Calif. 2017 Jul 1;123:76–88.

217. Garcia DA, Fettweis G, Presman DM, Paakinaho V, Jarzynski C, Upadhyaya A, et al. Power-law behavior of transcription factor dynamics at the single-molecule level implies a continuum affinity model. Nucleic Acids Res [Internet]. 2021 Feb 17 [cited 2021 Jun 28];(gkab072). Available from: https://doi.org/10.1093/nar/gkab072

218. Siersbæk R, Scabia V, Nagarajan S, Chernukhin I, Papachristou EK, Broome R, et al. IL6/STAT3 Signaling Hijacks Estrogen Receptor α Enhancers to Drive Breast Cancer Metastasis. Cancer Cell. 2020 Sep 14;38(3):412-423.e9.

219. He HH, Meyer CA, Chen MW, Jordan VC, Brown M, Liu XS. Differential DNase I hypersensitivity reveals factor-dependent chromatin dynamics. Genome Res. 2012 Jan 6;22(6):1015–25.

220. Braastad CD, Han Z, Hendrickson EA. Constitutive DNase I Hypersensitivity of p53-Regulated Promoters *. J Biol Chem. 2003 Mar 7;278(10):8261–8.

221. Funk CC, Casella AM, Jung S, Richards MA, Rodriguez A, Shannon P, et al. Atlas of Transcription Factor Binding Sites from ENCODE DNase Hypersensitivity Data across 27 Tissue Types. Cell Rep. 2020 Aug 18;32(7):108029.

222. Polach KJ, Widom J. Mechanism of Protein Access to Specific DNA Sequences in Chromatin: A Dynamic Equilibrium Model for Gene Regulation. J Mol Biol. 1995 Nov 24;254(2):130–49.

223. Huertas J, MacCarthy CM, Schöler HR, Cojocaru V. Nucleosomal DNA Dynamics Mediate Oct4 Pioneer Factor Binding. Biophys J. 2020 May 5;118(9):2280–96.

224. Eslami-Mossallam B, Schiessel H, van Noort J. Nucleosome dynamics: Sequence matters. Adv Colloid Interface Sci. 2016 Jun 1;232:101–13.

225. Hall DB, Wade JT, Struhl K. An HMG Protein, Hmo1, Associates with Promoters of Many Ribosomal Protein Genes and throughout the rRNA Gene Locus in Saccharomyces cerevisiae. Mol Cell Biol. 2006 Jan 5;26(9):3672–9.

226. Bowman GD, McKnight JN. Sequence-specific targeting of chromatin remodelers organizes precisely-positioned nucleosomes throughout the genome. BioEssays News Rev Mol Cell Dev Biol. 2017 Jan;39(1):1–8.

227. Dann GP, Liszczak GP, Bagert JD, Müller MM, Nguyen UTT, Wojcik F, et al. ISWI chromatin remodellers sense nucleosome modifications to determine substrate preference. Nature. 2017 Aug;548(7669):607–11.

228. Badis G, Chan ET, van Bakel H, Pena-Castillo L, Tillo D, Tsui K, et al. A Library of Yeast Transcription Factor Motifs Reveals a Widespread Function for Rsc3 in Targeting Nucleosome Exclusion at Promoters. Mol Cell. 2008 Dec 26;32(6):878–87.

229. Cairns BR, Lorch Y, Li Y, Zhang M, Lacomis L, Erdjument-Bromage H, et al. RSC, an Essential, Abundant Chromatin-Remodeling Complex. Cell. 1996 Dec 27;87(7):1249–60.

230. Zhang Q, Chakravarty S, Ghersi D, Zeng L, Plotnikov AN, Sanchez R, et al. Biochemical Profiling of Histone Binding Selectivity of the Yeast Bromodomain Family. PLOS ONE. 2010 Jan 26;5(1):e8903.

231. Cairns BR, Schlichter A, Erdjument-Bromage H, Tempst P, Kornberg RD, Winston F. Two Functionally Distinct Forms of the RSC Nucleosome-Remodeling Complex, Containing Essential AT Hook, BAH, and Bromodomains. Mol Cell. 1999 Nov 1;4(5):715–23.

232. Clapier CR, Kasten MM, Parnell TJ, Viswanathan R, Szerlong H, Sirinakis G, et al. Regulation of DNA Translocation Efficiency Within the Chromatin Remodeler RSC/Sth1 Potentiates Nucleosome Sliding and Ejection. Mol Cell. 2016 May 5;62(3):453–61.

233. Lorch Y, Zhang M, Kornberg RD. Histone Octamer Transfer by a Chromatin-Remodeling Complex. Cell. 1999 Feb 5;96(3):389–92.

234. Patel AB, Moore CM, Greber BJ, Luo J, Zukin SA, Ranish J, et al. Architecture of the chromatin remodeler RSC and insights into its nucleosome engagement. Wolberger C, editor. eLife. 2019 Dec 30;8:e54449.

235. Baker RW, Reimer JM, Carman PJ, Turegun B, Arakawa T, Dominguez R, et al. Structural insights into assembly and function of the RSC chromatin remodeling complex. Nat Struct Mol Biol. 2021 Jan;28(1):71–80.

236. Wittmeyer J, Saha A, Cairns B. DNA Translocation and Nucleosome Remodeling Assays by the RSC Chromatin Remodeling Complex. In: Methods in Enzymology [Internet]. Academic Press; 2003 [cited 2021 May 27]. p. 322–43. (Chromatin and Chromatin Remodeling Enzymes, Part C; vol. 377). Available from: https://www.sciencedirect.com/science/article/pii/S0076687903770207

237. Kurat CF, Yeeles JTP, Patel H, Early A, Diffley JFX. Chromatin Controls DNA Replication Origin Selection, Lagging-Strand Synthesis, and Replication Fork Rates. Mol Cell. 2017 Jan 5;65(1):117–30.

238. Prasad R, D'Arcy S, Hada A, Luger K, Bartholomew B. Coordinated Action of Nap1 and RSC in Disassembly of Tandem Nucleosomes. Mol Cell Biol. 2016 Aug 12;36(17):2262–71.

239. Lorch Y, Maier-Davis B, Kornberg RD. Chromatin remodeling by nucleosome disassembly in vitro. Proc Natl Acad Sci. 2006 Feb 28;103(9):3090–3.

240. Li M, Hada A, Sen P, Olufemi L, Hall MA, Smith BY, et al. Dynamic regulation of transcription factors by nucleosome remodeling. eLife [Internet]. [cited 2021 May 31];4. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4456607/

241. He S, Wu Z, Tian Y, Yu Z, Yu J, Wang X, et al. Structure of nucleosome-bound human BAF complex. Science. 2020 Feb 21;367(6480):875–81.

242. Dechassa ML, Zhang B, Horowitz-Scherer R, Persinger J, Woodcock CL, Peterson CL, et al. Architecture of the SWI/SNF-Nucleosome Complex. Mol Cell Biol. 2008 Oct;28(19):6010–21.

243. Bartholomew B. ISWI Chromatin Remodeling: One primary actor or a coordinated effort? Curr Opin Struct Biol. 2014 Feb;24:150–5.

244. Lieleg C, Ketterer P, Nuebler J, Ludwigsen J, Gerland U, Dietz H, et al. Nucleosome spacing generated by ISWI and CHD1 remodelers is constant regardless of nucleosome density. Mol Cell Biol. 2015 May;35(9):1588–605.

245. Simic R, Lindstrom DL, Tran HG, Roinick KL, Costa PJ, Johnson AD, et al. Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes. EMBO J. 2003 Apr 15;22(8):1846–56.

246. Qiu Y, Levendosky RF, Chakravarthy S, Patel A, Bowman GD, Myong S. The Chd1 Chromatin Remodeler Shifts Nucleosomal DNA Bidirectionally as a Monomer. Mol Cell. 2017 Oct 5;68(1):76-88.e6.

247. Deindl S, Hwang WL, Hota SK, Blosser TR, Prasad P, Bartholomew B, et al. ISWI Remodelers Slide Nucleosomes with Coordinated Multi-Base-Pair Entry Steps and Single-Base-Pair Exit Steps. Cell. 2013 Jan 31;152(3):442–52.

248. McKnight JN, Jenkins KR, Nodelman IM, Escobar T, Bowman GD. Extranucleosomal DNA binding directs nucleosome sliding by Chd1. Mol Cell Biol. 2011 Dec;31(23):4746–59.

249. Schwechheimer C, Rönicke F, Schepers U, Wagenknecht H-A. A new structure–activity relationship for cyanine dyes to improve photostability and fluorescence properties for live cell imaging. Chem Sci. 2018 Aug 8;9(31):6557–63.

250. Li H, Vaughan JC. Switchable Fluorophores for Single-Molecule Localization Microscopy. Chem Rev. 2018 Sep 26;118(18):9412–54.

251. Brouwer I, Lenstra TL. Visualizing transcription: key to understanding gene expression dynamics. Curr Opin Chem Biol. 2019 Aug 1;51:122–9.

252. Lenstra TL, Rodriguez J, Chen H, Larson DR. Transcription Dynamics in Living Cells. Annu Rev Biophys. 2016 Jul 5;45(1):25–47.

253. Rullan M, Benzinger D, Schmidt GW, Milias-Argeitis A, Khammash M. An Optogenetic Platform for Real-Time, Single-Cell Interrogation of Stochastic Transcriptional Regulation. Mol Cell. 2018 May 17;70(4):745-756.e6.

254. Paakinaho V, Presman DM, Ball DA, Johnson TA, Schiltz RL, Levitt P, et al. Single-molecule analysis of steroid receptor and cofactor action in living cells. Nat Commun. 2017 Jun 21;8(1):15896.

255. Senecal A, Munsky B, Proux F, Ly N, Braye FE, Zimmer C, et al. Transcription Factors Modulate c-Fos Transcriptional Bursts. Cell Rep. 2014 Jul 10;8(1):75–83.

256. Stavreva DA, Garcia DA, Fettweis G, Gudla PR, Zaki GF, Soni V, et al. Transcriptional Bursting and Co-bursting Regulation by Steroid Hormone Release Pattern and Transcription Factor Mobility. Mol Cell. 2019 Sep 19;75(6):1161-1177.e11.

8 Materials and Methods

This Section is taken and adapted from (138,202).

8.1 Expression and purification of Rap1-Halo

The Strep-MBP-TEV-Rap1-Halo construct (**Fig 14a**) was cloned into pACEBac1 (Geneva Biotech), and baculovirus particles were generated using the Geneva Biotech system per manufacturer's instructions.

For Rap1 expression, 1L cultures of Sf9 cells were grown to 2 - 2.5x10^6 cells/mL. Subsequently, the cells were infected with baculovirus, and the cultures were incubated for 3 days at 27° C before harvesting through centrifugation (1500 rcf, 4°C for 20 min). Supernatants were discarded, and pellets were resuspended in PBS, containing protease inhibitors (Roche) (10 mL PBS/L of culture), flash frozen and kept at -80°C.

For a typical purification of Rap1-Halo, 12-15 g of frozen pellets were thawed at room temperature with 36 mL of lysis buffer (200 mM KCl, 2 mM DTT, 100 mM Tris-HCl (pH 7.5), 50 mM MgOAc, 0.1% NP-40, Protease inhibitor cocktail (Roche), 1mM PMSF and 20 µL DNaseI (NEB)). Pellets were stirred with a magnetic stir bar until fully thawed and then kept on ice. The lysate was spun for 35 min at 35000 rpm at 4° C (Ti70 rotor, Beckman Coulter), and the supernatant was filtered through a 5 µm syringe filter (Millex, Millipore). The cleared lysate was loaded onto a Strep-Trap column (GE, AKTA system), pre-equilibrated with lysis buffer. The column was washed with storage buffer (200 mM KCl, 10 mM HEPES pH 7.6, 50 mM MgOAc, and 5 mM β-mercaptoethanol (βME)), and the protein was eluted with 5 x column volumes (CV) of elution buffer (storage buffer containing 2.5 mM desthiobiotin). Fractions containing Rap1 were identified by SDS-Page (**Fig 14c**), pooled and concentrated to ~ 500 µL total volume using Amicon 10k molecular weight cut-off (MWCO) centrifugal filters. The protein concentration was determined using UV spectroscopy. The MBP tag was subsequently removed by TEV protease digestion at 4°C (**Fig 14d**). For labelling, Janelia Fluor-549 HaloTag (Janelia, JF-549) was added at a protein to dye ratio of 1:1.5 followed by incubation for 1h. Labelled Rap1 was finally purified by size exclusion chromatography (SEC) using a Superose6 10/300 GL column (GE healthcare) in storage buffer using a flow rate of 0.4 mL/min (**Fig 14e**). Fractions were analysed using SDS-PAGE (**Fig 14e**), pure fractions were pooled, concentrated

(Amicon 10k MWCO filter), and protein concentrations were determined using UV spectrophotometry (at A280 and A571). Finally, labelling efficiency was calculated using the extinction coefficients for Rap1 (107'065 mol^{-1} cm^{-1}) and JF-549 (101'000 mol^{-1} cm^{-1}). Typical labelling efficiency was found to be >90%.

8.2 Expression and purification of recombinant histones

Histones were expressed and purified as described in (Kilic et al., 2015). Briefly, individual *wild-type* human histones were cloned into pet15b plasmid vectors and expressed in BL21 DE3 plysS cells. Cells were grown in LB media containing 100 µg/mL ampicillin and 35 µg/mL chloramphenicol at 37°C until the OD_{600} reached 0.6. Expression was induced by IPTG addition to a final concentration of 0.5 mM. After 3 h expression, cells were harvested by centrifugation and resuspended in lysis buffer (20 mM Tris pH 7.5, 1 mM EDTA, 200 mM NaCl, 1 mM βMe, Roche protease inhibitor) and frozen. Cells were lysed by freeze-thawing and sonication. Inclusion bodies were harvested by centrifugation. The inclusion body pellet was washed once with 7.5 mL of lysis buffer containing 1% Triton and once without. Inclusion body pellets were resolubilized in resolubilization buffer (6 M GdmCl, 20 mM Tris pH 7.5, 1 mM EDTA, 1 mM βMe) and dialyzed into urea buffer (7 M urea, 10 mM Tris, 1 mM EDTA, 0.1 M NaCl, 5 mM 1 mM βMe, pH 7.5). Histones were purified by cation exchange chromatography using a HiTrap SP HP 5 mL column (GE Healthcare). Fractions were analyzed by SDS-PAGE and pooled, followed by dialysis into water and lyophilization. Final purification was performed by preparative RP-HPLC. Purified histones were lyophilized and stored at -20°C until used for octamer refolding.

8.3 Large scale generation of recombinant plasmids

Plasmids containing recombinant DNA fragments for chromatin DNA assembly, which have been prepared previously (Kilic et al., 2018b) (**recP1**, **recP5**) or were newly generated using restriction digestion and ligation of previous fragments (**recP12** or **recP45**, **Fig 23b**), were transformed into DH5α cells (for sequence information see **Table S1**). Cells were cultured overnight in 6 L of 2xTY medium and harvested by centrifugation. For alkaline lysis, the cells were resuspended in 120 mL lysis solution I (50 mM glucose, 25 mM Tris pH 8, 10 mM EDTA). 240 mL lysis solution II (0.3 M NaOH, 1% SDS) was added and mixed by stirring. 240 mL lysis

solution III (4 M KAc, 2 M acetic acid) was added to neutralize the solution, which was left at 4° C for 15 min. After centrifugation, the supernatant was passed through Miracloth (Merck). Plasmid DNA was collected by isopropanol precipitation: 0.52 volume equivalents of isopropanol were added, followed by centrifugation at 11'000 x g for 20 min at 4° C. The DNA pellet was dissolved in TE 10/50 (10 mM Tris pH 7.5, 50 mM EDTA) in the presence of 100 units of RNAse A and digested for 2 h at 37° C. The buffer was adjusted to 2M KCl (10 mM Tris, 50 mM EDTA, and 2 M KCl) to perform SEC. The plasmid was then purified in the same buffer on an XK 50/30 column (GE Healthcare) containing a bed of 550 mL sepharose 6 Fast Flow (GE Healthcare). Eluted plasmid DNA was precipitated with isopropanol. The pellet was finally dissolved in TE 10 / 0.1 (10 mM Tris pH 7.5, 0.1 mM EDTA) and stored at -20° C.

8.4 Large scale restriction digest and purification of recombinant plasmids

Purified plasmid DNA was collected by isopropanol precipitation, and the DNA pellet was dissolved in Milli-Q H_2O. For a typical reaction, either 200 units of DraIII-HF (NEB) (for **recP12**) or 200 units of BsaI-HF (NEB) (for **recP45**) or 200 units of both DraIII-HF and BsaI-HF (NEB) (for **recP1** and **recP5**) were added to 200 pmol of plasmid DNA in 200µl 1x NEB CutSmart buffer. After 8-10 h digestion at 37°C, digestion progress was analyzed by gel electrophoresis on a 1% agarose gel (run in 1 x TBE running buffer, 100 V, for 50 min) to check completeness. If required, the digestion was pushed to completion by adding another 100 units of enzyme and incubating for a further 8-10 h at 37° C. Once the digestion was complete, 100 units of EcoRV-HF (NEB) was added and left 8-10 h at 37°C. Complete digestion was verified by electrophoresis as described above. If the digestion was not complete, an additional 50 units of the enzyme was added and left 8-10 h at 37° C. Once the digestion was complete, the desired chromatin DNA fragments were purified from the plasmid remnants through successive PEG precipitations. This involves adding 40% PEG 6000 to the digestion reactions until a final concentration of 5-6% PEG 6000 was reached. Additionally, the NaCl concentration was adjusted to 0.5 M. The sample was then spun at 20'000 x g at 4° C for 20 min. The supernatant was collected, and PEG 6000 was added to the supernatant to increase the final PEG % by increments of 0.5 %. The sample was then spun at 20'000 x g at 4° C for 20 min. This was repeated until a suitable purity was achieved. Finally, the chromatin DNA fragments were isolated using QIAquick PCR purification spin columns (Qiagen).

8.5 Oligonucleotide labelling

Fluorescently labelled oligonucleotides were generated as described in (Kilic et al., 2018b). Briefly, 5-10 nmol of single-stranded oligonucleotide, containing amino-modified C6 dT, was diluted in 25 µl 0.1 M sodium tetraborate, pH 8.5. 5 µl of a 5 mM stock of succinimidyl-ester modified fluorophore (Alexa 568, Alexa 647 or Cy3B) were added to the reaction mix and left shaking at room temperature for 4 – 8 hours. For a table enumerating all labelled oligonucleotides see **Table S1**.

Reaction progress was analyzed by RP-HPLC using a gradient from solvent A (95% 0.1M triethylammonium acetate (TEAA) pH 7, 5% ACN) to solvent B (70% 0.1M TEAA pH 7, 30% ACN) on a 3 µm 4.6x150 mm InertSustain C18 column (GL Sciences) over 20 min. More dye was added when required. For purification, the labelled DNA was ethanol precipitated (by adding 2.75 equivalents of cold ethanol, 0.3M NaOAc pH 5.2, followed by centrifugation at 20'000 x g at 4° C for 20 min) twice successively to remove the excess unconjugated dye. The DNA pellet was finally dissolved in 100 µl solvent A and purified by HPLC. The purified DNA was finally ethanol precipitated and dissolved in Milli-Q water to a concentration of 2.5 µM.

8.6 Production of labelled DNA fragments

Labelled DNA was prepared by PCR (fragments **P2, P3_S1**, **P3_S2**, **P3_S2p**, **P3_S12**, **P3_Rpl30, P3_Rpl30_S1** and **P4**, for sequences and labelling schemes, see **Table S1-2**). For a typical reaction, 96 x 50 µl PCR reactions in 1 x ThermoPol reaction buffer (NEB) were prepared using template (0.01 ng μL^{-1}), forward primer (0.250 µM), reverse primer (0.250 µM), dNTPs (0.2 mM, NEB) and Taq DNA polymerase (1.25 units, NEB). A typical program included an initial step of 12 s at 94° C, followed by 30 cycles of 12 s at 94° C, 12 s annealing at 58-65° C and 12 s extension at 72° C. Final extension was also done at 72° C for 12 s. PCR reactions were subsequently purified using QIAquick PCR purification spin columns (Qiagen).

About 0.33 nmol of PCR generated DNA (**P3_S1**, **P3_S2**, **P3_S2p**, **P3_S1S2, P3_Rpl30** and **P3_Rpl30_S1, Table S1**) was digested in 200 µl of 1 x CutSmart buffer using 100 units of BsaI-HF (NEB) and 100 units of DraIII-HF (NEB) for 8-10h at 37° C. The progress of the digestion was

analyzed on a 2% agarose gel (running conditions: 1 x TBE, 110 V for 50 min). Finally, the DNA fragments were purified using QIAquick PCR purification spin columns (Qiagen), and the concentration was determined by UV spectroscopy.

8.7 Ligation and purification of 1 x 601 DNA to biotin anchor

For the generation of nucleosome DNA for single-molecule experiments, an biotin containing anchor (**Anchor_rev, Table S2**) was annealed to its complementary strand containing a phosphorylated 5'- BsaI overhang (**P3_Anchor_fwd**, **Table S2**) and a 10-fold excess were added to 150-300 pmol (~20-40 µg) of digested PCR generated DNA (**P3_S1**, **P3_S2**, **P3_S2p** and **P3_S12)** in 100 µl 1x T4 ligase buffer (NEB). Upon complete ligation of digested DNA, excess biotin anchor was removed by PEG precipitation. Finally, the DNA fragments were purified using QIAquick PCR purification spin columns (Qiagen), and the concentration was determined by UV spectroscopy.

8.8 Mononucleosome (Nuc) formation

Nucleosomes (Nuc_S1, Nuc_S2, Nuc_S2p, Nuc_S1S2, Nuc_S1_FRET, Nuc_S2_FRET, Nuc_Rpl30, Nuc_Rpl30_S1, Nuc_Rpl30_S1_FRET, Nuc_Asy_+70 Cy3B_H2AN110C_A647_FRET_ Nuc_Sym_+70 Cy3B_H2AN110C_A647_FRET **Table S3**) were prepared following (Dyer et al., 2004). Typically, 1-5 µg of labelled and biotinylated DNA was combined with purified refolded octamers at experimentally determined ratios (1:1 to 1:2, DNA to histone octamer) in 10 µl TE (10 mM Tris-HCl pH 7.5, 1 mM EDTA) supplemented with 2 M KCl. After a 30 min incubation at room temperature, 10 µl TE was added and further incubated for 1 h. This was followed by sequential addition of 5 µl TE, 5 µl TE, and finally 70 µl TE with 1 h incubation periods between each addition to arrive at 0.2 M KCl. Samples were then spun at 20'000 x g for 10 min at 4 °C, and the supernatant was kept on ice. To determine the quality of MN assemblies, 0.8% Agarose 0.25 x TB gels were run at 90 V on ice for 90 min (**Figure S1**).

8.9 Electrophoretic mobility shift assays (EMSA)

EMSAs to determine Rap1 binding to DNA were done in single-molecule imaging buffer (IB, 50 mM HEPES pH 7.5, 130 mM KCl, 10% v/v glycerol, 0.005% v/v Tween 20, 2 mM Trolox, 3.2% w/v glucose), in the presence of 50 ng/µl poly-d(I-C) (Roche) and with 20 µl total volume. Typically, 200 nmol stocks of DNA and 3 µM stocks of Rap1-Halo were prepared and serially diluted to desired concentrations. Reactions were mixed by pipetting and left for 10 min at room temperature. Sucrose was added to a final concentration of 8%, and reactions were loaded onto 5% Polyacrylamide gels run in 0.5 x TBE at 100 V for 60 min. Images were taken using ChemiDoc MP (Biorad) (**Figure 15a**). ImageLab (Biorad) software was used for band quantification of bound and unbound fraction of DNA for densitometry quantifications. The data was analyzed in Origin (OriginLab) by non-linear curve-fitting using a sigmoidal function to determine K_d.

8.10 Convergent 3-piece convergent DNA ligation for the synthesis of 12x601 DNA

Once fragments **P12**, **P3** and **P45** have been digested and purified, test ligations are performed to optimize conditions for complete ligation, i.e. P12 with P3 and P2 with P45. These optimizations allow to reduce the amount of unligated products and simplify the PEG purification steps. Typically, an excess of the PCR generated piece is used for the first ligation between P12 and P3, i.e. 1 : 1.3 ratio of P12 : P3. This is because the difference in size between P12 and P123 is small; thus, any remaining unligated P12 is harder to purify out. However, for the ligation of P123 with P45, an excess P45 (both constructs P123 and P45 are similar in size) as P123 is more difficult to generate. P123 ligations are performed in 200 µl using optimized ratios of P12 to P3, in 1 x T4 Ligase buffer (NEB B0202) with 400 units of T4 ligase (NEB M0202). These reactions are usually left overnight at room temperature, after which a 1% agarose gel is run at 100 V for 50 minutes in 1 x TBE until all P12 has been ligated. If this is not the case, more T4 ligase can be added and more P3 if required. Upon complete P12 ligation, the P123 fragment is isolated via PEG precipitation (**Fig 23g).** A starting percentage of PEG from 5 – 6 % is recommended for fragments of this size and PEG steps of 0.5 %. Pellets containing the desired fragments are pooled and purified using QIAquick PCR purification

spin-column kits (Qiagen 28106). For the ligation between P123, P45 and biotin anchor, the following ratio is used, i.e. 1 : 1.3 of P1P2 : P3 and 5x excess of biotin anchor. Ligations are performed in 200 µl with 1 x T4 Ligase buffer (NEB B0202) with 400 units of T4 ligase (NEB M0202). These reactions are usually left overnight at room temperature, after which a 1% agarose gel is run at 100 V for 50 minutes in 1 x TBE until all P123 has been ligated. If this is not the case, more T4 ligase can be added and more P45 if required.

Upon complete ligation, the P12345 fragment is isolated via PEG precipitation (**Fig 23i**)**.** A starting percentage of PEG from 4 – 5 % is recommended for fragments of this size using PEG steps of 0.3 %. Pellets containing the desired P12345 fragments are pooled and purified using QIAquick PCR purification spin-column kits (Qiagen 28106).

8.11 Reconstitution of 12-mer chromatin fibres

Chromatin fibres (CH_S1, CH_S2, CH_601 **Table S1**) were reconstituted from singly/doubly-labelled and biotinylated 12x601 DNA and wild-type recombinantly purified human histone octamers. In a typical dialysis, 200-300 pM 12x601 DNA, 0.5-1 equivalents of MMTV DNA and reconstituted octamers (using experimentally determined DNA : octamer ratios) were added to a micro-dialysis unit (Thermo Scientific, Slide-A-Lyzer – 10’000 MWCO), then dialyzed in TE buffer (10 mM Tris pH 7.5, 0.1 mM EDTA pH 8.0) with a linear gradient from 2 M to 10 mM KCl for 16-18 h, and finally kept in TEK10 buffer (10 mM Tris pH 7.5, 0.1 mM EDTA pH 8.0, 10 mM KCl) for another 1 h. Chromatin assemblies were centrifuged at 21’000 x g for 10 min at 4° C, and the supernatant was then transferred to a fresh tube. The concentration and volume of the chromatin assemblies were determined using a UV spectrophotometer. Chromatin assembly quality was controlled by the appearance of MMTV nucleosomes and ScaI digestion of 12x assemblies. Digestion reactions were analyzed on a 0.8% agarose gel and 5% TBE polyacrylamide gel electrophoresis. All experiments were carried out at 4° C (**Fig 24B and Fig S1**).

8.12 Preparation of microfluidic chambers for sm-FRET/TIRF experiments

Cleaning, silanization and PEGylation of coverslips and glass slides were described previously (Kilic et al., 2015). Briefly, coverslips (24 x 40 mm, 1.5 mm thickness) and glass slides (76 x 26 mm with 2 rows of 4 holes drilled) were sonicated for 20 min in 10% Aconox, rinsed with Milli-Q water and the procedure repeated sequentially with acetone and ethanol. Both coverslips and glass slides were then placed in a piranha etching solution (25% v/v 30% H_2O_2 and 75% v/v H_2SO_4) for a minimum of 2 h. After thorough washing with Milli-Q H_2O, coverslips and slides were sonicated in acetone for 10 min, then incubated with 2% v/v aminopropyl-triethyl silane (APTES) in acetone for 15 min, and dried. Flow chambers were assembled from one glass slide and one coverslip separated by double-sided 0.12 mm tape (Grace Bio-labs) positioned between each hole in the glass slide, and the open ends were sealed with epoxy glue. Pipette tips were fitted in each of the 2 x 4 holes on each side of the silanized glass flow chambers as inlet reservoir and outlet sources and glued in place with epoxy glue. The glue was allowed to solidify for 30-40 min. Subsequently, 350 µL of 0.1 M tetraborate buffer at pH 8.5 was used to dissolve ~1 mg of biotin-mPEG(5000 kDa)-SVA, and 350 µL from this was transferred to 20 mg mPEG (5000kDa)-SVA. This was centrifuged and mixed to homogeneity with a pipette before 40-45 µL aliquots were loaded into each of the four channels in the flow chamber. The PEGylation reaction was allowed to continue for the next 2½-4 h, after which the solution was washed out with degassed ultra-pure water (Romil).

8.13 Single-molecule TIRF (sm-TIRF) co-localization microscopy measurements

Measurements were done according to (Kilic et al., 2015). Objective-type smTIRF was performed using a fully automated Nikon Ti-E inverted fluorescence microscope, equipped with an ANDOR iXon EMCCD camera and a TIRF illuminator arm, controlled by NIS-elements and equipped with a CFI Apo TIRF 100x oil immersion objective (NA 1.49), resulting in a pixel size corresponding to 160 nm. Laser excitation was realized using a Coherent OBIS 640LX laser (640 nm, 40 mW) and coherent OBIS 532LS laser (532 nm, 50 mW) on a custom setup laser bench. Wavelength selection and power modulation were done using an acousto-optical tunable filter (AOTF) controlled by NIS elements. Typical laser intensities in the objective used for measurements were 0.8 mW for both 532 nm and 640 nm laser lines. For all smTIRF experiments, flow channels were washed with 500 µL degassed ultrapure water (Romil),

followed by 500 µL 1 x T50 (10 mM Tris pH 8, 50 mM NaCl) and background fluorescence was recorded with both 532 nm and 640 nm excitation. 50 µL of 0.2 mg/mL neutravidin was then injected and incubated for 5 min and washed using 500 µL 1xT50. 50 pM of Alexa647 labelled DNA/mononucleosomes/12-mer chromatin assemblies were then flowed in for immobilization in T50 with 2 mg/mL bovine serum albumin (BSA, Carlroth) (25 x 50 µm imaging area was monitored using 640 nm excitation to check for sufficient coverage). 500 µL 1 x T50 was used to wash out unbound Alexa647 labelled DNA/mononucleosomes/12-mer chromatin assemblies. 50-100 pM JF-549 labelled Rap1-Halo (see table below for details) was flowed in using imaging buffer (50 mM HEPES pH 7.5, 130 mM KCl, 10% v/v glycerol, 0.005% v/v Tween 20, 2 mM Trolox, 3.2% w/v glucose, 1x glucose oxidase/catalase oxygen scavenging system and 2 mg/mL BSA). Images were recorded using the following parameters:

	Camera t_{on} (msec)	Camera t_{off} (msec)	Orange channel # frames	Far-red channel: # frames	n repeat
DNA	100	600	1	1	5000
Mononucleosome	100	0.3	199	1	40
12-mer chromatin fiber	100	0.3	199	1	40

Here t_{on} denotes the camera integration time, whereas t_{off} indicates interspersed time intervals of camera inactivity.

Each experiment was repeated several times (see **Table SX** for the number of repeats), using at least two independently produced chromatin preparations on two different days.

8.14 Photobleaching test for JF-549 Rap1-Halo

Slides were prepared as described in the preceding sections. However, no BSA was added to imaging buffer (50 mM HEPES pH 7.5, 130 mM KCl, 10% v/v glycerol, 0.005% v/v Tween 20, 2 mM Trolox, 3.2% w/v glucose, 1x glucose oxidase/catalase oxygen scavenging system). JF-549 labelled Rap1-Halo was flowed into the channel and nonspecifically adsorbed on the glass surface. Movies were recorded using continuous 532 nm illumination (t_{on} 50 msec and t_{off} 0.3

msec) using the indicated excitation laser powers (**Fig S1**). Absolute laser power was determined using a laser power meter at the objective.

8.15 Ensemble FRET measurements

All measurements were performed using a Fluorolog®-3 Horiba Jobin Yvon spectrofluorometer in T50 buffer (10 mM Tris pH 8, 50 mM NaCl) 60 µl total volume. Nucleosomes (final concentration of 25-30 nM) and Rap1 (0, 1, 2, 5, 10 equivalents) were mixed by pipetting in T50 buffer and left for 10 min room temperature to bind. Fluorescence emission spectra are taken from 585 nm to 700 nm (1 nm increments) using 578 nm as excitation wavelength. Spectra for DNA only, T50 only and donor only samples were taken. For a given sample, NaCl was added to 800 mM to observe nucleosome disassembly. FRET efficiency was calculated from donor emission:

$$E_{FRET} = 1 - \frac{F_{DA}}{F_D}$$

With F_{DA} denoting donor emission in the presence of acceptor and F_D denoting donor emission in the donor-only sample. Additionally, reactions were loaded onto 0.5x TBE 5% polyacrylamide gels to check binding.

8.16 Single-molecule FRET (smFRET) measurements for Chd1 remodelling

Measurement procedure. Flow cell preparation and nucleosome loading was performed as described in (Kilic et al., 2018b) and the preceding paragraphs. Experiments were performed in FRET imaging buffer (40 mM KCl, 50 mM Tris, 2 mM Trolox, 2 mM nitrobenzyl alcohol (NBA), 2 mM cyclooctatetraene (COT) 5 mM $MgCl_2$ and 3.2% glucose) supplemented with GODCAT (100x stock solution: 165 U/mL glucose oxidase, 2170 U/mL catalase). Experiments on chromatin remodelling with chd1 were performed with imaging buffer containing 40 mM KCl, and 0.1 mg/mL of BSA was added to prevent nonspecific binding to the glass surface. For Rap1 experiments, unlabeled Rap1-Halo was used.

smFRET data acquisition was carried out with a micro-mirror TIRF system (MadCityLabs) using Coherent Obis Laser lines at 405 nm, 488 nm, 532 nm and 640 nm, a 100x NA 1.49 Nikon CFI

Apochromat TIRF objective (Nikon), as well as an iXon Ultra EMCCD camera (Andor), operated by custom-made Labview (National Instruments) software.

For general smFRET imaging, a programmed sequence was employed to trigger the camera (at 800 EM gain) to acquire frames with alternating 532 nm and 640nm excitation at 100 ms time-resolution.

Injection of Chd1 and other components (1 mM ATP, 1nM Rap1) was triggered to start at 10 secs and at a flow rate of 200 ul / min until an injection volume of 100 ul was perfused.

Each experiment was repeated several times (see **Table S2** for the number of repeats), using at least two independently produced nucleosome preparations on two different days.

	Camera t_{on} (msec)	Camera t_{off} (msec)	Orange channel # frames	Far-red channel: # frames	Injection frame	n repeat
Asy/Sym Nucleosome	100	0.5	1	1	100	2000

8.17 Image processing, single-molecule trace extraction and trace analysis

Single-molecule trace extraction and trace analysis were done according to (Kilic et al., 2015) with some adjustments. Firstly, background subtraction was performed for all Rap1-Halo binding movies using a rolling ball background subtraction in ImageJ (using 50-pixel rolling ball size). Using a custom-built Matlab (Mathworks) program suite, DNA/nucleosome or chromatin positions were detected via a local maxima approach. Sequential images were aligned using the far-red channel to compensate for stage drift. Fluorescence intensities (in the orange channel) were extracted from the stack within a 2-pixel radius of the identified DNA peaks. Every detected spot in the orange channel was fitted with a 2D-Gaussian function to determine co-localization with immobilized DNA/chromatin. Peaks exceeding an experimentally determined PSF width for a single JF-549 molecule were excluded from further analysis. Extracted fluorescence traces were filtered using a forward-backwards non-linear filter (Chung and Kennedy, 1991) to reduce noise.

Residence times were determined using a semi-automatic procedure. Individual binding events were detected using a thresholding algorithm. Overlapping multiple binding events were excluded from the analysis. For each movie, cumulative histograms were constructed from detected bright times (t_{bright}) corresponding to bound Rap1 molecules, usually including data from ~100 individual traces. The cumulative histograms from traces corresponding to individual DNA / MN / chromatin fibres were fitted with either di- or tri-exponential functions:

$$y = \sum_{i=1}^{2} A_i \exp(-t / \tau_{off,i}) \quad \text{or} \quad y = \sum_{i=0}^{2} A_i \exp(-t / \tau_{off,i})$$

yielding nonspecific residence times $\tau_{off,0}$ or the specific residence times $\tau_{off,1}$ and $\tau_{off,2}$.

Cumulative histograms constructed from dark times (t_{dark}), in between binding events, were fitted with mono-exponential functions:

$$y = A \exp(-t / k_{on,app})$$

To obtain apparent on-rates. The detected on-rates contain both contributions from nonspecific and specific binding events. To calculate specific on-rates (k_{on}), the contributions from nonspecific events have to be filtered out. To this end, measured $k_{on,app}$ values were corrected using the amplitude contributions of nonspecific (A_0) and specific binding events (A_1, A_2).

$$k_{on,specific} = k_{on} \left(\sum_{i=1}^{2} A_i \Big/ \sum_{i=0}^{2} A_i \right)$$

8.18 Single-molecule FRET (smFRET) conformation analysis

smFRET data analysis. FRET reporting on chromatin conformation as a function of ionic strength or Rap1 binding was recorded as described above.

For FRET calculation, the orange and far-red channel detection efficiency ratio γ and donor dye bleed-through parameter β were determined using FRET traces presenting acceptor

bleaching. Maximum intensities in the 530 nm and 640 nm channels were then reported before and after photobleaching, and the following equation was used.

$$\beta = \frac{F_{A,bleach}}{F_{D,bleach}} \text{ and } \gamma = \frac{\Delta F_{A,bleach}}{\Delta F_{D,blea}}$$

β was found to be 0.058 ± 0.03 and γ was 0.398 ± 0.1.

For all recorded movies, background correction was performed in ImageJ using a rolling ball algorithm. Single-molecule kinetic trace extraction and analysis was performed in custom-written MATLAB software. Individual molecules were automatically detected in the initial acceptor image prior to donor excitation, and the same peaks were selected in the donor channel. Peaks that are (*i*) tightly clustered or (*ii*) above an intensity threshold of 8000 in the donor channel and 5000 in the acceptor channels indicating aggregation or (*iii*) do not appear in both donor and acceptor channels were excluded from the analysis. Selection criteria were similar to (Kilic et al., 2018b). Traces were included if they exhibited: (*i*) a single bleaching event, (*ii*) constant total fluorescence emission > 2000 counts from the combined donor and γ-corrected acceptor channel (*iii*) a constant baseline lasting for at least 2 s after donor bleaching, (i*v*) donor emission for at least 5 s and finally (*v*) the presence of acceptor dye. The last condition is verified as follows: If the donor dye bleaches first, acceptor emission must be detectable at the end of the experiment upon direct acceptor excitation. If the acceptor dye bleaches first, a significant increase is seen in the donor channel. From selected traces, donor (F_D) and acceptor (F_A) fluorescence emission intensity, FRET efficiency (E_{FRET}) was calculated as follows:

$$E_{FRET} = \frac{F_A - \beta F_D}{F_A - \beta F_D + \gamma F_D}$$

We determined the detection efficiency $\gamma = 0.398$ and the bleed-through $\beta = 0.058$ for the FRET pair Cy3B/Alexa647 with our experimental setup. These values were used to calculate E_{FRET} for the selected traces, and construct E_{FRET} histograms with a bin size of 0.02. E_{FRET} histograms of each trace of length > 5 s were normalized to total counts. Final histograms of each independent measurement were fitted using 3 Gaussian functions as follows:

$$\sum_i A_i e^{-\frac{(x-c_i)^2}{2\sigma_i^2}}$$

Where A_i is the amplitude or the height of the fitting peak, c_i is the position of the centre of the peak, and σ_i is the standard deviation which controls the width of the Gaussian peak. The integral area of each peak was calculated as follows:

$$\int_{-\infty}^{\infty} A_i e^{-\frac{(x-c_i)^2}{2\sigma_i^2}} dx = A_i \sigma_i \sqrt{2\pi}$$

Where indicated, low-FRET (LF), medium-FRET (MF) and high-FRET (HF) refer respectively to the centre of the Gaussian peak limited with $\boldsymbol{c_i} < 0.2$, $0.2 \leq \boldsymbol{c_i} \leq 0.4$, and $\boldsymbol{c_i} > 0.4$. The percentage of LF-population at compaction conditions, i.e. in high salt or presence of Mg^{2+}, indicates the fraction of uncompacted chromatin and hence reports chromatin assembly quality.

8.19 Statistical analysis

All results are presented as means with their standard deviation unless otherwise indicated. Pairs of experimental values were compared using two-sided, homoscedastic student t-tests with a confidentiality interval of 5%: a p-value below 0.05 was considered as statistically significant.

8.20 Nucleosome shift assays with RSC, Nap1 and Rap1

Purified RSC and recombinant yNap1 were used (for the purification, see (Kurat et al., 2017)). All reactions were performed in reaction buffer (10 mM Tris pH 7.4, 150 mM KCl, 3 mM $MgCl_2$, 0.1 mg/mL BSA) and a total volume of 50 µl. The following components were added in sequential order MNs (to give a 20 nM final concentration), yNap1 (10 eq. yNap1 : 1 eq. MNs), if required Rap1 (10 eq. Rap1 : 1 eq. MNs), RSC complex (0.2 eq. RSC : 1 eq. MNs) and finally ATP (1mM). Reactions were placed at 30°C and 10 µl were taken for each time point, to which was added a 3-fold molar excess of plasmid DNA (compared to nucleosomes) containing a Rap1 binding site and returned to 30° C for 5 min. Reactions were then placed on ice until glucose was added to make 8% final concentration and loaded onto commercial Criterion Precast Gel (Biorad) 5% TBE, 1mm, run in 1x TBE at 200 V for 35-45 min on ice. Gels were stained in GelRed and imaged using ChemiDoc MP (Biorad) **Fig 28c-f**). Leaving out Nap1 from

the reaction did not affect RSC remodelling (**Figure 28c**). Remodelling assays using MNs containing fluorescently labelled octamers were also performed (**Fig28 e)** using the same conditions as described above. To model the RSC displaced nucleosome, an asymmetric PCR generated *P3_S12_remodelled* **(Table S3**) DNA fragment was used. This DNA was reconstituted into a nucleosome and incubated with Rap1 for 10 min at 30° C in reaction buffer (10 mM Tris pH 7.4, 150 mM KCl, 3 mM $MgCl_2$, 0.1 mg/mL BSA), total volume of 10 µl. A 3-fold excess of plasmid DNA (compared to nucleosomes) containing a Rap1 binding site was added and returned to 30° C for 5 min. Reactions were then placed on ice until glucose was added to make 8% final concentration and loaded onto 5% polyacrylamide 0.5x TBE, 1.5 mm, run in 0.5x TBE at 120 V for 55-60 min on ice. Gels were stained in Gelred and imaged using ChemiDoc MP (Biorad) (**Fig 28**). For the sequential remodelling experiment, nucleosomes were incubated with RSC and Nap1 for 90 min as described above. At 90 min, the RSC reaction was stopped by the addition of 30 mM EDTA. Then, Rap1 was added for 5 min at 30 °C, followed by analysis on native PAGE (**Fig 28g**).

8.21 RSC sliding and MNase-seq

RSC sliding reactions were performed in reaction buffer (10 mM Tris pH7.4, 150 mM KCl, 3 mM $MgCl_2$, 0.1 mg/mL BSA) and a total volume of 70 µl. The following components were added in sequential order MNs (to make 20nM final concentration), Nap1 (10 Nap1 : 1 MN ratio), Rap1 (10 Rap1 : 1 MN ratio, for w/o Rap1 MQ water was used as a substitute), RSC complex (0.2 RSC : 1 MN ratio) and finally ATP (1 mM). Reactions were placed at 30°C for 90 min after which 10 µl was taken, and glucose was added to make 8% final concentration and loaded onto commercial Criterion Precast Gel (Biorad) 5% acrylamide, 1mm, run in 1xTBE at 200 V for 35-45 min on ice. Gels were stained in Gelred and imaged using ChemiDoc MP (Biorad) (**Fig 29a**). To the remaining 60 µl, 60 µl 50mM Tris-HCl pH 8 and 10x NEB MNase buffer (M0247S) (to make final 1x) was added. This 120 µl total sample was split into 3 x 40 µl aliquots, and to each either 6 units, 3 units or 1 unit of Mnase (M0247S) was added respectively and left to digest for 5 min at 37° C. To stop the reaction, an equal volume of stop buffer was added (200 mM NaCl, 30 mM EDTA, 1% SDS) and left on ice for 5 min. Finally, 10 µg of Proteinase K (Sigma P2308) was added and left for 1h at 60° C and DNA fragments were isolated using QIAquick PCR purification spin columns (Qiagen). For nucleosome only samples

(t0), reactions were performed directly in 1x NEB Mnase buffer (M0247S), Mnase and Proteinase K digestion, as well as DNA fragment purification, was performed as described for RSC assay nucleosomes (**Fig 29b**).

Following MNase digestion, DNA was purified using MinElute PCR Purification Kit (QIAGEN). The libraries were prepared using TruSeq ChIP Sample Preparation Kit (Illumina, Catalog IDs: IP-202-1012, IP-202-1024) according to the manufacturer's instructions. The libraries were sequenced on a HiSeq 4000 machine in 100 bp paired-end mode at the Genomics Platform of the University of Geneva (https://ige3.genomics.unige.ch/). Mapping of the sequencing data to the corresponding sequences was performed using Bowtie2 (sensitive end-to-end mode) on Galaxy (https://usegalaxy.org/). All densities were derived from read counts normalized to the total number of reads for each sample and BAM files was converted to bigWig files using bamCoverage and bigWig files converted to BedGraph format on Galaxy.

8.22 Expression and purification of Chd1

The Strep-MBP-TEV-Chd1 construct (**Fig 33a**) was cloned into pACEBac1 (Geneva Biotech) and baculovirus particles were generated using the Geneva Biotech system per manufacturer's instructions.

For Chd1 expression, 1L cultures of Sf9 cells were grown to 2 - 2.5x10^6 cells/mL. Subsequently, the cells were infected with baculovirus, and the cultures were incubated for 3 days at 27° C, before harvesting through centrifugation (1500 rcf, 4°C for 20 min). Supernatants were discarded and pellets were resuspended in PBS, containing protease inhibitors (Roche) (10 mL PBS/L of culture), flash frozen and kept at -80°C.

12-15 g of frozen pellets were thawed at room temperature with 36 mL of lysis buffer (200 mM KCl, 2 mM DTT, 100 mM Tris-HCl (pH 7.5), 50 mM MgOAc, 0.1% NP-40, Protease inhibitor cocktail (Roche), 1mM PMSF and 20 µL DNaseI (NEB)). Pellets, were stirred with a magnetic stir bar until fully thawed and then kept on ice. The lysate was spun for 35 min at 35000 rpm at 4° C (Ti70 rotor, Beckman Coulter) and the supernatant was filtered through a 5 µm syringe filter (Millex, Millipore). The cleared lysate was loaded onto a Strep-Trap column (GE, AKTA system), pre-equilibrated with lysis buffer. The column was washed with storage buffer (200 mM KCl, 10 mM HEPES pH 7.6, 50 mM MgOAc, and 5 mM β-mercaptoethanol (βME)) and the protein was eluted with 5 x column volumes (CV) of elution buffer (storage buffer containing

2.5 mM desthiobiotin). Fractions containing Chd1 were identified by SDS-Page (**Fig 33b**), pooled and concentrated to ~ 500 µL total volume using Amicon 10k molecular weight cut-off (MWCO) centrifugal filters. The protein concentration was determined using UV spectroscopy. The MBP tag was subsequently removed by TEV protease digestion at 4°C (**Fig 33c**). Chd1 was finally purified by size exclusion chromatography (SEC) using a Superose6 10/300 GL column (GE healthcare) in storage buffer using a flow-rate of 0.4 mL/min (**Fig 33d**). Fractions were analysed using SDS-PAGE (**Fig 33d**), clean fractions were pooled, concentrated (Amicon 10k MWCO filter) and protein concentrations were determined using UV spectrophotometry (at A280 and ε_{280} = 166 425 mol^{-1} cm^{-1}).

8.23 Chd1 sliding

Chd1 sliding reactions were performed in reaction buffer (10 mM Tris pH7.4, 150 mM KCl, 3 mM $MgCl_2$, 0.1 mg/mL BSA) and a total volume of 10 µl. The following components were added in sequential order Nucleosomes (to make 10nM final concentration), Rap1 (10 Rap1 : 1 MN ratio, for w/o Rap1 MQ water was used as substitute), Chd1 (1 Chd1 : 1 MN ratio) and finally ATP (1 mM). Reactions were placed at 30°C for 90 min after which 10 µl was taken and glucose was added to make 8% final concentration and loaded onto commercial Criterion Precast Gel (Biorad) 5% acrylamide, 1mm, run in 1xTBE at 200 V for 35-45 min on ice. Gels were stained in Gelred and imaged using ChemiDoc MP (Biorad) (**Fig 33e**).

9 Supplementary Figure

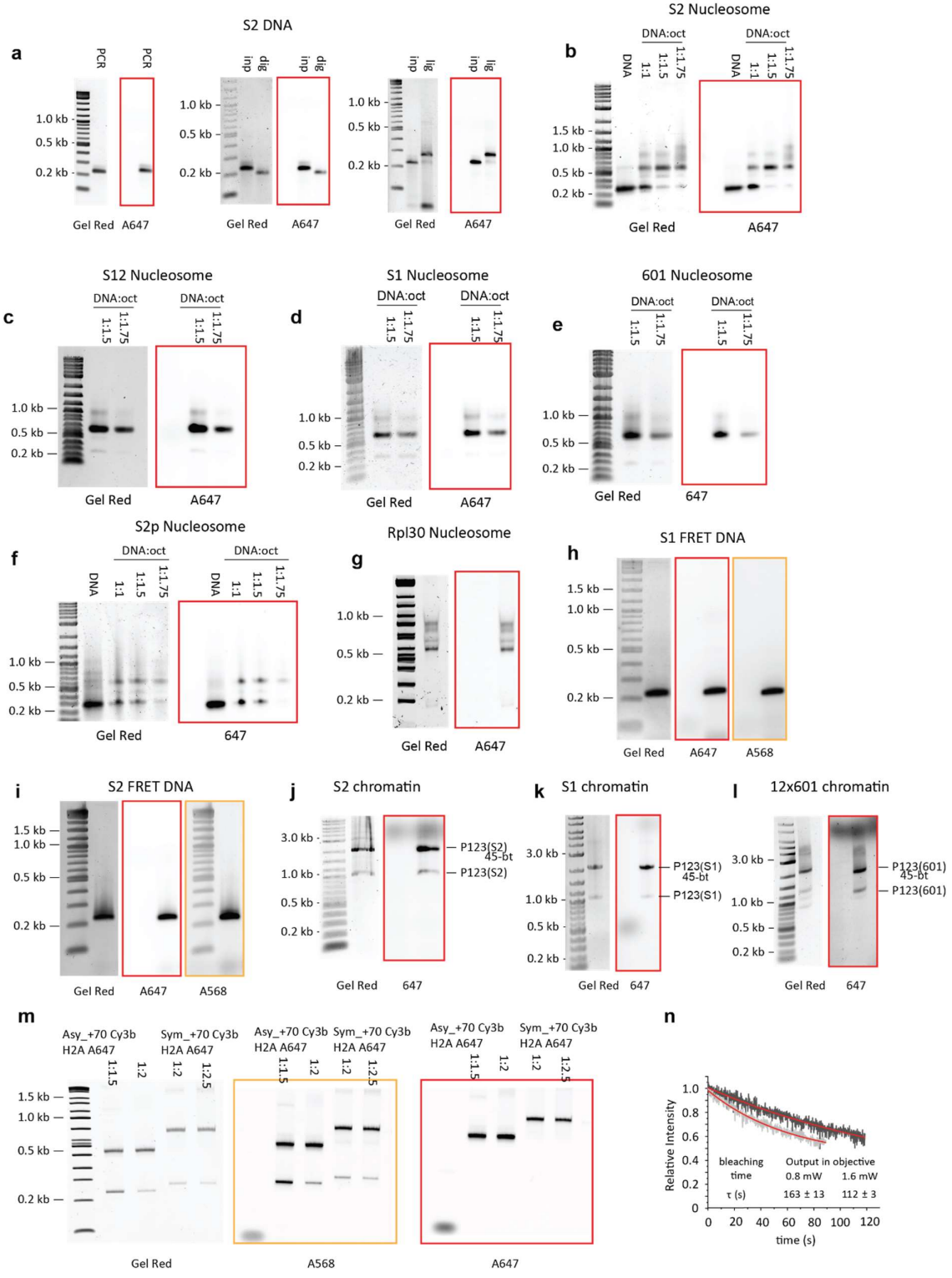


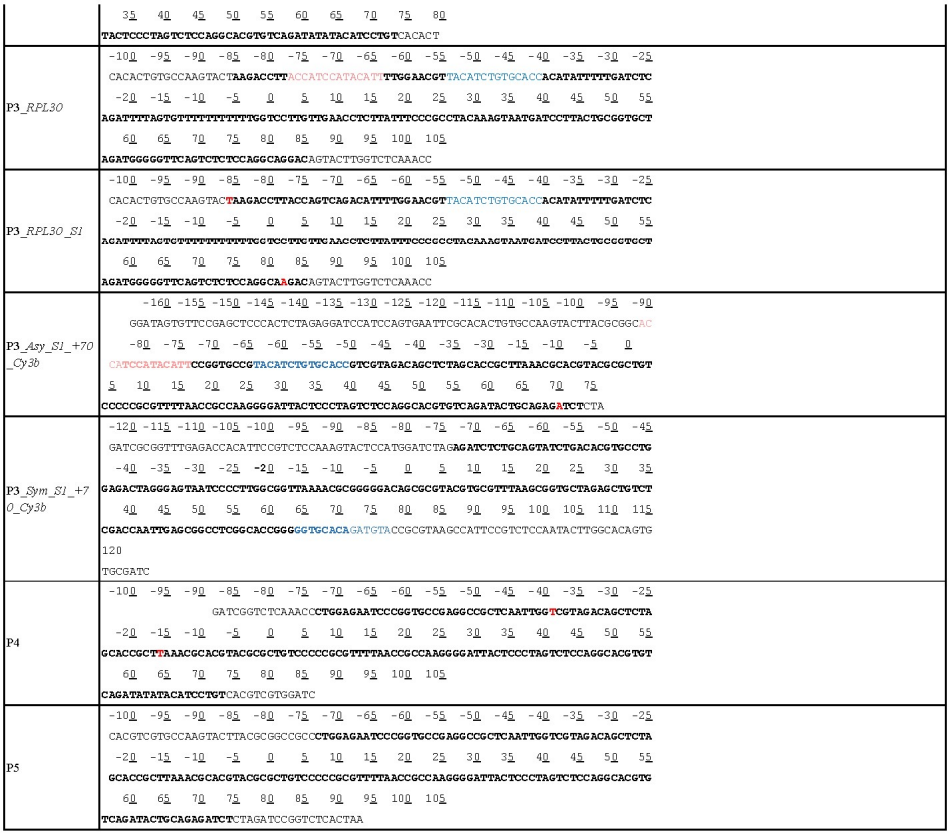
Figure S 1 a) PCR generation of S2 DNA, digestion of S2 DNA and ligation of S2 DNA with Biotin anchor. b) S2 nucleosome preparation c) S12 nucleosome preparation d) S1 nucleosome

preparation e) 601 nucleosome preparation f) S2p nucleosome preparation g) Rpl30 nucleosome preparation h) S1 FRET DNA i) S2 FRET DNA j) S2 chromatin DNA preparation k) S1 chromatin DNA preparation l) 12x601 chromatin DNA preparation m) Asymmetric +70 cy3b H2AN110C A647 and symmetric +70 cy3b H2AN110C A647 nucleosome preparation. n) Photobleaching test using JF549.

10 Supplementary Tables

Table S1 | Sequences of 1 x 601 pieces for recombinant and PCR-generated pieces.

Fragment	Sequence-
P1	-100 -95 -90 -85 -80 -75 -70 -65 -60 -55 -50 -45 -40 -35 -30 -25 CACTTGGTGGCGGCCGCCCTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTA -20 -15 -10 -5 0 5 10 15 20 25 30 35 40 45 50 55 GCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGT 60 65 70 75 80 85 90 95 100 105 CAGATACTGCAGAGATCTCTAGATCCATGGAGTACTTGGTCTCATAGC
P2	-100 -95 -90 -85 -80 -75 -70 -65 -60 -55 -50 -45 -40 -35 -30 -25 GATCGGTCTCATAGCCTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTA -20 -15 -10 -5 0 5 10 15 20 25 30 35 40 45 50 55 GCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGT 60 65 70 75 80 85 90 95 100 105 CAGATATATACATCCTGTCACACTGTGGATC
P3	-100 -95 -90 -85 -80 -75 -70 -65 -60 -55 -50 -45 -40 -35 -30 -25 CACACTGTGCCAAGTACTTACGCGGCCGCCCTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTA -20 -15 -10 -5 0 5 10 15 20 25 30 35 40 45 50 55 GCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGT 60 65 70 75 80 85 90 95 100 105 CAGATACTGCAGAGATCTCTAGATCCATGGAGTACTTGGTCTCAAACC
P3_*S1*	-100 -95 -90 -85 -80 -75 -70 -65 -60 -55 -50 -45 -40 -35 -30 -25 CACACTGTGCCAAGTACTTACGCGGCCGCCCTGGAGAATCCCGGTGCCGTACATCTGTGCACCGTCGTAGACAGCTCTA -20 -15 -10 -5 0 5 10 15 20 25 30 35 40 45 50 55 GCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGT 60 65 70 75 80 85 90 95 100 105 CAGATACTGCAGAGATCTCTAGATCCATGGAGTACTTGGTCTCAAACC
P3_*S1* +3	-100 -95 -90 -85 -80 -75 -70 -65 -60 -55 -50 -45 -40 -35 -30 -25 CACACTGTGCCAAGTACTTACGCGGCCGCCCTGGAGAATCCCGGTGCCGAGGTACATCTGTGCACCGTAGACAGCTCTA -20 -15 -10 -5 0 5 10 15 20 25 30 35 40 45 50 55 GCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGT 60 65 70 75 80 85 90 95 100 105 CAGATACTGCAGAGATCTCTAGATCCATGGAGTACTTGGTCTCAAACC
P3_*S1* +6	-100 -95 -90 -85 -80 -75 -70 -65 -60 -55 -50 -45 -40 -35 -30 -25 CACACTGTGCCAAGTACTTACGCGGCCGCCCTGGAGAATCCCGGTGCCGAGGCCGTACATCTGTGCACCGACAGCTCTA -20 -15 -10 -5 0 5 10 15 20 25 30 35 40 45 50 55 GCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGT 60 65 70 75 80 85 90 95 100 105 CAGATACTGCAGAGATCTCTAGATCCATGGAGTACTTGGTCTCAAACC
P3_*S2*	-100 -95 -90 -85 -80 -75 -70 -65 -60 -55 -50 -45 -40 -35 -30 -25 CACACTGTGCCAAGTACTTACGCGGCACCATCCATACATTCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTA -20 -15 -10 -5 0 5 10 15 20 25 30 35 40 45 50 55 GCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGT 60 65 70 75 80 85 90 95 100 105 CAGATACTGCAGAGATCTCTAGATCCATGGAGTACTTGGTCTCAAACC
P3_*S2p*	-100 -95 -90 -85 -80 -75 -70 -65 -60 -55 -50 -45 -40 -35 -30 -25 CACACTGTGCCAAGTACTTACGCGGCCGCCCTGGAGAATCCCGGTGCCGACCATCCATACATTGTCGTAGACAGCTCTA -20 -15 -10 -5 0 5 10 15 20 25 30 35 40 45 50 55 GCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGT 60 65 70 75 80 85 90 95 100 105 CAGATACTGCAGAGATCTCTAGATCCATGGAGTACTTGGTCTCAAACC
P3_*S12*	-100 -95 -90 -85 -80 -75 -70 -65 -60 -55 -50 -45 -40 -35 -30 -25 CACACTGTGCCAAGTACTTACGCGGCACCATCCATACATTCCGGTGCCGTACATCTGTGCACCGTCGTAGACAGCTCTA -20 -15 -10 -5 0 5 10 15 20 25 30 35 40 45 50 55 GCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGT 60 65 70 75 80 85 90 95 100 105 CAGATACTGCAGAGATCTCTAGATCCATGGAGTACTTGGTCTCAAACC
P3_*S12_remodelled*	-125 -120 -115 -110 -105 -100 -95 -90 -85 -80 -75 -70 -65 -60 -55 -50 TACTTACGCGGCACCATCCATACATTCCGGTGCCGTACATCTGTGCACCGTAGCCTTGGAGAATCCGGTGCCGAGGCC -45 -40 -35 -30 -25 -20 -15 10 -5 0 5 10 15 20 25 30 GCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGAT



601 or native RPL30 sequences indicated in bold. The labeled base pairs are indicated in red. The numbering is given as number of base-pairs relative to the dyad in the 601 sequence. Rap1 binding sites are indicated in blue for S1 and burgundy for S2.

Table S1 All 601 template sequences used

Table S2 | All kinetic parameters of Rap1 interacting with DNA, nucleosomes and chromatin.

	dissociation kinetics						binding kinetics		statistics
	dwell time (s)			Amplitude (%)			ıte constants (x 10^7 s-		
	$\tau_{off,0}$	$\tau_{off,1}$	$\tau_{off,2}$	A_0	A_1	A_2	k_{on}	$k_{on,specific}$	n exp.
DNA									
DNA *S2*	X	12.4 ± 4.5	451.5 ± 115*	X	35 ± 17	65 ± 17	6.3 ± 1.9	6.3 ± 1.9	4
Nucleosomes									
MN S1	0.6 ± 0.4	17.8 ± 10.8	116.4 ± 36.0‡	43 ± 5	27 ± 9	30 ± 7	7.2 ± 2.9	4.0 ± 2	4
MN S2	0.7 ± 0.2	8.4 ± 1.4	46.1 ± 3.0†	35 ± 6	54 ± 4.9	11 ± 4	11 ± 2.7	7.1 ± 2	4
MN S2*	0.3 ± 0.1	2.4 ± 0.4	7.7 ± 1.9	17 ± 5	51 ± 10	32 ± 8	18 ± 5	15 ± 5	5
MN no site	0.2 ± 0.1	3.5 ± 3.0	X	83 ± 10	17 ± 10	X	4.1 ± 3.4	0.5 ± 0.4	4
MN RPL30_S1	1.3 ± 0.2	10.4 ± 1.7	69.6 ± 8.3	58 ± 4	34 ± 5	8 ± 7	27 ± 5	11 ± 2	5
Chromatin									
CH *S1*	0.6 ± 0.1	3.2 ± 0.6	25.6 ± 4.0	73 ± 12.9	24 ± 11	4 ± 2	29 ± 7.7	7.1 ± 2	8
CH *S2*	0.6 ± 0.2	2.6 ± 0.6	16.8 ± 2.9	69 ± 10	27 ± 9	4 ± 2	33 ± 17	9.2 ± 6.6	6
CH no site	0.2 ± 0.1	1.4 ± 0.7	X	81 ± 2	18.7 ± 2	X	17 ± 4	3.0 ± 1.1	3

Reported are values uncorrected for photobleaching. For photobleaching correction, see **Star Methods**. For the longest time constants, corrected values are

*: The bleaching time of JF-549 is 168 s (Figure S4). Due to stroboscopic imaging, the DNA sample is only illuminated 14% of the time, extending the photon budget. If the experimental value is corrected for photobleaching, $\tau_{off,2}$ = 721 ± 183 s.

‡: If the experimental value is corrected for photobleaching, $\tau_{off,2}$ = 330 ± 102 s.

†: If the experimental value is corrected for photobleaching, $\tau_{off,2}$ = 63 ± 4.3 s.

Table S2

Table S3 | Overview of all chromatin DNA with different combinations of labels.

Experiment	Name	Backbone	Dye	Modification
Rap1 nucleosome binding	Nuc *S1*	P3_*S1*	Alexa647 (82)	3' biotin
	Nuc *S2*	P3_*S2*	Alexa647 (82)	3' biotin
	Nuc *S2p*	P3_*S2p*	Alexa647 (82)	3' biotin
	Nuc *S12*	P3_*S12*	Alexa647 (82)	3' biotin
	Nuc 601	P3	Alexa647 (82)	3‘ biotin
Rap1 chromatin binding	CH *S1*	P1P2 ; P3_*S1* ; P4P5	Alexa647 (82)	3' biotin
	CH *S2*	P1P2 ; P3_*S2* ; P4P5	Alexa647 (82)	3' biotin
	CH *NS*	P1P2 ; P3 ; P4P5	Alexa647 (82)	3' biotin
Nucleosome FRET	Nuc *S1* FRET	P3_*S1*	P3: Alexa647 (82), Alexa568 (-86)	3' biotin
	Nuc *S2* FRET	P3_*S2*	P3: Alexa647 (82), Alexa568 (-86)	3' biotin
	Nuc_Asy_Cy3B	P3_Asy_Cy3B	P3: Cy3b (70) H2AN110C(A647)	3' biotin
	Nuc_Sym_Cy3B	P3_Sym_Cy3B	P3: Cy3b (70) H2AN110C(A647)	3' biotin
Chromatin FRET	CH *NS* FRET	P1 ; P2 ; P3 ; P4 ; P5	P2: Cy3B (39), P4: Alexa647 (-39)	3' biotin
	CH *S2* FRET	P1 ; P2 ; P3_*S2* ; P4 ; P5	P2: Cy3B (39), P4: Alexa647 (-39)	3' biotin
Ensemble binding &	MN *RPL30*	P3_*RPL30*		
RSC experiments	MN *S1* /*S2*	P3_*S1S2*		

Table S3

11 CV

Mivelaz, Maxime Nicolas
Chemin de Montelly 21, 1004 Lausanne
Tel: 0798699581
E-Mail: maxime.mivelaz@hotmail.com

Nationality: Swiss
Date of Birth: 02/07/1990
Place of Origin: Poliez-Pittet, VD
Place of Birth: Hong-Kong
Marital Status: Single
Languages spoken: English (Native Speaker), French (Fluent)

Degrees

2006 - 2008	International Baccalaureat (IB) United World College of South East Asia, Singapore
2008 - 2009	Chinese Language Proficiency Beijing Language and Culture University, Beijing, China
2009 - 2012	Bachelors of Medicine University of Lausanne (UNIL), Switzerland (Not completed)
2012 - 2015	Bachelors of Science, Biochemistry University of Geneva (UNIGE), Switzerland
2015 - 2017 July	Masters of Chemical Biology National Center for Competence in Research (NCCR) in Chemical Biology at the University of Geneva
2017 – 2021	PhD in Chemistry Ecole Polytéchnique Fédérale de Lausanne (EPFL), Switzerland Laboratory of Biophysical Chemistry of Macromolecules Title: Dynamic Chromatin Binding and Remodelling of Pioneer Transcription Factors Completion November 2021

Publications

Chromatin Fiber Invasion and Nucleosome Displacement by the Rap1 Transcription Factor - M. Mivelaz et al., Mol. Cell, 2020, Vol. 77, Issue 3, DOI: 10.1016/j.molcel.2019.10.025

Observing protein interaction dynamics to chemically defined chromatin fibers by colocalization single-molecule fluorescence microscopy - M. Mivelaz et al., Methods, 2020 DOI: 10.1016/j.ymeth.2020.01.018