Functional Chromatin Extraction:

A method to study DNA accessibility in higher-order structures of chromatin

DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES DER NATURWISSENSCHAFTEN (DR.RER.NAT) DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN DER UNIVERSITÄT REGENSBURG

vorgelegt von

Uwe Schwartz

aus

Temeschburg, Rumänien

im Jahr 2021

Das Promotionsgesuch wurde eingereicht am:
23.04.2021

Die Arbeit wurde angeleitet von:
Prof. Dr. Gernot Längst

Unterschrift

(Uwe Schwartz)

Table of Contents

Abbreviations

°C	degree Celsius
A	adenine
Ad	Adenovirus
ANC	active nuclear compartment
ATP	adenosin-5'-triphosphate
bp	base pair
C	cytosine
C/EBP	CCAAT enhancer binding protein
CAGE	cap analysis of gene expression
cDNA	coding DNA
ChIP	chromatin immunoprecipitation
chr	chromosome
CREM	CAMP responsive element modulator
CTCF	CCCTC-binding factor
DAPI	4′,6-Diamidin-2-phenylindol
DNA	deoxyribonucleic acid
DNase	Deoxyribonuclease
dsDNA	double stranded DNA
EDTA	ethylenediaminetetraacetic
EM	electron microscopy
EN1	engrailed homeobox 1
ENCODE	encyclopedia of DNA elements
EZH2	enhancer of zeste homolog 2
FC	fold-change
FCE	Functional Chromatin Extraction
FCS	fetal calf serum
FISH	fluorescence in situ hybridization
FPKM	fragments per kilobase transcript per million mapped
G	guanine
h	hour
H1	Histone 1

H2A	Histone 2A
H2B	Histone 2B
H3	Histone 3
H4	Histone 4
HAd-C5	human Adenovirus type 5
hpi	hours post infection
HS	hypersensitive
IC	interchromatin compartment
IF	immunofluorescence
INC	inactive nuclear compartment
kbp/kb	kilo base pairs
LOESS	locally weighted scatterplot smoothing
log	logarithm
m	meter
MAPQ	mapping quality score
Mbp/Mb	mega base pairs
Mio	million
ML	major late promoter
MNase	Micrococcal nuclease
MOI	multiplicity of infection
MW	molecular weight
NFR	nucleosome free region
nm	nanometer
PBS	phosphate-buffered saline
polyA	polyadenylation
PRC2	polycomb-repressive complex 2
pVI	protein VI
pVII	protein VII
pVII-RL	protein VII repeat length
qPCR	quantitative polymerase chain reaction
R	correlation coefficient
RNA	ribonucleic acid
RPM	reads per million mapped reads
SDS	sodium dodecyl sulfate

T	thymine
TAD	topologically associating domains’
TAF-Ib	template activating factor I
TPM	transcripts per kilobase per million mapped reads
TSS	transcription start site
U	unit
µm	micrometer

1 Summary

Inside the cell nucleus DNA is compacted through the assembly with histones and other proteins into chromatin. The first layer of DNA packaging is the nucleosome core particle, which consists of DNA wrapped around a histone octamer. Nucleosome core particles are spaced by linker DNA forming a ”beads-on-a-string” structure. According to the textbook model nucleosome arrays are further regularly folded into distinct higher-order structures of chromatin. Since all DNA dependent processes require access to the DNA template, chromatin organization and folding into higher-order structures is thought to regulate genome activity.

This thesis investigates how chromatin is structurally organized inside the nucleus to modulate DNA accessibility. A high-throughput approach, called Functional Chromatin Extraction, was developed to analyze DNA accessibility in native chromatin. Therefore, chromatin was digested with different intensities directly inside the nucleus of living cells using the endonuclease MNase. DNA accessibility was assessed on both global and local scale by the differential release rates of nucleosomes from partially (low-MNase) and fully digested chromatin (high-MNase). Thorough analysis of the extracted nucleosomal DNA revealed, that AT rich nucleosomes are prone to over-degradation to sub-nucleosomal fragments in high-MNase. Therefore, nucleosomes of GC rich regions are overrepresented in high-MNase. In contrast, low-MNase results in a homogenous nucleosome distribution not affected by the DNA sequence, thereby obtaining an accurate representation of the global nucleosome landscape.

Surprisingly, after correcting for the sequence preferences of MNase, differentially accessible chromatin domains could not be identified. Euchromatin and heterochromatin exhibit similar accessibilities, suggesting that DNA in heterochromatin is in general available for small molecules, like transcription factors. Nevertheless, active regulatory sites, such as promoter and enhancer elements, reveal increased accessibility compared to other regions of the genome and are occupied by fragile nucleosomes showing, that DNA accessibility is modulated locally to regulate gene expression.

In summary, the results of this study indicate, that chromatin forms an accessible and dynamic polymer and domains of higher-order structures of chromatin do not exist in human cells.

In a second chapter, this thesis focuses on the chromatin architecture of Adenoviruses and dynamic changes during early infection. Similarly to eukaryotic genomes, adenoviral DNA in incoming virions is mainly associated with the structural protein VII (pVII) forming a nucleoprotein complex. However, little is known about the adenoviral chromatin organization and how it relates to viral gene activation during infection. Functional Chromatin Extraction

combined with transcriptome sequencing was applied during early infection of human cells. The viral DNA organization into pVII complexes was assessed, showing a defined and functional DNA packaging into nucleosome-like arrays. The chromatin structure of invading viruses correlates with the spatiotemporal activation of viral genes showing an open chromatin conformation with lower pVII densities at early gene loci.

Investigation of dynamic chromatin changes within the first four hours of infection, revealed viral chromatin de-condensation and nucleosome assembly preferentially at early gene loci. Remarkably, nucleosomes replace pVII molecules directly at the +1 site of early genes thereby resembling the structure of active host promoter. The time resolved analysis demonstrated, that remodeling of the viral chromatin precedes transcriptional activation and is a prerequisite to generate a transcription competent template.

2 Introduction

2.1 DNA packaging into chromatin

The biological information for development, growth and reproduction is stored in the genome of an organism. The genome consists of deoxyribonucleic acid (DNA), which is built up from repeating units called nucleotides. Nucleotides are composed of a phosphate group, a sugar molecule and a nitrogen containing base. The order of the four bases adenine (A), thymine (T), guanine (G) and cytosine (C) encodes the genetic information in the DNA polymer. In eucaryotic cells, DNA is stored in the cell nucleus. The genome in a diploid human cell consists of 6 billion base pairs (bp) of DNA split up into 23 chromosome pairs. In total the linear DNA polymer is ~ 2 m in length and has to be folded in order to fit inside the nucleus with a diameter of ~ 10 µm.

Proteins associated with DNA promote a 10,000 fold compaction of the genome and facilitate the spatial and temporal organization in the nucleus. The complex of DNA and proteins is called chromatin. Spread chromatin appears in electron micrographs as “beads on a string” showing the basic building unit of chromatin, the nucleosome (Figure 1A) (Olins et al., 1975).

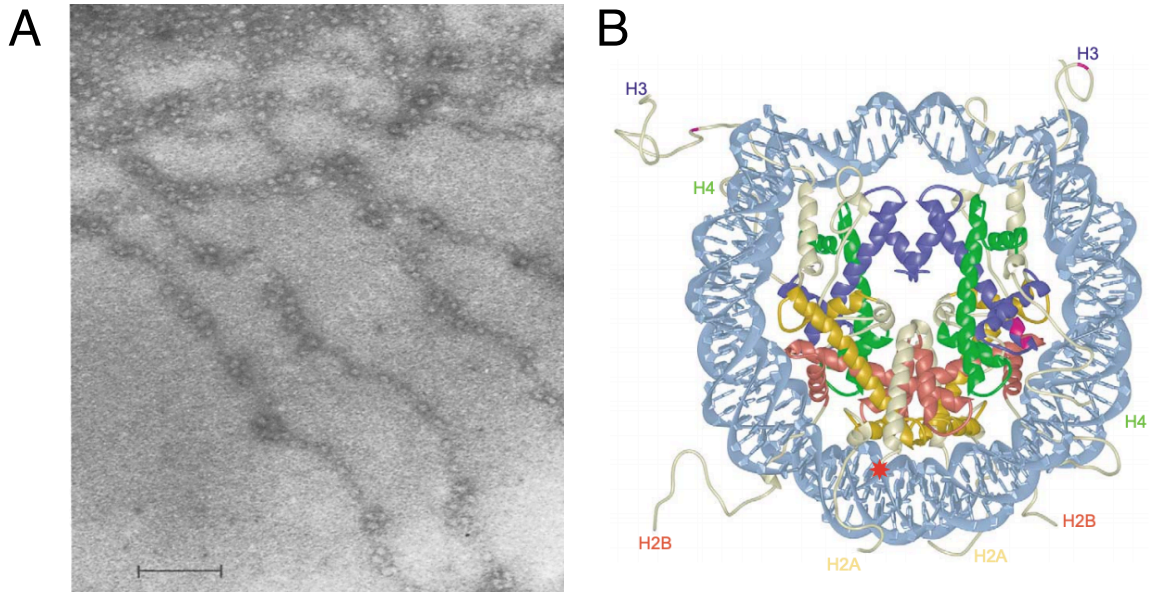


Figure 1: 10 nm fiber and structure of the nucleosome core particle

A: Electron micrograph of chromatin spread showing the “beads on a string structure” of the 10 nm fiber. Size marker: 0.1 µm (modified after Olins et al., 1975).
B: Front view of the nucleosome core particle. DNA (light blue) is wrapped around a histone octamer comprised of two each of histone H2A (yellow), H2B (red), H3 (dark blue) and H4 (green) (after Luger, 2003).

The nucleosome consists of the nucleosome core particle and the connecting linker DNA. In 1997 the 3-dimensional structure of the nucleosome core was resolved illustrating 147 bp of DNA wrapped around the central organizing protein unit, the histone octamer (Figure 1B) (Luger et al., 1997). The histone octamer is composed of two each of histone H2A, H2B, H3 and H4 forming a disc shaped particle (Kornberg, 1974). A human cell contains about 30 million nucleosomes establishing the first layer of DNA compaction, the 10 nm fiber.

2.2 Hierarchical folding of chromatin

To ensure the coordination of vital events, such as DNA transcription, repair, recombination and replication, the genome is spatially organized in the nucleus. In 1928, Emil Heitz observed different levels of DNA compaction in interphase moss cells and coined the terms hetero- and euchromatin (Passarge, 1979). Heterochromatin refers to parts of chromosomes, which remain condensed throughout interphase and are located at distinct areas of the nucleus, such as the nuclear periphery or the surrounding of the nucleolus (Figure 2). Heitz suggested that heterochromatin corresponds to genetically inert regions, whereas euchromatin unfolds in telophase and is connected to gene activity.

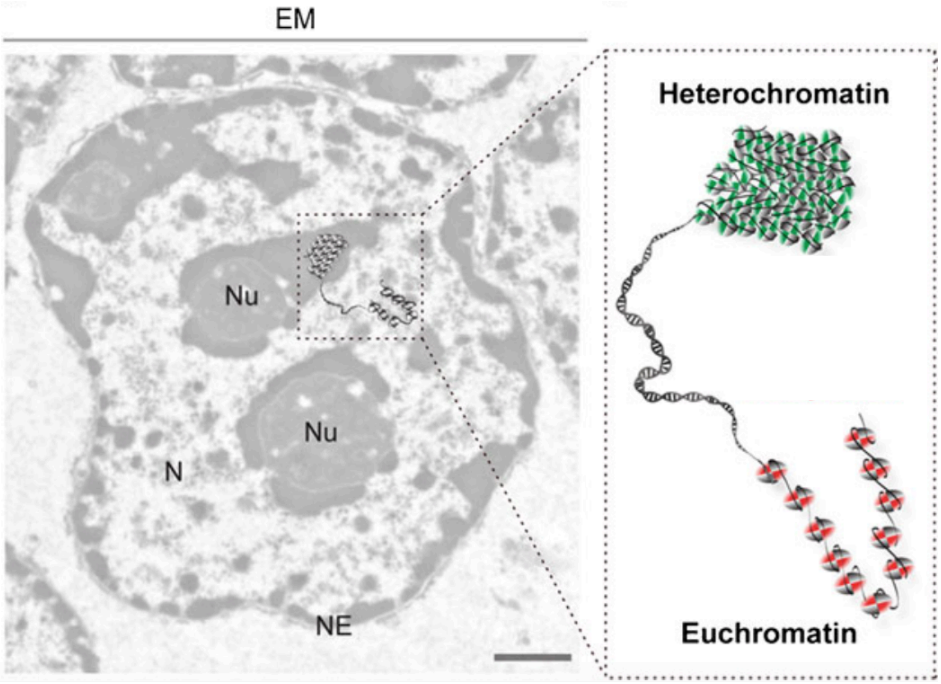


Figure 2: Hetero- and Euchromatin

Electron microscopy (EM) of a cell nucleus (N). Heterochromatin appearing as electron dense structures is concentrated at the nuclear envelope (NE) and the nucleolus (NE). In contrast euchromatin is further unfolded into a more open state and distributed within the nucleus (adapted after Jost et al., 2012).

Methods have been developed toward unraveling how chromatin is folded in three dimensions and which are the functional consequences of such folding. Transmission electron microscopy (EM) of purified nucleosome fibers revealed, that under certain conditions the 10 nm fibers are further condensed into regular higher order structures with a diameter of 30 nm (Finch and Klug, 1976) (Figure 3A). It was proposed that the nucleosome fiber folds into a solenoid composed of six successive nucleosomes every 11 nm (Figure 3B, left panel). The addition of Mg^{2+} and the presence of linker histone H1, which binds predominantly to the linker DNA, was required to stabilize the 30 nm fiber.

Later on an alternative folding model of a two start helix was proposed showing a zig-zag conformation (Woodcock et al., 1984) (Figure 3B, right panel). In addition, further structural variations of the 30 nm fiber have been suggested depending on the conditions used (van Holde and Zlatanova, 2007).

Based on the findings of the 30 nm fiber, it is thought that chromatin is hierarchically folded into discrete higher-order structures (Figure 3C). The hierarchical chromatin folding model proposes, that the 30 nm fiber assembles into helically folded 120 nm chromonema, which are further condensed into 300-700 nm chromatids, ultimately building up the mitotic chromosome (Belmont et al., 1999).

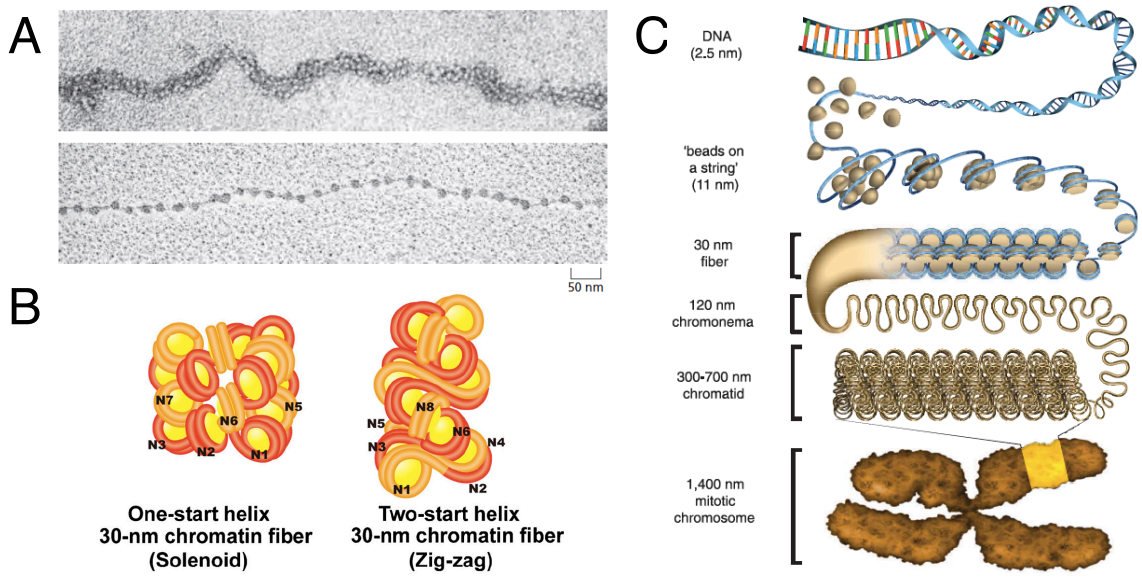


Figure 3: The 30 nm fiber and the hierarchical chromatin folding model

A: EM of the 30 nm fiber (top panel) and the 10 nm fiber (bottom panel) (after Alberts et al., 2002)
B: Model of the one-start helix (solenoid, left panel) and the two-start helix (zig-zag, right panel) illustrating possible nucleosome conformations in the 30 nm fiber. Labels (N1-N8) mark the positions from the first to the eight nucleosome (after Maeshima et al., 2014).
C: Schematic view of the hierarchical chromatin folding model (after Ou et al., 2017).

2.3 Chromatin organization *in vivo* without the 30 nm fiber

Even though, the 30 nm fiber could be readily assembled with reconstituted nucleosomal arrays (Dorigo et al., 2004), the structure of the 30 nm fiber could not be identified inside the nucleus of living cells and its existence *in vivo* was questioned. As cryo-EM of rapidly frozen and sectioned mitotic cells could not reveal any higher order periodic structures, such as the 30 nm fiber, Dubochet and colleagues suggested, that chromatin inside the nucleus behaves rather like a 'liquid drop', which is formed by aggregation of the 10 nm nucleosome fibers into a 'sea of nucleosomes' (Dubochet et al., 1986). The low concentration of cations and nucleosomes present in *in vitro* experiments might promote intra-fiber nucleosome interactions and simultaneously destabilize inter-fiber nucleosome contacts. Therefore, neighboring nucleosomes tend to stack together *in vitro*. In contrast, under physiological conditions nucleosome fibers interdigitate, thereby preventing the formation of higher order structures (Maeshima et al., 2019). Consequently, chromatin is not a static regular structure, as suggested by the 30 nm model, but rather disordered and interdigitated, like a 'polymer melt', enabling the dynamic and fluidic properties of chromatin.

2.4 Chromatin dynamics and accessibility

Chromatin movement and dynamics were monitored in live-cell imaging studies using fluorescent labeling techniques (Babokhov et al., 2020). In agreement with a liquid-like behavior of chromatin, chromatin is highly dynamic and exhibits local movements. Remarkably, high-resolution diffusion mapping revealed no differences of chromatin mobility in eu- and heterochromatic regions of the genome (Shaban et al., 2020). Only heterochromatin at the nuclear periphery, where movement is restricted by the close contact with the nuclear lamina, exhibits slower diffusion rates. The flexibility and steady movement of chromatin keeps DNA accessible for small molecules, even in condensed chromatin (Hihara et al., 2012). Molecules of as large as 90 nm in diameter experience unrestricted access to all sites within the nucleus and require a few seconds to roam the whole nucleus (Bancaud et al., 2009). In that way DNA directed processes, such as DNA replication, repair and transcription, are ensured in the entire genome, independent of chromatin compaction.

2.5 Higher-order chromatin architecture

Despite the irregularity of nucleosome fibers, chromatin is not distributed randomly in the interphase nucleus, but shows a higher-order organization. Chromosome painting using

fluorescence *in situ* hybridization (FISH) beautifully showed, that each chromosome occupies a defined volume of the nucleus, so called “chromosome territory” (Figure 4, left panel) (Cremer and Cremer, 2001). Interestingly, the radial position of the chromosomes correlates with its gene density, as gene rich chromosomes are mainly located in the center of the nucleus, whereas gene-poor chromosomes, are concentrated at the nuclear periphery (Boyle et al., 2001). With the development of methods to capture chromatin conformation, additional layers of higher-order chromatin organization were revealed. Chromatin forms two distinct compartments of several Mb in size, called A and B compartment (Figure 4, middle panel). The A compartment, corresponds to accessible and actively transcribed chromatin and is spatially segregated from the B compartment, which corresponds to closed and transcriptionally silent chromatin (Lieberman-Aiden et al., 2009). Folding of chromatin into large compartments could be well described by the fractal globule model (Grosberg et al., 1988). The fractal globule is formed by a polymer, which iteratively crumples into small globules, which then themselves associate into larger globules (Figure 4, right panel). In that way a DNA polymer can be highly condensed into a knot free conformation, which allows rapid unfolding and refolding of any genomic locus.

Consistent with the fractal globule model, the compartments are partitioned into smaller functional units of the genome, called ‘topologically associating domains’ (TADs) (Dixon et al., 2012). TADs are regions of the genome with a size range of ~ 100 kb to ~ 1 Mb, which exhibit higher DNA interaction frequencies within themselves relative to regions outside the given domain. Enhancer promoter interactions constitute a major class of long range DNA interactions and are limited to elements within the same TAD.

TADs emerge from a process, called DNA loop extrusion (Fudenberg et al., 2016). After cohesin, a ring-shaped complex, is loaded onto chromatin, a DNA loop is created, which increases in size by DNA extrusion in both directions (Kim et al., 2019). TAD boundaries are determined by CTCF, which acts as an extrusion barrier and cohesion is halted when it reaches DNA bound CTCF protein in opposing direction. Thus CTCF functions as an architectural DNA binding factor and shapes the genome topology.

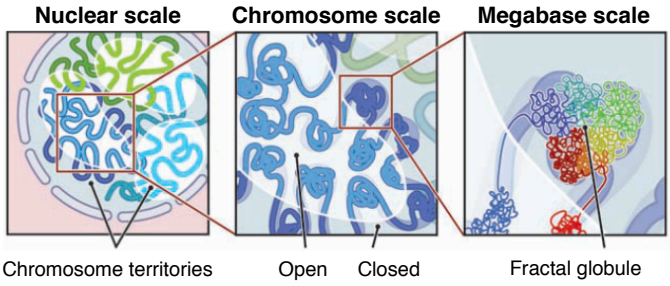


Figure 4: Higher order chromatin organization

Schematic overview of genome architecture at different scales. Chromosomes occupy distinct territories of the nucleus (left). The genome is spatially partitioned into an open and a closed chromatin compartment (middle). Chromosomes are folded into a series of fractal globules (right) (adapted after Lieberman-Aiden et al., 2009).

2.6 Structural and functional organization of the cell nucleus

The structural organization of the nucleus is connected with its functional tasks. The ANC-INC model integrates results from electron microscopy, fluorescence microscopy and correlation spectroscopy and aims to explain which role the nuclear organization plays in gene expression and cellular function (Cremer et al., 2020). According to the ANC-INC model the nucleus is partitioned into two co-aligned, compartments, an active nuclear compartment (ANC) and an inactive nuclear compartment (INC). The ANC is mainly comprised of decondensed chromatin at the periphery of chromatin domains, and an interchromatin channel system pervading the nuclear space, called the interchromatin compartment (IC). The DNA density in the IC is very low and the IC can been seen in analogy of a supply route, where nuclear distribution, import and export of macromolecules takes places. Furthermore, it is suggested that macromolecular machineries required for repair, transcription or replication are preferentially located in the IC and are excluded from the condensed inactive compartment due to their size. Nevertheless, the interaction between the ANC and INC is thought to be dynamic and chromatin loops or entire chromatin domains can rapidly change the compartments ensuring immediate gene activation, DNA replication or DNA repair.

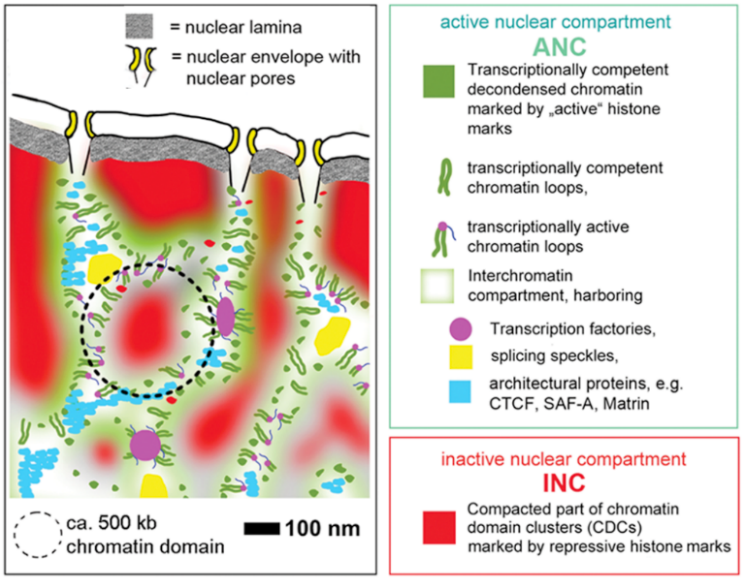


Figure 5: Functional nuclear architecture

The ANC-INC network model of nuclear organization. The interchromatin compartment, a 3 dimensional channel network pervades the nuclear space and functions as a supply route. Macromolecular machineries and functional protein clusters, such as transcription factories or splicing speckles, reside mainly in the interchromatin compartment and can access decondensed chromatin at the periphery of chromatin domains, whereas they are excluded from the more compacted inner chromatin domain (adapted after Cremer and Cremer, 2019).

3 Objectives

Inside the nucleus DNA is associated with proteins forming a complex called chromatin. For understanding the regulatory function of chromatin, it is essential to understand how the genome is structurally organized. This study focuses on the arrangement of nucleosomes along the DNA and the folding of the nucleosome fiber into higher order structures of chromatin.

3.1 Characterizing DNA accessibility in higher-order structures of chromatin

As first level of compaction DNA is packaged into nucleosomes to from a 11 nm nucleosome fiber. Classical models propose further a hierarchical folding of the nucleosome fiber into higher order structures of condensed chromatin. So far, it remains unclear, where in the genome higher order structures of chromatin are formed and how they might affect DNA dependent processes like transcription or replication. In this study a differential MNase digestion approach in native chromatin of human cells, called Functional Chromatin Extraction (FCE), is set up. This method aims to classify open and condensed chromatin domains and to annotate their genomic positions. A simplistic model simulating the stochastic cleavage of the endonuclease MNase on the chromatin fiber will be built to predict the preferential release of accessible domains in partially digested chromatin. Furthermore, MNase sequence preferences will be assessed and the obtained nucleosome profiles will be normalized to compensate for potential biases and to allow in-depth characterization. To get a better understanding of the structure-function relation of chromatin, chromatin accessibility will be correlated with histone tail modifications, chromatin states, DNase hypersensitive sites, transcription factor binding and gene expression.

3.2 Analyzing Adenovirus chromatin dynamics early in infections

DNA viruses utilize chromatin remodeling machineries and DNA packaging components of the host cell to initiate viral transcription and facilitate replication. While in eukaryotes DNA is wrapped around a histone octamer to form nucleosomes, adenoviral DNA in infecting virions is packaged mainly by the protamine-like protein VII (pVII) into nucleoprotein complexes. Early after nuclear entry histones are deposited onto incoming viral genomes. Until now pVII organization onto the viral genome is not analyzed in detail and it remains elusive how the viral chromatin adopts the cellular nucleoprotein structures to form a competent template for viral

gene transcription and genome replication. This study focuses on the dynamic changes of viral nucleoprotein structure from virus infection until the onset of transcription within the first four hours of adenovirus infection. Functional Chromatin Extraction (FCE) is applied to characterize pVII positioning and nucleosome assembly onto the viral genome with high spatial and temporal resolution. The relation of nucleosome assembly and loss of pVII will be assessed. Furthermore, preferential isolation of accessible chromatin in FCE allows to monitor Adenovirus chromatin decondensation during early phase of infection. In combination with RNA-seq and single-cell imaging, dynamic changes of viral nucleoprotein structure are correlated with nuclear entry and transcriptional activation.

4 Characterizing DNA accessibility in higher-order structures of chromatin

4.1 Results

4.1.1 Functional Chromatin Extraction: Isolation of mono-nucleosomal DNA from native chromatin using differential MNase hydrolysis

The endo-exonuclease MNase preferentially hydrolyzes accessible DNA as present in the linker region between nucleosomes, whereas hydrolysis of DNA wrapped around a histone octamer is restricted (Noll et al., 1975). Therefore, MNase is widely used to quantitatively excise nucleosomes from chromatin. In combination with subsequent DNA high-throughput sequencing (MNase-seq), nucleosome positions can be mapped with high accuracy (Zentner and Henikoff, 2014). Since standard MNase-seq experiments fully hydrolyze chromatin to mono-nucleosomal DNA, information about higher order chromatin organization is lost. To identify different levels of chromatin packaging as suggested by the hierarchical folding model a new protocol, termed Functional Chromatin Extraction (FCE), was established. In this approach different concentrations of MNase were applied to digest native chromatin. After mild permeabilization, HeLa cells were incubated either with 1000 U of MNase (high-MNase) quantitatively hydrolyzing chromatin to a major fraction of 80-90% mono-nucleosomes or 100 U of MNase (low-MNase) releasing 10-20% mono-nucleosomes from cellular chromatin (Figure 6A). Accordingly, if different levels of chromatin packaging exist in the cell, it would be expected, that preferentially active and open chromatin domains are released in low-MNase, whereas condensed chromatin structures are more resistant to MNase digestions and are only released using high-MNase (Figure 6B).

To test this hypothesis FCE was carried out in two independent experiments in HeLa cells. The mono-nucleosomal DNA band was excised from an agarose gel (Figure 6A) and after library preparation, DNA fragments were subjected to high-throughput paired-end sequencing on Illumina HiSeq platform. The quality of sequencing reads was assessed and read pairs defining both ends of the isolated DNA fragments were mapped to the reference human genome assembly. Thus, the size of the DNA fragments of FCE could be analyzed with base-pair resolution (Figure 6C). As expected, the length of DNA isolated from high- and low-MNase digested samples was on average close to 150 bp, which resembles the length of DNA wrapped around the histone octamer.

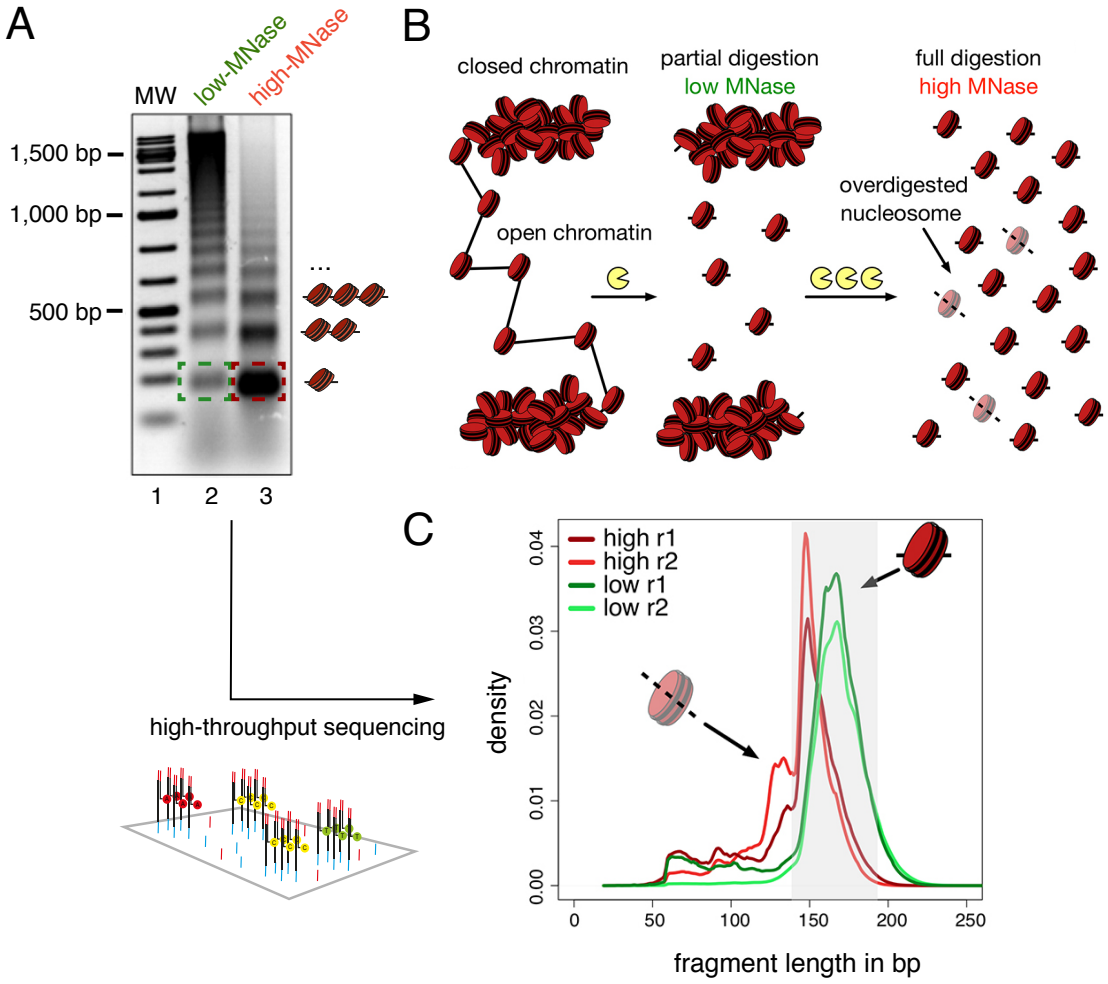


Figure 6: Functional Chromatin Extraction: Isolation of mono-nucleosomal DNA from native chromatin using differential MNase hydrolysis

A: Agarose gel showing the nucleosomal DNA ladder obtained from Functional Chromatin Extraction (FCE) in HeLa cells. Cells were incubated with either 100 U (low-MNase in lane 2) or 1000 U (high-MNase in lane 1) of MNase. Mono-, di- and tri-nucleosomal DNA fragments are indicated and the molecular weight marker is shown in lane 1. Dashed rectangles highlight the excised mono-nucleosomal DNA band which was subjected to high-throughput sequencing (adapted after Diermeier, 2014).
B: Scheme illustrating the expected results of FCE. Preferentially DNA in accessible chromatin is hydrolyzed in partial chromatin digestions, whereas folded and compacted chromatin structures are released at higher MNase concentrations.
C: Kernel density plot showing the fragment size distribution of isolated DNA in FCE. Fragment sizes corresponding to intact mono-nucleosomal DNA are highlighted by a shaded box (140-200 bp). Fragments derived from partially digested nucleosomal DNA are indicated.

High-MNase samples exhibited a defined and narrow peak around 150 bp, whereas the size distribution of low-MNase samples was broader and slightly shifting to larger fragments, indicating incomplete digestion of linker DNA in low-MNase. Interestingly, high-MNase samples were characterized by an additional peak at the subnucleosomal size, suggesting over-digested nucleosomes in this condition. For further analysis only fragments in the range between 140 – 200 bp, resembling full nucleosomes, were kept and longer or shorter fragments were removed from analysis. In total 393 million annotated mono-nucleosomes were obtained

using FCE, which enabled an in-depth analysis of DNA accessibility in context of chromatin organization.

4.1.2 *In silico* simulation of differential chromatin digestion at condensed chromatin

To simulate the stochastic cleavage events of MNase along the nucleosomal array, an *in silico* model was built. First, a nucleosome fiber of 9,000 bp in length comprising 50 regularly spaced nucleosomes was designed. One potential cleavage site was assigned to each linker and each nucleosome core with a distinct MNase cleavage probability $p_{cut}(linker)$ and $p_{cut}(nuc)$. The probability of an inter-nucleosomal MNase cutting event was set 10x more likely compared to an intra-nucleosomal cleavage, i.e. $p_{cut}(linker) = 10 \cdot p_{cut}(nuc)$, according to previous MNase footprinting studies (Cockell et al., 1983). Consequently, n cutting events were stochastically distributed onto the generated nucleosome fiber and selected cleavage sites were labeled (Figure 7A). Next, the nucleosome strand was fragmented according to the labels and the size of the fragments was deduced to select isolated mono-nucleosomes. The position and the number of the mono-nucleosomes was stored. This process was iterated until at least 1 million mono-nucleosomes have been collected.

To simulate the MNase digestion of varying enzyme concentrations the number of cutting events n at each template was adjusted according to the high-MNase and low-MNase conditions. Setting 70 cutting events (70% of all possible sites are cut) for high-MNase and 20 (20% of all possible sites are cut) for low-MNase resembled the fragmentation profile observed in the experiments best (Figure 7B, compared to Figure 6A). Running the simulation and plotting the distribution of collected mono-nucleosomes along the nucleosome fiber did not reveal any significant differences between low- and high-MNase (Figure 7C). The results indicate that there will be no preferential enrichment of mono-nucleosomes comparing high- and low-MNase, if the linker accessibility is homogenous in the whole genome.

Next, to model a densely packaged chromatin domain, a reduced inter-nucleosomal cleavage probability, $p_{cut}(linker, compact) = 0.8 \cdot p_{cut}(linker)$, was introduced and assigned to 25 linker in the center of the nucleosome fiber. The compacted chromatin domain was clearly visible in the distribution of isolated mono-nucleosomes of the low-MNase simulation and mono-nucleosomes of open chromatin regions are preferentially enriched (Figure 7D). In contrast, using the high-MNase simulation the compacted chromatin domain could not be detected.

In summary, the computer simulation confirmed the initial assumption, that comparing partially and fully digested chromatin should reveal the position of compacted chromatin domains.

Furthermore, the model shows, that mainly the low-MNase condition would be sensitive to differences in hierarchical chromatin organization and the commonly used quantitative chromatin digestion (high-MNase) would not allow to differentiate between various levels of chromatin compaction

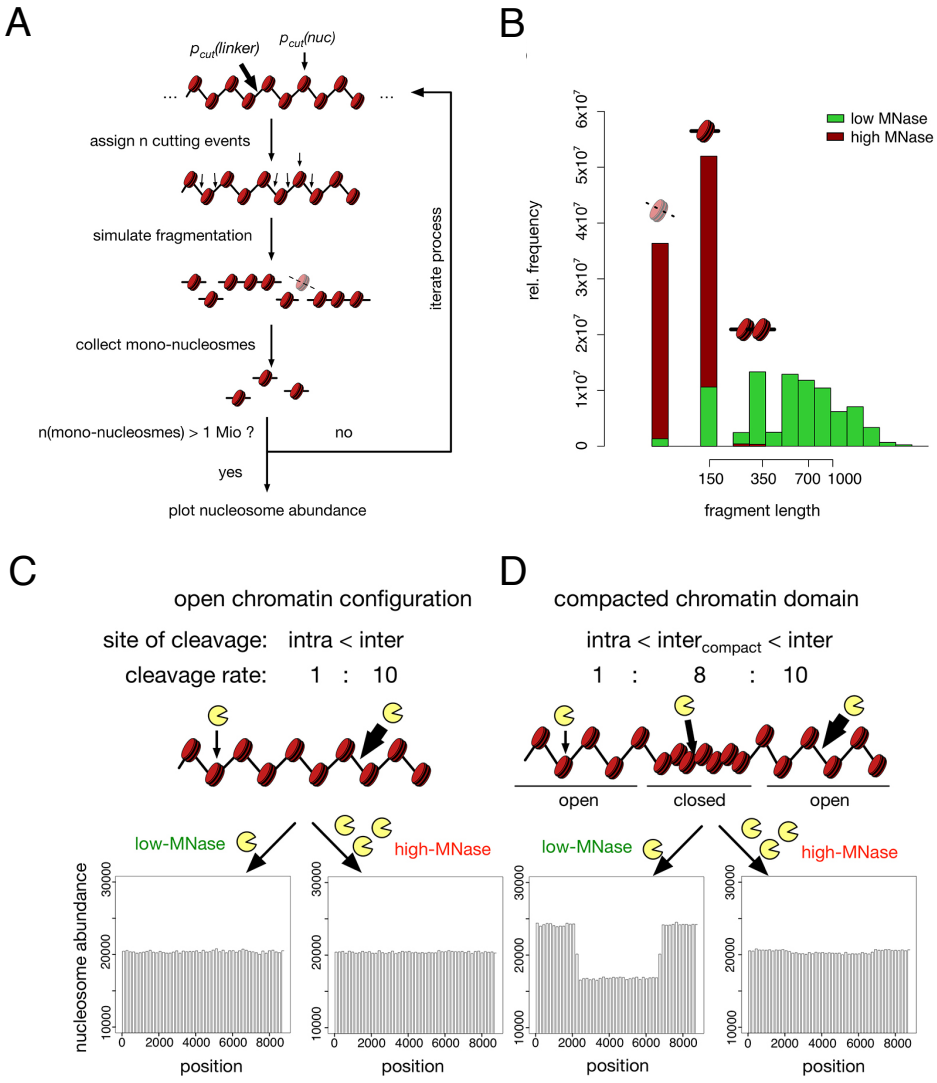


Figure 7: *In silico* simulation of a condensed chromatin domain

A: Scheme illustrating the steps of the *in silico* MNase digestion. The probability of intra-nucleosome cleavage was set to be 10x higher than the probability of inter-nucleosome cleavage ($p_{cut}(linker) = 10 \cdot p_{cut}(nuc)$). N sites of a nucleosome array were stochastically selected based on the assigned cleavage probabilities at nucleosomes ($p_{cut}(nuc)$) and at linker DNA ($p_{cut}(linker)$). The nucleosome array was fragmented at the selected sites and fragments representing mono-nucleosomes were collected. The process was iterated until 1 Mio mono-nucleosomes have been isolated.
B: Histogram showing the fragment size distribution after simulating n=70 (high-MNase, red bars) and n=20 (low-MNase, green bars) cutting events along the nucleosome array.
C: Simulation of high- and low-MNase digestions along an equally accessible nucleosome array. The probability of inter-nucleosome cleavage was held constant.
D: Simulation of high- and low-MNase digestions along a nucleosome array comprising a central condensed chromatin domain. The cleavage probability at the central 25 linker was reduced ($p_{cut}(linker, compact) = 0.8 \cdot p_{cut}(linker)$).
C-D: Bottom panel showing the distribution of nucleosomes either from the high- or low-MNase simulation. Each bar represents a nucleosome occupied position along the nucleosome array

4.1.3 Global differences in low- and high-MNase

According to the *in silico* simulation, regions of differential chromatin accessibility can be identified comparing the distribution of nucleosomes released with either low- or high-MNase. Hence, the genome wide distribution of extracted nucleosomes was evaluated (Figure 8A genome wide and Figure 8B along chr6). Overall both conditions showed a homogenous distribution across the genome (Figure 8A and B first two panels). At first glance, domains showing obvious dissimilarities between both conditions could not be identified. Therefore, the number of mono-nucleosomal fragments of either high- or low-MNase was counted in 1 Mb non-overlapping bins. To highlight differences the log fold-change was calculated, positive values corresponding to an enrichment of isolated nucleosomes in high-MNase and vice versa negative values in low-MNase (Figure 8A-B middle panel). Indeed, using this approach, domains spanning several Mbs could be clearly detected showing a preferential release of nucleosomes from chromatin with either high- or low-MNase.

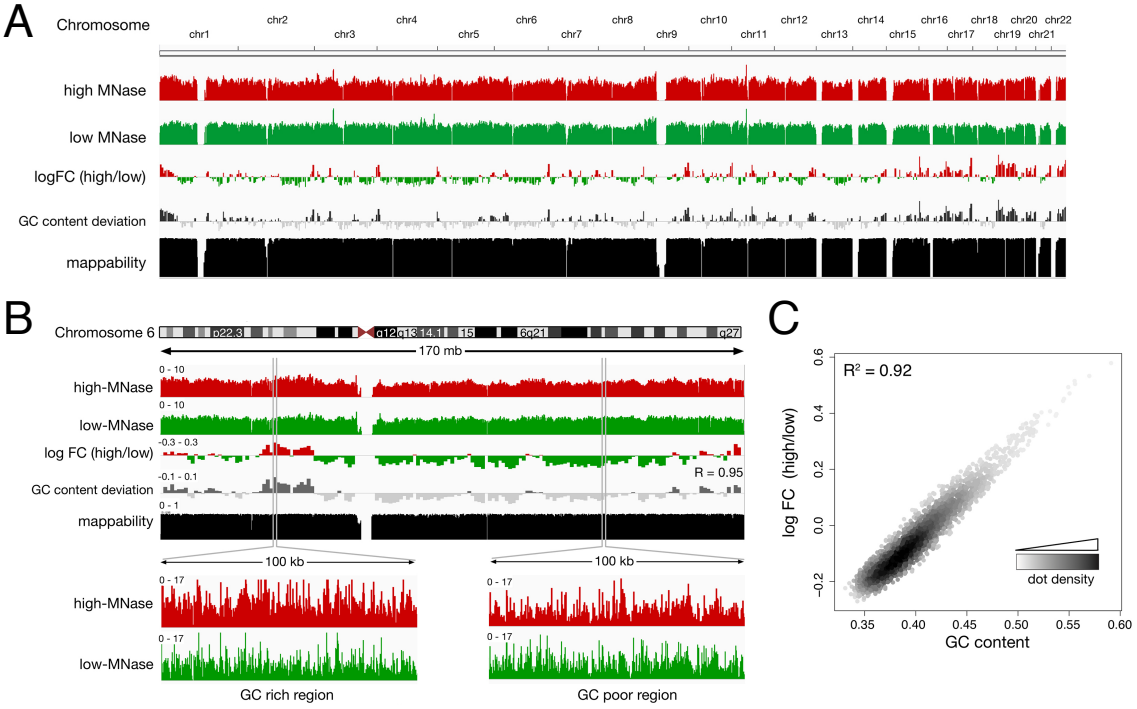


Figure 8: Global differences in low- and high-MNase

A-B: Genome browser snapshots showing the normalized mono-nucleosomal fragment counts that are obtained with either high- or low-MNase. The fragment frequencies were assessed in 1 Mb non-overlapping bins. Differences between high- and low-MNase are highlighted in the middle track as $\log_2$ fold changes (log FC). The GC content deviation from the genome wide average (40.5%) is shown in the fifth track. The last track displays the average mappability of 100mer sequences within the respective 1 Mb sized bin (Derrien et al., 2012). The position in the genome is indicated at the top
C: Scatterplot illustrating the genome wide correlation of GC content and differences in high- and low-MNase. Dots correspond to 1 Mb non-overlapping bins (as shown in A and B) and dot density is indicated by different shades of grey. The coefficient of determination (R^2) was obtained from linear regression analysis between GC content and the $\log_2$ ratio of high-MNase versus low-MNase (log FC (high/low)).

In agreement with the computer simulation, one plausible reason for the specific enrichment over large genomic regions could be domains with different chromatin accessibility confirming a hierarchical organization of chromatin with different layers of chromatin compaction.

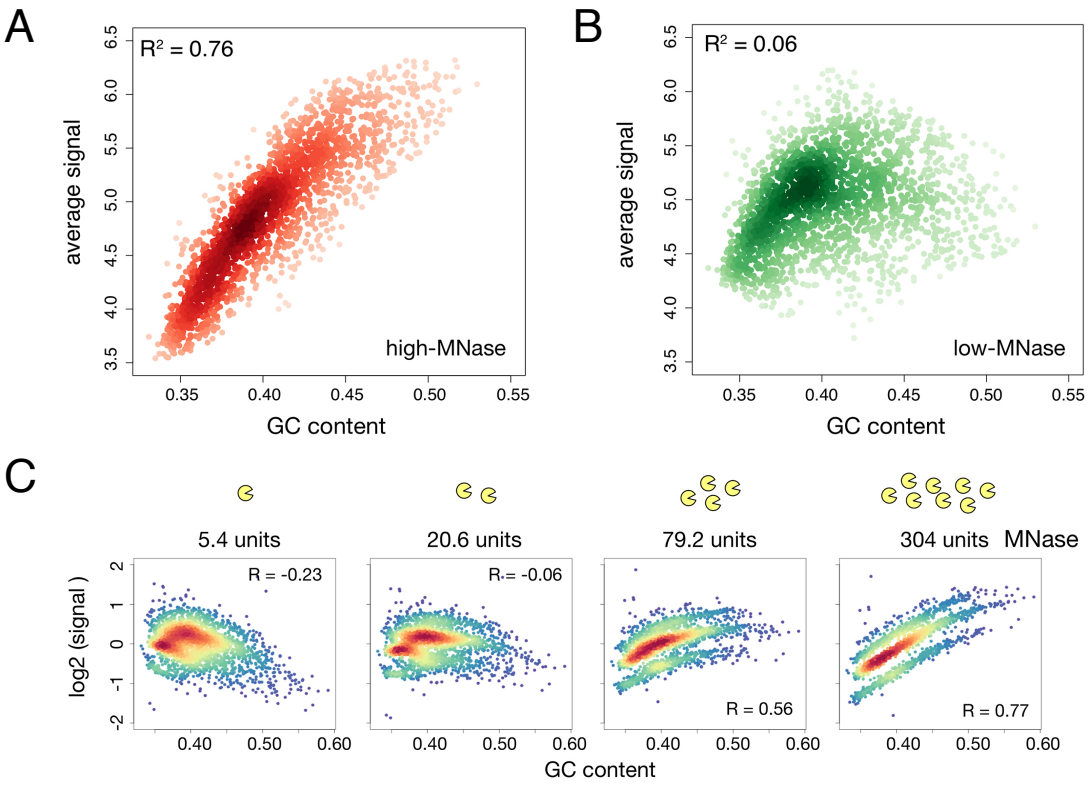
Since MNase has a known sequence preference preferentially hydrolyzing DNA at A/T nucleotides (Ishii et al., 2015), it was tested whether this attribute might bias the results and therefore the average GC content for each bin was assessed. Indeed, the differential isolation of nucleosomes in FCE exhibited a high correlation (R^2 = 0.92) to the variations of the underlying GC content (Figure 8C). Regions enriched for high-MNase released nucleosomes are GC rich and low-MNase enriched regions are AT rich.

Taken together, domains showing differential enrichment in FCE could be detected. However, comparison to GC content indicated that MNase sequence preferences might mask the classification of chromatin domains based on chromatin accessibility. Furthermore, in detail inspection of domains with either high or low GC content revealed a rather homogenous signal in low-MNase in contrast to high-MNase, which exhibited GC dependent variations (Figure 8B bottom panel), indicating a distinct impact of DNA sequence on the isolation of nucleosomes in high- and low-MNase.

4.1.4 MNase sequence preferences are prominent in high-MNase, but not in low-MNase

To assess the global dependency of DNA sequence on the results of FCE, high- and low-MNase samples were individually correlated with the GC content. Therefore, the number of nucleosome sized fragments in 1 Mb bins was calculated and plotted against the average GC content in that bin (Figure 9A and Figure 9B). In the high-MNase sample the number of extracted nucleosomes was well correlated with the GC content (R^2 = 0.76) (Figure 9A). In contrast, the low-MNase condition was mainly unaffected by DNA sequence (R^2 = 0.06), indicating that preferential hydrolysis of DNA at A/T nucleotides is enhanced with higher MNase concentration.

To further dissect the relationship of MNase concentration and DNA sequence composition, published MNase-seq data comprised of four different MNase concentrations were examined (Mieczkowski et al., 2016) (Figure 9C). The reanalysis confirmed the observation in FCE, that with increasing MNase concentration the extraction of mono-nucleosomes from GC rich regions is preferred. Interestingly, low-MNase conditions are less sensitive to variations in GC content.



Homo sapiens K562 cells; data taken from Mieczkowski et al. (2016)

Figure 9: MNase sequence preferences are prominent in high-MNase, but not in low-MNase

A-C: Scatterplots showing the correlation of GC content and MNase condition. Dots represent the average mono-nucleosome frequency in 1 Mb non-overlapping bins. Dot density is indicated by color. The coefficient of determination (R^2) was obtained from linear regression analysis between GC content and the normalized mono-nucleosome frequency in **(A)** high-MNase or **(B)** low-MNase.
C: Published MNase-seq data using various MNase concentrations in K562 cells were reanalysed (Mieczkowski et al., 2016) and the correlation coefficient R was determined. The corresponding MNase concentration is indicated at the top.

4.1.5 AT rich nucleosomes are over-digested in high-MNase

To test whether the observed sequence preferences in high-MNase would as well impact the annotation of nucleosome positions, nucleosomes of high- and low-MNase were called using the DANPOS software suite (Chen et al., 2013b) (Figure 10A, indicated by blue boxes). Overall the predicted nucleosome positions were highly concordant in high- and low-MNase confirming the reproducibility of this approach. Notably, less nucleosomes were predicted in AT rich regions for the high-MNase condition (Figure 10A) and the number of predicted nucleosomes correlated globally with the GC content ($R^2 = 0.76$) (Figure 10B). In contrast, the low-MNase condition showed a more regular distribution of nucleosomes across the whole genome (Figure 10D) independent of DNA sequence ($R^2 = 0.07$) (Figure 10C). Thorough comparison of high- and low-MNase revealed that nucleosomes are mainly depleted from AT rich sites in high-MNase condition (Figure 10A black arrows). The known AT preference of

MNase and the reduced capacity to call nucleosomes at GC poor regions suggests that AT rich nucleosomes are over-digested in high-MNase.

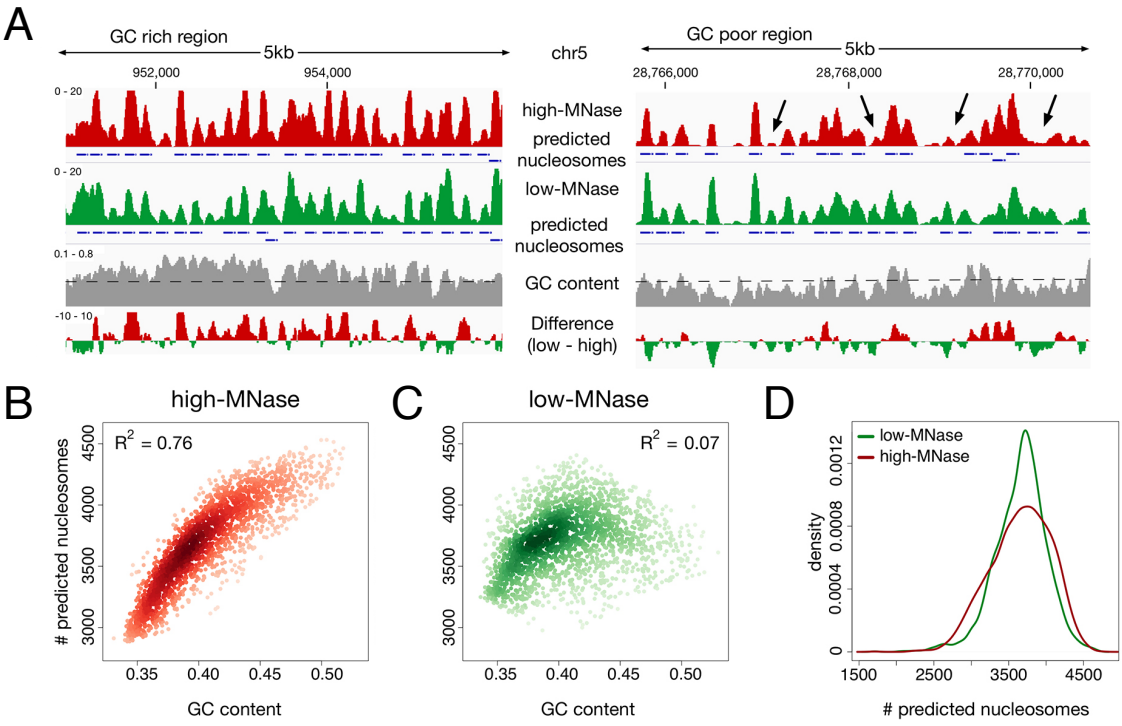


Figure 10: AT rich nucleosomes are over-digested in high-MNase

A: Genome browser snapshots illustrating the annotation of nucleosome positions at a GC rich and a GC poor locus. Nucleosome occupancy profiles are generated from the normalized mono-nucleosomal fragment counts of either high- or low-MNase digestions. Blue boxes display called nucleosome positions using the DANPOS2 toolkit (Chen et al., 2013b). GC content was calculated in sliding windows using 50 bp bin width and a step size of 10 bp. The genome wide average GC content is indicated by the dashed line.
B-C: Scatterplots showing the correlation of GC content and the number of predicted nucleosome positions. Dots represent the number of called nucleosome positions in 1 Mb non-overlapping bins. Dot density is indicated by color. The coefficient of determination (R^2) was obtained from linear regression analysis between GC content and the quantity of nucleosome prediction in (**B**) high-MNase or (**C**) low-MNase samples.
D: Kernel density plot illustrating the genome wide frequency distribution of nucleosomes detected in 1 Mb bins.

4.1.6 *In silico* simulation illustrates DNA sequence sensitivity of high-MNase

Since the *in silico* model did not account for the sequence preferences of MNase, it was further refined. According to the observation that mono-nucleosomes from AT rich regions are underrepresented in high-MNase (Figure 10) and the size distribution of high-MNase fragments exhibited an additional peak of over-digested mono-nucleosomes (Figure 6B) it is plausible, that AT rich nucleosomes are preferentially over-digested in high-MNase. To test this hypothesis, an increased intra-nucleosomal cleavage probability, $p_{cut}(nuc,AT+)$, was assigned to the 25 central nucleosomes of the *in silico* nucleosome fiber. Varying $p_{cut}(nuc,AT+)$ values

were tested, ranging from 1.5 to 3 fold increase relative to the standard intra-nucleosome cleavage probability, $p_{cut}(nuc)$. Using the settings to simulate the high-MNase condition and keeping the inter-nucleosome cleavage probability ($p_{cut}(linker)$) constant over all linker sites, nucleosomes from the AT rich region were depleted from the pool of isolated nucleosomes and the extent correlated with the applied $p_{cut}(nuc,AT+)$ values (Figure 11A). Remarkably, this effect was not visible in the low-MNase simulation (Figure 11B) and the nucleosome distribution was similar to the simulation without $p_{cut}(nuc,AT+)$ (Figure 7C). Taken together, the *in silico* model could mimic the results of the experiments and confirms that the commonly used high-MNase condition is sensitive to DNA sequence variations, whereas the low-MNase digestion is more robust and the distribution of extracted nucleosomes is unaffected by DNA sequence.

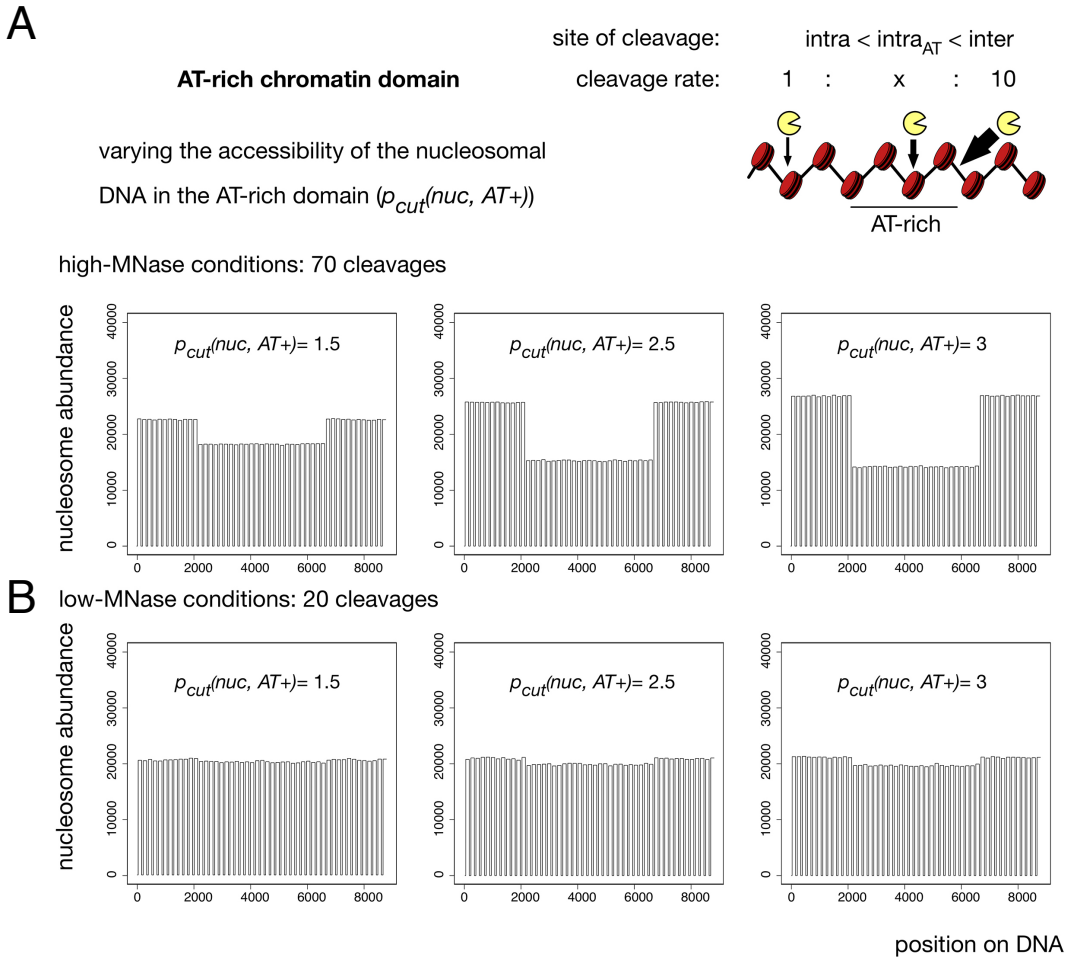


Figure 11: *In silico* simulation illustrates DNA sequence sensitivity of high-MNase

A-B: Simulation of high- and low-MNase digestions along an *in silico* generated nucleosome array comprising a central AT rich domain with varying accessibility of the nucleosomal DNA. The intra-nucleosomal cleavage probability at the central 25 nucleosomes occupying the AT rich DNA stretch was increased as indicated ($p_{cut}(nuc, AT+) = x \cdot p_{cut}(nuc)$). The probability of inter-nucleosome cleavage, $p_{cut}(linker)$, was held constant. Bar plots show the distribution of extracted nucleosomes either from the (**A**) high- or (**B**) low-MNase simulation. Each bar represents a nucleosome occupied position along the nucleosome array.

4.1.7 GC normalization corrects for MNase sequence preferences

Domains of differential chromatin accessibility might be masked by the dominant sequence preferences of MNase in the high-MNase condition. Therefore, it is required to normalize the data for the underlying GC content before comparing different MNase conditions. The normalized nucleosome frequency of each MNase experiment and the average GC content within 250 bp non-overlapping bins were assessed. Bins exhibiting low mappability and extreme GC content (< 30 % or > 70 %) were removed. Even after reducing the size of the bins from 1 Mb (Figure 8C) down to 250 bp, the differential MNase signal was well correlated with the underlying GC content (R=0.53) (Figure 12C, left panel). A locally weighted scatterplot smoothing (LOESS) was applied to model the global impact of the GC content on the individual MNase experiments (Figure 12A and B). The GC normalization factors were deduced from the LOESS regression and the nucleosome frequency of each bin was divided by the corresponding normalization factor. Afterwards, the normalized nucleosome frequencies of each experiment were no longer correlated with the GC content (Figure 12A and B, right panels). Consequently, this approach removed the GC bias from the ratio of low- and high-MNase (R=0.02) enabling the identification of differentially accessible chromatin domains (Figure 12C, right panel).

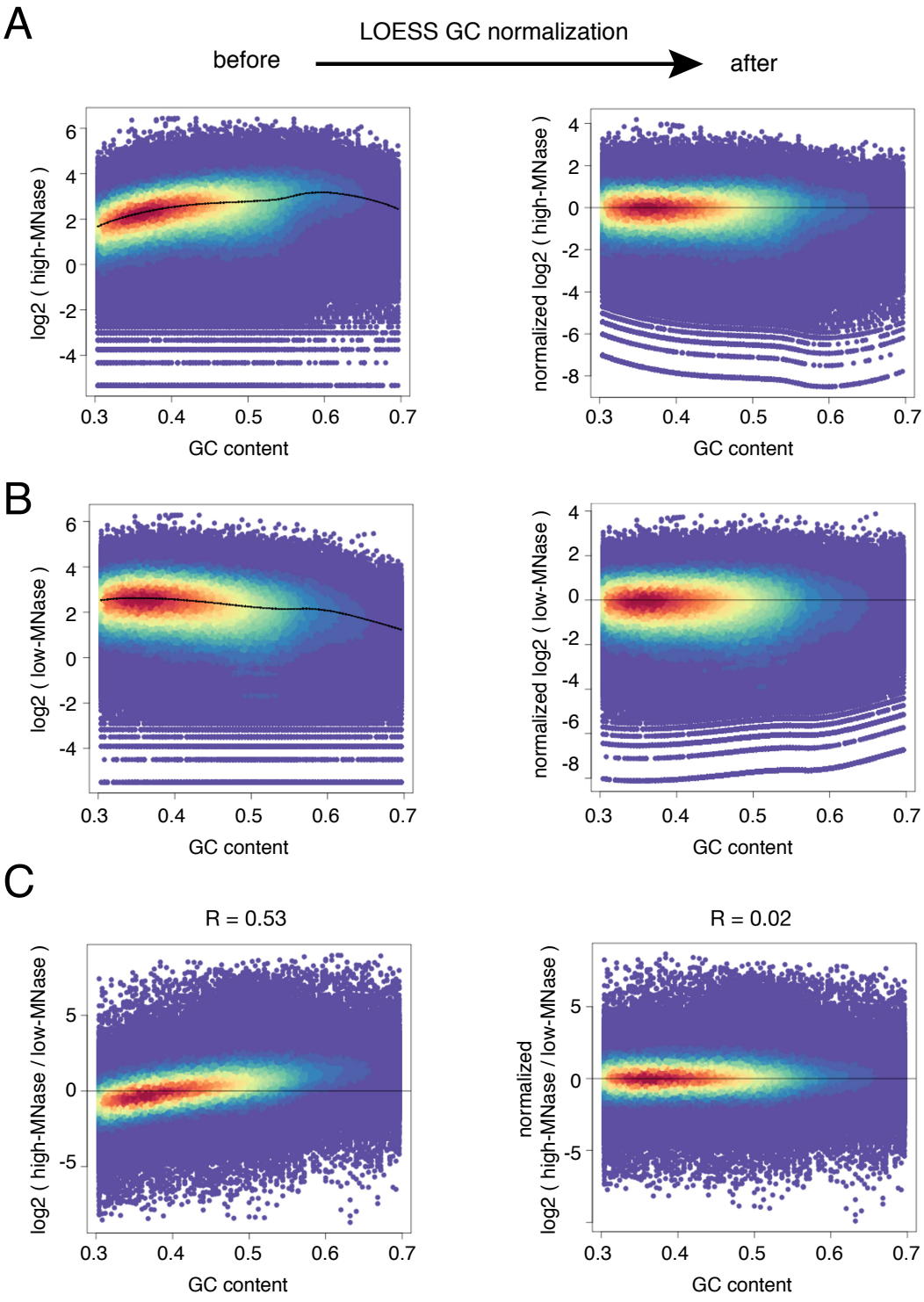


Figure 12: MNase GC normalization

A-B: Density scatter plots showing the correlation of (**A**) high- / (**B**) low-MNase profile with GC content before and after GC normalization. Dots represent the average frequency of nucleosome fragments in 250 bp bins. The GC dependency of each MNase experiment was estimated using a locally weighted scatterplot smoothing (LOESS) fit (left panel). The fitted LOESS curve is indicated as black line. GC normalization factors were deduced from the LOESS curve. The nucleosome fragment frequencies were divided by their corresponding normalization factors (right panel)
C: Correlation of MNase log2 fold-changes with GC content before and after the individual GC correction of high- and low-MNase experiments. The Pearson correlation coefficient is indicated at the top.

4.1.8 Repressed domains exhibit similar DNA accessibility as actively transcribed domains

To compare the chromatin accessibility of active and silent regions in the genome, the GC normalized FCE signal was correlated with characteristic histone modifications. Domains of facultative heterochromatin are associated with trimethylation at the 27th lysine residue of histone H3 (H3K27me3). A subunit of the polycomb-repressive complex 2 (PRC2), the histone methyltransferase EZH2, catalyzes the methylation of H3K27, which induces chromatin compaction and subsequently gene repression (Margueron and Reinberg, 2011). Therefore, this histone modification is thought to highlight chromatin domains associated with reduced DNA accessibility and according to the *in silico* simulation (Figure 7C, left panel) it is expected, that there nucleosomes are less efficiently released in low-MNase.

Mb-sized domains exhibiting high levels of H3K27me3 could be identified using ChIP-seq data from the ENCODE project (Figure 13A)(Davis et al., 2018; Dunham et al., 2012). However, H3K27me3 enriched domains did not show any specific enrichment in FCE. Even subdividing the domains into quartiles based on H3K27me3 intensity did not reveal any significant changes in FCE compared to regions lacking H3K27me3 indicating that chromatin of polycomb repressed domains is accessible. In contrast, acetylation of H3K27 (H3K27ac), which is mainly restricted to local sites in the genome and not spread over large domains, marks active enhancer and promoter. Thorough examination revealed that regions showing high levels of H3K27ac preferentially coincide with low-MNase enriched sites (Figure 13A, black rectangles). The global analysis confirmed this observation and showed a good correlation between H3K27ac density and chromatin accessibility (Figure 13B).

To further refine the analysis, annotated chromatin states from roadmap (Roadmap Epigenomics Consortium et al., 2015), which are defined based on the combinatorial pattern of several histone modifications (Ernst and Kellis, 2012), were used. In total 15 chromatin states are defined and the chromatin accessibility of each state was assessed calculating the average ratio of high-/low-MNase within (Figure 13C). In general, states mainly classified by negative values in FCE, as expected for compact and inaccessible chromatin, could not be identified. Even the heterochromatin associated states, like polycomb repressed domains, exhibit similar accessibility as euchromatic domains showing ongoing transcription (strong/weak transcription). Interestingly, local chromatin features as active transcription start sites (TSS) and enhancer showed the highest change in FCE and are located in chromatin with increased accessibility, which is preferentially released in low-MNase.

In summary, domains of inaccessible chromatin covering repressed and silent regions in the genome could not be detected in FCE. Furthermore, eu- and heterochromatin have similar accessibility indicating that chromatin is not organized globally in different levels of chromatin compaction. Remarkably, local chromatin features, like promoter and enhancer elements, are characterized with an increased accessibility in FCE.

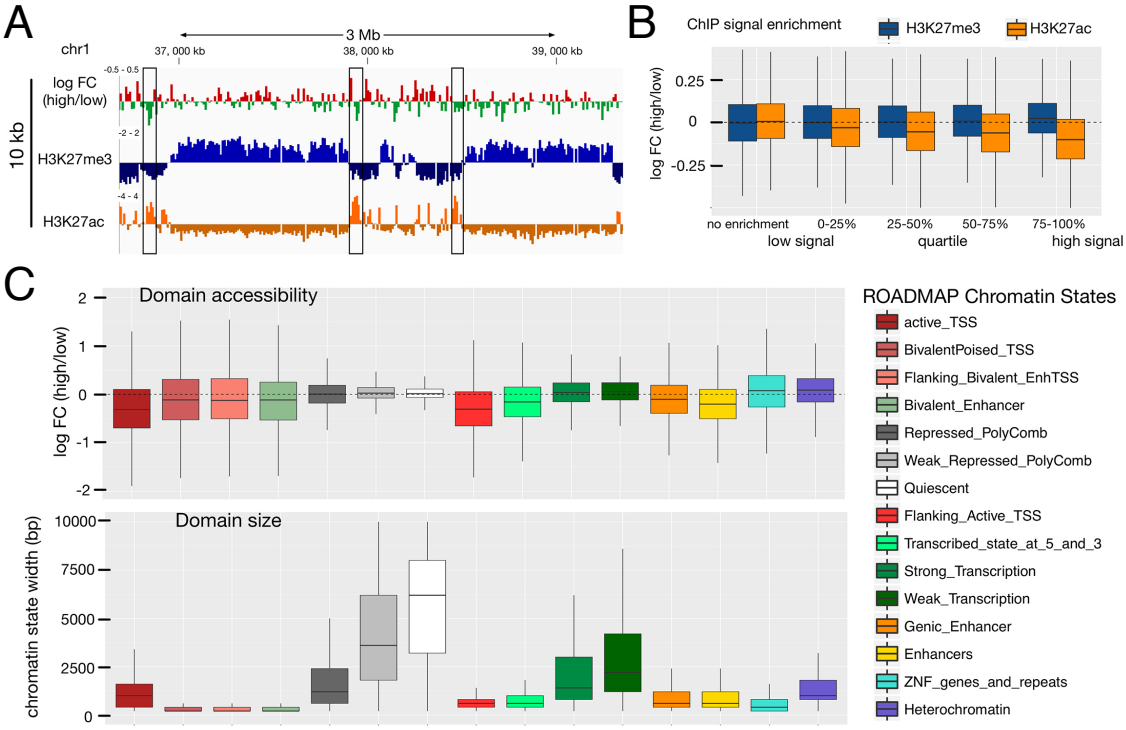


Figure 13: Repressed domains exhibit similar DNA accessibility as actively transcribed domains

A: Genome browser snapshot showing the GC normalized ratio of high-/low-MNase at H3K27me3 and H2K27ac marked regions. H3K27me3 (accession: ENCFF958BAN) and H3K27ac (accession: ENCFF311EWS) ChIP profiles are shown as log fold-change over input. The average signal was assessed in 10 kb non-overlapping bins and transformed on a logarithmic scale. Black rectangles highlight local sites with high levels of H3K27ac modification, which exhibit an accessible chromatin structure.
B: Boxplots showing the correlation of H3K27me3 (blue) and H3K27ac (orange) enrichment with chromatin accessibility. ChIP signal enrichment over input and the GC normalized differential MNase signal was assessed in 10 kb bins. First bins were split into two groups: (1) bins showing higher input signal and (2) bins enriched for the histone modification. Bins showing an enrichment over background of the respective modification were further splitted into quartiles on the basis of ChIP-seq intensities.
C: Boxplots illustrating the chromatin accessibility of annotated chromatin states. The differential MNase signal was assessed at 15 defined chromatin states in HeLa cells (Roadmap Epigenomics Consortium et al., 2015). The size distribution of each chromatin state is shown as boxplot in the bottom panel.

4.1.9 DNA accessibility is modulated at regulatory local sites

The gold standard to identify accessible and exposed DNA in chromatin is chromatin digestion using the endonuclease DNase I in combination with high throughput sequencing (DNase-seq). DNase I hypersensitive (HS) sites coincide with active enhancer and promoter (Crawford et al., 2006). The meta nucleosome profile of high- and low-MNase centered at DNase I HS sites

revealed, that nucleosomes around DNase I HS are hyper-accessible and more efficiently released from chromatin in low-MNase (Figure 14A and B). The enrichment of low-MNase at DNase I HS sites is not obtained due to the sequence bias of MNase as described above, since the GC content at the center of DNase I HS is high favoring the high-MNase condition (Figure 14B). Confirming the results of the chromatin state analysis that chromatin accessibility is modulated at local scale, the increased accessibility at the DNase I HS sites is not spread, but just spanning up to 3 nucleosomes.

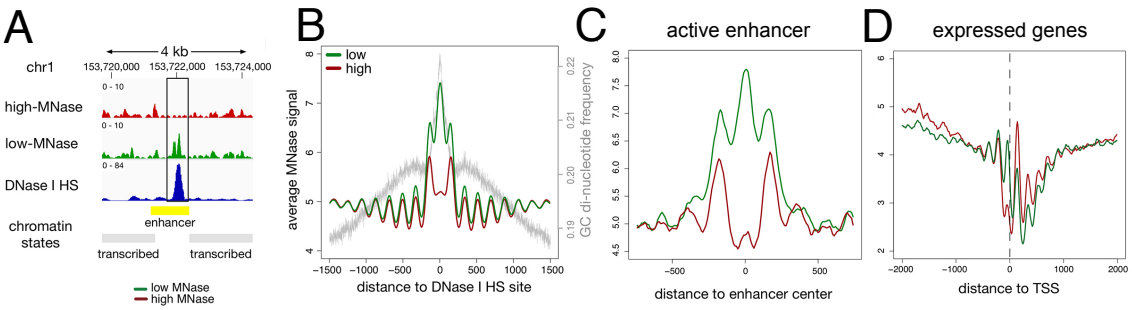


Figure 14: DNA accessibility is modulated at regulatory local sites

A: Genome browser snapshot showing nucleosome profiles of high- and low-MNase at a DNase I hypersensitive (DNase I HS) site. The chromatin state model for HeLa cells at this locus is indicated in the last track.
B-D: Average nucleosome profile of high- and low-MNase at regulatory sites.
B: DNase I HS sites in HeLa cells were obtained from ENCODE repository (accession: ENCFF692NCU). The grey line indicates the average GC di-nucleotide content.
C: Active enhancer in HeLa cells were identified by bidirectional transcription patterns using CAGE (Andersson et al., 2014)
D: Highly expressed genes were selected based on available RNA-seq data from HeLa cells obtained from ENCODE (accession: ENCFF000DNW). The 25% highest expressed genes were selected according to FPKM (fragments per kilobase transcript per million mapped) values.

Since the chromatin state analysis revealed an increased accessibility at enhancer and active TSS, the chromatin accessibility at those features was investigated in more detail. The nucleosome occupancy of high- and low-MNase was plotted around the center of active enhancers revealed by CAGE-seq in HeLa cells (Andersson et al., 2014; Chen et al., 2016). In agreement to the analysis of DNase I HS sites, chromatin is made hyper-accessible locally spanning 3 nucleosomes (Figure 14C). Chromatin accessibility is even more defined at the promoter of expressed genes showing increased accessibility just at the first nucleosome upstream of the TSS (Figure 14D). Previous MNase-seq studies reported a nucleosome depleted region (NDR) at this site of the promoter (Chereji et al., 2017). Commonly MNase conditions are used, which are comparable to the high-MNase digestion of this study. Accordingly, the nucleosome occupancy profile of the high-MNase experiment shows a gap suggesting a NDR at the promoter. However, a clear peak is visible at the same position in low-MNase indicating

a well-positioned nucleosome. At this site nucleosomal DNA might be partially unwrapped and made accessible for transcription machinery binding resulting in MNase fragile nucleosomes, which are over-digested in high-MNase, but maintained in low-MNase. This theory would be in agreement with a recent study mapping genome wide nucleosome positions using a chemical approach (Voong et al., 2016), which identified as well nucleosomes at the previously designated NDR.

Taken together, the in depth analysis showed that chromatin accessibility is modulated locally at regulatory elements to enable immediate access to DNA for transcription factor binding or transcription initiation.

4.2 Discussion

4.2.1 Differentially accessible domains of higher-order structures of chromatin do not exist in human cells

The presented study was initially designed to identify and locate domains of compacted chromatin in human cells. According to the text book model about chromatin organization, chromatin is hierarchically folded into ordered and densely packaged higher-order structures, which are thought to shield the underlying DNA from DNA interacting molecules (Alberts et al., 2002). Consequently, DNA packaged into higher order structures of chromatin should be less vulnerable to hydrolysis by nucleases. Hence, using differential MNase conditions, as introduced in FCE, to digest chromatin should uncover different levels of chromatin compaction. Surprisingly, applying FCE to HeLa cells did not reveal domains with increased or decreased DNA accessibility. Even, eu- and heterochromatin domains, characterized by distinct epigenomic features such as transcriptional activity and histone modifications, exhibit similar accessibility to MNase hydrolysis, arguing against hierarchical layers of DNA packaging as mechanism to regulate DNA accessibility and gene activity. This observation is consistent with a recent study using the restriction enzyme AluI in mouse hepatocytes (Chereji et al., 2019a), confirming the generality of the results obtained with MNase in the presented study.

Furthermore, super-resolution chromatin imaging confirmed the absence of regularly folded chromatin in large domains as proposed for the 30 nm fiber (Ou et al., 2017). The results of ChromEMT, a combination of electron microscopy tomography and selective DNA labeling, suggests that nucleosomes assemble into disordered chains with varying diameters of 5 to 24

nm. Therefore, Ou and colleagues propose that chromatin forms a disordered and granular chain, which is organized in different local nucleosome concentration densities and not in large scale chromatin structures. Consistently, a previous super-resolution nanoscopy (STORM) study reported heterogenous groups of assembled nucleosomes, termed "clutches", which arrange into disrupted chromatin fibers (Ricci et al., 2015). The view of chromatin as a heterogenous and disrupted fiber consisting of various local nucleosome arrangements is further supported by nucleosome resolution chromosome folding maps in yeast (Hsieh et al., 2015), which detected interspersed tri- or tetra-nucleosome folding motifs. Short nucleosome folding motifs could as well explain the higher frequency of close N+2 nucleosome contacts in heterochromatic regions as observed in RICC-seq without the formation of larger 30-nm fibers (Risca et al., 2017).

4.2.2 Chromatin forms an accessible and dynamic polymer

After correcting for the GC bias of MNase, the *in silico* simulation of an open chromatin structure resembled the experimental results obtained in FCE best (Figure 7C). Partial MNase digestions released nucleosomes from all areas of the genome with similar rates indicating that MNase diffuses equally well into all chromatin domains within the nucleus. The results of this study confirm fluorescence correlation spectroscopy (FCS) analyses, which showed that diffusion of small proteins is not correlated with chromatin density in interphase cells (Bancaud et al., 2009; Dross et al., 2009). Even in heterochromatin, chromatin is not dense enough folded to exclude small molecules suggesting that the linker DNA is in general accessible to transcription factors. Interestingly, theoretical modeling proposed that chromatin might function as a molecular sieve and larger complexes such as the transcription machinery may be expelled from domains of high chromatin density, thereby regulating gene activity (Maeshima et al., 2015). As MNase is a rather small molecule, it is intriguing, whether increasing the size of the enzyme would change its accessibility towards DNA in defined regions of the genome.
Live-cell imaging studies using LacO/LacI-GFP or CRISPR/Cas-based strategies to label genomic loci (Chen et al., 2013a; Robinett et al., 1996) demonstrated that chromatin is highly dynamic and constantly rearranging within the nucleus. In contrast to the static view of chromatin as proposed with the arrangement into regular 30 nm fibers, chromatin exists in a fluid-like state in living cells as presented in the polymer melt model (Maeshima et al., 2014). This dynamic property of chromatin is in well agreement with the findings of this study and ensures DNA accessibility, even in compacted chromatin (Hihara et al., 2012).

Considering recent insights from Hi-C experiments, which revealed the organization of chromatin into functional domains, so called topologically associating domains (TADs) (Dixon et al., 2012; Lieberman-Aiden et al., 2009), through chromatin loop extrusion (Fudenberg et al., 2016), chromatin forms larger-scale structures with a size range of ~100 kb to ~1 Mb.
Taken together, chromatin forms irregularly folded and highly dynamic domains, which can be penetrated by transcription factors. The chromatin domains are surrounded by a network of interchromatin channels where larger proteins and macromolecular aggregates, such as transcription factories or splicing speckles reside (Cremer et al., 2020), which cannot access the interior of the chromatin domains due to their size (Maeshima et al., 2015). Within this functional nuclear architecture, pioneering transcription factors can freely migrate into chromatin domains and bind their target DNA site. Due to the dynamic behavior of the chromatin, the transcription factor bound to its target site will relocate and show up at the periphery of the chromatin domain, which is the interface of the chromatin domain and the interchromatin compartment. At the domain periphery additional proteins are recruited to the target site and large protein complexes are formed, which prevent the relocation to the inner domain and hold the target site like a buoy at the surface. Assembly of the transcription machinery itself at the active DNA region, which is promoted by enhancer elements (Koch et al., 2011), could serve to anchor the site at the domain periphery and would be consistent with experiments showing that active RNA polymerase II constrains chromatin motion (summarized in (Babokhov et al., 2020)). In addition, histone modifications might change the physiochemical properties of chromatin and stabilize the location at the domain periphery by repulsion forces. As a result of the exposure to the open interchromatin compartment, chromatin at the domain periphery is hyper-accessible explaining the preferential release in low-MNase (Figure 14) or in other DNA accessibility assays, such as DNase-seq or ATAC-seq.

4.2.3 MNase as tool to map genome wide nucleosome architecture and nucleosome accessibility

Unexpectedly, thorough nucleosome position analysis of chromatin digested with either high- or low-MNase revealed, that the commonly used quantitative MNase digestion of chromatin to map nucleosome positions (Zentner and Henikoff, 2014) does not result in a representative picture of the actual nucleosome organization. MNase has a well-known preference to hydrolyze DNA at A/T nucleotides (Ishii et al., 2015). Comparing low- and high-MNase chromatin digestions revealed, that with increasing MNase concentration the AT bias of MNase becomes more pronounced and nucleosomes of AT rich regions are further digested to sub-

nucleosomal particles and therefore underrepresented in genome wide nucleosome maps (Figure 9). The increased sensitivity of AT rich nucleosomes towards MNase digestion is in line with a recent study in yeast, which confirmed the nucleosome particles using combinations of MNase digestion and histone ChIP (Chereji et al., 2017). Interestingly, the low-MNase condition of this study did reveal a more homogenous nucleosome distribution, which was not affected by DNA sequence. Moreover, nucleosome positions of low-MNase are highly concordant with nucleosome annotations of high-MNase in GC rich regions, where nucleosomes are not prone to over-digestion (Figure 10). Taken together, based on the findings of this study it is advisable to perform partial MNase chromatin digestions instead of the widely used full hydrolysis of chromatin to one major mono-nucleosomal fraction in order to map genome wide nucleosome positions.

At functional DNA elements such as active enhancer and promoter, an additional nucleosome sized peak was observed in the low-MNase condition, whereas a depletion of nucleosomal fragments in the high-MNase condition suggests a nucleosome free region (NFR) at those sites. Several studies showed that nucleosome sized peaks might arise in MNase-seq experiments from non-histone proteins binding to DNA (Henikoff et al., 2011; Kent et al., 2011; Mieczkowski et al., 2016) and it cannot be directly excluded, that the additional peaks at enhancer and promoter in this study originated from non-histone proteins bound to DNA. However, previous experiments in drosophila and mouse (Chereji et al., 2015; Ishii et al., 2015; Iwafuchi-Doi et al., 2016) reported so called fragile nucleosomes, which are nucleosomes with altered physical properties making them sensitive to MNase digestion. MNase chromatin fragmentations combined with subsequent histone ChIPs confirmed that histone associated structures upstream of the TSS and at enhancers exist, which can be obtained with mild MNase conditions, but are lost under higher MNase concentrations. Further supporting evidence that nucleosomes occupy the previously designated NFRs at the promoter comes from experiments using mutated histone H4, which triggers the cleavage of nucleosome DNA near the dyad in a chemical reaction (Voong et al., 2016). Even though, using this approach showed a clear nucleosome depletion immediately upstream of the TSS in *Saccharomyces cerevisiae* (Brogaard et al., 2012), conducted in *Schizosaccharomyces pombe* (Moyle-Heyrman et al., 2013) or mouse embryonic stem cells (Voong et al., 2016) the chemical nucleosome mapping revealed H4 binding at those sites indicating a species-specific regulation. The data of this study suggest, that the promoter of transcribed human genes harbor as well a fragile nucleosome. How the physiochemical properties of stable nucleosomes are changed to form fragile nucleosomes is a much debated active area of research. Fragile nucleosomes might be the

consequence of 1) active nucleosome remodeling, such as nucleosome unwrapping by the ATPase dependent remodeling complex RSC (Brahma and Henikoff, 2018; Kubik et al., 2015), 2) destabilized DNA-histone interactions of histone variants or post-translational modifications (Jin and Felsenfeld, 2007; Xi et al., 2011) or 3) alternative nucleosome structures (Rhee et al., 2014).

Recent studies employed MNase titration series (Chereji et al., 2019b; Mieczkowski et al., 2016) to characterize nucleosome stability and accessibility in greater detail. However it remains elusive and needs to be tested in more detail, whether using more than two MNase concentrations significantly improves the classification of nucleosome accessibility and fragility. The quantitative MNase-seq approach using 6 different MNase concentrations developed by Chereji and colleagues could reproduce in the drosophila model organism the findings of this study made in human cells and is in well agreement with the presented accessibility metric as log fold-change derived from two titrations points (Chereji et al., 2019b).

5 Analysing Adenovirus chromatin dynamics early in infections

5.1 Results

5.1.1 Time resolved analysis of Adenovirus genome localization during early phase of infection

H1299 lung epithelial cells were infected with human Adenovirus type 5 (Ad) at a MOI (multiplicity of infection) of 3,000. The used Ad strain exhibited a partial deletion at the E3 gene, but was replication competent. To obtain synchronized virus infections, cells were incubated with Ad and after 30 min the inoculum was replaced with fresh medium. The time point replacing the media is defined as time point zero in this study and cells were collected after 0, 0.5, 1, 2 and 4 hours post infection (hpi) for immunofluorescence (IF) analysis, transcriptome profiling using RNA-seq and FCE using differential MNase concentrations to digest chromatin (MNase-seq) (Figure 15A, see section 4.1.1 for details about FCE).

Ad enter the host cell through endocytosis (Meier and Greber, 2004). Upon acidification in early endosomes the viral capsid breaks up and viral proteins are released triggering the endosomal escape. Lysis of the endosome membrane is accomplished by exposure of the structural protein VI (pVI) (Wickham et al., 1993; Wiethoff et al., 2005). Accordingly, immunofluorescence analysis revealed that pVI is mainly detectable in the first 30 min of Ad infection and afterwards rapidly degraded (Figure 15B and C). pVI quantification shows that infections are synchronized and endosomal escape is accomplished within the first hour of infection.

After endosomal escape the virus is transported along the microtubule network towards the nucleus. Subsequently, the viral genome enters the nucleus through the nuclear pore complex (Trotman et al., 2001). The core protein VII (pVII) remains associated with the viral genome and the imported genomes can be visualized using pVII antibodies (Komatsu et al., 2015). The first viral genomes appear at 0.5 hpi as distinct pVII foci in the nucleus (Figure 15B). Quantification of nuclear pVII foci revealed, that most viral genomes have been imported 1 hpi (Figure 15D). At later time points the number of detected viral genomes increased only slightly and saturated.

Taken together, immunofluorescence analysis confirmed, that the applied infection system is synchronized and covers the time points before and immediately after nuclear entry of viral genomes.

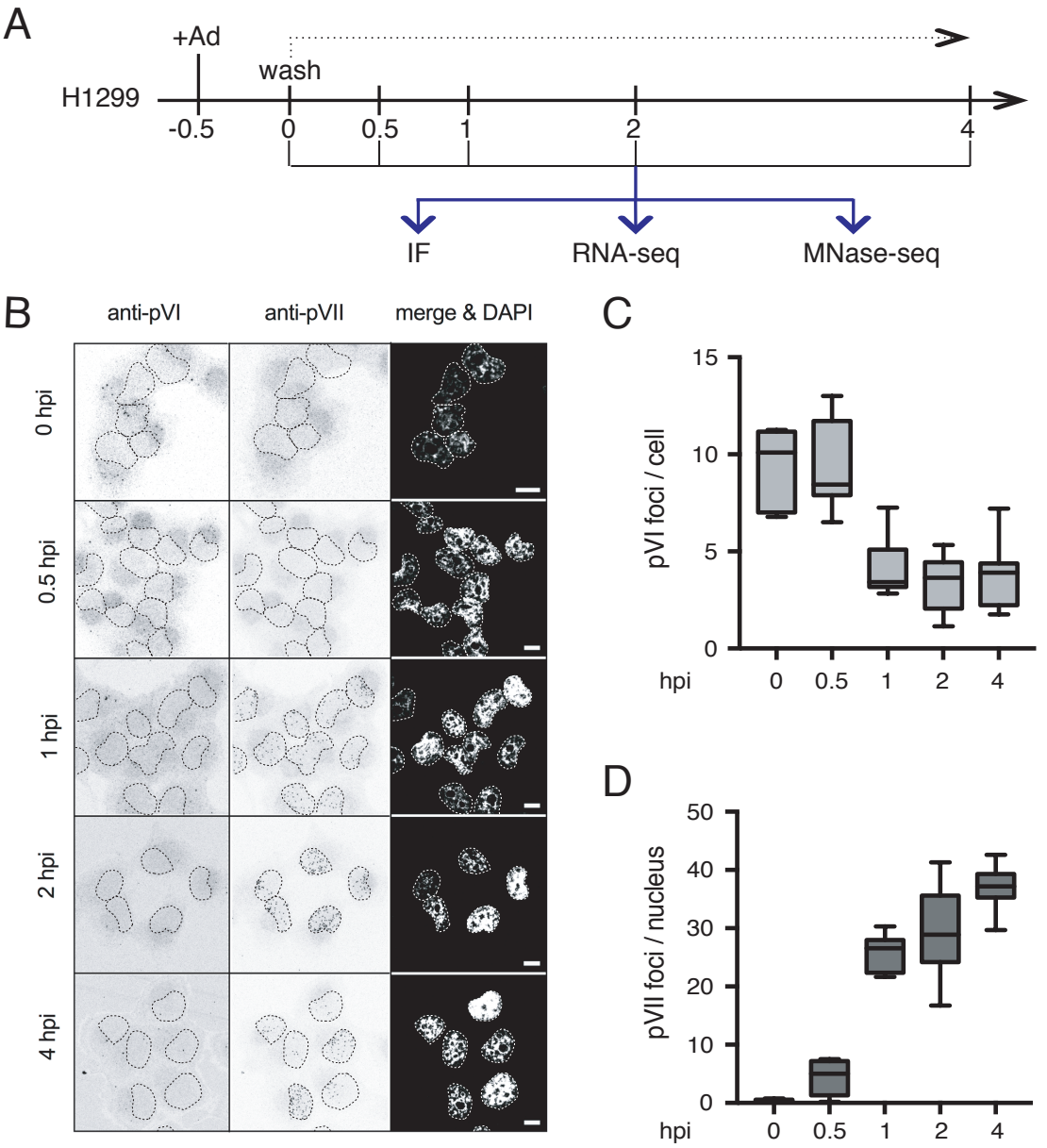


Figure 15: Time resolved analysis of Adenovirus genome localization

A: Scheme illustrating the experimental design. Adenoviruses (Ad) were added to H1299 cell cultures. After 30 min the cells were washed with fresh medium to remove unbound viruses. At the indicated timepoints (hpi = hours post infection) cells were collected and subjected to immunofluorescence (IF) analysis, transcriptome profiling (RNA-seq) and chromatin extraction (MNase-seq).
B: Representative confocal images showing pVI and pVII localization during infection time course. Staining using anti-pVI antibodies is shown in the left column and respectively pVII staining in the middle column. DNA was stained using DAPI and is shown in the right column. Nuclei are indicated by dashed lines. Scale bars indicate 10 µm.
C-D: Boxplots showing the quantification of (**C**) pVI foci per cell and (**D**) pVII foci per nucleus over time. pVI and pVII dots were quantified in at least 45 cells at each timepoint.

5.1.2 Initiation of viral transcription is accomplished 2 hours post infection

To monitor the initiation of viral transcription, RNA was isolated from infected cells. PolyA enriched RNA libraries were prepared and subjected to high-throughput sequencing. Reads

were simultaneously aligned to the human and viral genome. At 0, 0.5, and 1 hpi the isolated RNA was composed solely of host cell transcripts and the first significant amounts of viral transcripts could be detected 2 hpi (Table 1 and Figure 16A).

timepoint	replicate	sequenced reads	reads annotated to human and Ad genome	reads annotated to Ad genome	RPM annotated to Ad genome
0 hpi	R1	29,121,328	28,484,036	100	4
	R2	31,703,184	30,920,923	143	5
0.5 hpi	R1	23,814,664	22,986,503	32	1
	R2	23,296,828	22,503,152	32	1
1 hpi	R1	24,904,732	24,149,420	50	2
	R2	74,971,114	74,417,479	156	2
2 hpi	R1	27,062,105	25,188,583	382	15
	R2	27,789,187	26,455,494	462	17
4 hpi	R1	18,376,378	17,941,436	45,188	2519
	R2	18,670,627	18,303,871	53,942	2947

Table 1: RNA-seq statistics

Overview of obtained reads and alignment efficiency of RNA-seq data. RPM = reads per million mapped reads.

At 4 hpi viral transcription was markedly enhanced. Ad gene expression follows a strict spatial and temporal order (Nevins et al., 1979). First so called early genes (E1A, E1B, E3 and E4) at both ends of the genome are expressed, followed by more central located genes required for viral replication (E2A and E2B). Structural proteins are encoded in the center of the genome (L1 to L5) under control of the major late promoter (ML) and transcribed after the virus has replicated. Accordingly, at 2 hpi E1A and E3 were the first viral transcripts being detected (Figure 16A and B). E1A acts as a transcriptional activator and is, consistent with previous studies, the first Ad gene to be transcribed (Crisostomo et al., 2019). At 4 hpi, the virus exploits the host transcription machinery efficiently and transcripts from all early genes along with significant amounts of E2A transcripts accumulate.

In summary, Ad provide a transcription competent template within the first hours of infection and viral transcription initiates 2 hpi.

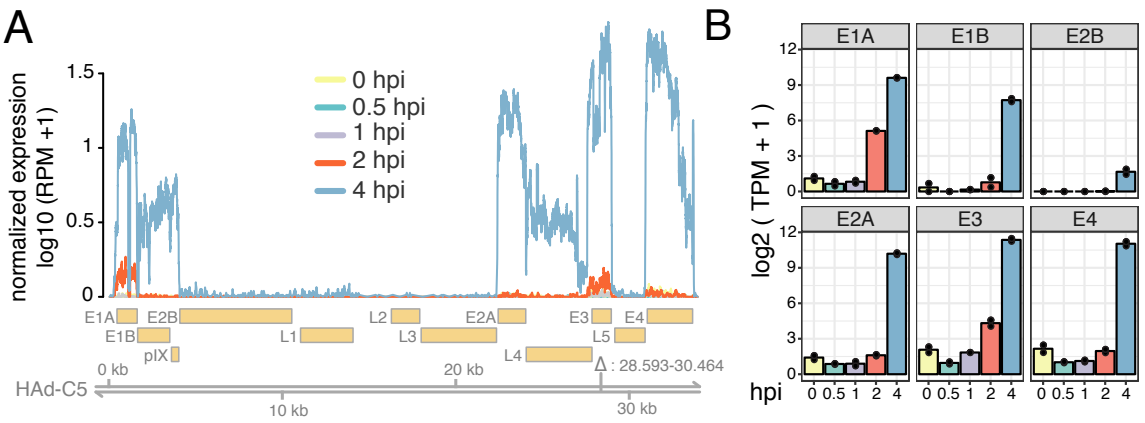


Figure 16: Initiation of viral transcription is accomplished 2 hours post infection

A: Profile showing viral RNA abundance at indicated timepoints. RNA-seq read coverage was normalized by sequencing depth to reads per million (RPM) and transformed on logarithmic scale after adding one pseudocount. The average profile of two independent experiments is shown. Gene annotation of human Adenovirus type 5 (HAd-C5) and genomic location are illustrated at the bottom. The deleted region (Δ) of the E3 gene is indicated.
B: Barplots showing viral transcript abundance over time. The number of reads per transcript was normalized by the gene length and sequencing depth to TPM (transcripts per kilobase per million mapped reads) and transformed on logarithmic scale after adding one pseudocount. Average of two independent replicates are shown and individual replicates are indicated by a black dot.

5.1.3 A subpopulation of nuclear Ad genomes are actively transcribed

Monitoring pVII exposure revealed that under the applied experimental conditions about 30 Ad genomes enter the nucleus in each cell (Figure 15D). In addition, transcriptome profiling showed that Ad genomes are actively transcribed producing polyadenylated gene transcripts (Figure 16). Abundance of viral transcripts increased over time, but it is still unclear whether this is due to an increasing number of viral genomes getting transcribed or the continuous transcription of a constant population of Ad. To monitor actively transcribing genomes at single cell level, RNAscope technology was employed. Due to a specific probe design, RNAscope allows to visualize *in situ* the location of viral transcripts with high spatial resolution and low background signal. Specific probes hybridizing to E1A mRNA, the first viral transcript produced during infection, were designed. After *in situ* hybridization precise E1A foci are visible in infected cells defining precisely the location of the transcripts (Figure 17A, left panel). At 1 hpi only very few cells exhibited a positive signal of E1A and consistent with the whole transcriptome analysis using RNA-seq, E1A transcripts could be detected as early as 2 hpi in most of the cells (Figure 17A and B). At 4 hpi E1A transcripts accumulated in the cells. Taken together time resolved IF analysis of E1A transcripts confirmed that viral transcription initiates at 2 hpi.

Detection of E1A transcripts and pVII associated viral genomes with high spatial resolution allows the identification of viral genomes being transcribed through colocalization of pVII and

E1A foci (Figure 17A). To determine the fraction of actively transcribed Ad genomes in host cells, the number of pVII foci colocalizing with nuclear E1A foci was divided through the total number of detected Ad genomes in the nucleus (Figure 17C). On average 20 % of all imported viral genomes became active and initiation of viral transcription was accomplished 2 hpi. The fraction of active viral genomes did not significantly increase further at 4 hpi, even though more E1A foci are detected confirming that the number of active viral genomes was not obtained due to random associations.

Ad genome import into the nucleus is fulfilled after 1 hour, however at this timepoint only few viral genomes are transcribed indicating that a lag phase between genome import and initiation of viral transcription exists.

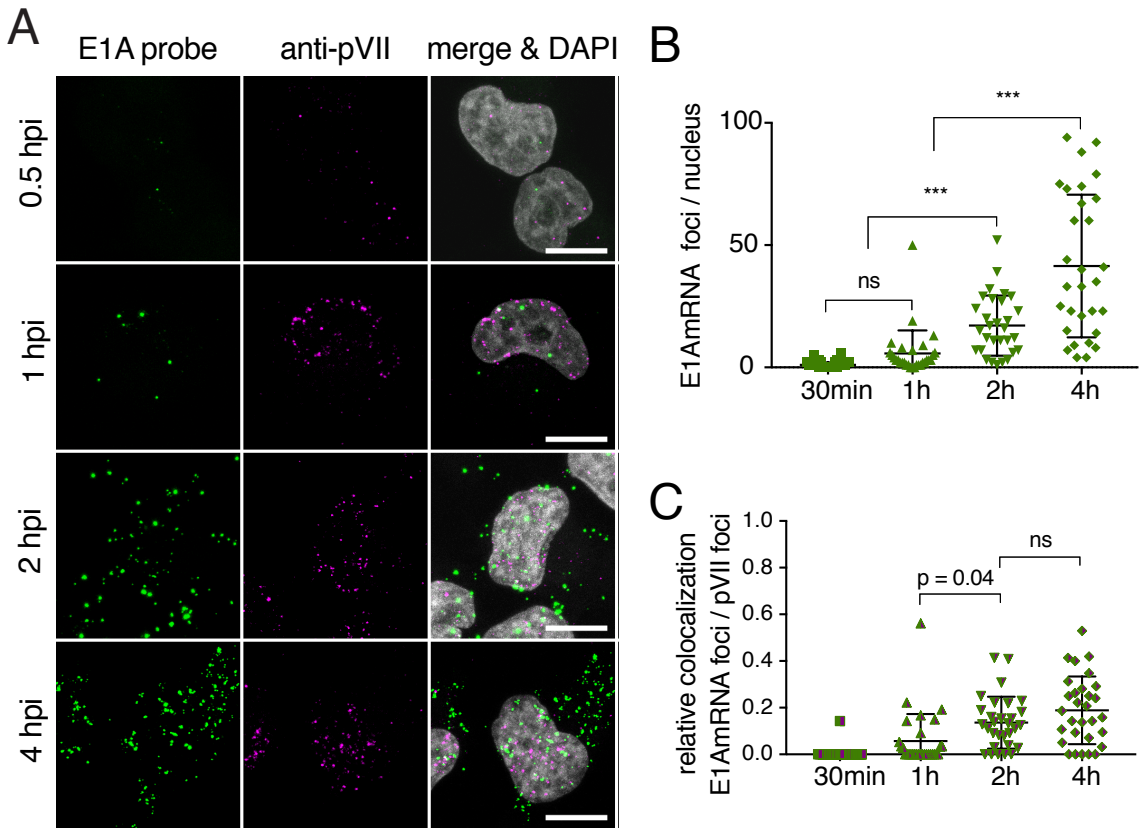


Figure 17: A subpopulation of nuclear Ad genomes are actively transcribed

A: Representative confocal images showing E1A transcripts and the localization of viral genomes during infection time course. Specific RNAscope probes were used to detect E1A transcripts (green dots) as shown in the left column. Staining of viral genomes using anti-pVII antibodies (magenta dots) is shown in the middle column. Merged images are shown in the right column and the nucleus is highlighted by DAPI staining. Scale bars indicate 10 µm

B: Quantification of E1A foci per nucleus over time. The number of E1A transcript foci was quantified in 30 nuclei per time point.

C: Quantification of E1A transcribing genomes. The number of viral genomes colocalizing with E1A transcript foci was quantified in 30 nuclei per time point and normalized by the total number of viral genomes per nucleus.

B-C: The mean and the standard deviation of at each time point is indicated. Statistical analysis was carried out using one-way ANOVA multi-comparison test (ns = not significant, *** : p-value < 0.001).

5.1.4 Functional chromatin extraction of Ad infected cells

Previous *in vitro* studies suggest that pVII condensed Ad genomes require active chromatin remodeling before initiation of viral transcription (Okuwaki and Nagata, 1998), which might explain the observed delay in transcription after nuclear import. To analyze spatiotemporal changes of Ad nucleoprotein structure FCE was applied and infected cells harboring host and viral chromatin were treated with various MNase concentrations (Figure 18A). MNase conditions were adjusted to obtain low- and high-MNase digestions as used in section 4.1.1 and DNA was isolated at 0, 0.5, 1, 2 and 4 hpi (Figure 18B). Note that the typical nucleosome ladder visible on agarose gels is caused by the host chromatin which represents the main fraction of digested chromatin.

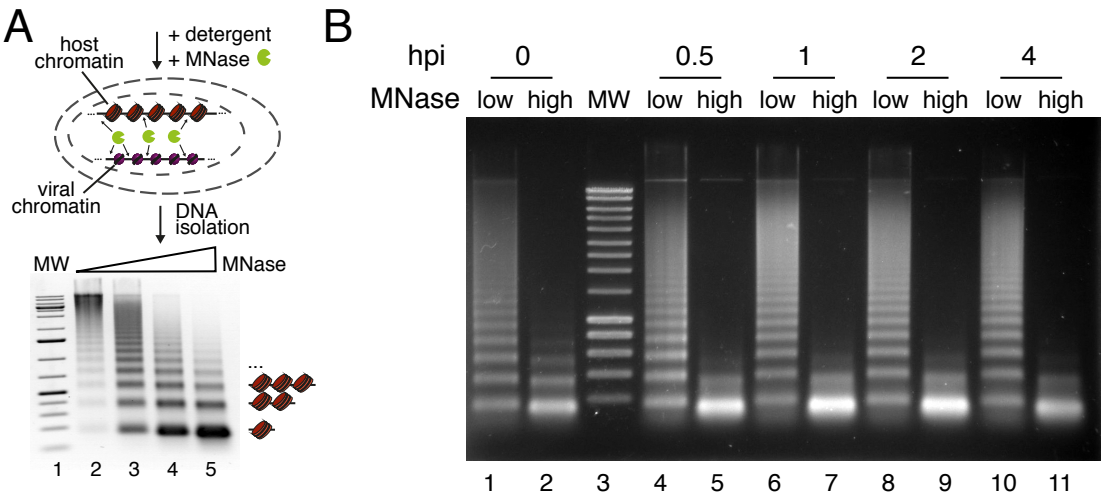


Figure 18: Functional chromatin extraction of Ad infected cells

A: Agarose gel showing the nucleosomal DNA ladders of MNase titration in Ad infected cells. Mono-, di- and tri-nucleosomal DNA fragments are indicated and molecular weight marker is shown in lane 1.
B: Agarose gel showing the nucleosomal DNA ladder obtained from Functional Chromatin Extraction (FCE) in Ad infected cells. H1299 cells were incubated with either 100 U (low-MNase shown in lane 1, 4, 6, 8 and 10) or 600 U (high-MNase shown in lane 2, 5, 7, 9, 11) of MNase at the indicated time points after Ad infection. The mono-nucleosomal DNA band was excised and subjected to high-throughput sequencing.

The mono-nucleosomal band was excised and subjected to high-throughput sequencing on Illumina HiSeq 2000 platform. In total 17 samples were sequenced yielding on average 223 Mio read pairs per sample (Table 2). The obtained reads were simultaneously aligned against the human (hg19) and the Ad (AY339865.1) genome reference. About 99.999% of the sequenced DNA fragments originated from the host cell and 0.001% could be annotated to the viral genome (Table 2). Nevertheless, assuming a fragment length of 150 bp, on average the samples exhibit a 700 fold coverage of the viral genome, allowing in depth analysis of viral nucleoprotein structure.

MNase condition	timepoint	replicate	sequenced fragments	fragments annotated to human genome	fragments annotated to Ad genome
high	0 hpi	R1	229,137,820	163,870,708	28,329
		R2	232,897,593	165,936,041	20,222
	0.5 hpi	R1	234,007,548	172,898,971	61,864
		R2	229,100,819	153,222,365	64,927
	1 hpi	R1	226,436,370	118,979,031	153,974
		R2	219,697,342	135,212,205	88,116
	2 hpi	R1	227,598,948	131,444,807	178,491
		R2	230,658,303	166,272,934	75,382
	4 hpi	R1	231,625,618	115,031,799	284,421
		R2	225,137,928	169,292,212	117,689
low	0 hpi	R1	212,388,716	156,852,024	86,603
	0.5 hpi	R1	218,965,404	153,473,717	205,834
		R2	216,033,646	165,531,283	135,786
	1 hpi	R1	218,399,432	157,676,688	200,931
		R2	216,415,263	142,234,881	541,881
	4 hpi	R1	212,947,907	178,933,555	445,892
		R2	221,396,121	182,403,786	197,983

Table 2: MNase-seq statistics

Overview of obtained read-pairs and alignment efficiency of MNase-seq data.

5.1.5 DNA sequence dictates positioning of pVII

In eukaryotes MNase chromatin digestions result in defined DNA fragment sizes representing monomer or multimers of nucleosomal DNA. After alignment of the sequenced DNA fragments the size distribution can be assessed with bp resolution and as expected the fragments annotated to the human genome exhibited the characteristic length of nucleosomal DNA (147 bp) (Figure 19A). In contrast, fragments derived from Ad chromatin were shorter and the size distribution was less defined. However, in high-MNase experiments, the MNase condition with the highest resolution, the viral DNA fragments exhibited constantly an average size of 70 bp indicating a defined protection from MNase digestion (Figure 19A and B). Most probably pVII particles protect associated DNA from hydrolyzation, which is in agreement with previous MNase footprinting studies (Daniell et al., 1981; Mirza and Weber, 1982). The differences in the fragment size distribution of pVII-DNA particles and nucleosomes suggest a different packaging structure.

It is known that nucleosomes prefer distinct DNA sequence pattern, which facilitate bending of DNA around the histone octamer (Kaplan et al., 2009). Hence, nucleosomes preferentially occupy positions in the genome exhibiting a regular phasing of A/T and G/C di-nucleotides. As A/T di-nucleotides form a narrow DNA minor groove and G/C di-nucleotides a wide DNA

minor groove, alternating oscillations of A/T and G/C di-nucleotides at every helical turn (≈ 10 bp) form an intrinsic curvature of DNA, which follows the path of the histone octamer surface and thereby increases the binding affinity (Widom, 2001). Plotting the di-nucleotide frequency of nucleosomal DNA revealed the favored DNA sequence pattern showing A/T or respectively G/C di-nucleotide repeats at 10.3 bp intervals along the extracted DNA fragments (Figure 19C). Interestingly, the same analysis of pVII nucleoprotein complexes revealed a similar di-nucleotide periodicity, but with shorter oscillations at 5.4 bp intervals. The observed DNA sequence pattern of pVII would result in a different curvature of DNA and would not favor the same packing mode as known for nucleosomes suggesting an alternative nucleoprotein structure. Nevertheless, the results indicate that pVII follows a specific positioning code, which might translate to a defined chromatin structure.

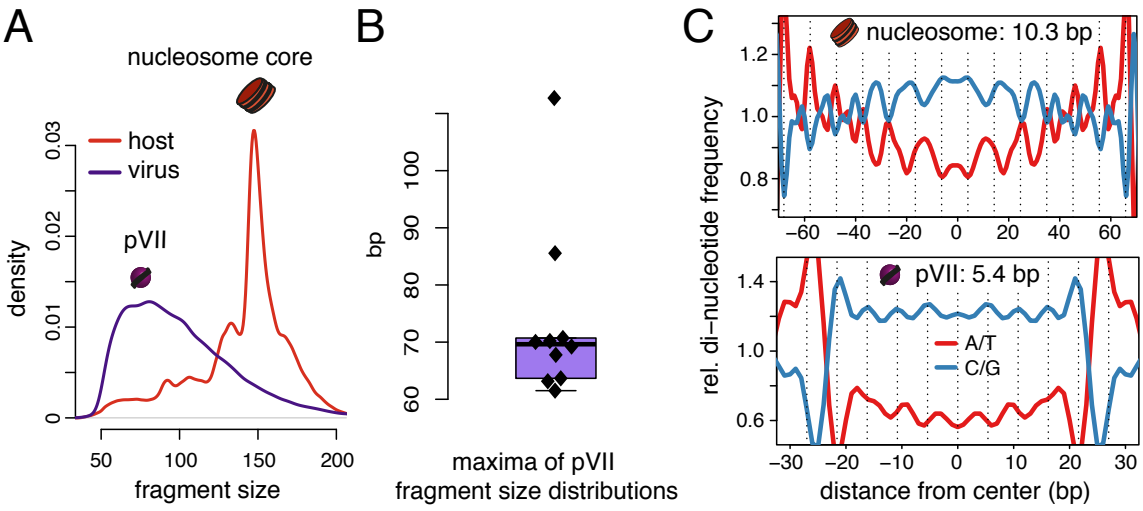


Figure 19: DNA sequence dictates positioning of pVII

A: Kernel density plot showing the fragment size distribution of isolated DNA after high-MNase chromatin extraction at 0 hpi. The distribution of viral DNA fragments representing pVII protected DNA are illustrated in purple and nucleosomal fragments derived from the human genome respectively in red.
B: Boxplot showing the main DNA protection size of pVII in high-MNase condition. The summit of the pVII fragment size distribution (black diamonds) was determined in high-MNase samples.
C: A/T- and G/C-di-nucleotide frequency at nucleosomes (upper panel) and pVII-nucleoprotein complexes (lower panel). Di-nucleotide frequency was calculated around the center of size selected fragments. The signal was smoothed using a Fast Fourier Transformation. The signal periodicity is indicated by dashed lines at every 10.3 (nucleosome) / 5.4 (pVII-nucleoprotein complex) bp.

5.1.6 Defined pVII organization provides functional chromatin architecture

Electron microscopy studies of pVII condensed Ad genomes suggest a nucleosome-like “beads-on-string” organization (Vayda et al., 1983). As nucleosomes in eukaryotes, DNA packaging by pVII might control the access to viral DNA. However, it is still unknown, how pVII is exactly organized along the viral genome and whether it occupies specific positions. To analyze

the viral nucleoprotein structure in incoming virions, pVII occupancy profiles at time point 0 hpi were generated using the well-established DANPOS2 toolkit (Figure 20A) (Chen et al., 2013b). At this timepoint the viral genomes have not entered the nucleus (Figure 15). Hence, it is unlikely that pVII has been actively remodeled, yet and it is assumed that timepoint zero represents the default nucleoprotein organization in virus particles.

Remarkably, defined peaks showing localized protection from MNase digestion imply a strict pVII organization, whereby pVII occupies distinct positions along the Ad genome (Figure 20A). The number of detected pVII peaks was constant across all timepoints confirming the reproducibility of this approach (Figure 20B). On average 238 pVII-DNA particles were detected, which is in agreement with recent projections from cryo-electron tomography (Pérez-Berná et al., 2015). Calculating the distance between adjacent pVII peak centers revealed a pVII repeat length (pVII-RL) of 97.3 bp, which is consistent with the observed size of pVII-DNA complexes of 70 bp plus an additional linker of 20-30 bp (Figure 20C).

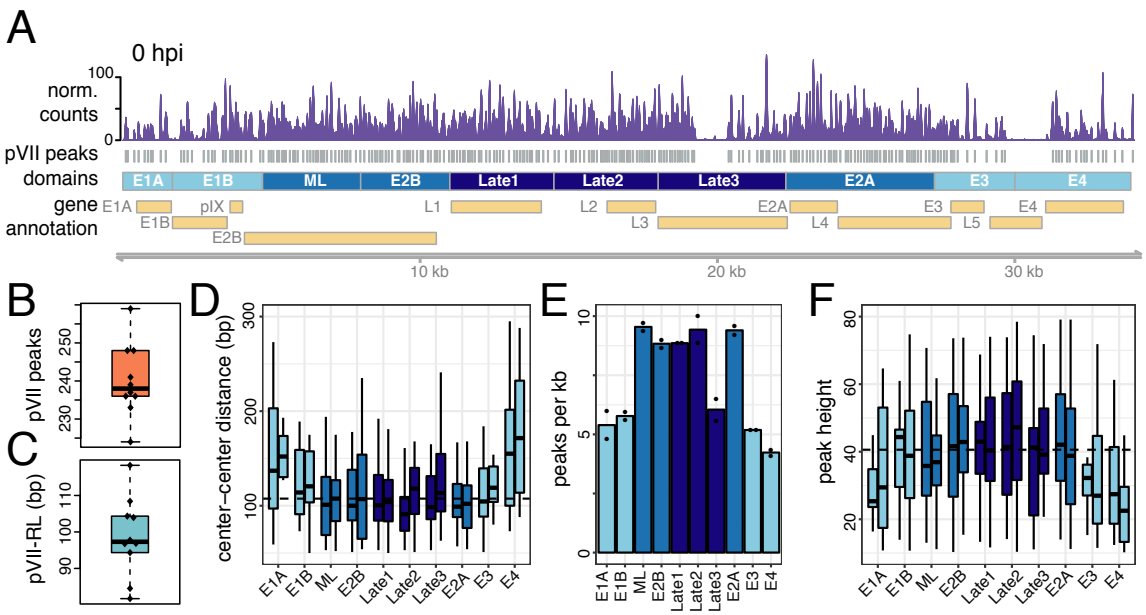


Figure 20: Defined pVII organization provides functional chromatin architecture

A: Genome browser snapshot showing pVII profile along the viral genome at 0 hpi. Coverage of pVII fragments extracted with high-MNase at 0 hpi is displayed in the first track. pVII positions were called using the DANPOS2 software suite (Chen et al., 2013b) and are displayed as grey boxes in the second track. The viral genome was partitioned into ten functional domains on basis of the gene annotation and the time point of transcription (from early (light blue) to late (dark blue) expression). Gene annotation of human Adenovirus type 5 (HAd-C5) and genomic location are illustrated at the bottom

B: Boxplot showing the number of pVII positions at the viral genome. pVII positions for all high-MNase samples were called (as shown in A). The total number of pVII positions for each sample are displayed as black diamonds.

C: Boxplot showing the predominant pVII repeat length (pVII-RL) of each time point. The pVII-RL was calculated as center to center distances of pVII peaks. The predominant pVII-RL of each experiment (black diamond) was deduced from the summit of the pVII-RL distribution.

D: Boxplots showing the pVII-RL in functional domains at 0 hpi. pVII peak center-to-center distances were calculated for high-MNase samples at 0 hpi. The results of two independent replicates are shown side-by-side.
E: Barplots showing the number of pVII positions per kb in functional domains. The number of pVII peaks was counted and normalized by the domain length for high-MNase samples at 0 hpi. The results of two independent replicates are indicated as black dots.
F: Boxplots showing the distribution of pVII peak heights in functional domains at 0 hpi. Coverage of pVII fragments at the peak summit was estimated for high-MNase samples at 0 hpi. The results of two independent replicates are shown side-by-side.

To explore how the defined chromatin structure relates to the spatiotemporally regulated transcription program of Ad, the genome was subdivided into 10 domains based on the location of the viral genes. The domains were adjusted to similar sizes and classified by the timepoint of transcription of the including gene (Figure 20A, from early (light blue) to late (dark blue) expressed). Indeed, pVII-DNA particles packaging early genes (E1A, E1B, E3 and E4) exhibited features of an open chromatin structure (Figure 20D-F). Due to a wider spacing of adjacent pVII nucleoprotein complexes (Figure 20D), less pVII occupied positions were detected at regions covering early genes (Figure 20E), thereby providing immediate access for binding of regulatory factors or transcription machinery. Furthermore, DNA at both ends of the genome was more accessible and less protected from MNase digestion as reflected by the average peak height indicating reduced pVII occupancy and stability at early genes (Figure 20F). Even though, the center of the Ad genome was in general tightly packaged by pVII, the Late3 region harbors surprisingly a DNA stretch without any detectable peaks indicating a special structural role of this region.
For the first time the pVII positioning onto the Ad genome could be mapped with high resolution. pVII organization suggests a functional chromatin structure facilitating the initiation of early gene transcription.

5.1.7 Nucleosomes assemble onto the viral genome 2 hpi

The high resolution maps of pVII positioning at the Ad genome allows to screen for temporal changes in viral nucleoprotein structure. Interestingly, starting at 2 hpi an additional peak appeared at the size distribution of viral DNA fragments (Figure 21A). Consistent with a previous study showing that histones are deposited onto incoming Ad genomes (Komatsu et al., 2011), the sizes resembled the MNase protection footprint of host nucleosomes indicating that full nucleosomes assemble onto the Ad genome. To estimate the fraction of nucleosomes relative to pVII-DNA particles, the size distributions of pVII fragments and host nucleosomes were combined with various weights (Figure 21B). The mixed distribution, which matched the actual size distribution best, was determined revealing that at 2 hpi about 8 % and at 4 hpi about

14 % of all the fragments released from viral chromatin represent fully assembled nucleosomes. Hence, at this time of infection pVII is not completely replaced by nucleosomes, suggesting that nucleosomes assemble only onto a subpopulation of Ad genomes or nucleosomes occupy a limited number of sites at the Ad genome.

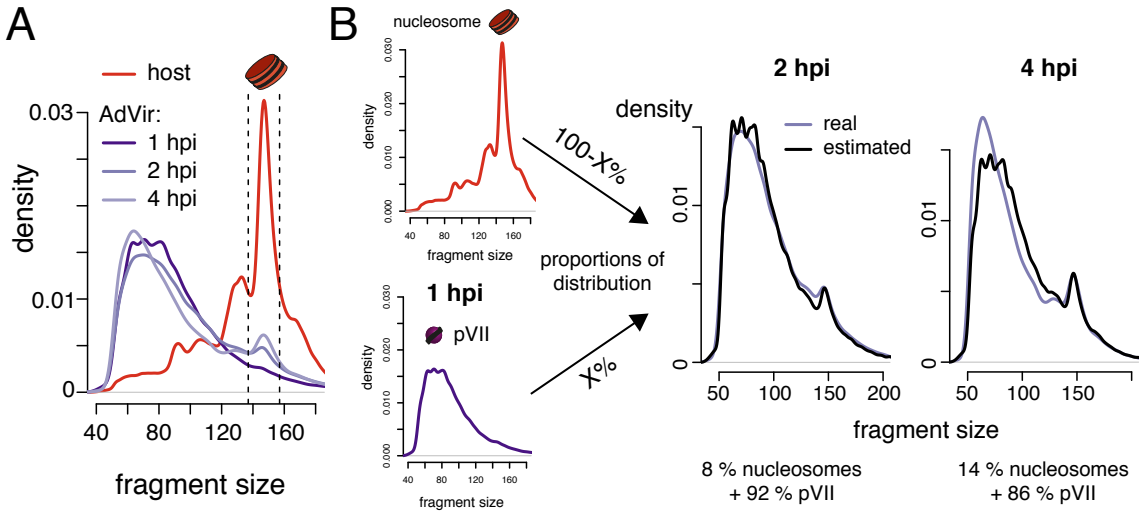


Figure 21: Nucleosomes assemble onto the viral genome 2 hpi

A: Kernel density plot showing the fragment size distribution of isolated DNA after high-MNase chromatin extraction from 1 - 4 hpi. The distribution of viral DNA fragments representing pVII protected DNA are illustrated in shades of purple and nucleosomal fragments derived from the human genome respectively in red. The dashed lines highlight nucleosome sized fragments (137 – 157 bp).
B: Simulation of a mixed distribution of pVII and nucleosome fragments to estimate the proportion of nucleosomes assembling onto viral genomes. The viral fragment size distribution at 2 and 4 hpi was reassembled by a mixed distribution composed of human nucleosomal fragments and viral pVII fragments at 1 hpi. The used proportions of fragments generating the best consensus distributions are denoted at the bottom.

5.1.8 Nucleosomes assemble preferentially at early transcribed genes

To determine the sites of nucleosome assembly, nucleosome occupancy profiles along the Ad genome were generated based on nucleosome sized fragments (137-157 bp indicated by dashed lines in Figure 21A) extracted from high-MNase experiments at 2 and 4 hpi. Profiles were normalized by the total number of fragments aligned to the viral genome in each experiment. To account for potential pVII background the same procedure was applied to fragments extracted before nucleosome assembly at 1 hpi. The nucleosome enrichment score was calculated representing the ratio between nucleosomal fragments at 2 or 4 hpi and the background signal at 1 hpi (Figure 22). Peaks exhibiting at least a 5-fold nucleosome enrichment defined sites of preferential nucleosome assembly (indicated as grey boxes in Figure 22). Using the described cut-off, 12 sites at 2 hpi and respectively 29 sites at 4 hpi were

identified as being occupied by nucleosomes. Sites of nucleosome assembly were reproducible between the two time points used and coincided with previously detected positions of histone H3 using ChIP-qPCR (indicated as green boxes in Figure 22) (Komatsu et al., 2011). Remarkably, sites of nucleosome assembly are distributed over the whole genome. Nevertheless, nucleosomes are concentrated at assembly hotspots covering early genes and the particular Late3 region.

In summary, nucleosomes are not deposited randomly onto the viral genome, but occupy distinct positions preferentially at early transcribed genes.

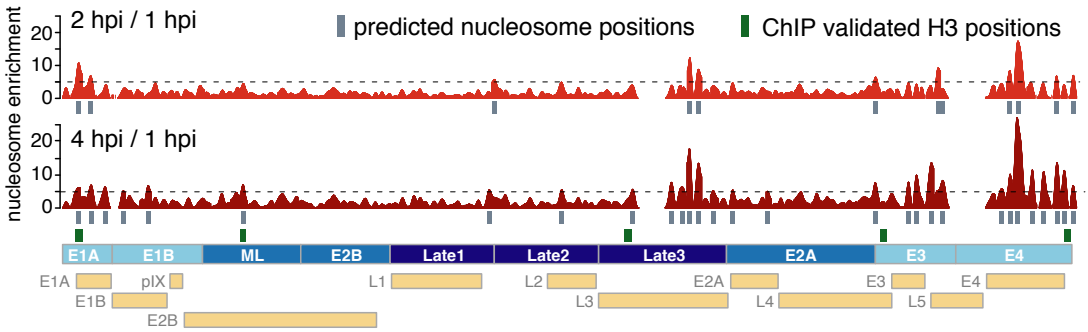


Figure 22: Nucleosomes assemble preferentially at early transcribed genes

Genome browser snapshot illustrating preferential sites of nucleosome assembly at the viral genome. The first two tracks represent the enrichment of nucleosomal sized fragments (137 – 157 bp) at 2 or 4 hpi over 1 hpi. An enrichment of at least 5 fold was used as threshold (dashed lines) to determine nucleosome positions (grey boxes). Green boxes show H3 associated sites, which have been validated in ChIP-qPCR experiments (Komatsu et al., 2011). Functional domains and gene annotation of the human Adenovirus type 5 (HAd-C5) are illustrated at the bottom.

5.1.9 pVII is replaced by nucleosomes at the +1 position of early genes

Histones might be deposited onto accessible DNA free of pVII or alternatively replace pVII from viral genomes. To assess the relation of pVII organization and nucleosome assembly, the average pVII occupancy profile centered over all detected nucleosome positions was plotted (Figure 23A). Strikingly, the meta profile of pVII exhibited a defined peak at newly assembled nucleosome positions indicating a replacement of pVII-DNA complexes by nucleosomes.

Further, in depth characterization of nucleosome positions at early genes confirmed that nucleosomes are formed at previously pVII occupied sites (Figure 23B and C). Interestingly, nucleosomes occupy the position directly downstream of the TSS at the E1A and E4 gene (highlighted in Figure 23B and C). This position is reminiscent of active genes in eukaryotes, where the first nucleosome downstream of the TSS, also known as +1 nucleosome, is strongly positioned (Schones et al., 2008). In addition the region upstream of the TSS of E1A or E4 is

free or exhibits weak nucleoprotein occupancy allowing the binding for regulatory factors and thereby resembling a *bona fide* host promoter. Remarkably, the appearance of the +1 nucleosome at the E1A and E4 gene timely coincided with the initiation of transcription indicating that both processes are functionally interconnected.

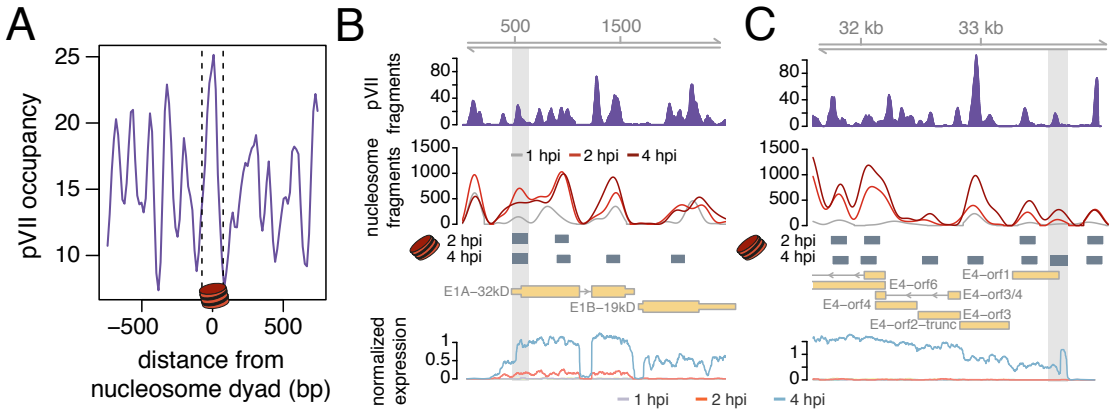


Figure 23: pVII is replaced by nucleosomes at the +1 position of early genes

A: Initial distribution of pVII around sites of nucleosome assembly at 0 hpi. The area occupied by the nucleosome is indicated by dashed lines

B-C: Genome browser snapshots highlighting replacement of pVII by nucleosomes at the +1 position of (**B**) the E1A and (**C**) the E4 gene. The first track shows initial pVII occupancy at 0 hpi. The second track displays the profile of nucleosome sized fragments (137 – 157 bp) at the indicated time points. Grey boxes indicate sites of nucleosome assembly. The last track shows the RNA-seq profile at the indicated time points. The +1 position of E1A and E4 is highlighted.

5.1.10 low-MNase reveals time resolved changes of viral chromatin accessibility

To investigate how the nucleoprotein structure of Ad genomes changes during infection, the high resolution pVII positioning maps of different timepoints were compared (Figure 24). Correlating the high-MNase treated samples did not reveal any significant changes of nucleoprotein structure between different timepoints. The low variation between the samples could be explained due to a dominant population of transcriptionally inactive viral genomes, which present about 80% of all viral genomes in the host (see section 5.1.3). Most likely inactive viral genomes do not undergo active remodelling in the host and the nucleoprotein structure represents the default state as observed in incoming virions (Figure 20) explaining the high correlation to timepoint zero.

As shown in section 4.1.9 partial chromatin digestions preferentially extract nucleoprotein structures from accessible and regulatory sites. Hence, analyzing temporal changes of nucleoprotein structure using low-MNase samples would reveal genomic sites that alter their

accessibility. Indeed, in contrast to the high-MNase condition, temporal changes of the nucleoprotein organization could be detected in low-MNase samples (Figure 24B)

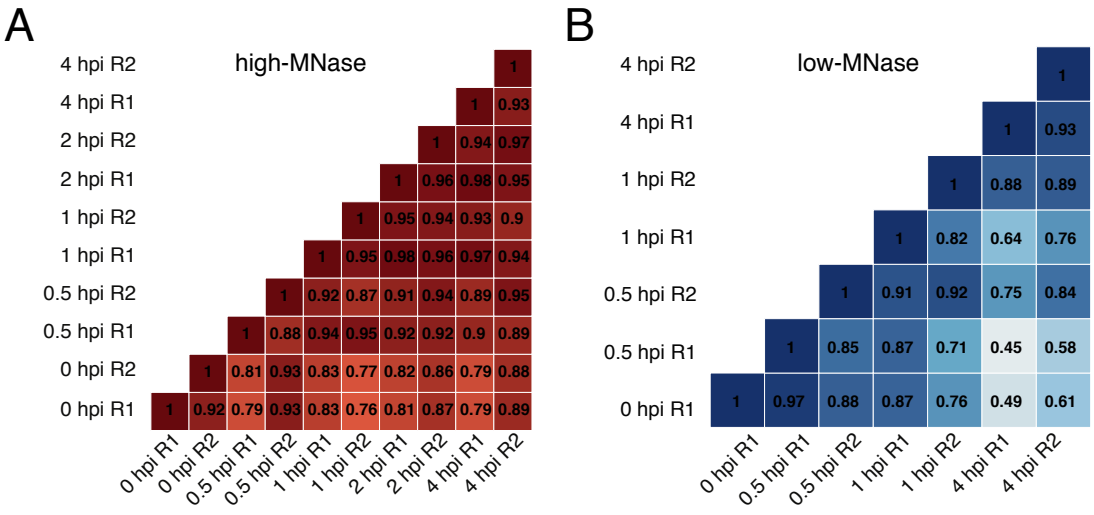


Figure 24: low-MNase digestion reveals dynamics of viral chromatin

A-B: Pearson correlation analysis of genome wide pVII occupancy profiles from (**A**) high-MNase or (**B**) low-MNase chromatin extractions. The sample-wise correlation coefficient is shown and the correlation strength is indicated by different color shades.

5.1.11 Viral chromatin de-condensation at early genes precedes gene activation

Examination of low-MNase nucleoprotein maps revealed specific changes in peak intensities over time (Figure 25A and whole genome shown in Figure 25C) suggesting localized de-condensation of viral chromatin. In order to quantify chromatin accessibility changes, a linear regression analysis was implemented. Therefore, the normalized fragment counts were assessed for each timepoint in 300 bp overlapping bins. The slope of the fitted linear regression line was used as measure for DNA accessibility changes (Figure 25B). Distinct genomic regions are associated with a positive accessibility score suggesting chromatin de-condensation at those sites (Figure 25C). In contrast, a negative accessibility score highlights sites of viral chromatin remaining condensed. Remarkably, chromatin opens up preferentially at early genes indicating a functional coherence with viral transcription (Figure 25D). Already at 1 hpi the nucleoprotein structure is altered at early genes preceding initiation of viral transcription (Figure 25C). The accessibility score at the E1A gene is modest and not as pronounced as at other early genes. As this region exhibits an open chromatin configuration while entering the host (see section 5.1.4) it might not require further de-condensation already providing an accessible template for transcription.

Taken together, the results of the chromatin accessibility analysis suggest that chromatin de-condensation is required to enable initiation of viral transcription at early genes.

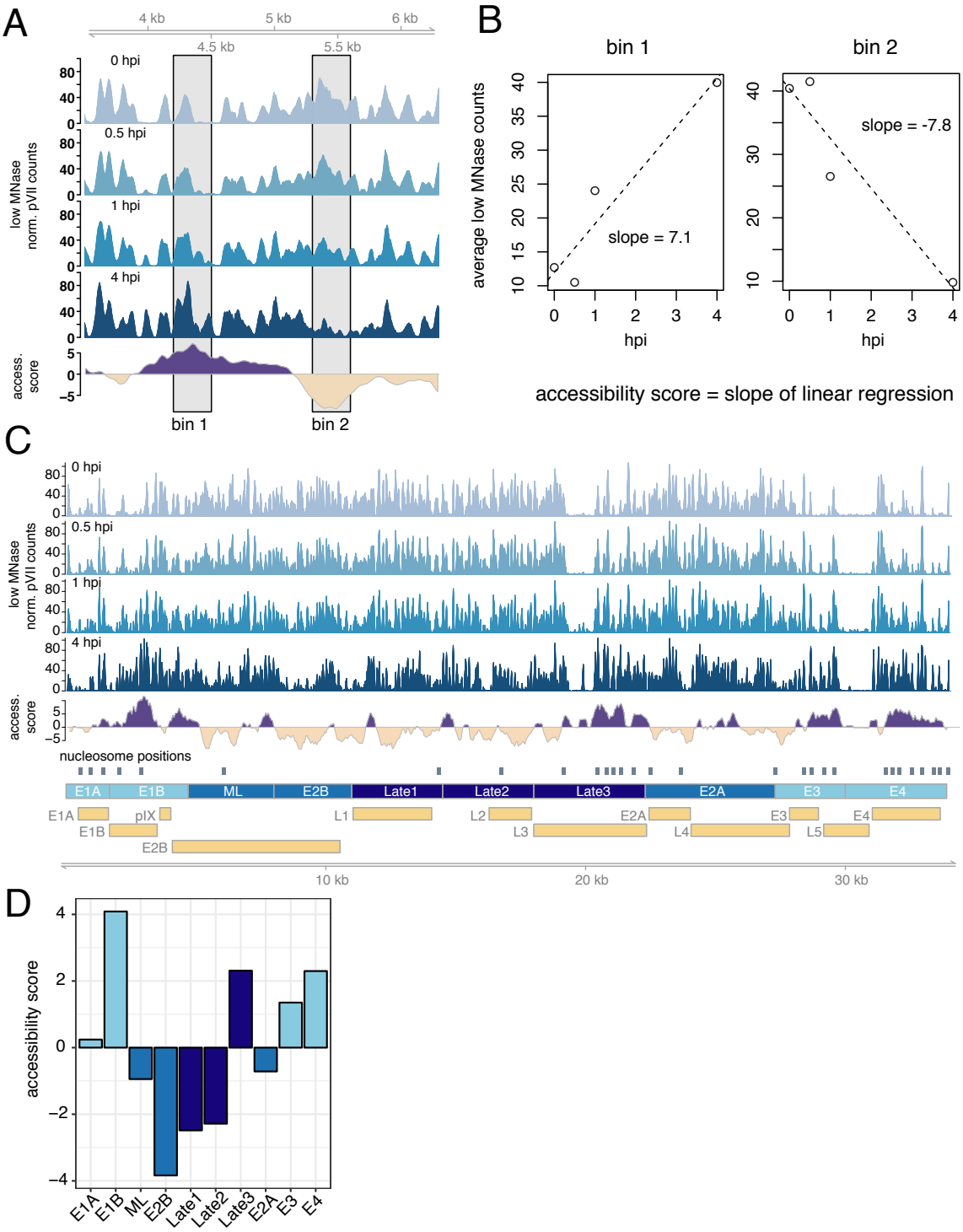


Figure 25: Viral chromatin de-condensation at early genes precedes gene activation

A-B: Quantification of viral chromatin de-condensation over the first 4 hours of infection. Calculation of the accessibility score at two regions is shown based on low-MNase generated pVII profiles
A and C: Genome browser snapshots illustrating the dynamics of chromatin accessibility (**A**) at a 3 kb wide locus or (**C**) genome wide. Normalized pVII fragment profiles obtained with low-MNase conditions are shown at the indicated timepoints. The average profile of two independent chromatin extractions is

shown, except at time point 0 hpi, where only one experiment was carried out. The accessibility score was assessed within sliding windows of 300 bp bin width and a step size of 20 bp using a linear regression fit (as shown in **B**).
B: Linear regression analysis to measure chromatin de-compaction at two exemplary sites. The normalized number of pVII fragments obtained in low-MNase was assessed in 300 bp bins. A linear regression line was fitted to the fragment counts over the infection time course. The accessibility score was deduced from the slope of the regression line.
D: Bar plot showing the average accessibility score in functional domains.

5.1.12 The Late3 region exhibits features of an active enhancer

Surprisingly, the Late3 region was as well associated with a positive accessibility score, again implying a functional role early in infection. In addition, a new peak appeared in the middle of the pVII depleted Late3 region at 4 hpi (Figure 25 C). It is unlikely that this newly assembled peak represents either a stable pVII-DNA particle or a nucleosome bound site, since it was not detected in the high-MNase condition. As mild MNase digestions are able to recover DNA interactions with non-histone proteins (Kent et al., 2011; Mieczkowski et al., 2016), the indicated peak most likely resembles DNA binding of a transcription factor. Potential binding sites for host transcription factors EN1, CREM, C/EBP-alpha and C/EBP-beta were identified in the center of the peak (Figure 26A). The RNA-seq data was used to analyze the expression of the indicated transcription factors during infection. Intriguingly, transcription of EN1, CREM and C/EBP-beta was upregulated in the host at 4 hpi (Figure 26B) supporting their relevance in virus infection.

In summary, before viral replication the Late3 region is not transcribed, accessible and potentially bound by transcription factors. Those are characteristic features of active enhancer elements and therefore the Late3 region might contribute to the coordination of viral transcription processes as a distal regulatory element.

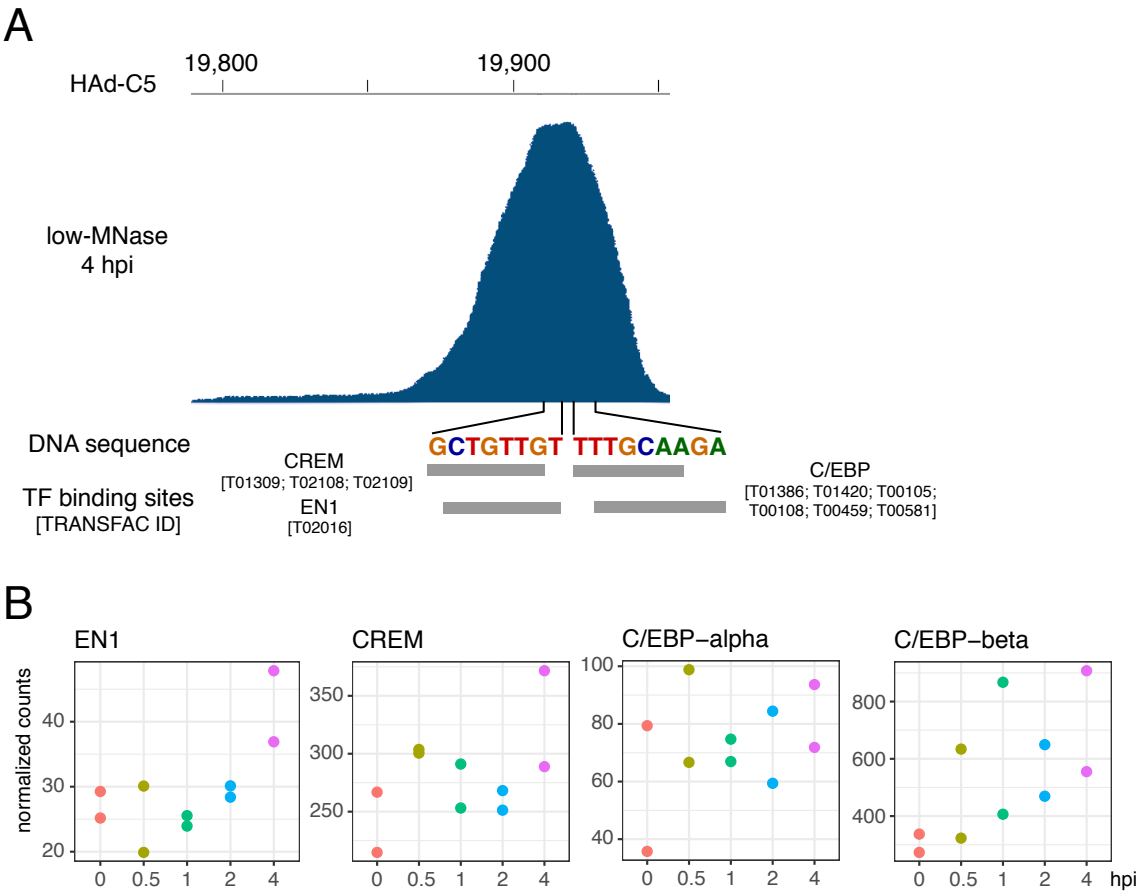


Figure 26: Transcription factor binding analysis at the Late3 region

A: Transcription factor motif analysis of a MNase sensitive peak located at the Late3 region. The locus from 19,800 to 19,950 of the Ad genome is shown. Profile illustrates the number of fragments obtained in low-MNase at 4 hpi. The DNA sequence in the center of the peak was used to search for known DNA binding motifs of human transcription factors in the TRANSFAC data base. Location, DNA sequence and corresponding TRANSFAC ID of the motifs are indicated.
B: Expression dynamics of human transcription factors during the infection time course. RNA abundance of selected transcription factors was estimated from RNA-seq data at indicated time points. The normalized counts of two independent experiments at each time point are shown.

5.1.13 Nucleosomes assemble preferentially at de-condensed chromatin

Early genes and the Late3 region were classified as preferential sites of nucleosome assembly (see section 5.1.8) and simultaneously chromatin de-condenses during infection at those sites (section 5.1.11), indicating a functional link (Figure 24C). To exclude, that this observation is not a random coincidence the accessibility score of each nucleosome position was assessed and the cumulative score over all nucleosome positions was deduced. Next, the cumulative score of randomly distributed nucleosome positions was quantified in 10,000 simulations. The actual nucleosome positions result in a cumulative accessibility score (red dashed line in Figure 27), which is distinct from the main distribution of the randomly sampled sites and is only in very few simulations reached (2 out of 10,000; p-value = 0.0002), suggesting that nucleosome assembly takes place specifically at accessible regions of the viral genome.

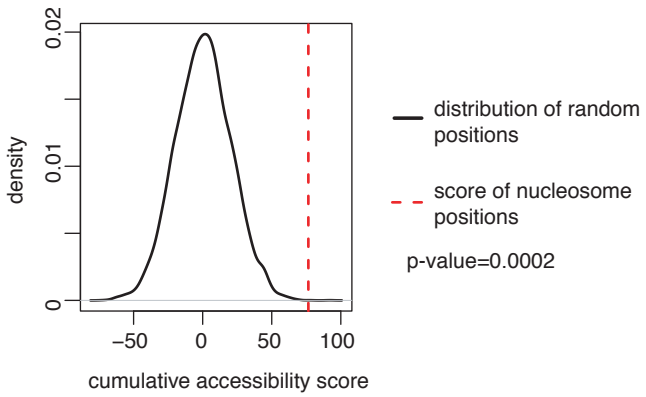


Figure 27: Nucleosomes assemble preferentially at de-condensed chromatin

Kernel density plot showing the distribution of the cumulative accessibility score of randomly shuffled nucleosome positions. The predicted nucleosome positions were randomly shuffled along the Ad genome 10,000 times and the cumulative accessibility score was calculated. The actual cumulative accessibility score of the predicted nucleosome positions is indicated by a red dashed line.

5.2 Discussion

5.2.1 Adenoviruses exhibit a defined pVII organization

Ad genome packaging by pVII plays a fundamental role in regulating viral gene expression (Giberson et al., 2012), however it is still unclear how pVII is organized along the genome. This study reveals for the first time a detailed position map of pVII-nucleoprotein complexes, which gives insights into the structure-function relation of the pVII nucleoprotein architecture. Structural details about the pVII-DNA complex conformation are still missing and it is suggested that 30-250 bp of DNA are protected by pVII multimers consisting of 2-6 molecules (Corden et al., 1976; Daniell et al., 1981; Mirza and Weber, 1982; Vayda et al., 1983). The MNase approach of this study revealed a distinct fraction of DNA fragments from 60 to 70 bp in size originating from the viral genome and indicating repeating units of tightly associated nucleoproteins protecting the viral DNA from continuous nuclease digestion. As only pVII and the terminal protein, which is covalently bound to the 5' ends of the Ad genome, remain associated with the viral DNA after nuclear entry (Giberson et al., 2012), the observed fragmentation pattern can be attributed to stable pVII nucleoprotein complexes placed along the viral genome. Supporting evidence comes from electron microscopy studies which propose that pVII packages the viral genome into a nucleosome-like "beads-on-a-string" structure (Mirza and Weber, 1982). Consistently, simultaneous host chromatin digestion resulted mainly in 147 bp long DNA fragments, which corresponds to the size of DNA wrapped around the histone octamer. In comparison to the defined peak of the host nucleosomes, the size distribution of pVII protected DNA fragments is rather broad, which could be the result of

either 1) a heterogenous association of pVII with DNA or 2) too mild MNase conditions. The MNase digestion used was optimized to isolate mono-nucleosomes, since the degree of DNA protection and compaction by pVII was not known *a priori*. Higher MNase concentrations could better dismantle partial protection of the interconnecting linker DNA, which might arise from tight associations of neighboring pVII complexes. Analogously, the fragment size distribution of the mono-nucleosome fraction is less resolved in low-MNase compared to high-MNase in human cells (see section 4.1.1).

Nevertheless, annotation of DNA fragments obtained in the high-MNase approach revealed a unprecedented genome wide map of specific pVII positions. The defined nucleoprotein structure is already detectable at early timepoints of infection, when the virus has not entered the cell nucleus yet, suggesting that the viral genomes enter the cell with a defined chromatin organization already present in the virus particles. Considering that only 20% of the viral genomes entering the nucleus are transcriptionally active during the time course of this study, which leads to a dominant background signal of inactive viruses, the pVII profile does not substantially change over time in the high-MNase condition. Consequently, the 10 high-MNase samples can be treated as replicates showing the initial chromatin configuration in the virion.

About 238 pVII occupied positions can be annotated onto the Ad genome. The number of pVII particles identified in the MNase-seq approach is in well agreement with previous estimations from electron microscopy (Mirza and Weber, 1982) or cryo-electron tomography (Pérez-Berná et al., 2015) studies predicting 200 or 230 pVII nucleoprotein complexes. Taking into account, that quantitative profiling of viral proteins revealed about 500 copies of pVII associated with one Ad genome (Benevento et al., 2014), the presented data suggests, that pVII binds as homodimer to DNA, which is consistent with the findings of previous biochemical studies (Chatterjee et al., 1985; Everitt et al., 1975).

On average the pVII repeat length is about 100 bp, which reflects the observed length of pVII protected DNA plus a DNA linker of 30 to 40 bp. Remarkably, spacing of pVII particles varies along the viral genome and especially early expressed genes exhibit a wider pVII spacing and are occupied by less stable pVII nucleoprotein complexes indicating an accessible chromatin configuration at early expressed genes. Consistently, Komatsu and colleagues showed that optimizing the degree of pVII packaging enhances the expression of a reporter plasmid after transfection (Komatsu et al., 2011). Taken together, the high resolution of the pVII positioning map obtained through this study reveals a highly organized and functional DNA packaging of the viral genome, which allows prompt activation of early genes.

It is still unclear how the specific positioning of pVII along the Ad genome is coordinated. In eukaryotes nucleosome positioning is determined by the combination of DNA sequence, active nucleosome remodeling and DNA binding of non-histone proteins (Struhl and Segal, 2013). Like nucleosomes, the pVII DNA complex displays characteristic oscillations of A/T and G/C di-nucleotides, indicating a preference to distinct DNA sequences. However, the di-nucleotide periodicity of 5.4 bp is shorter and would not conform to the intrinsic curvature of DNA, which follows the path around the histone octamer surface. Hence, the observed DNA sequence preference of pVII particles points to an alternative nucleoprotein structure, where the DNA forms a spiral around the pVII core. Similarly, it was proposed that sperm DNA is compacted by protamins in a toroidal structure (Wykes and Krawetz, 2003). Interestingly, pVII exhibits stretches of arginine residues that are similar to arginine domains in protamins, indicating similar biochemically properties (Cao and Sung, 1982; Sato and Hosokawa, 1984). In summary, the reported data suggests that pVII positioning is affected by DNA sequence. However, it is unlikely, that such a complex and defined nuclear architecture can be achieved only by DNA sequence composition and other processes like active pVII remodeling during virus maturation could further contribute to establish the defined DNA packaging.

5.2.2 Dynamics of viral chromatin during early infection

Spatiotemporal analysis of the viral genome localization revealed that after 30 min the first viral genomes have entered the host cell nucleus. Nevertheless, viral transcripts were initially detected at 2 hpi showing a lag phase of about 1 hour between nuclear entry and initiation of viral transcription. The delay of viral gene transcription indicates that the condensed Ad core is not a suitable template for the host transcription apparatus and the chromatin structure needs to be adjusted. This observation is consistent with *in vitro* studies showing that the pVII condensed genome is restricted to transcription and requires active chromatin remodeling to stimulate viral gene transcription (Okuwaki and Nagata, 1998). The epigenetic enzyme TAF-Ib has been identified as transcriptional activator, which remodels the Ad nucleoprotein structure, thereby increasing the accessibility of the transcription machinery to the template DNA.

Using the low-MNase approach, which releases preferentially accessible chromatin, chromatin de-condensation of activated viral genomes can be monitored. As gross changes in accessibility are already detected at 1 hpi, the temporal analysis confirms, that remodeling of Ad chromatin precedes active Ad transcription. Noteworthy, chromatin de-compaction takes mainly place at early genes, emphasizing the link between chromatin remodeling and transcription initiation.

Interestingly, in addition to the early gene regions the not transcribed Late3 domain exhibited as well increased DNA accessibility. Active chromatin rearrangements at the central region might be required to allow binding of regulatory transcription factors. In line with this hypothesis an additional peak is showing up at the initially pVII depleted area at 4 hpi in the low-MNase data set. It is unlikely, that this peak represents a pVII particle or a newly assembled nucleosome, since no peak was present at the same position in the high-MNase samples. The low-MNase conditions are capable to resolve weak protein-DNA interactions as known for transcription factors, which would be lost under high-MNase treatment (Henikoff et al., 2011; Kent et al., 2011). Intriguingly, binding motifs for transcription factors, whose expression is upregulated during infection, coincide with the center of the peculiar peak. It would be interesting to test, whether this region might act as an enhancer element and whether deleting or replacing the Late3 region would affect early gene transcription.

Under the experimental conditions of this study transcription initiates only at a subpopulation of imported viral genomes. And in contrast to the low-MNase data set, the pVII profiles using high-MNase, which represent the population average chromatin structure, did not significantly change over time indicating that the chromatin structure of most viral genomes is not altered after nuclear import. It remains an intriguing question, why only a subpopulation of genomes are remodeled and transcriptionally active, and how the viral genomes are selected to be transcribed. Activation of a subpopulation of viral genomes, which produce the required components for viral replication, could be a sneaky way to infiltrate the host cell without triggering an immediate defense response. Extending the time frame of this study could reveal, whether the inactive viral genomes are used later on for genome replication.

5.2.3 Nucleosomes replace pVII at specific positions

The fate of pVII after nuclear genome import is still a topic of debate and it is controversial, whether pVII is evicted through a transcription dependent process (Chen et al., 2007; Komatsu et al., 2011). This study shows, that the fragmentation profile of viral chromatin is altered 2 hours after infection and an additional fraction of nucleosomal fragments was appearing. Nucleosomes were not randomly distributed along the Ad genome, but occupied distinct positions, favoring early transcribed genes. The timepoint and the genomic location of nucleosome assembly at the viral genome coincide with the findings of previous histone ChIP-qPCR experiments (Komatsu et al., 2011). Noteworthy, nucleosomes preferentially assemble at initially pVII occupied positions suggesting a replacement of pVII proteins by histones. Even though most of the nucleosomes are located at early gene regions, some nucleosomes reside in

late transcribed domains, especially at the Late3 region, pointing to a transcription independent exchange mechanism. Interestingly, pVII particles were replaced by nucleosomes at the +1 position of early genes. The well-positioned +1 nucleosome directly downstream of the TSS, is a characteristic feature at active gene promoters in eucaryotes (Schones et al., 2008). Furthermore the region upstream of the newly assembled +1 nucleosome is pVII or nucleosome depleted indicating that the stereotypical chromatin landscaped is formed at the viral gene promoter enabling immediate gene activation. In the setting of this study nucleosome assembly and onset of viral transcription were detected simultaneously (E1A at 2 hpi and E4 at 4 hpi). However, while the RNA abundance of E1A increased greatly at 4 hpi, as well as the number of nucleosomal fragments at other positions, the assembly of the E1A +1 nucleosome was completed at 2 hpi and did not significantly change at 4 hpi implying, that nucleosome assembly is preceding viral gene transcription.

In agreement with previous pVII and histone re-ChIP assays (Komatsu et al., 2011) the presented MNase-seq data indicates that both pVII and cellular histones are associated with the same viral genome forming a "chimeric" chromatin structure in early phases of infection. pVII is exchanged for histones at distinct positions, however at this point it can only be speculated which machinery guides the replacement of pVII by histones and what determines the site specific recruitment. Since TAF-Ib is a known remodeling factor for pVII-DNA complexes and has a histone chaperone activity (Okuwaki and Nagata, 1998), it is possible that TAF-Ib drives histone binding on viral chromatin. However, knock-down experiments of TAF-1b did not reveal any contribution on the binding level of histones on viral chromatin (Komatsu et al., 2011). Giberson and colleagues showed that the histone variant H3.3 is preferentially associated with Ad DNA during early phases of infection (Giberson et al., 2018). In contrast to the histone variant H3.1, H3.3 is expressed throughout all phases of the cell cycle and is deposited through a replication-independent mechanism (Tagami et al., 2004). Known H3.3 histone chaperones are DAXX and HIRA. Interestingly, after fertilization HIRA contributes to H3.3 deposition on incoming male pro-nuclear DNA (Loppin et al., 2005), which is associated with protamins. As pVII exhibits sequence similarities to protamins (Sato and Hosokawa, 1984), it would be an elegant way to mimic the sperm chromatin structure and utilize host cell machineries, such as HIRA, to activate viral genes. Noteworthy, experiments using helper-dependent Ad vectors showed a reduction of histone association accompanied with attenuated transgene expression upon HIRA knockdown (Ross et al., 2011). However, HIRA knockdown did not affect histone deposition onto wild type Ad DNA (Giberson et al., 2018).

The involvement of DAXX, another H3.3 specific histone chaperone, in histone deposition during infection is unlikely, since the activity of DAXX correlates with viral gene repression and degradation of the DAXX/ATRX complex is triggered by viral proteins during early phases of infection (Schreiner et al., 2013).

The results of the presented study are coherent with an activating role of nucleosome assembly at the viral genome, as 1) nucleosomes are preferentially formed at de-compacted regions of the viral genome, 2) the minor fraction of nucleosomes corresponds to the small subpopulation of actively transcribing genomes and 3) nucleosomes form a chromatin structure typical for active genes at early gene promoter.

Taken together the presented study uncovered the adenoviral nucleoprotein structure with high resolution, shows the dynamics of Ad chromatin structure during early infection and how it is related to transcriptional activation. The findings of this study and further exploration could aid in the development of efficient Ad vectors with sustained transgene expression.

6 Materials and Methods

6.1 Materials - Characterizing DNA accessibility in higher-order structures of chromatin

6.1.1 Annotation and publicly available data

Name	Source	Identifier
Human genome assembly (Feb. 2009)	UCSC https://hgdownload.soe.ucsc.edu/downloads.html	GRCh37 / hg19
Chromatin states, 15-state model	ROADMAP https://egg2.wustl.edu/roadmap/web_portal/chr_state_learning.html	E115
100mer CRG mappability track	ENCODE https://hgdownload.soe.ucsc.edu/downloads.html	Align100mer
50mer GEM mappability track	(Derrien et al., 2012) http://boevalab.inf.ethz.ch/FREEC/	out50_hg19
H3K27ac ChIP-seq of HeLa-S3	ENCODE https://www.encodeproject.org/	ENCFF311EWS
H3K27me3 ChIP-seq of HeLa-S3	ENCODE https://www.encodeproject.org/	ENCFF958BAN
DNase-seq of HeLa-S3	ENCODE https://www.encodeproject.org/	ENCFF692NCU
RNA-seq of HeLa-S3	ENCODE https://www.encodeproject.org/	ENCFF000DNW
Enhancer elements of HeLa-S3	(Chen et al., 2016)	N/A

6.1.2 Software and online tools

Name	Source	Version
bowtie2	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml	2.3.4
samtools	http://www.htslib.org/	1.9
bedtools	https://bedtools.readthedocs.io/en/latest/index.html	2.27.1
picard	https://broadinstitute.github.io/picard/	2.18.6

wigtoBigWig	https://github.com/kevlim83/ucsctools	1.04
DANPOS2	https://sites.google.com/site/danposdoc/	2.2.2
HOMER	http://homer.ucsd.edu/homer/	4.10
IGV	http://software.broadinstitute.org/software/igv/home	2.3 - 2.6
Control-FREEC	http://boevalab.inf.ethz.ch/FREEC/	8.7
GCparser	https://github.com/uschwartz/FCE_HeLa	1.0
R	https://cran.r-project.org/	3.1.2 – 3.6.0
seqinr	https://cran.r-project.org/	3.4-5
rtracklayer	https://www.bioconductor.org/	1.32.1
RColorBrewer	https://cran.r-project.org/	1.1-2
ggplot2	https://cran.r-project.org/	2.1.0 - 3.1.0
LSD	https://cran.r-project.org/	4.0.0
GenomicFeatures	https://www.bioconductor.org/	1.22.0
colorRamps	https://cran.r-project.org/	2.3

6.1.3 Data and code availability

The raw and processed MNase-seq data has been submitted to NCBI Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE100401. The code and additional files used to analyze the sequencing data are available at GitHub: https://github.com/uschwartz/FCE_HeLa.

6.2 Methods - Characterizing DNA accessibility in higher-order structures of chromatin

6.2.1 Experimental procedures

6.2.1.1 Functional chromatin extraction

Differential MNase hydrolysis of HeLa cells (ATCC® CCL-2) and subsequent library preparations were carried out by Sarah Diermeier (for details of the protocol see (Diermeier, 2014)). Briefly, HeLa cells were cultivated in DMEM/10% FCS on 15 cm cell culture dishes at 37°C with 5% CO_2. At 70-80% confluency cells were washed with PBS and incubated with permeabilization buffer containing 0.2% NP40 and freshly added 100U/1000U of MNase to perform low-/high-MNase digestions. Stop buffer containing EDTA and 1% SDS was used to stop DNA hydrolysis by the MNase nuclease. Subsequently RNAs and proteins were degraded

by the addition of RNase A and proteinase K. After DNA isolation, DNA fragments were separated by size on 1.1% agarose gels and the mono-nucleosomal DNA band was excised. DNA sequencing libraries were prepared according to manufacturer's instructions using the NEBNext DNA Library Prep Master Kit (New England Biolabs). Paired-end sequencing was carried out on Illumina HiSeq2000 platform at the EMBL GeneCore facility.

6.2.2 Data analysis

6.2.2.1 Genome alignment and postprocessing

Paired-end reads obtained from MNase chromatin digestions in HeLa cells and subsequent high-throughput sequencing were aligned to the human reference assembly (hg19) using the bowtie2 aligner (Langmead and Salzberg, 2012). Local alignments were performed using the *--very-sensitive-local* option and read pairs mapping in wrong directions or unexpected distances, e.g. different chromosomes, were removed specifying the option *--no-discordant*. Further quality filtering was carried out using the samtools package (Li et al., 2009). Reads exhibiting a mapping quality score (MAPQ) below 20 were removed and only properly mapped read pairs were kept for further analysis. The *CollectInsertSizeMetrics* script from the Picard software suite (https://github.com/broadinstitute/picard) was used to analyze the fragment size distributions. Fragments representing intact mono-nucleosomal DNA with a size of 140-200 bp were extracted and used for downstream analysis.

6.2.2.2 Nucleosome frequency maps

The DANPOS2 package (Chen et al., 2013b) was used to obtain genome wide nucleosome frequency maps. First, read-pairs were shifted to the fragment center and then extended to 70 bp. Each sample was normalized to 200 million fragments using the option *-c* to account for variations in sequencing depth. High numbers of clonal reads due to ambiguous regions in the reference genome (blacklisted regions) or PCR overamplification were corrected using the option *--clonalcut 1e-2*.

Copy number variations in the used HeLa cell line were called using the Control-FREEC tool (Boeva et al., 2012) and subsequently used to normalize the nucleosome occupancy profiles. DNA extracted from sonicated chromatin following single-end high-throughput sequencing was used to determine CNVs. The raw reads were aligned to the human genome using bowtie2 aligner with the specification *--very-sensitive-local* and subsequently mapped reads were sorted using samtools. Following specifications were used to run the Control-FREEC algorithm: *ploidy = 3*; *window = 500000*; *step = 50000*; *minMappabilityPerWindow = 0.80*;

breakPointThreshold = 3. Furthermore, a hg19 mappability track (available at http://boevalab.com/FREEC/), which was designed for single-end reads of 50 bp length and allowing up to two mismatches, was integrated to the analysis.
For visualization of nucleosome frequency distributions, the data tracks were loaded into the IGV genome browser.

6.2.2.3 Differential MNase analysis

The average nucleosome fragment frequency of the high-MNase / low-MNase normalized profile was assessed in non-overlapping bins of constant sizes (ranging from 250 bp to 1 Mb). The average GC content from the corresponding bin was extracted from the fasta genome sequence files using the R package seqinr. Mappability scores were derived from the 100mer CRG alignability track (Derrien et al., 2012) for each bin. Only bins exhibiting a mappability score above 0.9 on average were used for further analysis. Coefficients of determination (R^2) or correlation coefficients (R) were determined by linear regression analysis using the *lm* function in the R environment. Beforehand, Cooks distance measurements were used to remove highly influential data points applying a Cooks distance threshold of 0.001.

6.2.2.4 Nucleosome positioning analysis and meta-profiles

Nucleosome positions were called on the CNV normalized nucleosome frequency maps using the *dpos* script of the DANPOS2 software package with default settings. The DANPOS2 function *profile* was used generate nucleosome meta-profiles centered at DNase I HS sites (ENCODE accession: ENCFF692NCU), active enhancer (Chen et al., 2016) and at the TSS of expressed genes (ENCODE accession: ENCFF000DNW)

6.2.2.5 GC normalization of nucleosome frequency maps

To estimate the impact of GC content on the individual MNase experiments a locally weighted scatterplot smoothing (LOESS) was applied. Therefore, a LOESS curve was fitted to the nucleosome frequency versus the GC content in 250 bp bins using the R function *loess*. The degree of smoothing was set to 0.025 using the option *span*. The curve fit was deduced by interpolation of data points in a distinct segment specifying the option *control=loess.control(surface="interpolate", cell=0.1)*. The normalization factor at a distinct GC content was deduced from the LOESS curve. Bins exhibiting an extreme average GC content of more than 70% or less than 30% were removed from the analysis. Subsequently the

nucleosome frequency in the 250 bp bins was divided by the corresponding GC normalization factor.

6.2.2.6 Histone modifications and chromatin states

To correlate nucleosome accessibility with histone modifications in repressed and active genomic regions, H3K27ac (ENCFF311EWS) and H3K27me3 (ENCFF958BAN) ChIP-seq tracks were downloaded from the ENCODE repository. The GC normalized nucleosome frequency of high-MNase and low-MNase samples, as well as the read density of the H3K27 histone modifications, were assessed in 10 kb bins.

Chromatin state annotations (HeLa, accession E115), which were deduced from combinatorial interactions between different chromatin marks in their spatial context, were obtained from the roadmap epigenome project (Roadmap Epigenomics Consortium et al., 2015). The genome was segmented into bins according to the chromatin state classification. The GC normalized nucleosome frequency of high-MNase and low-MNase samples was assessed and the log2 ratio was calculated.

6.2.2.7 Simulations of FCE

An *in silico* model predicting the release of nucleosomes from compacted and open chromatin using different MNase concentrations was designed. To simulate the endo-nuclease cleavage of a nucleosome array composed of 50 regularly spaced nucleosomes, an index *PosCut* was created containing the positions of one potential cleavage site for each linker and each nucleosome. The index was designed as follows: The cleavage site of the first linker was set to position 0 ($posCut(linker)_1 = 0$) and subsequent linker are regularly spaced by a distance of 180 positions ($posCut(linker)_n = 180 \cdot (n-1)$). The same procedure was applied to create the intra-nucleosome cleavage index, starting at position 90 ($posCut(nuc)_1 = 90$; $posCut(nuc)_n = 90 + 180 \cdot (n-1)$). To distinguish intra- and inter-nucleosome cleavage events by size of the obtained fragments, every second nucleosome cleavage index was shifted 5 positions downstream, accordingly $posCut(linker)_n - posCut(linker)_{n-1} = 180$ and $posCut(nuc)_n - posCut(nuc)_{n-1} = 180 \pm 5$. This resulted in a cleavage index along the nucleosome array:

$PosCut = [\ posCut(linker)_1 = 0, posCut(nuc)_1 = 85, posCut(linker)_2 = 180, posCut(nuc)_2 = 270, \ldots, posCut(nuc)_{50} = 8910, posCut(linker)_{51} = 9000\ ]$

Next, each cleavage site (*posCut*) was assigned with a cleavage probability (p_{cut}) as indicated in the text. According to the p_{cut} values the cleavage positions were sampled and according to the number of simulated cutting events (n), n cleavage positions were randomly drawn from this probability distribution without replacement. To simulate high-MNase digestions, n = 70 was chosen and n = 20 for low-MNase simulations, respectively. The collected cleavage positions were sorted and the distance between neighboring cleavage sites was used to deduce the fragment length. Finally the fragment length was trimmed by 30 bp and fragments of 150 bp length, representing mono-nucleosomes, were collected. The simulation was iterated until at least one million mono-nucleosomal fragments have been collected and the coverage of released nucleosomes along the nucleosome array was plotted.

6.3 Materials - Analysing Adenovirus chromatin dynamics early in infections

6.3.1 Annotation and publicly available data

Name	Source	Identifier
Human genome assembly (Feb. 2009)	UCSC https://hgdownload.soe.ucsc.edu/downloads.html	GRCh37 / hg19
Human adenovirus C serotype 5 assembly	NCBI https://www.ncbi.nlm.nih.gov/nuccore	AY339865.1
Human gene annotation	GENCODE https://www.gencodegenes.org/	Release 25 (GRCh37)

6.3.2 Software and online tools

Name	Source	Version
bowtie2	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml	2.3.4
STAR	https://github.com/alexdobin/STAR	2.6.1c
samtools	http://www.htslib.org/	1.9
bedtools	https://bedtools.readthedocs.io/en/latest/index.html	2.27.1
picard	https://broadinstitute.github.io/picard/	2.19.2
subread	http://subread.sourceforge.net/	1.6.4
wigtoBigWig	https://github.com/kevlim83/ucsctools	1.04
DANPOS2	https://sites.google.com/site/danposdoc/	2.2.2
HOMER	http://homer.ucsd.edu/homer/	4.10
IGV	http://software.broadinstitute.org/software/igv/home	2.6.1-2.8.2
PROMO	http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3	3.0.2
R	https://cran.r-project.org/	3.6.0
rtracklayer	https://www.bioconductor.org/	1.32.1
RColorBrewer	https://cran.r-project.org/	1.1-2
ggplot2	https://cran.r-project.org/	3.1.0
GenomicFeatures	https://www.bioconductor.org/	1.32.3
nucleR	https://www.bioconductor.org/	2.14.0

DESeq2	https://www.bioconductor.org/	1.20.1
Gviz	https://www.bioconductor.org/	1.29.0

6.3.3 Data and code availability

The raw and processed MNase-seq and RNA-seq data has been submitted to NCBI Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE136550. The code and additional files used to analyze the sequencing data are available at GitHub: https://github.com/uschwartz/AdVir .

6.4 Materials - Analysing Adenovirus chromatin dynamics early in infections

6.4.1 Experimental procedures

6.4.1.1 Cell culture and infection

Cell culture and adenovirus infections were carried out by Tetsuro Komatsu and Floriane Lagadec in the lab of Harald Wodrich at the University of Bordeaux. Briefly, H1299 cells (ATCC® CRL-5803) were cultivated in DMEM supplemented with 10 % FCS and 1% Penicillin/Streptavidin at 37°C and 5% CO_2. For infections, cells were incubated with 3000 physical particles/cell of E3-deleted HAd-C5 virus for 30 min. Subsequently cells were washed with PBS to remove unbound viruses. Afterwards cells were incubated with fresh medium at 37°C and collected at different time points for further analysis.

6.4.1.2 Immunofluorescence

Immunofluorescence (IF) staining and data evaluation was carried out by Tetsuro Komatsu. Briefly, cells were grown on 15 mm glass cover slips and fixed in 4% paraformaldehyde/PBS at different time points after infection. First cover slips were incubated in IF buffer consisting of PBS, 10% FCS and 0.1% Saponin for 15 min. Subsequently, IF buffer was replaced by antibody solution and cells were incubated with primary antibodies (mouse anti-pVI and mouse anti-pVII) for 1 h at room temperature. Next, cells were washed with PBS and incubated with secondary antibody (mouse Alexa 647) for 1 h at room temperature. Unbound secondary antibodies were removed by washing with PBS, ddH_2O and absolute ethanol. Finally, cells were air-dried and mounted in DAKO mixed with DAPI (1:1000).

6.4.1.3 RNAscope

RNAscope experiments and data evaluation was carried out by Floriane Lagadec. Briefly, experiments were performed according to the manufacturer's (ACD Bio) protocols from the RNAscope® Multiplex Fluorescent Assay. Cells were grown on 15 mm glass cover slips and fixed in 4% paraformaldehyde/PBS at indicated time points after infection. Protease III diluted 1:10 in PBS was added to the cells and after incubation for 15 min at room temperature, cells were washed three times with PBS. Subsequently, E1A RNAscope probes were hybridized for 2 h at 40°C in a humidified chamber. After washing with PBS, hybridization with signal amplifier probes was carried out following manufacturer's instructions. Finally, cells were washed with PBS and incubated in IF buffer for 15 min at room temperature.

6.4.1.4 RNA library preparation and sequencing

RNA isolation was carried out by Tetsuro Komatsu and RNA sequencing libraries were prepared and sequenced at the EMBL GeneCore facility in Heidelberg. TRI Reagent (Molecular Research Center, Inc.) was used to isolate total cellular RNA from infected cells at indicated time points after infection. Sequencing libraries were prepared using TruSeq RNA Sample Preparation Kit v2 (Illumina) according to the manufacturer's protocol. Briefly, RNAs were enriched for polyA tails using oligo(dT) beads, cleaved to 200-300 bp fragments, reverse transcribed to cDNA using random hexamer primers and sequencing adapters were ligated to the dsDNA fragments. Single-end RNA sequencing was carried out on Illumina HiSeq2000 platform.

6.4.1.5 Functional chromatin extraction

MNase digestions of infected cells and subsequent library preparation were carried out by Claudia Huber. Briefly, chromatin of infected cells was partially/completely digested using 100U/600U of MNase for 4/5 min at 37°C (see section 6.2.1). DNA was isolated and the mono- and sub-nucleosomal bands were excised from 1.3% agarose gels. Next, sequencing libraries were prepared using NEBNext DNA library prep Master Mix Kit (New England Biolabs) according to the manufacturer's protocol. Paired-end sequencing was carried out on Illumina HiSeq2000 platform at the EMBL GeneCore facility.

6.4.2 Data analysis

6.4.2.1 Adenovirus annotation

The DNA sequence and the gene annotation of the Human Adenovirus C serotype 5 (accession AY339865.1) was obtained from the NCBI nucleotide archive. The used virus strain exhibits a partial E3-deletion, therefore nucleotides 28,593-30,464 were excised from the complete Ad reference sequence and the corresponding gene annotation. Furthermore, the annotation of the protein coding genes was converted to GTF data format for subsequent analyses. The amended versions of the Ad annotation are available at the github project page (https://github.com/uschwartz/AdVir).

6.4.2.2 RNA-seq analysis

Spliced transcript alignment was conducted using the STAR aligner (Dobin et al., 2013). First a combined reference genome index consisting of the modified Ad sequence (see section 6.4.2.1) and the human genome assembly (hg19) was created using the *genomeGenerate* script. Known splice junctions of human transcripts (GENCODE version 25) were integrated to the genome index using the option *–sjdbGTFfile* and the length of the genomic sequence around the annotated junction was specified with *–sjdbOverhang* option to 49 bp. RNA-seq reads were aligned to the reference using modified STAR settings. To adapt the alignment to the well characterized splicing features of human transcripts, following standard ENCODE options were used: *--outFilterType BySJout*; *--outFilterMultimapNmax 20*; *--alignSJoverhangMin 8*; *--alignSJDBoverhangMin 1*; *--alignIntronMin 20*; *--alignIntronMax 1000000*. Furthermore, to avoid unspecific read splicing during alignment, only alignments containing canonical junctions were kept using the specification *--outFilterIntronMotifs RemoveNoncanonical* and *--outSAMstrandField intronMotif*. The maximum number of recorded multiple alignments was restricted to one using *--outSAMmultNmax* option and multiple highest score alignments were ranked randomly specifying *--outMultimapperOrder Random*. For visualization of RNA-seq coverage along the genome, normalized wig files were created setting the STAR options: *--outSAMtype BAM SortedByCoordinate*; *--outWigType wiggle*; *--outWigStrand Unstranded*. Finally, wig files were converted to bigwig format using the wigToBigWig script from UCSCtools package (https://github.com/kevlim83/ucsctools).

The featureCounts function from the subread software package (Liao et al., 2014) was used to quantify the abundance of RNA reads mapping to viral and human genes. Following options were used: *-t exon*; *-g gene_id* and *-s 0*. The read counts for each gene were normalized to reads

per million mapped reads (RPM) or transcripts per million mapped reads (TPM) (Wagner et al., 2012):

$$RPM\ of\ a\ gene = \frac{Number\ of\ reads\ mapped\ to\ a\ gene \times 10^6}{Total\ number\ of\ mapped\ reads}$$

$$TPM\ of\ a\ gene = \frac{number\ of\ reads\ mapped\ to\ a\ gene \times 10^6}{total\ number\ of\ mapped\ reads \times\ length\ of\ gene\ in\ bp}$$

6.4.2.3 Functional Chromatin Extraction analysis

Paired-end reads obtained from MNase chromatin digestions of infected cells and subsequent high-throughput sequencing were simultaneously aligned to the Ad (see section 6.4.2.1) and human (hg19) reference assembly using bowtie2 (Langmead and Salzberg, 2012). First a combined reference genome index was created using the *bowtie-build* script. Local alignments were conducted using the *--very-sensitive-local* option and read pairs mapping in wrong directions or unexpected distances, e.g. different chromosomes, were removed specifying the option *--no-discordant*. Further quality filtering was carried out using the samtools package (Li et al., 2009). Reads exhibiting a mapping quality score (MAPQ) below 30 were removed and read pairs properly mapped to the Ad genome were extracted for further analysis. The *CollectInsertSizeMetrics* script from the Picard software suite was used to analyze the fragment size distribution (https://github.com/broadinstitute/picard).

6.4.2.4 pVII positioning

Since the size distribution of pVII fragments exhibited its maximum between 60 and 70 bp, fragments shorter than 140 bp were selected for further analysis to avoid fragments originating from pVII di-particles. The DANPOS2 software suite was used to annotate pVII positions along the Ad genome. First, pVII read-pairs were shifted to the fragment center and a sequencing-depth normalized coverage profile was calculated using the *dpos* function with following options: *-u 1e-5*; *-a 1*; *-z 5*; *-p 0.001*; *--extend 35* and *--mifrsz 0*. Between sample normalization was carried out using the *wiq* function. Thereby, a quantile normalization is applied to each sample and the profiles were normalized to the coverage distribution of the first replicate at 0 hpi from the low-/high-MNase digestion series. Finally, the *dpos* script was executed at the normalized pVII occupancy profiles to call pVII positions using following specifications: *-jd 50*; *-q 10*; *-z 5*; *-a 1*.

6.4.2.5 Rotational positioning

The periodic occurrence of A/T- and G/C-containing di-nucleotides at human nucleosomes is correlated with DNA bending into minor and major groove and can be visualized in MNase-seq experiments. Therefore, nucleosomal fragments mapped to the human genome of 147 bp in size were extracted from the first replica at 0 hpi of the high-MNase samples. The *annotatePeaks.pl* function from the HOMER software package (Heinz et al., 2010) was executed to calculate the A/T and G/C di-nucleotide frequencies around the fragment centers. Next, the di-nucleotide frequencies were smoothed using the Fourier transformation (*filterFFT* function with the option *pcKeepComp=0.2*) of the nucleR Bioconductor package (Flores and Orozco, 2011). To account for the nucleosome symmetry and un-stranded orientation, the mirrored di-nucleotide profile was added and the total frequency was divided by 2. The same analysis was repeated to study the rotational positioning of pVII, except that 49 to 51 bp sized fragments, which aligned to the virus genome, were selected from all high-MNase samples.

6.4.2.6 Detection of nucleosome assembly sites

To specify the locations of nucleosome assembly onto the viral genome, fragments of the typical nucleosome size between 137 – 157 bp were extracted from the high-MNase samples at 1, 2 and 4 hpi. The enrichment of nucleosome sized fragments at distinct positions over time was analyzed using the Bioconductor package nucleR. First, the positions of the selected fragments were loaded in the R environment and converted to coverage profiles, which were normalized by sequencing depth to fragments per million Ad mapped reads. Next, the built-in *filterFFT* function was used to smooth the profile specifying the option *pcKeepComp=0.007*. Afterwards, the *peakDetection* function was applied to the nucleosome occupancy profiles at 2 or 4 hpi in order to call nucleosome positions of 147 bp width. The coverage profile at 1 hpi was considered as background of possible contaminations of pVII di-particle fragments. To obtain nucleosome positions with high probability, the signal ratio to the background was calculated and a stringent cut-off of at least 5-fold enrichment was used to define sites of nucleosome assembly.

6.4.2.7 Dynamics of Ad genome accessibility

To characterize alterations in Ad chromatin accessibility during early infection, the relative release efficiency of pVII was analyzed using low-MNase samples. Analogously to the high-MNase samples (see section 6.4.2.4), normalized pVII profiles were generated from pVII fragments released under mild MNase conditions. Here, variations of pVII frequency over time

would reflect changes in the accessibility of the pVII linker DNA (see section 4.1.9). To score the chromatin accessibility the average pVII fragment coverage was calculated using 300 bp bins sliding by 20 bp steps along the Ad genome. The average pVII frequency in a particular bin was plotted versus the corresponding time points and a linear regression line was fitted to the data. Alterations of chromatin accessibility over time were scored by the slope of the linear fit.

6.4.2.8 Transcription factor motifs

To identify potential transcription factors, which might bind to the Ad genome at the Late3 region, the DNA sequence in the center (position 19,890-19,940) of the newly formed peak in the low-MNase samples at 4 hpi was extracted. The TRANSFAC (version 8.3) database was screened for known transcription factor binding motifs in human and mouse matching sub-sequences of the query DNA using the PROMO web server. The motif search was conducted using a maximum dissimilarity rate of 2% and motifs with a recognition sequence shorter than 5 nucleotides were removed. Finally, the RNA-seq data was used to test, whether the corresponding transcription factor was expressed in H1299 cells in the experimental setup. The default DESeq2 analysis (Love et al., 2014) was carried out and motifs were removed, if the corresponding transcription factor was not considered for statistical analysis due to independent filtering.

7 References

Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al. (2014). An atlas of active enhancers across human cell types and tissues. Nature *507*, 455–461.

B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, P.W. (2002). Molecular Biology of the Cell, 4th edition tle.

Babokhov, M., Hibino, K., Itoh, Y., and Maeshima, K. (2020). Local Chromatin Motion and Transcription. J. Mol. Biol. *432*, 694–700.

Bancaud, A., Huet, S., Daigle, N., Mozziconacci, J., Beaudouin, J., and Ellenberg, J. (2009). Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. EMBO J. *28*, 3785–3798.

Belmont, A.S., Dietzel, S., Nye, A.C., Strukov, Y.G., and Tumbar, T. (1999). Large-scale chromatin structure and function. Curr. Opin. Cell Biol. *11*, 307–311.

Benevento, M., Di Palma, S., Snijder, J., Moyer, C.L., Reddy, V.S., Nemerow, G.R., and Heck, A.J.R. (2014). Adenovirus composition, proteolysis, and disassembly studied by in-depth qualitative and quantitative proteomics. J. Biol. Chem. *289*, 11421–11430.

Boeva, V., Popova, T., Bleakley, K., Chiche, P., Cappo, J., Schleiermacher, G., Janoueix-Lerosey, I., Delattre, O., and Barillot, E. (2012). Control-FREEC: A tool for assessing copy number and allelic content using next-generation sequencing data. Bioinformatics *28*, 423–425.

Boyle, S., Gilchrist, S., Bridger, J.M., Mahy, N.L., Ellis, J.A., and Bickmore, W.A. (2001). The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. Hum. Mol. Genet. *10*, 211–219.

Brahma, S., and Henikoff, S. (2018). RSC-Associated Subnucleosomes Define MNase-Sensitive Promoters in Yeast. Mol. Cell 1–12.

Brogaard, K., Xi, L., Wang, J., and Widom, J. (2012). base-pair resolution. 2–7.

Cao, T.M., and Sung, M.T. (1982). A protamine-like domain in basic adenovirus core protein. Biochem. Biophys. Res. Commun. *108*, 1061–1066.

Chatterjee, P.K., Vayda, M.E., and Flint, S.J. (1985). Interactions among the three adenovirus core proteins. J. Virol. *55*, 379–386.

Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S., et al. (2013a). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell *155*, 1479–1491.

Chen, J., Morral, N., and Engel, D.A. (2007). Transcription releases protein VII from adenovirus chromatin. Virology *369*, 411–422.

Chen, K., Xi, Y., Pan, X., Li, Z., Kaestner, K., Tyler, J., Dent, S., He, X., and Li, W. (2013b). DANPOS: Dynamic analysis of nucleosome position and occupancy by sequencing. Genome Res. *23*, 341–351.

Chen, Y., Pai, A.A., Herudek, J., Lubas, M., Meola, N., Järvelin, A.I., Andersson, R., Pelechano, V., Steinmetz, L.M., Jensen, T.H., et al. (2016). Principles for RNA metabolism and alternative transcription initiation within closely spaced promoters. Nat. Genet. *48*, 984–994.

Chereji, R. V., Ocampo, J., and Clark, D.J. (2017). MNase-Sensitive Complexes in Yeast: Nucleosomes and Non-histone Barriers. Mol. Cell *65*, 565-577.e3.

Chereji, R. V., Eriksson, P.R., Ocampo, J., Prajapati, H.K., and Clark, D.J. (2019a). Accessibility of promoter DNA is not the primary determinant of chromatin-mediated gene regulation. Genome Res. *29*, 1985–1995.

Chereji, R. V., Bryson, T.D., and Henikoff, S. (2019b). Quantitative MNase-seq accurately maps nucleosome occupancy levels. Genome Biol. *20*, 198.

Chereji, V., Kan, T., Grudniewska, M.K., Romashchenko, V., Berezikov, E., Zhimulev, I.F., Guryev, V., Morozov, V., and Moshkin, Y.M. (2015). Genome-wide profiling of nucleosome sensitivity and chromatin accessibility in Drosophila melanogaster. 1–16.

Cockell, M., Rhodes, D., and Klug, A. (1983). Location of the primary sites of micrococcal nuclease cleavage on the nucleosome core. J. Mol. Biol. *170*, 423–446.

Corden, J., Engelking, H.M., and Pearson, G.D. (1976). Chromatin-like organization of the adenovirus chromosome. Proc. Natl. Acad. Sci. U. S. A. *73*, 401–404.

Crawford, G.E., Holt, I.E., Whittle, J., Webb, B.D., Tai, D., Davis, S., Margulies, E.H., Chen, Y.D., Bernat, J.A., Ginsburg, D., et al. (2006). Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS). Genome Res. *16*, 123–131.

Cremer, M., and Cremer, T. (2019). Nuclear compartmentalization, dynamics, and function of regulatory DNA sequences. Genes, Chromosom. Cancer *58*, 427–436.

Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat. Rev. Genet. *2*, 292–301.

Cremer, T., Cremer, M., Hübner, B., Silahtaroglu, A., Hendzel, M., Lanctôt, C., Strickfaden, H., and Cremer, C. (2020). The Interchromatin Compartment Participates in the Structural and Functional Organization of the Cell Nucleus. BioEssays *42*, 1900132.

Crisostomo, L., Soriano, A.M., Mendez, M., Graves, D., and Pelka, P. (2019). Temporal dynamics of adenovirus 5 gene expression in normal human cells. PLoS One *14*, e0211192.

Daniell, E., Groff, D.E., and Fedor, M.J. (1981). Adenovirus chromatin structure at different stages of infection. Mol. Cell. Biol. *1*, 1094–1105.

Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., Hilton, J.A., Jain, K., Baymuradov, U.K., Narayanan, A.K., et al. (2018). The Encyclopedia of DNA elements (ENCODE): Data portal update. Nucleic Acids Res. *46*, D794–D801.

Derrien, T., Estellé, J., Sola, S.M., Knowles, D.G., Raineri, E., Guigó, R., and Ribeca, P. (2012). Fast computation and applications of genome mappability. PLoS One *7*, 30377.

Diermeier, S.D.H. (2014). Local and global analysis of the structure-function relation in chromatin.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: Ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R.R., and Richmond, T.J. (2004). Nucleosome arrays reveal the two-start organization of the chromatin fiber. Science (80-.). *306*, 1571–1573.

Dross, N., Spriet, C., Zwerger, M., Müller, G., Waldeck, W., and Langowski, J. (2009). Mapping eGFP Oligomer Mobility in Living Cell Nuclei. PLoS One *4*, e5041.

Dubochet, J., Adrian, M., Schultz, P., and Oudet, P. (1986). Cryo-electron microscopy of vitrified SV40 minichromosomes: the liquid drop model. EMBO J. *5*, 519–528.

Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J., Kaul, R., et al. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Ernst, J., and Kellis, M. (2012). ChromHMM: Automating chromatin-state discovery and characterization. Nat. Methods *9*, 215–216.

Everitt, E., Lutter, L., and Philipson, L. (1975). Structural proteins of adenoviruses. XII. Location and neighbor relationship among proteins of adenovirion type 2 as revealed by enzymatic lodination, immunoprecipitation and chemical cross-linking. Virology *67*, 197–208.

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

Flores, O., and Orozco, M. (2011). nucleR: a package for non-parametric nucleosome positioning. Bioinformatics *27*, 2149–2150.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of Chromosomal Domains by Loop Extrusion. Cell Rep. *15*, 2038–2049.

Giberson, A.N., Davidson, A.R., and Parks, R.J. (2012). Chromatin structure of adenovirus DNA throughout infection. Nucleic Acids Res. *40*, 2369–2376.

Giberson, A.N., Saha, B., Campbell, K., Christou, C., Poulin, K.L., and Parks, R.J. (2018). Human adenoviral DNA association with nucleosomes containing histone variant H3.3 during the early phase of infection is not dependent on viral transcription or replication. Biochem. Cell Biol. *96*, 797–807.

Grosberg, A.Y., Nechaev, S.K., and Shakhnovich, E.I. (1988). The role of topological constraints in the kinetics of collapse of macromolecules. J. Phys. *49*, 2095–2100.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Henikoff, J.G., Belsky, J.A., Krassovsky, K., Macalpine, D.M., and Henikoff, S. (2011). Epigenome characterization at single base-pair resolution.

Hihara, S., Pack, C.G., Kaizu, K., Tani, T., Hanafusa, T., Nozaki, T., Takemoto, S., Yoshimi, T., Yokota, H., Imamoto, N., et al. (2012). Local Nucleosome Dynamics Facilitate Chromatin Accessibility in Living Mammalian Cells. Cell Rep. *2*, 1645–1656.

van Holde, K., and Zlatanova, J. (2007). Chromatin fiber structure: Where is the problem now? Semin. Cell Dev. Biol. *18*, 651–658.

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

Ishii, H., Kadonaga, J.T., and Ren, B. (2015). MPE-seq , a new method for the genome-wide analysis of chromatin structure. 1–9.

Iwafuchi-Doi, M., Donahue, G., Kakumanu, A., Watts, J.A., Mahony, S., Pugh, B.F., Lee, D., Kaestner, K.H., and Zaret, K.S. (2016). The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. Mol. Cell *62*, 79–91.

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

Jost, K.L., Bertulat, B., and Cardoso, M.C. (2012). Heterochromatin and gene positioning: Inside, outside, any side? Chromosoma *121*, 555–563.

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

Kent, N.A., Adams, S., Moorhouse, A., and Paszkiewicz, K. (2011). Chromatin particle spectrum analysis : a method for comparative chromatin structure analysis using paired-end mode next-generation DNA sequencing. Nucleic Acids Res. *39*.

Kim, Y., Shi, Z., Zhang, H., Finkelstein, I.J., and Yu, H. (2019). Human cohesin compacts DNA by loop extrusion. Science (80-.). *366*, 1345–1349.

Koch, F., Fenouil, R., Gut, M., Cauchy, P., Albert, T.K., Zacarias-Cabeza, J., Spicuglia, S., De La Chapelle, A.L., Heidemann, M., Hintermair, C., et al. (2011). Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. Nat. Struct. Mol. Biol. *18*, 956–963.

Komatsu, T., Haruki, H., and Nagata, K. (2011). Cellular and viral chromatin proteins are positive factors in the regulation of adenovirus gene expression. Nucleic Acids Res. *39*, 889–901.

Komatsu, T., Dacheux, D., Kreppel, F., Nagata, K., and Wodrich, H. (2015). A method for visualization of incoming adenovirus chromatin complexes in fixed and living cells. PLoS One *10*, e0137102.

Kornberg, R.D. (1974). Chromatin structure: A repeating unit of histones and DNA. Science (80-.). *184*, 868–871.

Kubik, S., Bruzzone, M.J., Jacquet, P., Falcone, J.L., Rougemont, J., and Shore, D. (2015). Nucleosome Stability Distinguishes Two Different Promoter Types at All Protein-Coding Genes in Yeast. Mol. Cell *60*, 422–434.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Liao, Y., Smyth, G.K., and Shi, W. (2014). FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.

Lieberman-Aiden, E., Van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science (80-.). *326*, 289–293.

Loppin, B., Bonnefoy, E., Anselme, C., Laurençon, A., Karr, T.L., and Couble, P. (2005). The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus. Nature *437*, 1386–1390.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*.

Luger, K. (2003). Structure and dynamic behavior of nucleosomes. Curr. Opin. Genet. Dev. *13*, 127–135.

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

Maeshima, K., Imai, R., and Tamura, S. (2014). Chromatin as dynamic 10-nm fibers. 225–237.

Maeshima, K., Kaizu, K., Tamura, S., Nozaki, T., Kokubo, T., and Takahashi, K. (2015). The physical size of transcription factors is key to transcriptional regulation in chromatin domains. J. Phys. Condens. Matter *27*, 064116.

Maeshima, K., Ide, S., and Babokhov, M. (2019). Dynamic chromatin organization without the 30-nm fiber. Curr. Opin. Cell Biol. *58*, 95–104.

Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. Nature *469*, 343–349.

Meier, O., and Greber, U.F. (2004). Adenovirus endocytosis. J. Gene Med. *6*, S152–S163.

Mieczkowski, J., Cook, A., Bowman, S.K., Mueller, B., Alver, B.H., Kundu, S., Deaton, A.M., Urban, J.A., Larschan, E., Park, P.J., et al. (2016). MNase titration reveals differences between nucleosome occupancy and chromatin accessibility. Nat. Commun. *7*, 11485.

Mirza, M.A., and Weber, J. (1982). Structure of adenovirus chromatin. Biochim. Biophys. Acta *696*, 76–86.

Moyle-Heyrman, G., Zaichuk, T., Xi, L., Zhang, Q., Uhlenbeck, O.C., Holmgren, R., Widom, J., and Wang, J.P. (2013). Chemical map of Schizosaccharomyces pombe reveals species-specific features in nucleosome positioning. Proc. Natl. Acad. Sci. U. S. A. *110*, 20158–20163.

Nevins, J.R., Ginsberg, H.S., Blanchard, J.M., Wilson, M.C., and Darnell, J.E. (1979). Regulation of the primary expression of the early adenovirus transcription units. J. Virol. *32*, 727–733.

Noll, M., Thomas, J.O., and Kornberg, R.D. (1975). Preparation of Native Chromatin and Damage Caused by Shearing. Science (80-.). *187*, 1203–1206.

Okuwaki, M., and Nagata, K. (1998). Template activating factor-I remodels the chromatin structure and stimulates transcription from the chromatin template. J. Biol. Chem. *273*, 34511–34518.

Olins, A.L., Douglas Carlson, R., and Olins, D.E. (1975). Visualization of chromatin substructure: ν bodies. J. Cell Biol. *64*, 528–537.

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

Passarge, E. (1979). Emil Heitz and the concept of heterochromatin: Longitudinal chromosome differentiation was recognized fifty years ago. Am. J. Hum. Genet. *31*, 106–115.

Pérez-Berná, A.J., Marion, S., Chichón, F.J., Fernández, J.J., Winkler, D.C., Carrascosa, J.L., Steven, A.C., Šiber, A., and San Martín, C. (2015). Distribution of DNA-condensing protein complexes in the adenovirus core. Nucleic Acids Res. *43*, 4274–4283.

Rhee, H.S., Bataille, A.R., Zhang, L., and Pugh, B.F. (2014). Subnucleosomal structures and nucleosome asymmetry across a genome. Cell *159*, 1377–1388.

Ricci, M.A., Manzo, C., García-Parajo, M.F., Lakadamyali, M., and Cosma, M.P. (2015). Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. Cell *160*, 1145–1158.

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

Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–329.

Robinett, C.C., Straight, A., Li, G., Willhelm, C., Sudlow, G., Murray, A., and Belmont, A.S. (1996).

In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. J. Cell Biol. *135*, 1685–1700.

Ross, P.J., Kennedy, M. a, Christou, C., Risco Quiroz, M., Poulin, K.L., and Parks, R.J. (2011). Assembly of helper-dependent adenovirus DNA into chromatin promotes efficient gene expression. J. Virol. *85*, 3950–3958.

Sato, K., and Hosokawa, K. (1984). Analysis of the interaction between DNA and major core protein in adenovirus chromatin by circular dichroism and ultraviolet light induced cross-linking. J. Biochem. *95*, 1031–1039.

Schones, D.E., Cui, K., Cuddapah, S., Roh, T.Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic Regulation of Nucleosome Positioning in the Human Genome. Cell *132*, 887–898.

Schreiner, S., Bürck, C., Glass, M., Groitl, P., Wimmer, P., Kinkley, S., Mund, A., Everett, R.D., and Dobner, T. (2013). Control of human adenovirus type 5 gene expression by cellular Daxx/ATRX chromatin-associated complexes. Nucleic Acids Res. *41*, 3532–3550.

Shaban, H.A., Barth, R., Recoules, L., and Bystricky, K. (2020). Hi-D: Nanoscale mapping of nuclear dynamics in single living cells. Genome Biol. *21*, 95.

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

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). Histone H3.1 and H3.3 Complexes Mediate Nucleosome Assembly Pathways Dependent or Independent of DNA Synthesis. Cell *116*, 51–61.

Trotman, L.C., Mosberger, N., Fornerod, M., Stidwill, R.P., and Greber, U.F. (2001). Import of adenovirus DNA involves the nuclear pore complex receptor CAN/Nup214 and histone H1. Nat. Cell Biol. *3*, 1092–1100.

Vayda, M.E., Rogers, A.E., and Flint, S.J. (1983). The structure of nucleoprotein cores released from adenovirions. Nucleic Acids Res. *11*, 441–460.

Voong, L.N., Xi, L., Sebeson, A.C., Xiong, B., Wang, J.-P., and Wang, X. (2016). Insights into Nucleosome Organization in Mouse Embryonic Stem Cells through Chemical Mapping. Cell *167*, 1555-1570.e15.

Wagner, G.P., Kin, K., and Lynch, V.J. (2012). Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci. *131*, 281–285.

Wickham, T.J., Mathias, P., Cheresh, D.A., and Nemerow, G.R. (1993). Integrins alpha v beta 3 and alpha v beta 5 promote adenovirus internalization but not virus attachment. Cell *73*, 309–319.

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

Wiethoff, C.M., Wodrich, H., Gerace, L., and Nemerow, G.R. (2005). Adenovirus protein VI mediates membrane disruption following capsid disassembly. J. Virol. *79*, 1992–2000.

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

Wykes, S.M., and Krawetz, S.A. (2003). The structural organization of sperm chromatin. J. Biol. Chem. *278*, 29471–29477.

Xi, Y., Yao, J., Chen, R., Li, W., and He, X. (2011). Nucleosome fragility reveals novel functional states of chromatin and poises genes for activation. Genome Res. *21*, 718–724.

Zentner, G.E., and Henikoff, S. (2014). High-resolution digital profiling of the epigenome. Nat. Publ.

Gr. *15*, 814–827.

8 Acknowledgments

Without the support of many people this thesis could not have been completed.

First I would like to express my deep gratitude to my supervisor Prof. Dr. Gernot Längst for his outstanding guidance throughout the years. I am very grateful for all his contributions, ideas and funding to keep the project going and making it to an inspiring and successful PhD experience. He provided excellent support while giving me the opportunity to develop and follow my own mind-set. Furthermore, I am thankful for all the courses and meetings I was able to attend and the numerous scientific contacts and collaborations I was introduced to.

Furthermore, I would like to thank my mentor Prof. Dr. Rainer Merkl and Prof. Dr. Harald Wodrich for fruitful discussions and advises. Special thanks to Harald Wodrich for welcoming me in Bordeaux and introducing me to his lab.

I owe everything to the wet lab scientists, Claudia Huber, Sarah Diermeier, Rodrigo Maldonado, Elisabeth Silberhorn, Tetsturo Komatsu, Floriane Lagadec, for making the experiments and providing the data. Without their amazing work, this thesis would not have been possible.

It is a great pleasure to thank the whole team of the House of the Ribosome for creating a stimulating an inspiring work atmosphere. Thank you for all the help, support, discussions and cake sessions.

Besonderen Dank gehört meinen Eltern. Sie haben mir eine freie Wahl ermöglicht und mich auf dem langen Weg bedingungslos und ununterbrochen Unterstützt. Der stete Rückhalt hat mir viel bedeutet und Sicherheit in allen Lebenslagen gegeben.

Und das wichtigste zum Schluss, der Dank an meine bezaubernde Frau Kai. Sie war stets an meiner Seite auf dieser langen Reise, vom Anfang bis zum Ende. Ohne die vielen schönen Erlebnisse mit Ihr, kann ich mir nicht vorstellen, dass ich bis zum Schluss durchgehalten hätte. Und ich freue mich sehr mit ihr die Zukunft zu bestreiten.