3D organization of the chromatin fiber

Lama Soueidan

THÈSE

En vue d'obtention du grade de
Docteur de l'Université de Lyon, délivré par l'École Normale Supérieure de Lyon
Discipline: Science de la Vie

Laboratoire de Biologie Moléculaire de la Cellule, LBMC

Ecole Doctorale de Biologie Moléculaire Intégrative et Cellulaire, BMIC
Présentée et soutenue publiquement le 13 Février 2015
Par
Lama SOUEIDAN

3D organization of the chromatin fiber

Directeur de thèse: Dimitar Anguelov
Apres l'avis de: Dr Marc Lavigne et Dr André Verdel.

Devant la commission d'examen formée de:

Dr. Dimitar Anguelov	Directeur de thèse
Dr. Marc Lavigne	Rapporteur
Dr. Stefan Dimitrov	Examinateur
Dr. André Verdel	Rapporteur
Dr. Gael Yvert	Président

Thèse préparée au sein du
Laboratoire de Biologie Moléculaire de la Cellule
Ecole Normal Supérieure de Lyon
46, allée d'Italie
69364 Lyon Cedex 07
FRANCE

Abstract in English

Title: 3D organization of the chromatin fiber

The local chromatin state plays a crucial role in all fundamental DNA-templated processes, such as transcription control, DNA replication or repair, signaling, etc. The precise 3D organization of the second level folding, the so-called 30 nm fiber, has been a matter of intense speculations and debates over the past 40 years. Two competing models have been proposed on the basis of *in vitro* data, the solenoid and zigzag arrangements. In the solenoid model, consecutive nucleosomes interact with each other and follow a helical trajectory with bent DNA linker. In the zigzag model, alternate nucleosomes interact with each other with straight, twisted or coiled linker DNA. During my thesis, I developed a new biochemical approach, called ICNN (Identification of the Closest Neighbor Nucleosomes), allowing a direct "visualization" of the neighboring nucleosomes within H1-dependent compacted chromatin. We showed that within H1-compacted regular nucleosomal array, N±2 nucleosomes are the nearest neighboring interaction partners of any arbitrary nucleosome N. This finding provides an unambiguous evidence for the zigzag two-start helix conformation of the 30 nm fiber. Furthermore, this organization remains independent on the DNA linker length, demonstrating that the nucleosome repeat length (177-227 bp) does not affect the chromatin structure and therefore cannot be a reason for chromatin structural heterogeneity as suggested in the literature. Chromatin structure and dynamics might be affected by the incorporation of histone variants. Our ICNN experiments with H2A.Z arrays showed no difference of folding between H2A.Z- and H2A-containing fibers. This finding suggests that the H2A.Z-specific transcriptional regulation involves mechanisms other than chromatin folding. CENP-A is a hallmark for centromeric chromatin that is indispensable for cell division. Published crystals structure showed that CENP-A-containing nucleosomes are more "open" than conventional ones due to the shorter α-$N_{CENP\text{-}A}$ helix. Besides, recent results in our lab showed that H1 does not stably bind to the CENP-A nucleosomes and that no stem organization of the linker is formed (not published). Our ICNN study at the chromatin level confirmed the absence of H1-induced folding of a regular CENP-A array. Interestingly, replacement of CENP-A with the α-N_{H3}-CENP-A mutant restored proper H1 binding and folding of the fiber into the usual zigzag conformation, as does the conventional H3-containing fiber.

Résumé de thèse

Titre de thèse: Organisation 3D de la fibre de chromatine

L'état local de la chromatine joue un rôle crucial dans tous les processus génétiques comme le contrôle de la transcription, de la réplication et de la réparation de l'ADN, de la signalisation, etc... La structure précise de l'organisation 3D du deuxième ordre de compaction de la chromatine, aussi appelé fibre de 30 nm, a été le sujet d'intenses débats durant les 40 dernières années. En se basant sur les données *in vitro*, deux modèles concurrents se distinguent: le solénoïde et le zigzag. Dans le modèle solénoïde, des nucléosomes consécutifs interagissent pour former une trajectoire hélicoïdale avec courbure de l'ADN de liaison. Dans le modèle zigzag, deux hélices d'empilement de nucléosomes se forment, reliées par l'ADN de liaison qui peut être droit ou tordu. Durant ma thèse, j'ai développé une nouvelle approche expérimentale biochimique, appelée ICNN (Identification of the Closest Neighbor Nucleosome), permettant de déchiffrer les interactions entre nucléosomes voisins au sein de la fibre et ainsi de définir sous quelle forme la chromatine se compacte. Nous avons démontré que dans une fibre compactée H1-dépendante, les nucléosomes N+2 sont les plus proches voisins d'un nucléosome N arbitrairement choisit. Ces résultats montrent, sans ambiguïté, l'organisation zigzag de la fibre de 30 nm. De plus, cette organisation reste indépendante de la longueur de l'ADN de liaison, démontrant ainsi que la longueur des nucléosomes (177-227 bp) n'affecte pas la structure de la chromatine et ne peut donc pas être responsable de l'hétérogénéité structurale décrite dans la littérature. La dynamique et la structure de la chromatine peuvent être affectées par l'incorporation de variants d'histones. Nos expériences utilisant des assemblages H2A.Z ne montrent aucune différence de repliement entre les fibres contenant H2A.Z et H2A. Ces données suggèrent que le mécanisme moléculaire de régulation de la transcription spécifique de H2A.Z n'est pas lié au repliement de la chromatine. CENP-A est un élément de la chromatine centromérique indispensable à la division cellulaire. La structure par diffusion des rayons-X dans des cristaux montrent que les nucléosomes contenant CENP-A sont plus ouverts que le nucléosomes conventionnels à cause de l'hélice α-N_{CENP-A} plus courte. Parallèlement, des résultats récents de notre laboratoire ont montré que H1 ne se lie pas d'une façon stable aux nucléosomes contenant CENP-A, ne conduisant à aucune organisation de l'ADN de liaison (données non publiées). Notre étude au niveau de la

chromatine a confirmé l'absence de repliement de la fibre contenant CENP-A en présence de H1. De manière intéressante, le remplacement de CENP-A par le mutant α-N_{H3}-CENP-A restaure la bonne liaison de H1 et le repliement de la fibre avec une conformation habituelle en zigzag comme dans une fibre contenant H3.

Acknowledgements

I am deeply indebted to Dr. Dimitri Anguelov for giving me the opportunity to work on this challenging project and for providing the excellent environment that was necessary to complete my PhD and successfully use the wide range of techniques.

I want to express my gratitude to the group of people that were collaborating with me on the project of chromatin fiber structure: Dr. Stefan Dimitrov for his enthusiastic knowledgeable discussions and his influence on my thesis. Dr Manu Shukla for teaching me the chromatin biochemistry and for giving me the necessary tools to continue what he started.

I would like to thank, Dr. André Verdel, Dr. Marc Lavigne and Dr. Gael Yvert for their time and efforts to evaluate the thesis.

I would like to extend my gratitude to all the members of the Anguelov's team past and present. In particular I would like to thank my favored indians Dr. Charles Richard john Lalith (Richy baba) and Dr. Imtiaz Nisar Lone (mimi) for the worm welcome they gave me when I first joined the team and for the great fun times we had in the algeco. My special thanks to Rama, my lab drinking buddy, for all the stories we shared about our drunken nights. My thanks to Ogi, Hervé and Elsa for all their help and time during my PhD.

During my four years stay in ENS I met some wonderful people that I enjoyed spending time in ENS and in bars. Special thanks goes to Fanny for the great time we spent in ENS, specially playing badminton and drinking coffees. I can't thank her enough for the best night of my life, the night I ran 14 km at 2am on an icy pathway and the night that I fell more than 7 times. I can't thank you enough for all the corrections you made and all the "ca se dit pas!!!". I also would like to thank Lydia, for all the "pause clope" that we did, for all the sweets that we ate (without anybody knowing) and for the great time we are having together every Wednesday (I tried to make a sentence with hydratation in it and I couldn't). A special thanks goes to Paulina and Xuan Nhi, for making this last year sweet with our afternoon coffee and lunch breaks and specially for listening to 'bad luck' stories all of last year. I also would like to thank Mehrnaz, Loan and Marion one of the first people I met once I started in ENS, for the wonderful time we spend together and for hicking and skiing trips. To Mehrnaz I would like to add, don't be

afraid we are going to all be here for you. Special mention goes to Juliana for all the time we spent in the “Lapin blanc” and hope I’ll see you soon in Japan (if you invite me). I can end the ENS chapter without thanking, mulumulu ducon (alias maxime) and Etera for being the main consumers and providers for ‘Bar Lamoush’.

I owe ‘monsieur le bricoleur’ (alias Mathieu) a big thank you, for being present for me when I needed help, especially for repairing my bike and my door but all also for all the fun night and yummy burger we ate together. A special thanks to caro (Dufaure de citre), for the brunches and for the ‘confession intime’ discussion we had and hope we continue doing.

I would like also to thank all the people I met during my college days. Starting by Johanna, I would like to thank you and your family for accepting me into your family during Christmas days and for the time we spend at the riverside. I’m really proud to have as my friend. I would like also to thank Fatima, for all the time she been late to movies, restaurant and parties. I hope I’ll be there to see become doctor soon.
A special thanks to Julie, for all the times you have been mean to me (specially with lorraine) and for all the times you’ve been nice to me and for all great summer week end we had at your parent’s house. I would also like to thank Lolo for making me a chartreuse addict, for the great friend you have been during all this years and for introducing to your ‘copinette group’, Marie-Audrey, Laure and Caro (Jaillet). I would like also to thank Elodie, Anais and Laura for the great support they have been during the past years; I know I can count on them whenever they are needed.

A special thought to my childhood friends, to be more precise to my teenagehood friends. Elma, we’ve been friends for a long time, we’ve been through good and bad times, and I’m glad I can still count you as close friend. I would like to thank Claire, also known as claira mama, for making my visits to Lebanon fun and enjoyable for the past 10 years.

Last but not least, I would like to thank my family, who has been there for me all along, who sent me to France at 17 to have a better life. I dedicate this thesis to my father Issa, to my mother Aida, to my two brothers, Maher and Samer and finally to my

older sister Liliane. A special mention goes to Maher, who put up with me during 8 long years. A special thanks also my beloved sister, I don't where will I be if you weren't there for me each second of my life. Thank you for teaching me to be the strong, independent and open-minded woman that I am today. I hope one day, I can repay 1% of what you've given me. And finally I would like to thank my mom, who had the courage to send her youngest daughter 3000 Km away allowing me to have a better life and a better education.

Again, I would like to thank you all deeply, for your presence and influence on me. My life would have not been the same without you.

Table of Contents

Figures:

Tables:

Lists of Abbreviations:

2D: 2 Dimensions

3D: 3 Dimensions

aa : Amino-Acid

AFM: Atomic Force Microscopy

ATP: Adenosine-5'-TriPhosphate

bp: base pair

CAF1: Chromatin Assembly Factor 1

CEMOVIS: Cryo-Electron Microscopy of Vitreous Sections

CENP-A: Centromere Protein A,

CH: Carboxy Terminal Helix

CTD: C-Terminal Domain

CTF: Contrast Transfer Function

DiSCO: Discrete Surface Optimization

DNA: DeoxyriboNucleic Acid

DSB: Double Strands Break

DTT: Dithiothreitol

EM: Electron Microscopy

ES: Embryonic Stem

FRAP: Fluorescence Recovery After Photobleaching

GD: Globular Domain

GH1: Globular Domain H1

H1: Histone H1

H2A: Histone H2A

H2A.Bbd: Histone: H2A Barr body-deficient

H2B: Histone H2B

H3: Histone H3

H4: Histone H4

HATs: Histone Acetyltransferases

HDAC: Histone deacetylases

HILS: histone H1-like Protein in Spermatids 1

HMT: Histone Methyltransferase

HOX: Homeotic genes

HP1: Heterochromatin Protein 1

ICNN: Identification of Closest Neighbor Nucleosomes

IR: Infra-Red

MNase: Micrococcal Nuclease

RNA: Ribonucleic Acid

mRNA: Messenger RNA

Nap1: Nucleosome Assembly Protein-1

NASP: Nuclear Autoantigenic Sperm Protein

NBP: Nucleosome Binding Protein

NCP: Nucleosome Core Protein

NMR: Nuclear Magnetic Resonance

NRL: Nucleosome Repeat Length

NTD: N-Terminal Domain

PCR: Polymerase Chain Reaction

PTM: Post-Translational Modification

RNAi: RNA Interference

rRNA: ribosomal RNA

SELEX: Systematic Evolution of Ligands by Exponential Enrichment

SMC: Structural Maintenance of Chromosomes

Ub: ubiquitin

UV: Ultra Violet

WT: Wild Type

Scope of the thesis

Regulation of eukaryotic genomes in a chromatin context is a fundamental issue in biology. Highly conserved histone proteins function as building blocks to package eukaryotic DNA into repeating nucleosome units that are folded into a higher-order structure. The nucleosomes, consisting of approximately 147 bp of DNA wrapped around a histone octamer and an additional 10-90 bp of linker DNA complexed with linker histone H1, are the structural repeating unit of chromatin. Nucleosomes are spaced quite regularly and form fibers of 30 nm in diameter, referred to as the 30 nm chromatin fiber. The nucleosome is a barrier to numerous vital cell processes that require access to free DNA. The cell uses incorporation of histone variants, histone modifications and ATP-dependent chromatin remodeling complexes to overcome this nucleosome barrier and allow chromatin-templated processes.

A variety of distinct models have been proposed for the arrangement of nucleosomes and linker DNA in this "30 nm chromatin fiber". The dynamics and the structure of chromatin are directly involved in control of gene expression and many other nuclear processes. The folding of nucleosomes into the chromatin fiber has remained a controversial issue due to the heterogeneity and the flexibility of the native material. Two main classes of helical models have been proposed: the one-start solenoid models with nucleosomes arranged side by side and coiling into an helix, and the two-start models, with nucleosomes zigzagging back and forth to form a two-stranded, helical arrangement.

To address these questions, multitudes of techniques are employed to study the dynamics of nucleosome and chromatin structure that will provide the molecular basis for cellular processes.

I. Introduction

Chapter 1: DNA organization in eukaryotic cells.

1.1 Chromatin overview

Stuffing the long strands of chromosomal DNA into the eukaryotic nucleus requires the DNA to compact in length approximately 10-50.000 folds. Indeed, eukaryotic cells contain ~$5x10^9$ DNA base pairs (bp) in a 5-7μm nucleus and the DNA molecules that code the human genome should measure ~2 meters if they were laid end to end. This vast amount of DNA is packed by the histone proteins into a hierarchical structure called the chromatin. The word "chromatin" has been derived from the Greek word "khroma" meaning colored based on its staining ability in dye. The dynamics and organization of this chromatin influences many function of the genome. Its primary functions can be divided into 4 categories, 1) to package DNA into a smaller volume to fit in the cell, 2) to reinforce the DNA to allow mitosis, 3) to prevent DNA damage and 4) to control gene expression and DNA replication.

1.2 Chromatin history

The history of chromatin began in late 19th century when biologist W. Flemings suggested the name 'chromatin' [1] while working on nuclear division (mitosis was another word suggested by him)(1880). Flemings, influenced by H. Zacharias microscopy studies of protease-digested nuclei, wrote: *" ... in the view of its refractile nature, its reactions, and above all its affinity to dyes, is a substance which I have named chromatin..." [2].*

During this time F. Miescher developed methods for the isolation of nuclei from pus leukocytes and described what he then called 'nuclein' as a strong phosphorus-rich acid [3]. Later on, in 1884 A. Kossel, who was working with Miescher in Hoppe-Seyler's laboratory, described the 'histone' in the acidic extract from erythrocytes nuclei [4].

The first half of the 20th century was great for the field of genetics but showed no advances in the understanding of the chromatin structure. During this time we witnessed the rediscovery of the Mendel principle in 1900 by H. de Vries, the development of the gene theory and the principle of linkage in 1910 by T.H. Morgan. In 1928, F. Griffith identified the 'transforming principle' by infecting mice with 2 strains of pneumococcus (*Streptococcus pneumoniae*) and proving the transfer of the genetic

information. Later on, in 1944 O. Avery, C. MacLeod and M. McCarty proved that DNA was the molecule responsible for bacterial transformation.

The discovery of polytene chromosomes in *Drosophila* and of gene localization in 1933 by E. Heitz and H. Bauer, T. Painter and in 1935 by C. Bridges provided experimental material for exploring the chromatin structure. Inspired by this work, D. Mazia conducted in 1941 a series of analyses on the *Drosophila* salivary gland polytene chromosomes and on plant chromosomes using proteases and nucleases suggesting that both the salivary gland and plant chromosome are composed of a continuous framework and a matrix that occupies a considerable volume [5]. Most of the investigators were back then agreeing that chromosomes and chromatin formed the structural basis of the genes and that the histone's primary function was carrying genetic information [6].

Major structural biology discoveries were made during the second half of the 20th century due to the development of X-ray imaging techniques. The discovery of the protein α-helix in 1951 by L. Pauling [7] followed by the famous discovery of double helix structure of DNA published in 1953 by J. Watson & F. Crick [8], M. Wilkins [9], and R. Franklin & R. Gosling [10] were turning points in the chromatin field. Another key advance in the field was done by G. Zubay and P. Doty who were able to prepare a soluble chromatin molecule [11]. In the 60s, J. Gall demonstrated that the backbone of the chromatid is a continuity of one single molecule of DNA [12]. The fractioning of histones by E. Johns [13] and the discovery of association between histone modification and chromatin transcription by V. Allfrey [14] were among the important advances made during this decade on the chromatin.

The first electron microscopy images performed on chicken erythrocyte nuclei in 1973 by A. Olins & D. Olins [15] and C. Woodcock [16] showed the 'beads on a string' structure. 6 months later, R. Kornberg published his chromatin model, based on nuclease assays and histone cross-linking, showing that about 200 bp of DNA formed complex with four histone pairs [17]. In 1975, this chromatin subunit received the name of 'nucleosome' [18]. The discovery of the nucleosome represented a 'quantum jump' in the understanding of the chromatin structure. During the next 30 years a race took place to uncover the structure of chromatin and the 30 nm fiber (Figure1).

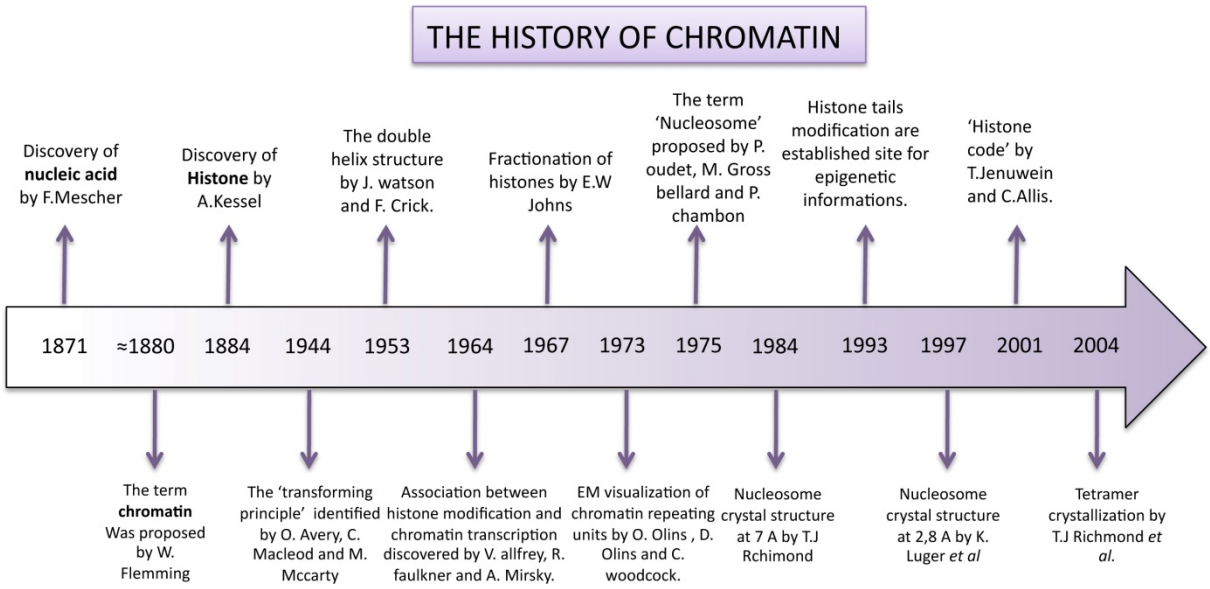


Figure 1: Historical diagram representing major discoveries of chromatin studies. Image adapted from [19].

1.3 DNA packaging

In order to fit into the nucleus, DNA must be packaged into a highly compacted structure known as chromatin. In the first step of this process DNA is condensed into an 11 nm fiber that represents an approximate 6-fold level of compaction [20]. This is achieved through nucleosome assembly.

The nucleosome is the smallest structural component of chromatin, and it is produced through interactions between DNA and histone proteins. A histone octamer is formed from the histones H2A, H2B, H3 and H4, although in some cases other histone variants may also be found in the core (e.g. H2A.Z, MacroH2A, H2A.Bbd, H2A.lap1, H2AX, H3.3, CENP-A and others). Additionally, a 147 bp DNA segment then wraps itself around the histone octamer 1.75 times, thus completing the formation of a single nucleosome. As shown in figure 2, the nucleosome is part of a wider process, whereby multiple nucleosomes form in a linear fashion along the DNA molecule. This ultimately produces the 11 nm fiber, which is traditionally described, based on its appearance, as "beads on a string". Here, adjacent nucleosomes are connected via "linker DNA", which is usually bound to the H1 histone and is between 20-80 bp long. Additionally, flexible histone tails extend away from nucleosomal DNA and can interact with other nucleosomes, stabilizing more complex 3D structures [21]. Of note, nucleosomes are dynamic entities and can undergo spontaneous sliding, "splitting" or even complete dissociation. The level of compaction attained through the formation of the 11 nm nucleosome fiber is insufficient to package the whole genome into the nucleus. Instead, this fiber forms the

basis for other higher-order chromatin structures that are established through additional folding and bending events. Despite the extensive knowledge already gained on the structure of the 11 nm nucleosome fiber the intermediate chromatin structures commonly known as the 30 nm fiber is a matter of debates. Studies over the past decades have emphasized the critical importance of chromatin components whose nature and spatial organization are sources of information that contribute to cellular function and identity. Changes in chromatin structure have been found to play a fundamental role in the regulation of multiple genomic processes, from gene expression to chromosome segregation and the maintenance of genome integrity and stability.

Two popular models were proposed based on *in vitro* data: the solenoid and the zigzag. In each case, the 11 nm nucleosome fiber undergoes additional folding to form a 30 nm fiber [22, 23]. In the one-start solenoid model, bent linker DNA sequentially connects each nucleosome core, creating a structure where nucleosomes follow each other along the same helical path [23, 24]. Alternatively, in the two-start zigzag model, straight linker DNA connects two opposing nucleosome cores, creating the opposing rows of nucleosomes that form so called "two-start" helix. In zigzag model, alternate nucleosomes become interacting partners [22, 25]. Interestingly, some studies offer a model, in which intermediate 30 nm fibers contain both the solenoid and zigzag conformations [26].

One aspect shared by most of the models for higher-order chromatin organization is the dynamic existence of decondensed loops among more compact chromatin structures. In most cases, higher-order chromatin has to be decondensed so that the transcription machinery can gain access to the genes [27, 28].

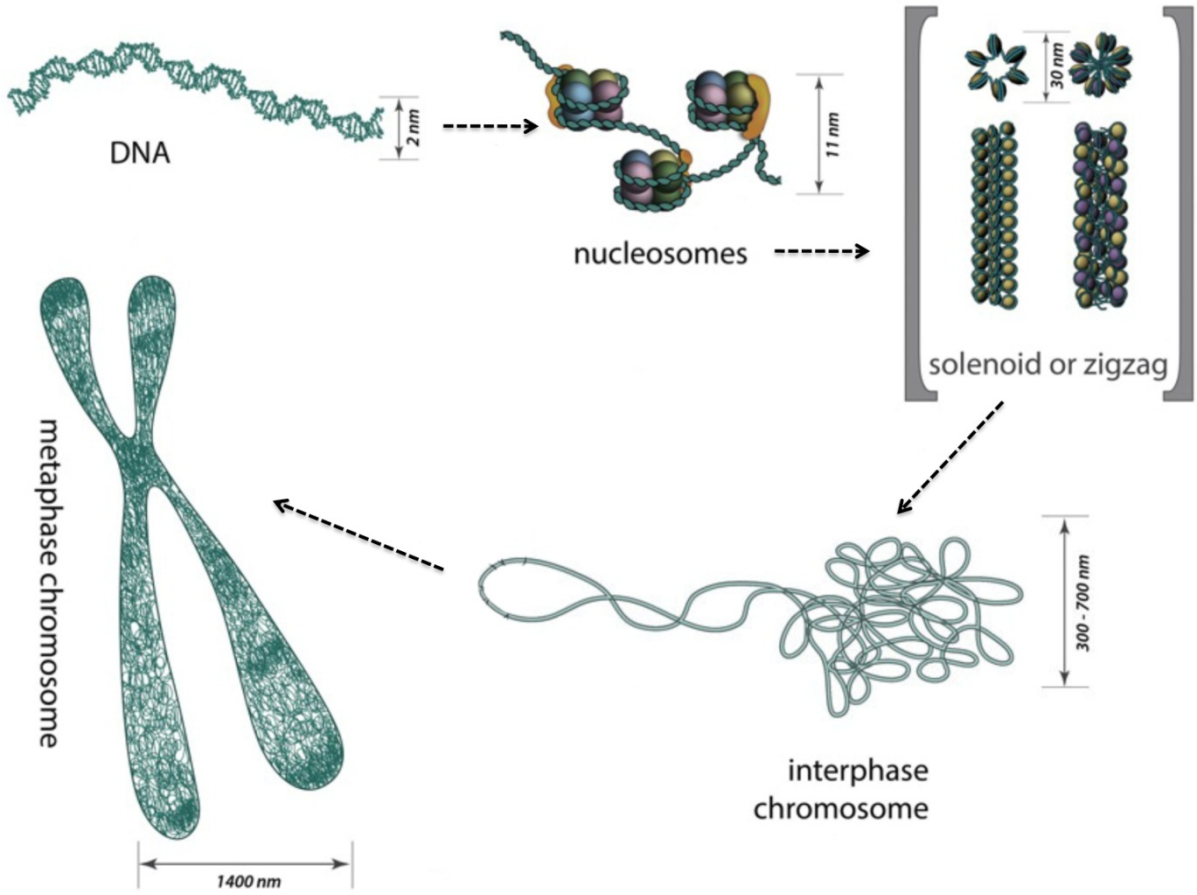


Figure 2: Organization of eukaryotic chromatin fibers. The lowest level of DNA packaging is the nucleosome, in which two superhelical turns of DNA are wound around the outside of a histone octamer. Nucleosomes are connected to one another by short stretches of linker DNA. At the next level of organization, the string of microsomes ('beads on a string') is folded into a fiber about 30 nm in diameter, and these fibers are then further folded into higher-order structures. At levels of structure beyond the nucleosome details of folding are still uncertain. Figure adapted from [29]

1.4 Types of chromatin

Interphase chromatin can be divided into two domains, euchromatin and heterochromatin. These can be distinguished in a cytological manner by differential staining [30]. The weakly and strongly stained regions are the euchromatin and the heterochromatin, respectively (Figure 3A). These two vary on different genetic levels such as gene activity, histone modifications and nucleosome packaging [31]. They might also have a different higher-order packaging [32, 33] and nuclear organization [34].

1.4.1 Heterochromatin

Heterochromatin was initially defined as the portion of the genome that retains deep staining with DNA specific dyes as the dividing cell returns to interphase from

metaphase. Heterochromatin regions consist of repetitive DNA, including satellite sequences and middle repetitive sequences related to transposable elements and retroviruses. These regions are typically gene-poor, replicate in late S-phase and generally have a reduced frequency of meiotic recombination [35]. Heitz first introduced it in 1928, by identifying regions of mitotic chromosomes that retained a compact structure during interphase. Autoradiographic studies using tritium-labeled uridine ([^{3}H]UdR) provided convincing evidence that RNA synthesis is only possible in the diffuse state, whereas condensed chromatin is transcriptionally inactive [36]. Two reasons were established for inactive chromatin:

1. The rare coding DNA present in these regions is temporarily and reversibly inactivated.
2. The majority of this region composed of non-coding DNA is incapable of transcription.

The heterochromatin is spread everywhere in the chromosome in small regions but resides in the centromere, telomeres and the Barr body of the inactivated X chromosome [45,37].

However, heterochromatin plays an important role in gene expression during development and differentiation [38]. Many components are necessary in generating heterochromatin, such as repetitive DNA sequences, methylation of histone H3 lysine 9, heterochromatin protein 1(HP1) and RNA interference (RNAi) [39]. The post-translational modifications of the key lysine residues on histone protein H3 determine the formation of different chromatin structures. The methylation of N-terminal lysine residues of histone H3 is among the most studied histone modification. The lysines 4,9,27,36 and 79 can be modified by methytransferases and demethylases. Transcription and chromatin structure are affected by mono- or di-methylation of these residues, while epigenetic regulation is attributed to tri-methylation (H3K4me3, H3K9me3, H3K27me3, H3K36me3, H3K79me3) (Figure 3C) [40, 41].

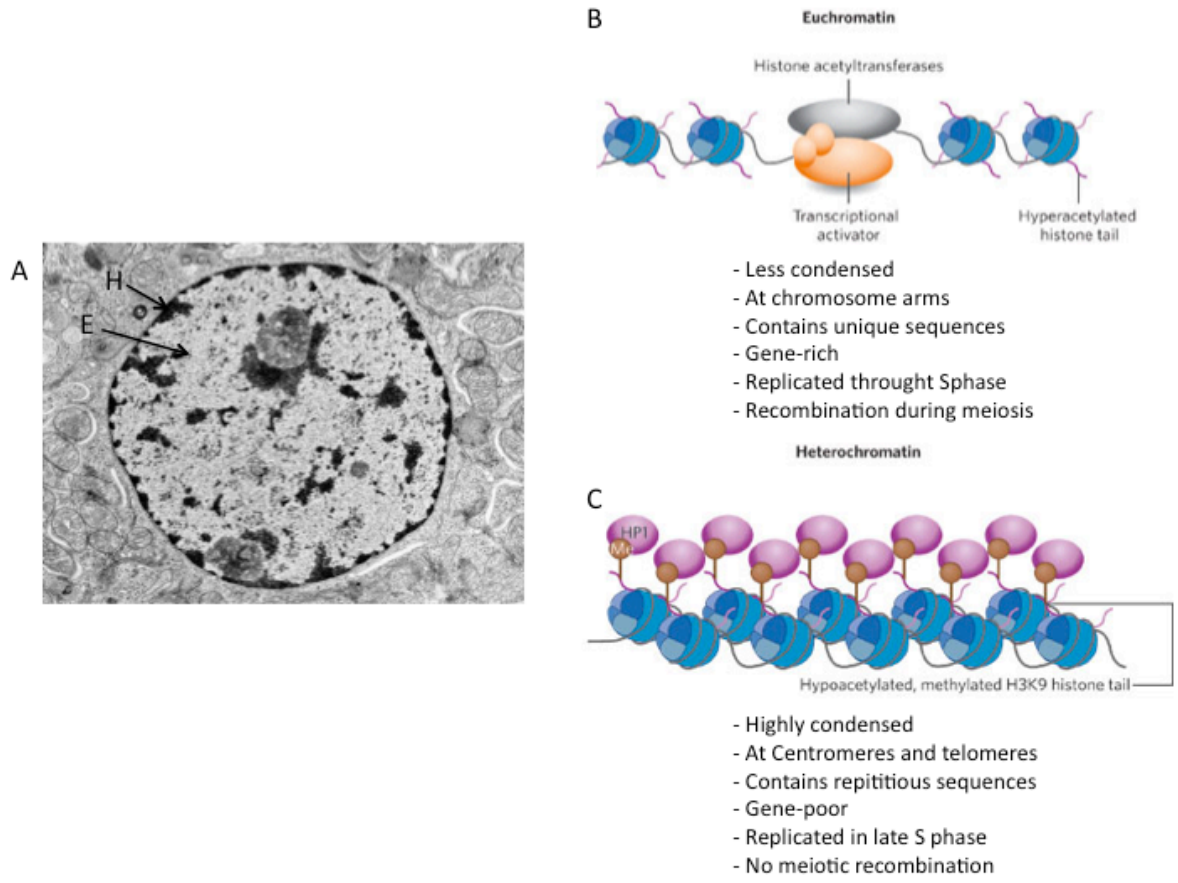


Figure 3: euchromatin and heterochromatin properties. (A) The distribution of euchromatin (E) and heterochromatin (H) in a normal thymus lymphocyte. (B-C) Characteristics of euchromatic and heterochromatic regions. Images adapted from [42].

Three different kinds of condensed chromatin have been observed within interphase nuclei: tissue-specific condensed euchromatin, facultative heterochromatin and constitutive heterochromatin.

1.4.1.1 Tissue-specific condensed euchromatin

Although almost all cells in an organism contain the same DNA, only a part of the available information within a genome is expressed in any given cell type. This selectivity depends on the inactivation of some euchromatic components of the genome in a particular cell. The condensation happens at specific differentiation states as part of a tissue-specific transcriptional control system. This condensation state is highly variable from one cell type to another within a given species. This condensation state is reversible [43], but once it is dispersed, it does not automatically adopt a transcriptionally active form as euchromatin.

1.4.1.2 Facultative heterochromatin

The facultative heterochromatin is found mainly at the developmentally-regulated loci, next to rRNA genes or at the transcriptionally active region of euchromatin [34]. It plays a key role in normal cell lineage development, cell differentiation by somatic methylation and inactivation of germline-specific genes [44].

Furthermore, the facultative heterochromatin might be involved in allelic exclusion, genomic imprinting and inactivation of the X chromosome, the gene loci of immunoglobulin and T receptors α and β [45, 46]. For example, the inactivation of the X chromosome in female mammals can occur because one X chromosome is packed as facultative heterochromatin and is thus silenced while the other packed as euchromatin and expressed [47]. Proteins like Polycomb-group proteins and non-coding genes such as Xist regulate the formation of facultative heterochromatin.

1.4.1.3 Constitutive heterochromatin

These are the regions to which Heitz originally gave the name heterochromatin. They are mostly composed of non-coding and largely repetitive DNA such as clusters of satellite sequences and transposable elements of centromeres, pericentric foci and telomeres [34]. These transposable elements are highly mutagenic because they target coding genes for insertion causing chromosome breakage and genome rearrangement [48]. This DNA is permanently non transcribable rather than simply repressed. The constitutive heterochromatin plays an essential role in keeping genomic integrity by preventing abnormal chromosome segregation, recombination and DNA replication.

1.4.2 Euchromatin

Euchromatin is lightly packed chromatin and highly concentrated in active gene. It is compacted during cell division and relaxed during interphase. The replication of this chromatin occurs during S-Phase and the recombination during meiosis. It is also been shown that the condensation coincides with the lack of synthesis of mRNA during mitosis. Euchromatin is subjected to post-translational modifications (PTMs) such as histone methylation and acetylation (Figure 3B).

Some of these modifications have been suggested to have a global effect (indirect) and others affect mostly the boundary between euchromatin and heterochromatin. The molecular mechanisms responsible for different gene silencing in euchromatin are still largely unknown [49].

1.4.3 Centromeres

Mitosis is the process by which eukaryotic cells divides to produce two daughter cells that contain the same number of chromosomes as the parent cell. The centromere plays a crucial role in mitotic segregation. The genes found along the centromere are highly divergent while the genes found in the chromosomes arms are highly conserved. The main function of the centromere is to build the kinetochore at meiosis and mitosis, which will serve as the physical connection of the chromosomes to the microtubule-based spindle [50]. However, the repetitive DNA typically found in this region is required for neither the identity nor the function of the centromere. The architecture and the scale of centromeric chromatin vary between different eukaryotic species. However, all centromeres contain the histone H3 variant CENP-A, a centromere-specific protein [51, 52]. The centromere identity is defined by the presence of CENP-A in the chromatin [53].

Even though the word "centromere" is derived from the greek centro ("central") and mere ("part"), chromosomes can be classified into four types based on the location of the centromere: metacentric, submetacentric, acrocentric and telocentric (Figure 4).

Figure 4: Chromosome classification. On the basis of the location of the centromere, chromosomes are classified into four types: metacentric, submetacentric, acrocentric, and telocentric. Figure adapted from [54].

1.4.4 Telomeres

Telomeres are ribonucleoprotein structures that protect the end of linear chromosomes from recognition as double-stranded breaks and activation of DNA damage response. They are G-rich and short repeat sequences that comprise the physical termini of chromosome [55]. The number of repeats and the length of telomeres are subject to regulation and influence biological processes such as aging and cancer [56]. Telomere length is controlled by a homeostatic mechanism that involves telomerase and conventional replication machinery. Telomerase, a telomere-specific protein, is a specialized reverse transcriptase that synthesizes the G-tail using intrinsic RNA template and thus prevents premature shortening of the chromosomes [57].

Chapter 2: The nucleosome

2.1 Nucleosome overview

As mentioned in the previous chapter, the genetic information of eukaryotic cells is stored in a 2 meter-long DNA molecule. At the first level of compaction, two super-helical turns of DNA are wrapped around a protein assembly of 8 histone molecules to form the nucleosome core particle. Hundreds of thousands of nucleosomes are further organized into multiple higher levels.

Due to its packaging, the structure and the accessibility of DNA in living cells is different from that of linear, naked DNA. This is of fundamental importance for our understanding of all processes that use DNA as a substrate, such as transcription, replication, DNA repair, etc. Eukaryotic cells developed complex mechanisms to modulate and control chromatin dynamics

2.2 DNA structure

The DNA molecule has a right-handed double-stranded helical structure in which the two strands run opposite and interwined to each other. Each helix is a polymer of four basic nucleotides. The polymer backbone is composed of the alternating sugar (deoxyribose)-phosphate units, attached to which are four types of heterocyclic nitrogenous bases namely Adenine, Thymine, Guanine and Cytosine, which hold the two strands together. Adenine from one strand forms two hydrogen bonds with Thymine form the other stand, while Guanine forms three hydrogen bonds with the Cytosine from the opposite strand. The DNA double helix also has two different-sized "grooves": a major groove and a minor groove (Figure 5A). DNA double helical structure has been classified into three types namely B-DNA, A-DNA and Z-DNA. B-DNA is the most abundant form of DNA commonly found under physiological conditions in a cell. In this structure, the helix takes a turn every 3.4 nm, and the distance between two neighboring base pairs is 0.34 nm. Hence, there are about 10 pairs per turn. In a solution with higher salt concentrations or with alcohol added, the DNA structure may change to an A form, which is still right-handed, but makes a turn every 2.3 nm and there are 11 base pairs per turn. A-DNA forms are present under some biological conditions that are not yet well understood. Another DNA structure is called the Z form, because its bases seem to zigzag. Z-DNA is left-handed and also narrower than the other two types of DNA. One turn spans 4.6 nm, comprising 12 base pairs. The DNA molecule with alternating G-C

sequences in alcohol or high salt solution tends to have such structure. Z-DNA has been found in synthetic short segments of DNA.

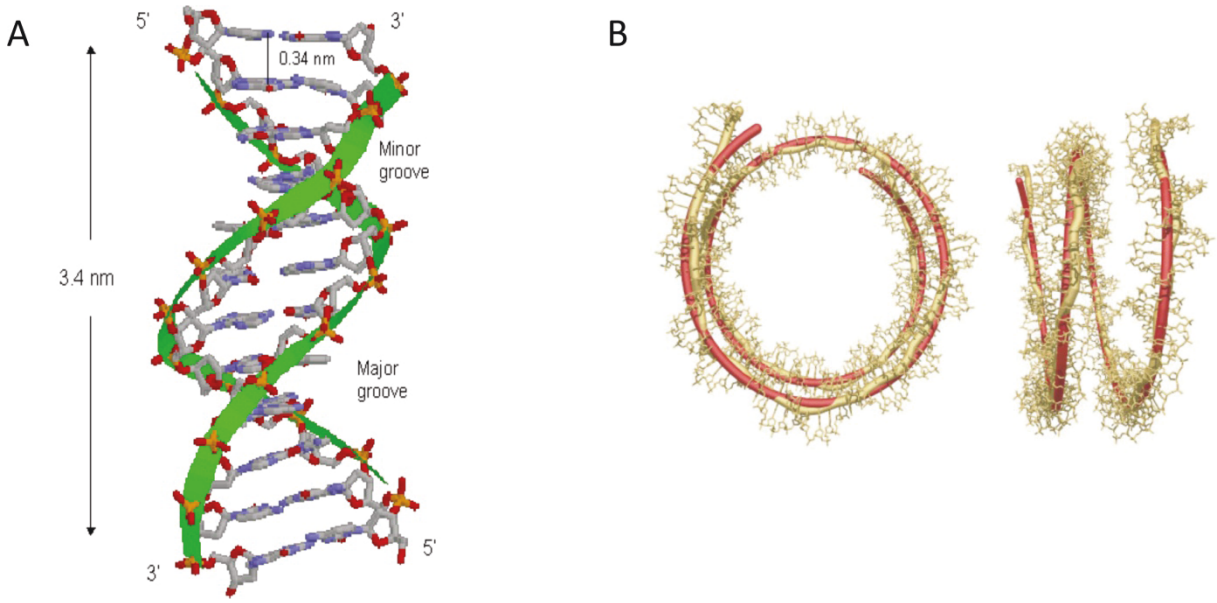


Figure 5: Structure of the DNA. (A) A cartoon of DNA double helix showing DNA major and minor groove. (B) Structural alignment of the NCP147 (gold) and best-fit, ideal superhelix (red) paths. The NCP147 DNA structure is superimposed (gold). The left view is down the superhelix axis and the right view is rotated 90° around the pseudo-two-fold axis (vertical). Figure adapted from [58]

The 1.9- -resolution crystal structure [59] of the nucleosome core particale (NCP) containing 147 bp DNA reveals the conformation of nucleosomal DNA with high accuracy. An atomic-level description of DNA conformation in the nucleosome core and comparaison with naked DNA is essential to an understanding of chromatin properties such as nucleosome position and mobility. The NCP DNA double helical structure is generally a B-form as judged by the phosphate coordinates. It has a radius of 41.9 and a pitch of 25.9 , however the DNA superhelix is not bent uniformly due to these three reasons: 1) flexibility of the DNA, 2) local structural features of the DNA sequence and 3) the presence of histone octamer (Figure 5B).

2.3 Core histone proteins

The histones are found in large quantities in the nuclei of almost all eukaryotic cells. Five types of histones can be distinguished, the linker histone H1 and four core histones H2A, H2B, H3 and H4 [13]. The complete set of core histones is essential for cell viability.

These are found in the same equal molar stoichiometry in all eukaryotic organisms [60] and their sequences are highly conserved in evolution.

2.3.1 Sequences and domain structure of histone

The core histones are divided into three types of structural domains: the histone fold, the histone fold extensions and histone tails. α-helix, loops and coil elements are the main secondary structure in histones (Figure 6).

1. Histone fold: this central region of 70 aa (amino-acid) of the protein chain contains a 3 to 4 turn α-helix (α1), a loop of 7-8 aa, an 8-turn α-helix (α2), a loop of 6 aa and a final 2 to 3 turn α-helix (α3).

2. Histone fold extensions: It contains less uniform structural elements. It extends N-terminally from the histone-fold domain of H3 (αN) and C-terminally from the histone fold domains of H2A and H2B (αC).

3. Histone tails: The N-terminal part of all core histones is made up of random-coil elements of different lengths (16aa for H2A to 44 aa for H3). In the context of a single nucleosome core, these tails are flexible and unstructured [20]. H2A also has also a 10 aa C-terminal tail that varies between H2A variants. These regions are also subjected to many posttranslational modifications such as acetylation, methylation, phosphorylation, ubiquitination, sumoylation, biotinylation, glycosylation, parpylation and ADP-ribosylation, which will be discussed later on.

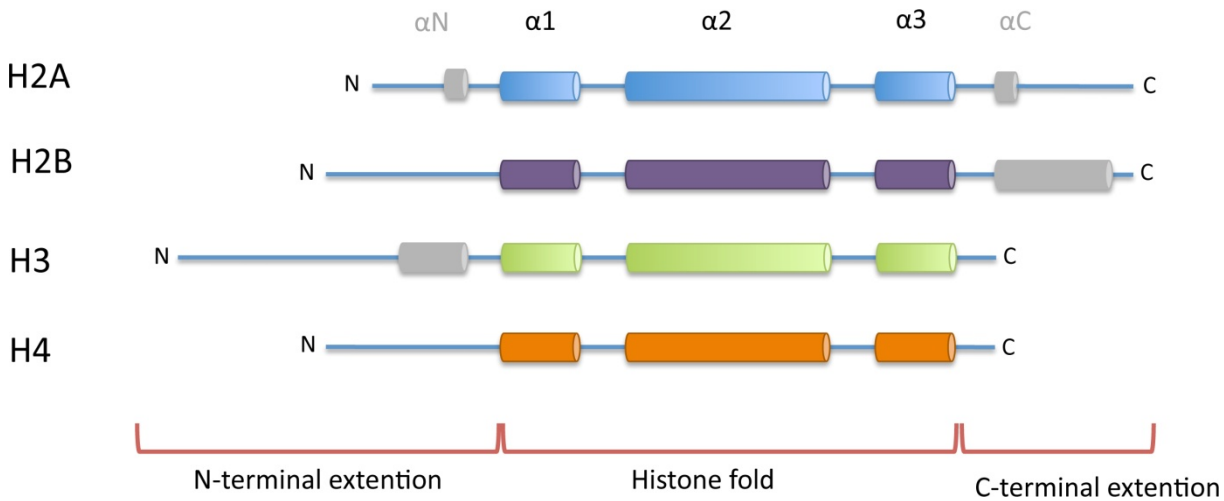


Figure 6: Schematic representation of all the core histones. The boxes indicate the helices of the histone fold domain and histone fold extensions. Figure is inspired from [61]

2.3.2 Histone octamer

The histones assemble in heterodimer pairs, H2A and H2B on one hand and H3 and H4 on the other hand. The dimers are formed through histone-specific interactions, called the handshake motif [62]. The L1-α2-L2-α3 regions will interact in head-to-tail orientation through extensive hydrophobic contacts between helices. The short terminal helices fold and rotate over the central helix, which will cause for the terminal helices to interlink and for central helices to overlap (Figure 8). Furthermore, β-bridges are formed due to loop interaction: the C-terminal loop (L1) of H2A interacts with the N-terminal loop (L2) of H2B, while the C-terminal loop of H3 interacts with the N-terminal domain of H4. These bridges will serve as primary docking sites on the histone surface.

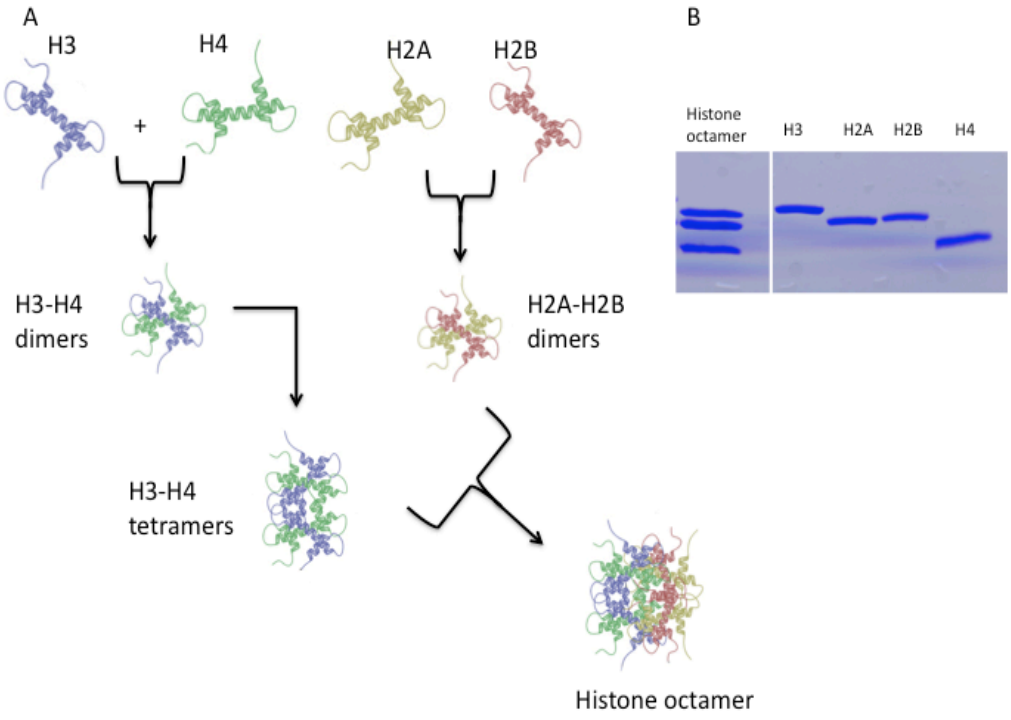


Figure 7: Histone octamer assembly. (A) A histone octamer is the eight protein complex found at the center of a nucleosome core particle. It consists of two copies of each of the four core histone proteins (H2A, H2B, H3 and H4). The octamer assembles when a tetramer, containing two copies of both H3 and H4, complexes with two H2A/H2B dimers. Figure adapted from [63]. (B) SDS electrophoresis gels of the histone octamer and all histone cores.

The central and C-terminal helices (α2-α3) of the two H3 within the H3-H4 dimers interact together and thus form a four-helix pack. These interactions result in the formation of the H3-H4 tetramer. The hetero-tetramer has the shape of a twisted open horseshoe [64] and plays a role in nucleosome positioning. In addition, it exists as a soluble complex in physiological ionic strength solutions and binds much more strongly

to the DNA than the H2A-H2B dimers. To obtain a fully folded histone octamer, two units of H2A-H2B dimers bind to the opposite side of the H3-H4 tetramer.

2.4 Nucleosome core particle (NCP)

The nucleosome core particle has been determined by high resolution X-ray crystallography [59, 65] (Figure 8B). The NCP comprises 147 DNA bp and the histone octamer and forms 1.67 turns of a superhelix. The H3-H4 tetramers occupy the central 60 bp region whereas each of the H2A-H2B dimer binds to the 30bp adjacent to the central region. The remaining 13 bp on each end of the DNA are held by the H3 αN helix. (Figure 8A). The orientation of the histone octamer favors the positioning of the arginine side chains in the minor grooves of the B-DNA which results in a left-handed superhelix.

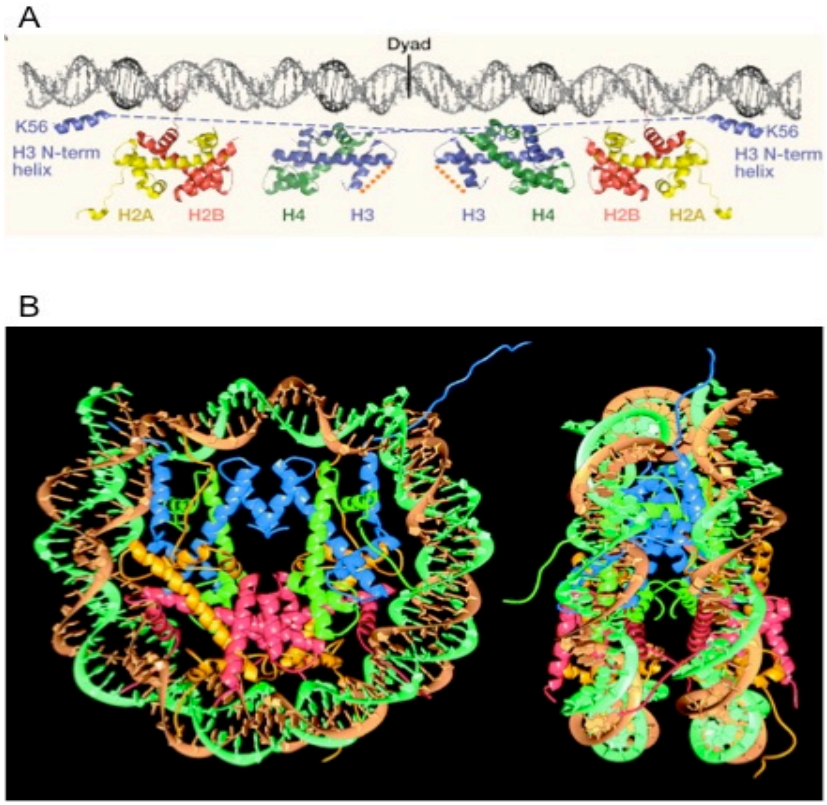


Figure 8: The organizationof the nucleosome core particle (A) Unraveling of the nucleosomal DNA to indicate which regions of the 147 bp of DNA are organized by which histone proteins. The orange dotted line indicates the H3/H4 tetramerization interface. The N-terminal alpha helix of H3 interacts with a different DNA gyre from that with which the remainder of the H3/H4 dimer interacts. (B) The nucleosome core particle: The views are down the DNA superhelix axis for the left particle and perpendicular to it for the right particle. Ribbon traces for the 146-bp DNA phosphodiester backbones (brown and turquoise) and eight histone protein main chains (blue: H3; green: H4; yellow: H2A; red: H2B. (Protein Data Bank (PDB): 1KX5). Figures adapted respectively from [66] and [65].

The interactions between the histone octamer and the DNA are divided into three groups:

1. Charge neutralization of the acidic DNA phosphate groups
2. Hydrophobic interactions.
3. Hydrogen bonds, mainly between main-chain amide group and phosphate oxygen atoms.

The flexible histone tails represents another means of DNA-protein or protein-protein interaction. Indeed, the N-terminal of both H2B and H3 are random coil segments that pass through the DNA gyres and are reported to interact with the linker DNA [65, 67].

2.5 Nucleosome positioning

As mentioned earlier, nucleosomes are arranged into regularly spaced arrays, with the length of the linker region connecting nucleosomes. It has long been known that the nucleosomes favor some position in the genome. Thanks to high-resolution genome-wide analysis, a common pattern has been discovered: the nucleosomes are absent at almost all enhancer, promoter and terminator regions [68, 69].

The term "nucleosome positioning" is used to indicate where nucleosomes are located with respect to the genomic DNA sequence. The sequence-based mapping approaches identify the positions of individual nucleosomes in a single cell at a specific time, which might overlook the dynamic process of nucleosome positioning. Nucleosome occupancy has a critical biological role because nucleosomes inhibit the access of non-histone protein to the DNA. Transcription factors have different modes of bindings depending on their ability to be inhibited by nucleosomes. For example, a class of transcription factors such as FoxA and GATA are bound to their target sequence located at the nucleosome occupied site. These classes of factors are supposed to recruit nucleosome remodelers, thus opening up the local chromatin and allowing the binding of other transcription factor that otherwise would be inhibited by the nucleosome [70]. Several factors can influence nucleosome positioning *in vivo*, including DNA sequence preferences of the nucleosomes themselves, DNA methylation, histone variants, PTMs, chromatin remodelers, DNA-binding proteins and higher-order structure.

Nucleosome positioning can vary from an almost perfect positioning, in which a nucleosome is located at a given 147 bp stretch in all DNA molecules in a substantial

fraction of cell population, to no defined positioning, in which nucleosomes are located at all possible genomic positions with equal frequency across a cell population.

The histone octamer exhibits some DNA sequence specificity, but unlike DNA binding proteins that achieve specificity by direct and strong interactions between few DNA base pairs and amino acids, the specificity of nucleosome formation resolves around the ability for a given 147 bp sequence to bend around the histone octamer [71]. *In vitro*, many DNA sequences have been tested for the ability to position nucleosomes and the highest-affinity sequence so far identified is the artificial 601 element isolated by SELEX [72]. The 601 sequence has been used numerous *in vitro* studies because it assembles into a very stable, highly positioned nucleosome. The rDNA is another nucleosome positioning sequence is widely used in *in vitro* experiments [73]. Although displaying lower positioning capacity that the 601, it has the advantage to be a natural DNA sequence.

New regulators, called chromatin remodelers, are emerging as important controls of gene expression. They form a complex chromatin remodeling machinery that uses the energy of ATP hydrolysis to control the nucleosome positions. In other words, some chromatin remodeling complexes can establish a specific nucleosome positioning patterns that defines accessibility of DNA and with it the 'on' or 'off' states of nucleosome-dependent gene expression [74].

2.6 The linker histone

The protein known these days as the linker histone was initially described as the lysine–rich nuclear protein. It was separated from the other major nuclear proteins by ion-exchange chromatography [75, 76]. While studying the nucleosome organization within the chromatin, the linker histone was found to be binding to the DNA between nucleosomes and helps chromatin compaction.

2.6.1 Sequence and domain structure of linker histone

Even though a lot of biochemical similarities were found between linker and core histones, they differ in function, architecture and evolutionary origin. Structurally, the canonical linker histones of higher eukaryotes can be separated into three domains: a short N-terminal (20 aa) a long lysine-rich C-terminal (100 aa), and a non-polar central globular domains (80 aa), while the two tails are unstructured [77]. The linker histone family is highly diverse, in that 11 stage and species-specific variants have been

discovered so far. They all differ in sequence, molecular weight, biochemical/biophysical and immunochemical properties [78].

2.6.1.1 The Globular Domain of histone H1 (GH1)

The central domain belongs to the 'winged helix' family of DNA-binding proteins. The structure of the GH1 was resolved by crystallographic studies on the closely related linker histone H5 (H1 variant in avian erythrocytes) [79]. Even though the GH1 has been internally located in the 30 nm fiber, a lot of controversy lies around its exact location within the nucleosome.

The GH1 adopts a mixed α/β fold, with three α-helices forming the core of the domain and β-strands (Figure 9A). α1 and α2 are separated by β-strands, followed by α3 and 2 anti-parallel β-strands and a loop are formed by the extension of these 2 β-strands. 2D NMR experiment on GH5 and chicken histone GH1 showed a conserved 3D structure of the globular domain among linker histones [79]. Even though the linker histone seems to be involved in stabilization of a higher-order of chromatin structure, it does not seem to be essential for DNA compaction.

For the past 30 years, efforts have been directed towards understanding the location of globular domain on either native or reconstituted nucleosomes. A limited digestion of the whole chromatin by microccocal nuclease (MNase) releases an intermediate particle containing a 166 bp nucleosome and the linker histone, often called chromatosome [80]. Several models have been postulated to decipher the exact location of the globular domain on either native or reconstituted nucleosome substrates (Figure 9) [81]. The first such model was proposed in 1986 and states that the GH1 binds to 10 bp of the entering and 10 bp of the exiting DNA (linker DNA) of the nucleosome in such a way that it is placed near the dyad axis in a symmetrical manner (Figure 9A). This model was validated by DNase I and •OH radical foot printing on the nucleosome dyad [82, 83]. However, Zlatanova and coworkers challenged this model by proposing asymmetrical GH1 binding model. Asymmetrical model suggests that GH1 binds to 20 bp of linker DNA on either entering or exiting DNA (Figure 9C) [84]. A third mode called as "bridging" model" proposes that the linker histone interacts with the dyad and with only one free DNA arm (either entering or exiting) (Figure 9B) [85]. Fluorescence recovery after photobleaching (FRAP) studies suggested the presence of only two DNA binding sites in the globular domain. One of the two binding sites fits within the major groove close to the dyad axis and the other within the minor groove on the linker DNA close to the NCP

[85]. One of the reasons for the controversial models could be partly attributed to the way H1 is deposited on the reconstituted nucleosomes. The above-mentioned *in vitro* studies used salt dialysis to deposit H1 on the nucleosomes. However, this method might leads to improper assembly of H1 [86]. Another reason that could contribute to the controversy is the use of poorly positioned nucleosomes. Indeed, the widely used 5S rDNA was shown to exhibit several translational positions, which in turn would interfere with the mapping of histone H1 nucleosomal DNA contacts.

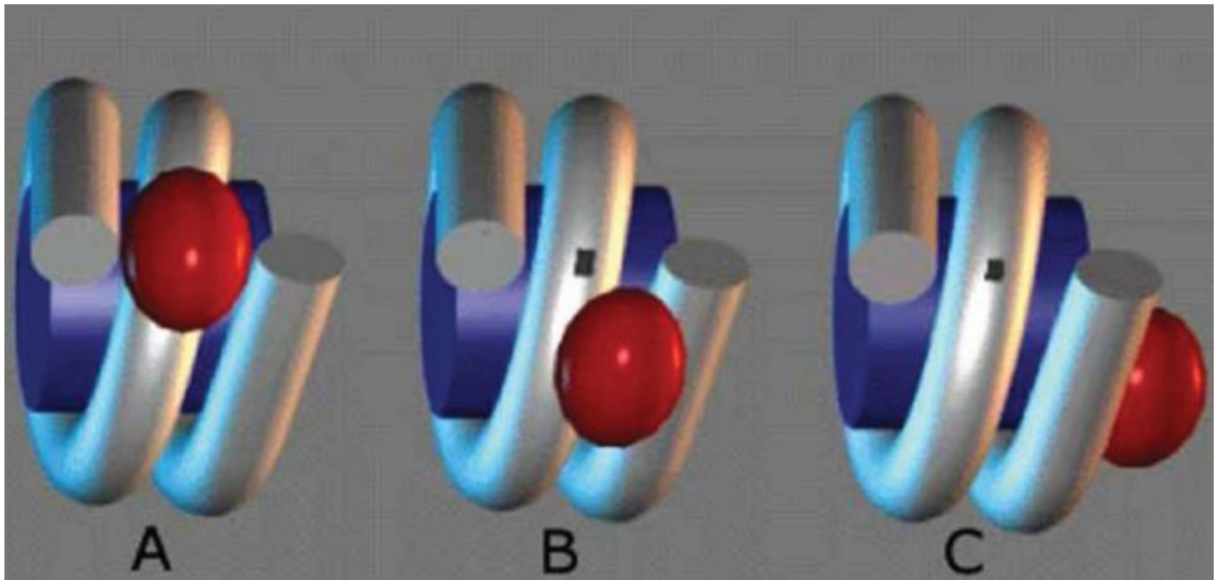


Figure 9: Three major models showing the binding of globular domain to nucleosome. (A) Symmetrical model. (B) Bridging model. (C) Asymmetrical model. In grey, Nucleosomal DNA, in bleu, histone octamer and in red, linker histone H1.

2.6.1.2 The C-terminal Domain (CTD)

The C-terminal domain is very poorly conserved between different linker histone subtypes and species. The CTD is lysine-rich region with a repeat sequence S/TPXK leading to an even distribution of around 30-50 net positive charges [87, 88].

The CTD is involved in the folding of the nucleosome arrays into chromatin fibers [89] and is required for high affinity binding to chromatin *in vivo* [90]. The data collected on the binding affinity of histone H1 subtypes to the chromatin are somewhat contradictory.

Early work showed a difference in the binding affinity of the H1 subtypes to the nucleosome arrays and thus variability in their ability for inducing chromatin compaction [91]. In another study where H1 subtypes were purified from mouse liver and testis and their binding affinity tested with reconstituted mono-nucleosomes, the

binding affinity was shown to be identical for all subtypes except H1b [92]. The methods of H1 preparation could be the reason for these contradictory results. Indeed, it has been shown that the acid-extracted H1 may have altered folding ability of the GH1 and CTD domains, which is not the case of salt-extracted proteins.

However, the specific sites or residues within the H1 CTD required for chromatin binding remain undefined. It is believed that the excess of positive charges on the CTD neutralizes the negative charges on the DNA and helps with its condensation. The exact mechanism of chromatin stabilization by the CTD is still unclear. Even if it is disordered in solution, it can adopt a segmental α-helical form upon interaction with DNA or in presence of secondary structure stabilizers (such as trifluorethanol and $NaClO_4$) (Figure 10) [93]. These results indicate that CTD binding to linker DNA within chromatin can induce folding in this specific area leading to the hypothesis that the depletion of the CTD leads to the loss of H1-dependent chromatin compaction.

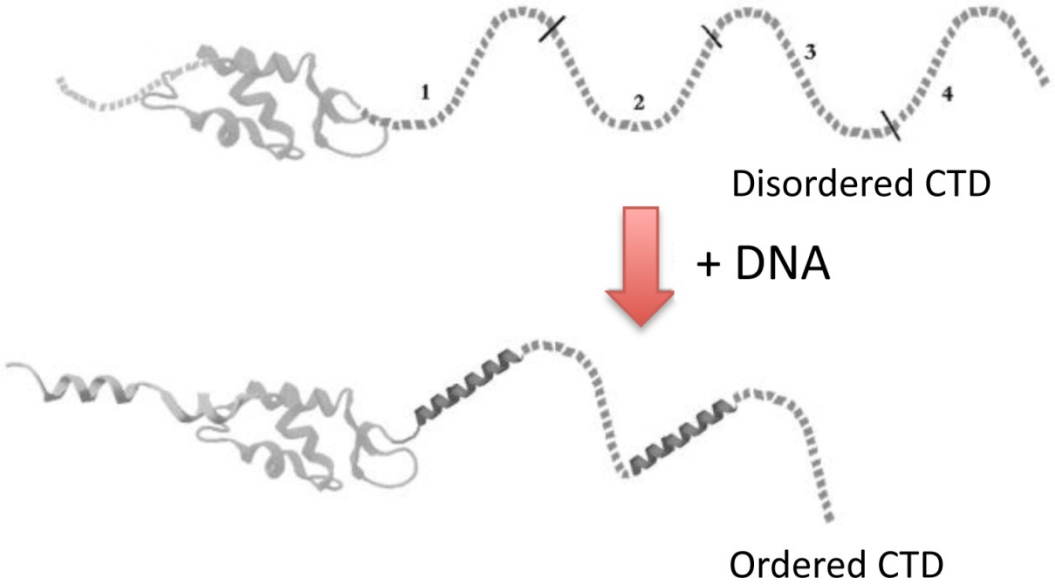


Figure 10: The H1 CTD is disordered in the absence of interacting partners. Upon interaction with DNA or other targets in chromatin, the domain adopts a secondary structure. Figures adapted from [94]

2.6.2 Functional roles of the linker histones

In vitro studies on H1-containing chromatin showed a strong inhibition of transcription [95]. Furthermore, remodeling by SWI/SNF seems to be modulated by the linker histone H1 by partially inhibiting the remodeling activity [35].

However, *in vivo*, the concept of H1 as a general repressor has been disproved by experiments showing almost no transcriptional differences when H1 is depleted in lower eukaryotes possessing a single H1 gene. Genetic 'knockout' experiments in higher

eukaryotes have been complicated by the presence of compensatory factors such as H1 alleles or variants [96].

Interestingly, studies showed a direct relation between the ratio of linker histone H1 and the nucleosome repeat length. For example, neuronal chromatin has a 0.45 H1 ratio and a162 bp NRL while glial chromatin has a 1.04 H1 ratio and a 201 bp NRL (reviewed in [97]). Alongside other data, it is stipulated that the role of histone linker is more architectural than regulatory.

2.6.3 The dynamics of H1

Measuring the dynamics of a biological molecule in a living cell was a very important task. It became possible with the advances in imaging technologies and the discovery of fluorescent proteins. FRAP is one of these techniques used to study the spatial and temporal dynamics of protein binding and/or diffusion. After photobleaching by micro-irradiation, the time required for H1-GFP to achieve recovery of fluorescence is determined by the binding affinity of H1 to the nucleosome. FRAP done on H1.2 showed a recovery time of approximately 1-2 min, which is rather fast, compared to the core histone showing little or no recovery [98, 99]. However, compared to transcription factor or other chromatin-binding proteins, which have a time recovery of 25 s, it is rather slow. These FRAP experiments of H1 described it as a highly dynamic and mobile protein rather than a statically bound protein.

Moreover, interrogations were issued concerning the domain responsible for the low binding affinity. Is it the globular domain or the C-terminal domain? In order to answer this question, a binding kinetics of histone H1 was performed using the FRAP experiments and mathematical modeling. As shown in figure 11, three different pools of histone H1 exist: a rapidly diffusing pool, a low-bound pool and finally a strong-bound pool [100].

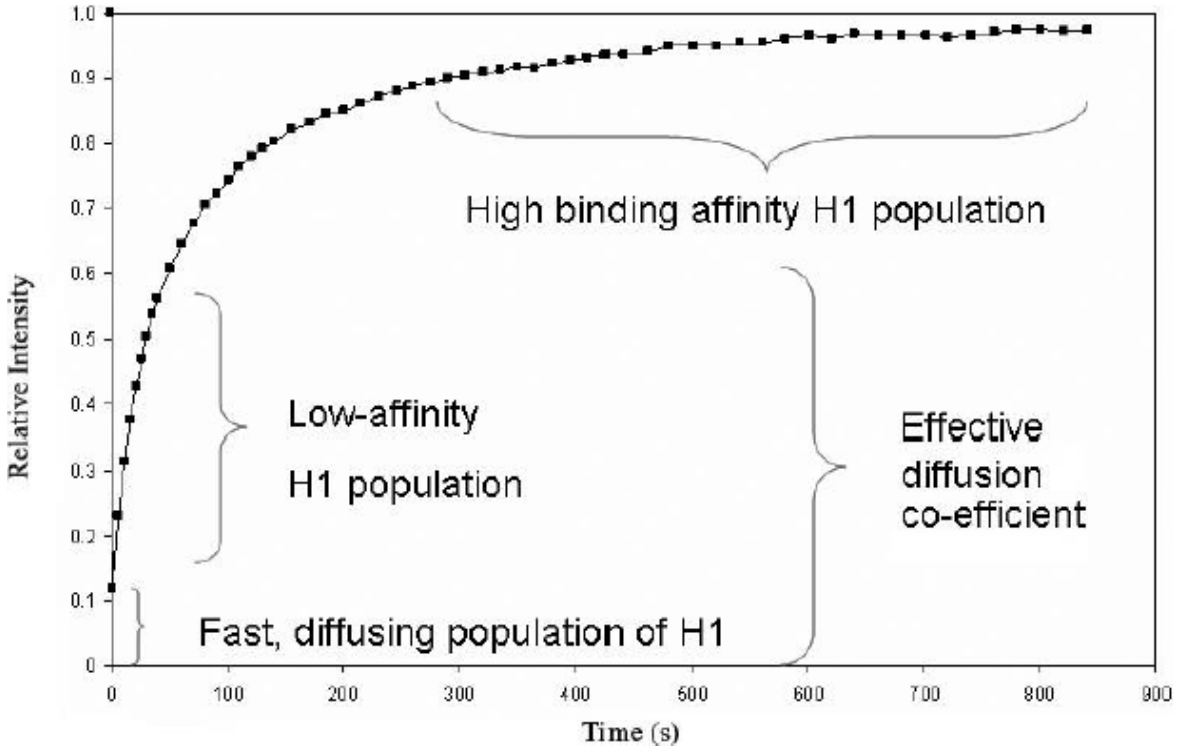


Figure 11: A typical FRAP curve for H1 binding to chromatin showing multiple populations. Figure adapted from [101]

Based on this distribution, 3 models were suggested to determine the domain responsible for the low affinity pool. The first model designates the CTD to be key for the low affinity binding by interacting first with the linker DNA. It will then allow a stable and efficient binding of the globular domain [85]. If the CTD fails to connect with the linker DNA, H1 binding would be compromised. The second model suggested that the GH1, through nonspecific electrostatic interactions, mediates the first contact, thus allowing the CTD to acquire a 3D structure capable of high affinity binding. To confirm this model, FRAP experiments were done on a mutated CTD. When Ser/Thr sites were mutated to Glu, Ala and Lys, most of the deviations in the curves were seen after 65-90 s (high binding phase), while the phase prior to 50 s (low affinity phase) was found unchanged [90]. Nevertheless, some inconsistency in the results does not validate this second model. Indeed, the Thr152Glu mutation showed alteration in the low affinity phase as well as in the high affinity phase, which suggested that both domains might be responsible for the low affinity binding. These findings led to a third model where the GH1 and CTD of H1 make the contacts simultaneously with linker DNA through an electrostatic clamp (Figure 12).

The eviction of H1 from its stable binding site is yet another unclear process.

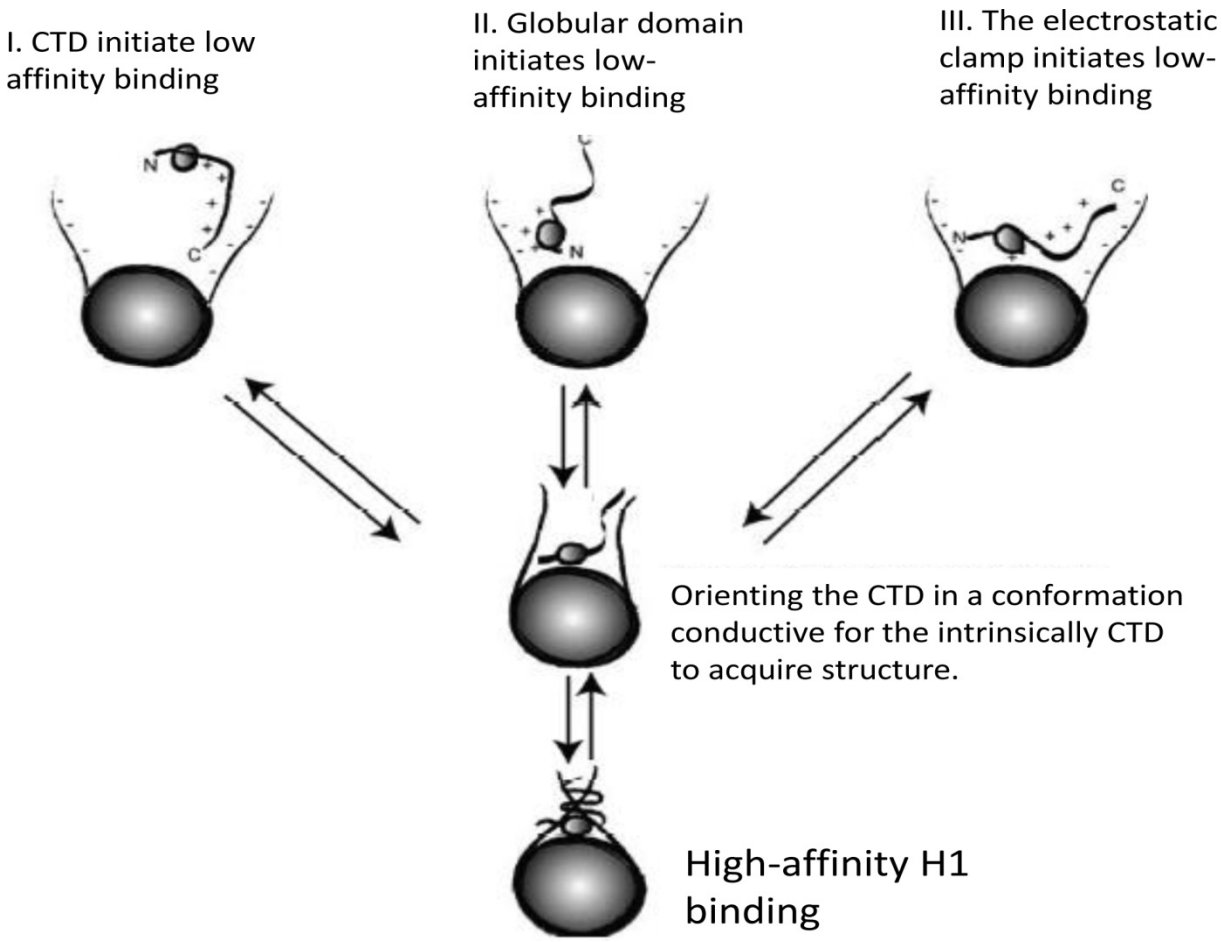


Figure 12: Different models for the reversible association of histone H1 with the nucleosome. In model I, the C-terminal domain (CTD) associates nonspecifically through electrostatic interactions with the linker DNA. In model II, the globular domain initiates a low-affinity binding interaction between the linker DNA and histone H1. In model III, both the globular domain and the CTD associate with the linker DNA to form an electrostatic clamp. Figure adapted from [102]

2.6.4 The linker histone per nucleosome stoichiometry

The linker histones H1 are involved in both the nucleosome structure, as mentioned earlier and in the formation of a higher-order chromatin structure. The 30 nm fiber was first described as a solenoid arrangement of a chain of nucleosomes in the presence of H1. This solenoid fiber was also detected in the absence of H1 at an increased ionic strength. However, this fiber was not as condensed and regularly oriented as the fiber in presence of H1 [103]. At low ionic strength, the loss of H1 leads to a decondensation of the chromatin to a more open and randomly folded beads-on-a-string form [104]. Based on the changing chromatin conformation in presence of mono- or divalent ions and on the polyelectrolyte theory of Manning, it was concluded that the mechanism of compaction is primarily electrostatic, with H1 playing a key role [105]. These results were confirmed *in vivo*, when the loss of H1 in *Tetrahymena* led to an increase of the

nuclear volume, consistent with expectations based on the disruption of the nuclear electrostatic balance [106]. The presence of several H1 subtypes complicates the studies of H1 stoichiometry. A thorough quantification of H1 stoichiometry indicated a 1 H1 per nucleosome stoichiometry in lymphocyte and glial nuclei and 0.8 H1 per nucleosome in liver nuclei [107]. Moreover, in erythrocyte nuclei, which contain the 2 subtypes H1 and H5, it was shown of 0.9 H1 and 0.4 H5 per nucleosome that added up to a ratio of 1.3:1. However, the generalization of 1 histone H1 per nucleosome has persisted in the years. Besides, it is not very clear whether the quantification of the H1/H5 levels is as precise as claimed and whether many of these *in vivo* reports are not rather measurement artifacts.

In vivo, mouse embryonic stem (ES) cells null for three H1 genes showed dramatic chromatin changes such as increasing of nucleosome spacing, reduced compaction and a decrease in histone modification. However, this depletion showed that the expression of very little number of genes was affected. The affected genes are all normally regulated by DNA methylation [108].

2.6.5 The linker histone isoforms

The H1 family of linker histones is the most divergent class of histones protein. Eleven subtypes have been identified in the human genome. This family of histones can be subdivided according to different criteria. Linker histone variants differ in their timing of synthesis, rate of synthesis, turnover rates, phosphorylation status and ability to compact chromatin. Broadly, they can be classified into three groups based on their expression pattern:

1. Genes expressed during the S-phase of the cell cycle (histone H1.1 to H1.5)
2. Genes with variable modes of expression in somatic cells (histones H1.0 and H1x)
3. Genes expressed in germ cells (histones H1t, H1T2, H1LS and H1oo)

The genes of linker histones have been found to exist either clustered (H1.1 to H1.5 and H1t) or solitary (H1.0, H1x, H1T2, H1LS and H1oo) and their distribution in the genome was found highly conserved between the human, mouse and rat genomes.

H1 protein (and synonyms)	H1 gene (and synonyms)	Gene locus (chromosome)	Length (amino acids)	Expression (cell, tissue)	Reference
H1.1 (H1a)	HIST1HA (H1F1)	6p21.3-22	214	Ubiquitous	[109]
H1.2 (H1c)	HIST1HC (H1F2)	6p21.3-22	212	Ubiquitous	[110]
H1.3 (H1d)	HIST1HD (H1F3)	6p21.3-22	220	Ubiquitous	[111]
H1.4 (H1e)	HIST1HE (H1F4)	6p21.3-22	218	Ubiquitous	[111]
H1.5 (H1b)	HIST1HB (H1F5)	6p21.3-22	225	Ubiquitous	[112]
H1t	HIST1H1T (H1FT)	6p21.3-22	206	Spermatocytes	[113]
H1T2	H1FNT (HANP1)	12q13.11	233	Spermatids	[114]
H1oo (H1foo)	H1FOO	3q21.3[f]	345	Oocytes	[115]
HILS1 (Hils1)	HILS1	17q21.33	230	Spermatids	[116]
H1x (H1X, H1.X)	H1FX	3q21.3	212	Ubiquitous	[117]
H1.0 (H1°)	H1FV (H1F0)	22q13.1	193	Differentiated cells	[118]

Table 1: The H1 histone family in mammals. Table inspired from [119]

2.6.5.1 Cell-cycle dependent H1

The five main histone H1 genes (H1.1 to H1.5) share the particular characteristics as the majority of histone genes. They are located in the major cluster of chromosome 6, they lack introns, they have short 5' and 3' non coding sequences and their mRNA ends with a hairpin loop involved in the regulation of their expression [120]. The gene transcription of H1.1 is restricted to the proliferation phase, where the cells undergo successive divisions. Furthermore, H1.1 has the highest turnover among the different variants, with a half-life of five days. Reported studies showed that the level of H1.1 decreases from 5% to 0.5% during the post-natal development of cortical neurons in rat brain and is replaced by the H1.4 variant [121].

H1.2 exhibits the highest levels of expression compared to the other somatic variants. It has been shown that H1.2 interacts with a group of protein (YB1 and PURα) and forms a

stable repressor complex that regulates p53-mediated trans-activation [122]. The H1.2 complex represses p53 via a direct H1.2/p53 interaction causing a blockage of the p300-mediated acetylation of chromatin. In addition to its nuclear role, H1.2 translocates to the cytoplasm in response to apoptotic stimuli such as DNA damage. Moreover, reducing or depleting H1.2 increases cellular resistance to apoptosis [123]. H1.2, H1.3 and H1.4 are present in non-dividing and quiescent cells of almost all tissues. H1.3 and H1.4 seem to be missing in active chromatin. H1.5 is highly expressed in the heterochromatin regions and at the nuclear periphery and unlike H1.2, H1.3 and H1.4, it has a very low level in quiescent cells.

2.6.5.2 Germ-cell line specific H1 histones

It should be kept in mind that the H1t gene is also located on chromosome 6 and has a 3' hairloop end but this protein is only expressed in morphologically distinct spermatocyte chromatin [124]. It has been reported that the H1t binds less tightly than the somatic H1 variant, causing the chromatin to be more sensitive to nuclease. This property of H1t keeps the chromatin in a decondensed state during meiosis which facilitates spermatogenesis events such as recombination or protein transition [125].

H1LS is a histone H1-like protein expressed only in elongating and elongated spermatids. It has several biochemical properties that are similar to those of linker histones including aggregating chromatin and binding to mononucleosomes. Since H1LS is expressed in late spermatids (that do not contains the core histones), it might use another mechanism for chromatin condensation that the linker histones. In addition, based on the expression pattern of H1LS, it was suggested that it participates in chromatin remodeling and regulates transcription during spermiogenesis [116].

Oocyte-specific linker histone H1oo is the mammalian homologue of the oocyte-specific linker histone B4 in frogs and cs-H1 in sea urchin. It is expressed in early stages, as early as germinal vesicle, metaphase II, polar bodies formation and two-cell stage embryo, but it disappears in the 4-8 cell embryonic stages. The functions of this protein are unclear but it might play a key role in control of gene expression during oogenesis and embryogenesis through perturbation of the chromatin structure [126].

2.6.5.3 Non cell cycle dependent H1

H1.0 and H1c genes are both solitary genes, located respectively on the 22nd and 3rd chromosomes. Their genes are intronless and they both express polyadenylated mRNAs.

The expression of this replication-independent variant of H1 is linked to cessation of DNA synthesis [127]. Although having the shortest C-terminal tail, H1.0 binds to the chromatin with a moderate affinity compared to some others H1 variants with longer C-terminal tails [98].

However, their patterns of expression are different. Indeed, H1.0 gene is expressed in only some cell types and the protein H1.0 plays a role in terminal differentiation [128]. As for the H1x gene, it is ubiquitously expressed and the level of the protein remains stable throughout the cell cycle [129].

2.7 The histone tail and higher-order structure

Each histone has an N-terminal tail domain that extends to the exterior of the nucleosome and contacts extra-nucleosomal constituents in chromatin (Figure 13). Together with the C-terminal tail of H2A, the count goes up to 10 tails projecting outside the nucleosome through the superhelical gyres [65]. The N-terminal tails of H2B and H3 go between the superhelical gyres whereas the N-terminal tails of H2A and H4 go under or over them. The tails present a low affinity to the nucleosome core DNA. In absence of DNA, they adopt an unstructured random coil conformation, which subjects them to a higher proteolytic cleavage compared to the histone fold domains [130]. They are highly charged positively due to the predominance of lysine and arginine residues. The histone tails main implication was reported to be regulation of transcription *in vivo*. However, it was noted that they play a small role in nucleosome stability and accessibility to DNA binding factors. The involvement of the tails in higher-order chromatin structure was first noticed during experiments involving proteolytic removal of the tail domain. Indeed, early experiments on native chromatin extracted from chicken erythrocytes nuclei showed that the digestion of the tails resulted in unfolded chromatin. Once these tails were replaced by an extraneous polypeptide, the chromatin acquired back its original compaction under specific ionic strength conditions [131].

These findings were later on confirmed on native and reconstituted chromatin. Indeed, the self-association of oligonucleosomes in presence of divalent cations (Mg^{2+}) showed that the H3/H4 tail mediates the compaction directly and not through the electrostatic interaction that might occur. Furthermore, hybrid nucleosome array where either H2A/H2B dimers or H3/H4 tetramers lacked their N-terminal tail did not show any compaction even in high salt concentration. These results combined bring us to the

conclusion of the equal importance of all the core histone tails [132]. On the contrary, in presence of divalent cations, nucleosomes containing only H3/H4 tails were able to oligomerize unlike nucleosomes containing only H2A/H2B tails [133]. Due to the importance of histone tails in the chromatin structure, it is also important to note the effect of the post-translational modifications that these tails encounter on the chromatin. Indeed, as mentioned earlier, the histones tails are subjected to many modifications such as methylation, acetylation, ubiquitination, phosphorylation, etc. (Figure 13).

The histone tails-induced interactions can be divided into two categories: intra-nucleosome interactions where the tails interact with DNA/protein of their own nucleosomes and the inter-nucleosome interactions, where the tails mediate the interactions between different nucleosomes within condensed chromatin. The core histone tail domains can however participate in both kinds of interactions simultaneously. Understanding how tails interact within and between nucleosomes is essential for understanding chromatin folding and genetic regulation.

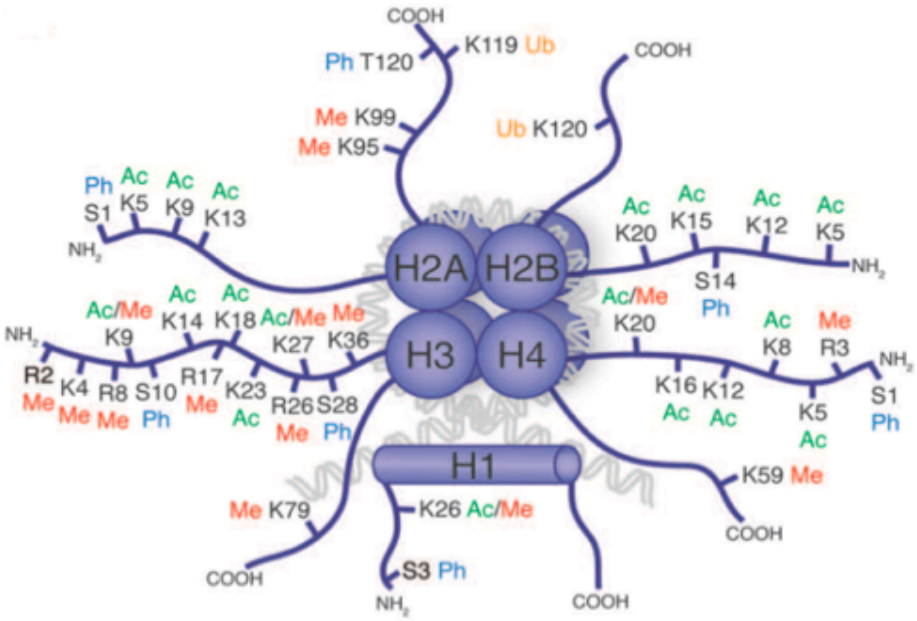


Figure 13: Schematic drawing of a nucleosome with the four canonical histones (H2A, H2B, H3 and H4) and the linker histone H1. The covalent PTMs [methylation (Me), acetylation (Ac), ubiquitination (Ub), and phosphorylation (Ph)] are highlighted on the N- and C-terminal tails of each histone. Figure adapted from [134].

Throughout the years some intra-nucleosome contacts of tail domains have been characterized. Using high-intensity UV laser crosslinking it was shown that the core histone N-termianal tails, especially H3 and H4 interact with linker DNA as well in cellular chromatin as in reconstituted nucleosomes and these interactions persists upon acetylation [67, 135]. It was also shown that the C-terminal tail of histone H2A interacts with the DNA near the nucleosome dyad in the NCP and also near the edge of the nucleosome core region when linker DNA is present [136, 137]. The N-terminal tail of H2A has been reported to bind to the nucleosome core DNA as well, at two positions around 40 bp from each side of the dyad [138]. In addition to H2A, the N-terminal tail of H4 presents a similar kind of binding, where the H4 tail binds to core DNA at a distance of 1.5 helical turn on both sides of the dyad [139].

Moreover, it has been demonstrated that the H3 tail binds within the nucleosome DNA when the nucleosome array is in beads-on-string conformation, but when the array undergoes salt-dependent folding and oligomerization the tails engage in inter-nucleosome interactions. Since H3 contributes to both types of interactions, crosslinking studies showed that only 20% of these interactions are inter-nucleosome whereas the 80% left are intra-nucleosome interactions. Furthermore, the exposed regions of H3 tails contribute to the inter-nucleosome interactions, while the regions near the histone fold are associated with the intra-nucleosome interactions. The same studies performed in presence of H1 showed that the histone H1 does not affect these interactions but does however stabilize the chromatin structure [140].

While the nucleosome crystal structure gave a lot of information on the intra-nucleosome contacts of the tail domains, this information was not complete. Indeed, the lack of linker DNA in the crystal of the nucleosome and the presence of a very high divalent cation concentration may affect the actual interaction within the native chromatin. Interestingly, on the first X-ray structure of the nucleosome, it was noted that the H4 tail domain of one nucleosome interacts with the surface of the H2A/H2B dimer of a nearby nucleosome [65].

2.8 Histone modifications

The four core histones, the linker histone and their variant are all likely to be subjected to post-translational modifications (PTMs), which play a crucial role in chromatin dynamics. PTMs are particularly abundant on the N-terminal histone tails, but they also

exist within the core and on the C-terminal tails. The histone code hypothesis is proposed in order to explain and understand the pattern of PTMs and their biological consequences. This hypothesis states that PTMs act through two distinct mechanisms:

- Regulating DNA accessibility by favoring inter-nucleosome contacts.
- Being used as a docking site to initiate biological processes.

The histones can be subjected to a wide range of PTMs such as acetylation, phosphorylaton, methylation and ubiquitination (Figure 14).

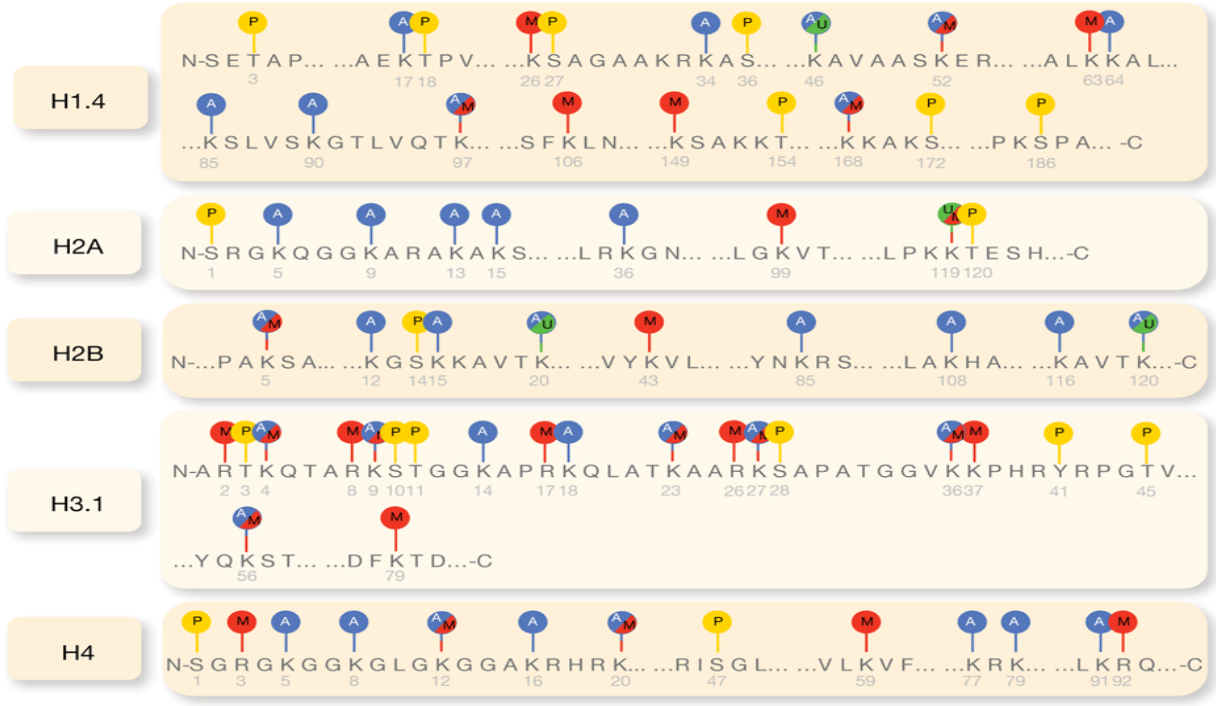


Figure 14: Histone post-translational modifications. All histones are subject to post-translational modifications, which mainly occur on histone tails. The four post-translational modifications are depicted in this figure: acetylation (blue), methylation (red), phosphorylation (yellow) and ubiquitination (green). The numbers in gray under each amino acid represent their position in the protein sequence. Figure adapted from [141].

Distinct histone amino-terminal modifications can generate synergistic or antagonistic interaction affinities for chromatin-associated proteins, which in turn dictate dynamic transitions between transcriptionally active or transcriptionally silent chromatin states. The combinatorial nature of histone amino-terminal modifications thus reveals a "histone code" that considerably extends the information potential of the genetic code. Hyperacetylated oligonucleosomes exhibit a reduced tendency to fold into secondary and tertiary structures compared to non-acetylated arrays [142]. For example, the substitution of Lys by Glu, which mimics the acetylation of the histone tails, inhibits the

array condensation [143]. Interestingly, the H3 acetylation only disturbed DNA wrapping stability whereas H2A acetylation showed different outcomes depending on the other histone modifications. This phenomenon can be described as the acetylation-dependent histone-histone interactions [143].

2.8.1 Histone acetylation

All core histones can be acetylated but the majority of acetylations occur on lysines in the N-terminal of H3 and H4. It occurs during DNA replication and when genes are activated.

This process is performed by a class of enzymes called Histone Acetyl Transferases (HATs). HATs can be cytoplasmic or nuclear. The cytoplasmic HATs acetylate histones prior to nuclear organization and chromatin assembly, whereas the nuclear HATs acetylate histones in order to affect transcription or other DNA-dependent processes. Biochemical analysis of a variety of proteins and protein complexes involved in transcription activation allowed the identification of a large number of HATs [144]. *Tetrahymena* p55 was the first nuclear HAT identified that provided a link between histone acetylation and transcriptional activation [145]. Recently, a connection between HATs and transcriptional elongation has been made. Indeed, coding regions need to be poorly acetylated to prevent aberrant initiation of transcription. Through a complex mechanism of methylation, a Histone DeAcetylase Complex (HDAC) is recruited and specifically deacetylates coding regions [146]. H4K16 plays many major roles, for example, maintaining boundaries between euchromatin and heterochromatin in yeast. In *Drosophila*, it is also involved in the process of dosage compensation (equalization of gene expression between male X chromosome and female X chromosome). It was also determined that the H4K16 inhibits the formation of higher-order chromatin [147]. The single acetylation of the Lys 16 seems to have the most important effect on chromatin folding compared to other modifications. The nucleosome-nucleosome interactions decrease once Lys 16 is acetylated, which contributes to the unfolding of the compacted chromatin. A sedimentation experiment was performed on 12x177 bp nucleosome arrays containing wild-type (WT) H4 and K16Ac H4 octamers. Arrays containing WT H4 in presence of 1 mM $MgCl_2$ formed compact fiber, sedimenting at 54S whereas the arrays containing acetylated H4 were unable to compact, thus sedimenting at 44S. Moreover, *in vivo* studies consisting in separating chromatin fractions via Micrococcal nuclease (MNase) digestion into $MgCl_2$-soluble and -insoluble component showed that

H4K16Ac is present in higher amount in soluble fraction, corresponding to decondensed chromatin, suggesting its involvement in chromatin compaction [147]. The exact mechanism by which this acetylation alters the chromatin folding is still unknown. The loss of the lysine positive charge due to acetylation can plays a role but it reduces by 10% only the positive charges in the tail domain of histone H4.

Histone acetylation is involved in yet another DNA-dependent process; DNA damage repair. The HAT protein TIP60 activates the DNA repair damage pathway by acetylation and activation of the ATP kinase [148]. Moreover, H3K56 acetylation is required for inactivation of the DNA damage checkpoint after DNA repair is complete, thus allowing cells to re-enter the cell cycle [149].

2.8.2 Histone phosphorylation

Histone phosphorylation has important functions in mitosis and gene regulation. Abnormal histone phosphorylation has been observed in many cancers but it remains unclear whether histone phosphorylation is the cause or the effect. The four core histone, histone variants and H1 are phosphorylated on either the N-terminal or C-terminal sides. It has been demonstrated that H1 and H3 are phosphorylated at different times during cell cycle whereas H2A and H4 are phosphorylated at stable rate throughout the cell cycle [150]. The cell-cycle dependent phosphorylation of H1 and H3 hits its highest level during the M-phase and its lowest level during the G1-phase. *In vivo* studies showed that the phosphorylation of linker histone H1 tails influence polynucleosome folding. Indeed, phosphorylation of N- or C-terminal ends of linker histone H1 leads to the unfolding of the chromatin which in turn regulates the access of transcription factors [151].

H3 phosphorylation increases dramatically during mitosis while H3 is dephosphorylated upon anaphase entry. H3S10 phosphorylation has been proved to weaken the association of the H3 tail to the DNA, which might promote DNA binding of other factors [152].

The centromere-specific histone H3-like variant (CENP-A) is phosphorylated by aurora B during prophase and dephosphorylated by protein phosphatase 1γ 1(PP1γ1) during anaphase. Absence of this phosphorylation is reported to delay the terminal stages of cytokinesis in HeLa cells [153].

H2A histone variants are also phosphorylated, for example H2AX is rapidly phosphorylated at Ser139 in response to double-strand break. Phosphorylated H2AX

appears to decondense the chromatin, recruit factors involved in DNA repair and favor the assembly of DNA repair complexes [154].

Histidine phosphorylation of histone H4 occurs on histone residues H18 and H75. A clear relationship has been established between an enhanced activity of the H4 histidine kinase and regeneration of damaged liver in porcine [155]. However, no mammalian histone H4 histidine kinase has been characterized yet.

2.8.3 Histone methylation

Histones H2B, H3 and H4 are the only histone modified by methylation. Lysine and arginine residues at the N-terminal tail of H3 and H4 are methylated. These modifications do not affect the charge of the protein but they increase its hydrophobicity and reduce its ability to form hydrogen bonds.

Histone methyltransferases (HMTs) are classified into two families the histone lysine methyltransferases and the histone arginine methyltransferases, based on substrates and structural motifs. Each HMT can methylate multiple substrates.

H4K20 is one methylation site that might alter the nucleosome structure because of its location in the basic region that binds to the DNA. The level of H4K20 methylation varies during the cell cycle, reaching maximum levels at G1 and mitosis and decreasing at S phase. In contrast, H4K16 acetylation reaches maximum levels at mid S phase and decreases to lower levels at G1 and mitosis phases. These observations, alongside other studies, showed that H3K20 methylation inhibits H3K16 acetylation and/or vice versa [156].

Histone H3 may be methylated at lysines 4, 9, 27, 36 and 79. It has been shown that transcribed condensed regions were highly enriched in acetylated H3 and di-methylated H3K4. On the contrary, repressed condensed regions are poor in acetylated histones and enriched in methylated H3K9. These two domains can be separated by an insulator element, which supports the histone code hypothesis that each histone modification has a distinct biological consequence [157]. This methylation is responsible for recruiting HPI to heterochromatic regions.

2.8.4 Histone ubiquitination

Histone H2A was the first identified protein to be modified by ubiquitin in cells [158]. Together with H2B, they are the most abundant ubiquitinated proteins in the nucleus.

Multiubiquitination of intracellular proteins is required for selective degradation by the proteasome through the ubiquitin-dependent pathway. However, monoubiquitinated histones are not degraded by the proteasome and used for intercellular signaling as for most monoubiquitinated proteins. The most common ubiquitinated form of histone is monoubiquitination, where a single ubiquitin is added to a lysine residue (Lys 119 for H2A and Lys 123 (yeast) or Lys 120 (vertebrate) for H2B). In addition, H3 and H4 are poly-ubiquitinated by ubiquitin ligase complex after UVC irradiation, but the biological function is still unclear [159]. H2A(ub) is reported to be associated to gene silencing and H2A ubiquitin ligases were found in transcription repressor complexes. Furthermore, H2A deubiquitin ligase is required for gene activation, which provides another evidence of the gene silencing role of H2A(ub) [160]. On the other hand, H2B (ub) is implicated in gene activation through various mechanisms, including promoting other histone modifications and elongation by DNA polymerase II. For example, over-expression of H2B ubiquitin ligase in mammalian cells leads to an increase in H2B(ub), which in turn leads to an increase in methylation of H3K4 and H3K79 and thus to over-expression of the HOX gene [161].

The precise mechanism by which ubiquitination contributes to gene regulation, DNA repair and other biological processes is still poorly understood.

The classical interpretation of the effect of epigenetic modifications, implying that a single modification is independent and has a predictable and stable outcome, has now evolved. A new model has been established based on the multitude of crosstalks between specific histone modifications. The main assay for crosstalk detection involves site-specific antibodies that specifically recognize the modification coupled to mass spectroscopy analysis, but this requires a prediction of possible interactions. However, the use of microarrays and next generation sequencing will increase the crosstalk identification.

2.9 Histone Variants

Even though histones are among the slowest evolving proteins, they have non-allelic variants that can have significant differences in primary sequence. Some of them have different biophysical characteristics that might alter the properties of the nucleosomes, while others have specific locations in the genome. In general, histone variants are present in single copies in a gene and are expressed throughout the cell cycle. Some of

these variants are replacements histones that can be substituted to histones during development and differentiation. Histone H4 does not appear to have a known sequence variant but variants were reported for all three remaining conventional histones.

Core histone	Histone variant	Species	Chromatin effect	Function
H3	CENP-A	Ubiquitous		Epigenetic marker of the centromere
	H3.3	Ubiquitous	Open chromatin	Transcription
H2A	H2A.Z	Ubiquitous	Open/closed chromatin	Transcription/double strand break repair
	H2AX	Ubiquitous	Condensed chromatin	Double strand break repair/meiotic remodeling of sex chromosome
	MacroH2A	Vertebrate	Open chromatin	Gene silencing/ X chromosome inactivation
	H2A.Bbd	Vertebrate	Chromatin condensation	Epigenetic mark of active chromatin.
H2B	SpH2B	Sea urchin	Chromatin condensation	Chromatin packaging

Table 2: Histone variants and their functions. Table inspired from [162]

2.9.1. Histone H2A variants

The H2A is among the largest family of identified histone variants with some universal variants found in almost all organisms such as H2A.Z and H2AX (Table2). The C-terminal tail is where sequence diversification between variants is found, regarding length as well as amino acid sequence.

2.9.1.1 H2A.Z Histone Variants

Histone H2A.Z is the best-studied histone variant; it shares roughly 60% of homology with canonical H2A within the same species. H2A.Z is important to many organisms such as mouse, fly, *Tetrahymena* etc… Surprisingly, despite the divergence in sequence, H2A.Z nucleosome structure shows high similarity to the canonical one [163]. The significant differences are in L1 loop, which is important to interactions between the H2A-H2B dimers and the DNA within the nucleosomes. These differences of structure led to the hypothesis of a homotypic H2A.Z nucleosome, because the presence of both types of

histone in one nucleosome would disturb the particle. However, *in vivo* and *in vitro* experiments showed the presence of a heterotypic H2A.Z nucleosome in *S.cerevisiae*, fly and human. The docking domain of the C-terminal moiety displays a sequence difference, suggesting a possible alteration of the interaction with the linker histone H1. Finally, an increased acidic patch was observed in H2A.Z nucleosome suggesting a difference in internucleosomal interactions [164]. Different methodologies used to understand the properties of the H2A.Z-containing nucleosome might be the reason of contradictory results over the stability of this nucleosome. The biological function of H2A.Z has been extensively studied revealing a role in the regulation of transcription, DNA repair, heterochromatin formation, chromosome segregation and mitosis. Although it is not essential to yeast cells, it has been reported that the absence of H2A.Z is lethal to some species such as *Tetrahymena thermophila*, *Xenopus laevis*, *Drosophila melanogaster* and mice [165]. Genome-wide studies showed a non-random pattern of H2A.Z distribution [166]. Indeed, in euchromatin, H2A.Z presence peaked at the 5' end of many genes, enhancers and insulators. Furthermore, in budding yeast H2A.Z nucleosomes are present in flanking regions of transcription start sites. In yeast cells, the presence of H2A.Z at gene promoters is inversely correlated to transcription levels, and by contrast in *Drosophila* and human cells, the presence of H2A.Z is positively correlated to transcription [166]. Non-acetylated H2A.Z is found in the heterochromatin regions, distributed all over hundreds of kilobases. The deposition of H2A.Z is catalyzed by ATP-dependent nucleosome remodeling complexes such as Swr1 complex in budding yeast [167]. It has been demonstrated that the interaction between Swr1 complex and H2A.Z is essential for the deposition. Since the mechanism by which H2A.Z is targeted by these complexes need to be resolved, it would be interesting to check if these complexes will remain associated with H2A.Z when it is incorporated within the chromatin. Recently, an answer was partially given suggesting that histone chaperone ANP32E evicts H2A.Z/H2B dimer by interacting with a short region of the docking domain of H2A.Z [168].

2.9.1.2 H2AX Histone variant

H2AX has an extended C-terminal moiety characterized by a unique SQE (Ser-Gln-Glu) motif, which is targeted for post-translational phosphorylation in response to DNA damage. Genome-wide analysis showed that H2AX is evenly distributed and represents 2 to 25% of the mammalian histone H2A pool [169]. Phosphorylated H2AX accumulates

at the site of high-intensity laser-induced double strands break (DSB), generating structures called foci [169]. In response to a double strand break, an accumulation of repair proteins within the damaged foci has been noticed by fluorescent microscopy. Two contradictory role have been assigned to H2AX on one hand that H2AX foci are essential for recruiting the DNA repair proteins [170], while on the other hand H2AX foci could be essential for the retention of these proteins rather than for recruiting [171]. Emerging data indicate that H2AX may have other functions than repair. These data showed that H2AX is phosphorylated independently of DSB and that these modifications are under cell cycle regulation. However, the exact function of these foci is not known.

2.9.1.3 MacroH2A

MacroH2A is a vertebrate-specific variant, which has two distinct domains. The N-terminus that is similar to H2A, sharing 64% identical amino acids and a large C-terminus that has no similarity to other histones. The C-terminal domain is lysine-rich and includes a random coil. MacroH2A-containing nucleosomes show the same structural similarity to canonical nucleosomes and wrap the same amount of DNA [172]. However, due to a four amino-acid difference within the L1 loop, macroH2A-containing nucleosomes are more stable in structure and form more compact chromatin [173]. The main role attributed to microH2A in mammalian cells was transcriptional repression, based on its enrichment in the transcriptionally inactive X-chromosome (Xi) and its association to inactive alleles of imprinted genes. However, further studies showed macroH2A enrichment in inducible genes and bivalent genes in stem cells where it remains bound before and after transcription, leading to the hypothesis that macroH2A is necessary for establishing a chromatin environment for gene transcription. Many hypotheses have been proposed to explain the repression mechanism of macroH2A. On one hand it was suggested that macroH2A directly prevents the binding of transcription factors and chromatin remodeling [174], and on the other hand, it was suggested that macroH2A inhibits the p-300-dependent histone acetylation *in vitro*, and indirectly represses the transcription [175].

2.9.1.4 H2A.Bbd

The histone H2A.Bbd (Barr body deficient) was first identified from a search of human ESTs (expressed sequence tag) with distant homology to H2A [176]. Indeed, H2A.Bbd

has only 48% identity with conventional histone H2A in humans. H2A.Bbd lacks both the C-terminal tail and the end of docking domain that characterizes a typical H2A family histone. Chromatin fractioning showed that H2A.Bbd is present in nucleosomes with histone H4 acetylated on the Lys-12. Photobleaching studies showed a faster recovery of H2A.Bbd than H2A in the nucleus, which indicates a higher mobility. Furthermore, stability studies on H2A.Bbd nucleosomes led to the conclusion that the nucleosome is more open and less stable than H2A nucleosomes [177]. H2A.Bbd-containing nucleosomes protect only ~120bp of DNA resulting in open nucleosomes that facilitate transcription [178, 179, 180]. Besides, H2A.Bbd display a testis specificity and its low abundance in somatic cells has made it extremely difficult to find the function for this protein [181].

2.9.2 Histone H2B variants

The few documented H2B variants are capable of completely replacing the major H2B subtypes, and they have specialized functions in chromatin compaction and transcription regulation. The sperm-specific H2Bs of sea urchins have a long N-terminal tail that is highly charged. This tail is involved in chromatin condensation, which implies that SpH2B plays an important role in the chromatin packaging in sperm [182].

2.9.3 Histone H3 variants

Histones H3 have five variants: H3.1, H3.2, H3.3, CENP-A and 3.1t. H3.3 and CENP-A are the two major histone H3 variants. The distribution of the H3 variants among eukaryotes is varied, for example *S. cerevisiae* only contains an H3 variant that resembles to H3.3 and Cse4 (CENP-A). H3.1 may play a role in silencing and gene activation once it is dimethylated at lysine 9 and acetylated at lysine 4. H3.2 is also implicated in gene silencing through di- and trimethylation of lysine 27. H3.1 and H3.2 are canonical H3 proteins, highly expressed in S-phase and incorporated into the chromatin during DNA replication.

2.9.3.1 Histone H3.3

The histone H3.3 differs from H3.1 by 5 amino acids, is expressed throughout the cell cycle and is incorporated into the chromatin during DNA replication or independently [183]. H3.3 is linked to active transcription by modifications such as the tri-methylation of lysine 4. Despite little effect on the mononucleosome structure, recent findings

demonstrated that H3.3 incorporation leads to an open chromatin that promotes transcription [184]. Moreover, H3.3 co-localizes with H2A.Z at the promoters of active genes.

The 5 amino acid difference between canonical H3s and H3.3 might play a direct role in nucleosome dynamics and PTM pattern. For example, the presence of serine instead of alanine at position 31 leads to a probable phosphorylation site during mitosis. Although H3.3 is present at lower levels in dividing cells, the levels increase significantly during terminal differentiation and contribute to more than 50% of the total amounts of H3 in the cell [185].

H3.3 has a different mechanism of deposition than the canonical H3s. H3.1 is deposited by histone chaperone CAF1, whereas the H3.3 is deposited by histone chaperone HIRA [186].

2.9.3.2 CENP-A, the centromere-specific histone

CENP-A is a centromere-associated protein that is required to build a fully functioning kinetochore. It was purified alongside the nucleosome core proteins, which indicated that it forms a complex with the core histones [187]. Homologues of CENP-A were identified from yeast to mammals, where their loss was found lethal. The homology between CENP-A and H3 is essentially found in the α-helical carboxy-terminal histone-fold domain. The N-terminal tail of CENP-A is highly variable between species and it is mostly required to recruit kinetochore proteins to the centromere [188]. Furthermore, in mammalians cells, CENP-A is overexpressed throughout the cell cycle where it is deposited in non-centromeric regions. These mislocalized CENP-A will recruit CENP-C and other kinetochore components which will drive the alignment and segregation of chromosomes.

The CENP-A-H4 tetrasome has been found to be more rigid that the H3-H4 tetrasome, and the fact that CENP-A-containing nucleosomes are more prone to unwrapping might allow easier removal from non-centromeric chromatin [189]. CENP-A nucleosome is structurally distinct from the canonical nucleosome which might expose the CENP-A nucleosome on surface of the centromere. The crystal structure of the CENP-A nucleosome showed that the histone octamer wraps the DNA in a left-handed orientation and covers only 121 bp vs 147 bp for canonical nucleosome. As mentioned before the αN helix is important for the orientation of the DNA in the canonical H3 nucleosome, which might explain the missing 13 bp on each end of the DNA in the

crystal structure of the CENP-A nucleosome (Figure 15) [190]. The crystal structure also revealed that CENP-A contains 2 extra positively charged amino acid residues (Arg and Gly) residues in the loop 1, whose depletion demonstrated that they are important for a stable CENP-A retention at the centromere but not for targeting CENP-A to the centromere. All these differences between CENP-A and H3 nucleosomes might play a crucial role in the centromeric chromatin architecture [190]. The only PTM of CENP-A known to date in human cells is the phosphorylation of the Ser 7 in the N-terminal tail, the absence of which results in a delay in cytokinesis [191].

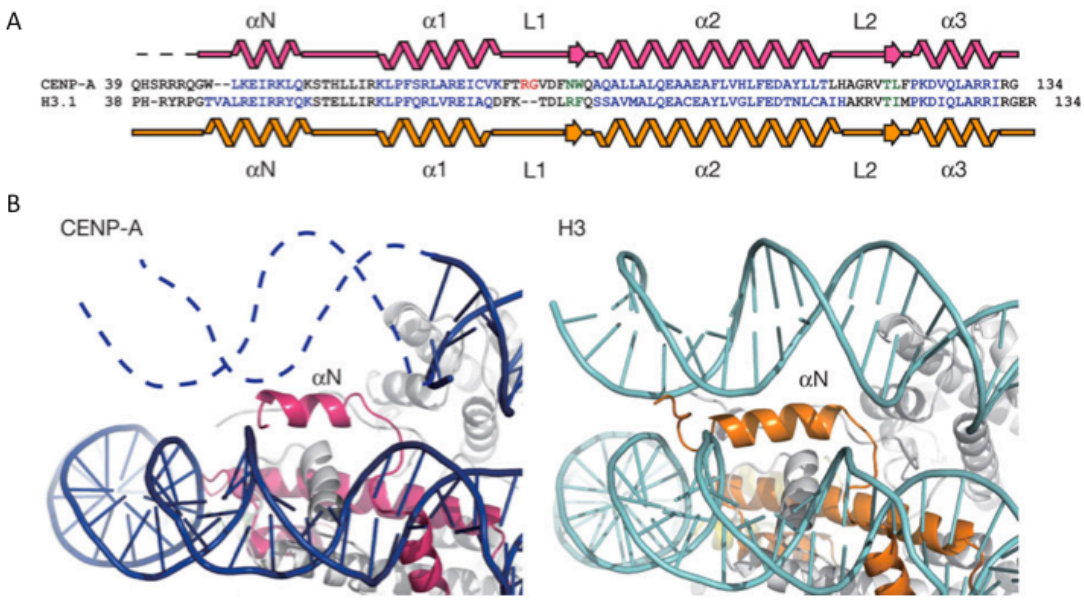


Figure 15: Structure of the DNA entrance and exit of the human CENP-A nucleosome (A) Secondary structure of CENP-A in the nucleosome. The sequences of human CENP-A and H3 are aligned with the secondary structure. (B) Close-up views of the αN helices and the DNA edge regions of the CENP-A (left panel) and H3 (right panel) nucleosomes. The dashed lines correspond to the DNA region that is not visible in the crystal structure. H3 is represented in orange and CENP-A is represented in magenta. Figure adapted from [190]

2.10 Histone chaperones

The term histone chaperone was first used to describe the role of nucleoplasmin in the prevention of histone DNA aggregation during nucleosome assembly. However, they are more generally used to define a group of proteins that bind to histones and regulate nucleosome assembly [66]. In general, histone chaperone can be classified as either H3-H4 or H2A-H2B chaperones based on their preferential histone binding. However, most of the histone chaperones bind to both H3-H4 tetramer and H2A-H2B dimer. In addition to the canonical histones, variant histones have probably unique histone chaperones

that recognize and deposit them specifically. Each histone chaperon can participate in four distinct steps of the nucleosome assembly.

1. Histone protein importers. For example, Nap1 transports the newly synthesized histone from the cytoplasm to the nucleus [192].
2. Maintaining histone supply. For example, NASP acts as a histone reservoir and regulates histone supply under stress conditions [193].
3. Regulating histone-modifying enzymes. For example, ASF1 serves as a bridge between histone sand the PTM enzymes [194].
4. Depositing histones during nucleosome assembly [195].

Table 3 summarizes the major histone chaperones and their role in nucleosome assembly.

Histone chaperone	Target	Function
Anti-silencing factor 1 (Asf1)	H3-H4	Histone import; histone transfer to CAF-1 and HIRA; regulation of H3K56ac
Chromatin assembly factor 1 (CAF-1)	H3.1-H4	H3.1–H4 deposition; (H3–H4)2 formations
Death domain associated protein (Daxx)	H3.3-H4	H3.3–H4 deposition at telomeric heterochromatin.
DEK	H3.3-H4	Regulation of H3.3–H4 incorporation and maintenance of heterochromatin
Histone cell cycle regulation defective homolog A (HIRA)	H3.3-H4	Deposition of H3.3-H4 at genic regions
Nuclear autoantigenic sperm protein (NASP)	H3-H4	Histone supply and turnover
Regulator of Ty transposition (Rtt106)	H3-H4	Formation and deposition of (H3–H4)2 tetramer
Holliday junction recognition protein (HJURP)	CENP-A-H4	Regulation of incorporation of the H3 variant CENP-A
Facilitates chromatin transcription (FACT)	H3-H4, H2A-H2B, H2AX-H2B	Deposition and exchange of H3–H4, H2A-H2B, H2AX–H2B
Nucleosome assembly protein 1 (Nap1)	H3-H4 and H2A-H2B	H2A–H2B nuclear import and deposition
Chaperone for H2A.Z–H2B (Chz1)	H2A.Z-H2B	H2A.Z–H2B deposition
Aprataxin-PNK-like factor (APLF)	Core histones and macroH2A-H2B.	Regulation of macroH2A.1 incorporation during DNA damage

Table 3: Histone chaperones and their functions during nucleosome assembly. Table inspired from [196]

Nap1 shares sequence homology with a large class of histone chaperones, including Nap1-like proteins (NAPL) in human cells and Vps75 in yeast. *In vitro* studies show that Nap1 has the same affinity for H3-H4 tetramers and H2A-H2B dimers but *in vivo* Nap1 preferentially binds to H2A-H2B. These contradictory findings suggest that there is a regulatory mechanism for the interaction between Nap1 and H2A-H2B dimers [192]. Like all histone chaperones, Nap1 intervenes at various steps of H2A-H2B deposition and exchange. It starts by mediating the interaction between the importin Kap114 and H2A-H2B, which allows the dimers to move from the cytoplasm to the nucleus. Then, alongside ACF (ATP-dependent Chromatin remodeling Factor), they help the deposition of H3-H4 and H2A-H2B onto the DNA for the nucleosome formation. Finally, Nap1 disrupts non-productive histones-DNA interactions [197]. Nap1 was also identified as a chaperone of linker histone B4 in *Xenopus laevis* [198].

The crystal structure of yeast Nap1 shows a α-helix fold that is mainly responsible for dimerization and a β-sheet, which is similar to other histone chaperone proteins. yNap1 exists as stable dimers and self-associated oligomers in solution. No crystal data on higher eukaryotic Nap1 have been resolved yet [199].

In addition to nucleosome assembly, Nap1 is implicated in transcriptional regulation and cell cycle regulation. In *S. cerevisiae*, the expression of 10% of all genes is affected by the depletion of *nap1* gene, which indicates transcriptional regulation. yNap1 was found to interact with B type cyclin (clb2), kinase Gin4 and NBP (Nap1 binding protein), which suggests that Nap1 participates in the control of mitotic events [200].

2.11 Chromatin remodelers

Chromatin remodeling is an enzyme-assisted and ATP-dependent histone or nucleosome mobilization, which influences local chromatin structure to facilitate or prevent protein accessibility, which is required to initiate DNA-templated reactions. These enzymes are called chromatin remodelers and they play an important role in maintaining the promoters either in permissive state or in non-permissive state [74]. Accordingly, remodelers have been shown to modulate transcription, replication and DNA repair. There are four families of chromatin remodelers, SWI/SNF (switching defective/sucrose non-fermenting), ISWI (imitation Switch), CHD (Chromodomain Helicase DNA binding) and INO80 (Inistol requiring 80). All four share some properties

such as ATP hydrolysis but at the same time possess some unique domains in their catalytic ATPases and a unique set of associated proteins [201].

The SWI/SNF remodelers can slide and eject nucleosomes and their functions are correlated with nucleosome disorganization and promoter activation [201]. They have domains that bind acetylated tails promoting their targeting or activity in promoters undergoing activation [202]. The ISWI family remodelers carry out nucleosome recognition which often promotes repression [203]. They generally remodel nucleosomes that lack acetylation at H4K16, confining their activity at transcriptionally inactive regions [204]. The CHD family of remodelers utilizes a number of recruitment mechanisms that include binding to sequence-specific transcription factors, histone marks, methylated DNA and poly (ADP-ribose) [205]. These remodelers are implicated in transcription activation as well as repression [206]. Finally the INO80 family has been implicated in transcription and DNA repair. It has a 'split' ATPase domain with a long insertion present in the middle of the ATPase domain.

The ATP-dependent remodeling complexes have been shown to have the ability to alter and rearrange the nucleosomes in a way that increases their accessibility. However the mechanism of how ATP hydrolysis is coupled to disruption of histone-DNA contacts is still debatable.

Chapter 3: The chromatin structure

3.1 The structure of the 30 nm fiber

The so-called "30 nm fiber" is the second state of chromatin compaction after the extended "string-on-beads" conformation (Figure 2). Structural information is essential to understand the molecular mechanisms involved in biochemical processes. For the past 30 years, elucidation of the chromatin and chromosome structure has relied on many approaches such as biochemical, physico-chemical, high-resolution structural (X-ray crystallography and neutron scattering) and direct imaging techniques (fluorescence, electron and atomic force microscopy). The chromatin subunits were first visualized by microscopy and then detected by enzyme digesting techniques. Chromatin containing H1 were shown to fold into a 250 diameter fiber when subjected to increased NaCl molarity [103].

3.1.1 Chromatin fiber models

Based on the experimental evidence of the past 30 years, several models for the folding of nucleosomes into the chromatin fiber have been proposed. They can be divided into two main classes based on the number of helical starts, the one-start helix and the two-start helix. The solenoid model, the best example for one-start helix, was first proposed by Finch and Klug in 1976 [23] whereas the two-start model, resulting in a zigzag arrangement, was first proposed by Worcel in 1981 [207] and improved by Woodcock and colleagues in 1984 [22]. Twisting and coiling of the nucleosome stacks can produce different forms of zigzag models.

3.1.1.1 The solenoid model

The solenoid model is a single start helix with a pitch of 11 nm and a mass of 6 nucleosomes per 11 nm resulting in roughly 6 nucleosomes/turn. The nucleosomes have their faces tilted around 20° from the fiber axis. Although the DNA linker length might change, the diameter of the helix stays 30 nm. The nucleosomes are arranged with the dyad pointing into the center of the fiber, which means both the linker DNA and linker histone H1 are also located at the center (Figure 16A) [24]

The super coiled linker model is another one start helix where the linker follows a coiled path similar to the DNA wrapped around the octamer. In this case, the dyad position is dependent on the actual linker length (Figure 16 B) [22].

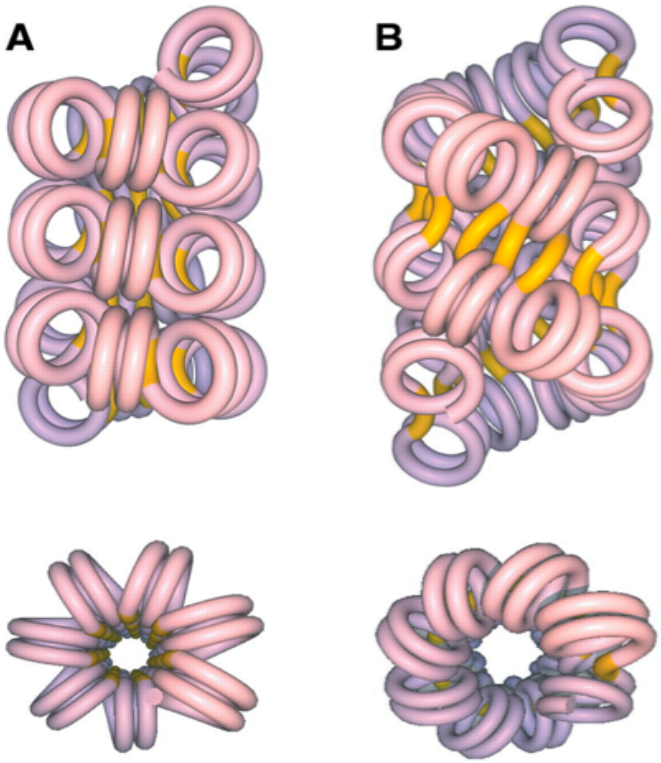


Figure 16: Models for the solenoid 30 nm chromatin fiber. (A) The solenoid model presented by Finch and Klug. (B) The coiled linker model, continuously bent linker model. Figure adapted from [208]

3.1.1.2 The zigzag model.

The most detailed zigzag model is the cross-linker structure of condensed chromatin. It forms a two-start left-handed helix and it is derived from the appearance of a zigzag form chromatin at low salt concentration and condensing into two strands nucleosomes. The diameter as well as the mass per unit is dependent on the linker length, for example the minimum pitch of one nucleosome is 25 nm, which corresponds to 12.5 nm spacing between nucleosomes. Until now, different models of crossed linker models have been suggested based on the number of helical gyres n from the helical repeat (e.g n = 1, 2, 3 or 5) (Figure 17A) [209].

The twisted ribbon model is another zigzag-based model, which suggests that the fiber is formed by coiling of a ribbon-like structure into a hollow cylinder. In this model, the helical pitch increases and the mass per unit decreases with increasing linker length. However the fiber diameter is independent of the linker length and stays constant. Since the center is hollow, the linker DNA and the histone H1 occupy the same perimeter as the nucleosome cores (Figure 17B) [22].

More recent observations of fibers by electron tomography in situ reconstitutions as well as atomic force microscopy (AFM) describe an irregular chromatin fiber. It is based on a loose array of nucleosomes with an underlying zigzag arrangement. This model does not define symmetry or a specific geometry compared to the other models. The nucleosomes are not in contact with each other and the structure is completely dependent on the rigidity of the linker DNA. It is obvious that H1 plays an important role in determining the linker DNA path (Figure 17C) [210].

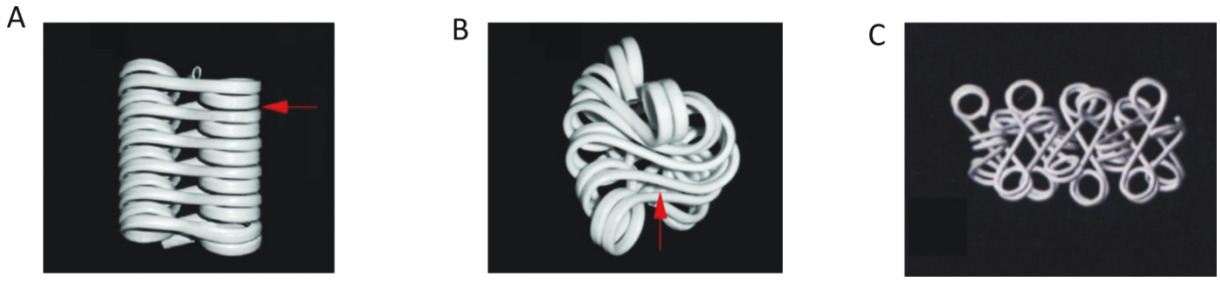


Figure 17: Models for Zigzag 30nm chromatin fiber. (A) A zigzag chromatin fiber with crossed linker DNA. (B) A zigzag chromatin model with twisted linker DNA (C) Irregular zigzag chromatin fiber. Figure adapted from [26]

3.2 Zigzag vs solenoids

Throughout the years, many techniques were used to solve the 30 nm chromatin structure. But due to the experimental limitation of these techniques, no common model has been agreed on. First images of the chromatin were obtained after isolating interphase nuclei from rat thymus, rat liver and chicken erythrocytes. The chromatin fibers exhibited spherical particles called v bodies that measured around 60 to 80 in diameter and were connected by a thin filament (Figure 18A) [15]. The first structure to be proposed was a solenoid containing 6 nucleosomes per turn and a small pitch angle suggesting a single-start structure. This suggestion was based on purified chromatin in presence of divalent cation Mg^{2+} or Histone H1 (Figure18B) [23]. These results were contested shortly after by the suggestion of a zigzag helical ribbon model, where two nucleosomes were connected by a relaxed spacer DNA (Figure18C) [207]. Later on, the form of the linker DNA within the established zigzag model was also subject of controversy. EM images and X-ray scattering data of purified chromatin of different species with different linker DNA length supported the two crossed-linker models with left handed two-start helix (Figure 18D) [209]. The limits of the technologies used were

the main reason for this multitude of models, and the debate on 30 nm structures was becoming more and more important.

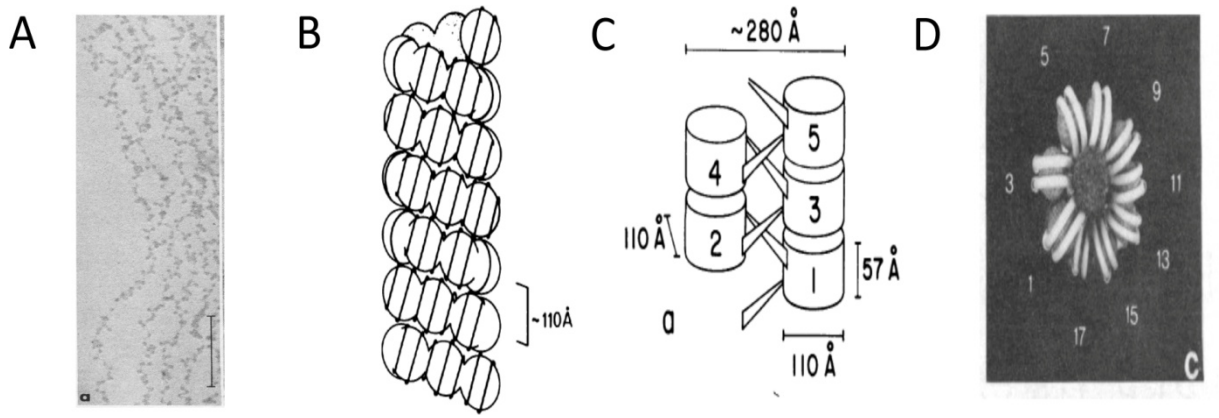


Figure 18: The 30 nm fiber structure 1. (A) Rat thymus chromatin, positively stained. (B) schematic diagram showing the folding of the nucleofilament into a solenoid. (C) Helical ribbon of stacked nucleosomes, nucleosomes are drawn as flat cylinders to show the symmetrical inter-nucleosome contacts. (D) A view of the twisted-ribbon model, the model has a helical repeat of 18 nucleosomes, pitch of 32 nm, dameter of 30 nm and a central hole of 8,5 nm diameter. Figures adapted from [15], [23], [207] and [209] respectively.

3.2.1 *In silico* chromatin models

3D modeling of the chromatin was first introduced in the 80s where the models suggested were compatible with EM observation of isolated fiber and measurement of the linking DNA [207]. The first geometrical model suggested that the geometry of native chromatin fiber extracted from nuclei can be described using two angles. The first angle (α) is the angle between the entry and exit linker of a single nucleosome and the second angle (β) is the angle between consecutive nucleosomes. These two angles are affected by the variable DNA length, which leads to an irregular fiber (Figure 19A) [211]. This theoretical model was used for further investigation of the chromatin fiber in order to achieve a better understanding of the folding, however, *in vivo* experiments showed no evidence that the DNA linker is straight. More *in silico* models were suggested throughout the years using sophisticated 3D analysis such as Monte-Carlo simulation or the Discrete Surface Optimization (DiSCO). These models were suited to fit a fiber compaction into a 30 nm fiber but none allowed the fiber to reach compaction as the one seen for metaphasic chromosomes. A simple explanation to these results might be that not all the known parameters of chromatin fiber were taken into consideration. For example, the integration of the electrostatic charges of a nucleosome to DiSCO analysis proved to be more useful for chromatin dynamics than chromatin structure (Figure 19B). The most reasonable explanation is that the first DiSCO model did not consider the

effect that H1 might have on linker DNA [212]. The elucidation of the histone tail contribution to the chromatin structure and their involvement in nucleosome-nucleosome interactions was a starting point to new features to account in chromatin models. Integrating these results in a new mesoscopic theoretical model, suggested that the chromatin folds into an irregular zigzag with a dominant interaction between nucleosome N and N+4 [213].

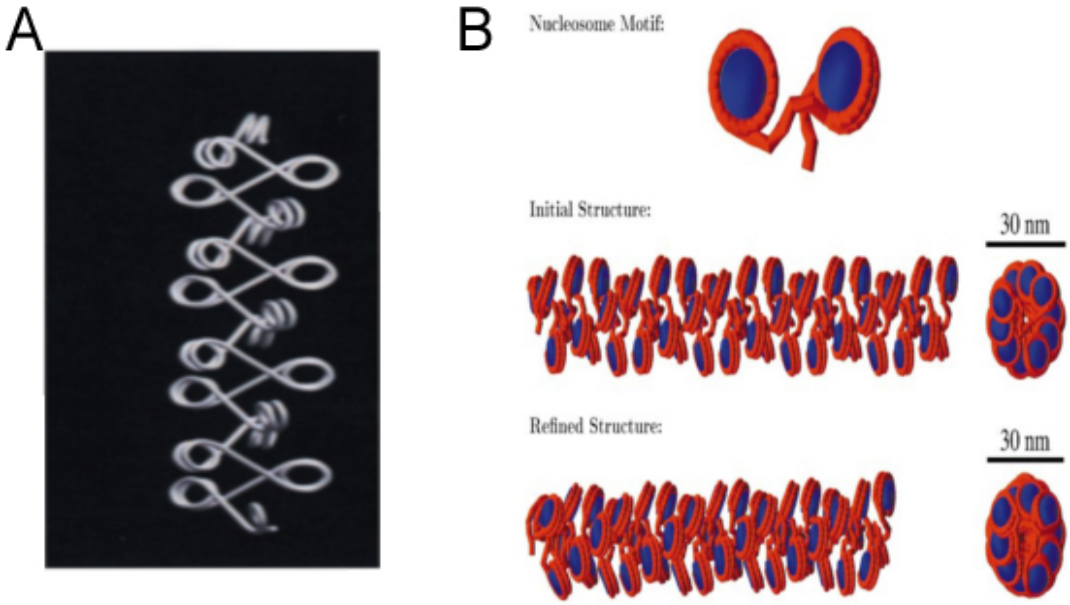


Figure 19: The 30 nm fiber structure 2. (A) An irregular symmetric chromatin generated with fixed linker length (46bp). (B) Upper panel shows the nucleosome-folding motif obtained for a condensed dinucleosome. The middle panel shows the 48 nucleosome system constructed from repeating the nucleosome motif and lower panel shows the same chromatin with integration of electrostatic charges. Figures adapted from [211] and [212] respectively.

3.2.2 Electron microscopy and X-ray crystallography

It is important to note that at the beginning most of the studies were done on purified chromatin or DNA from nuclei but the production of the "601 sequence" by Jonathan Widom group allowed much more extended experiments for determining the 30 nm fiber structure [72]. A SELEX experiment was carried out to isolate and characterize the highest nucleosome affinity members of a large set of DNA fragments. The affinity of $5x10^2$ different chemically synthetic random DNA molecules of 220 bp was tested for the histone octamer. The highest affinity 147bp sequence selected out of the random DNA molecules showed by far stronger nucleosome positioning properties than natural sequences already known. Since then, the 601 sequences have been extensively used because of their strongest nucleosome positioning properties, enabling reconstitution of a single nucleosome over a long DNA stretch containing a single copy of this high-affinity

sequence. In particular, this property makes the 601 sequence a unique tool to reconstitute long arrays of strongly positioned nucleosomal repeats with variable linker lengths [72].

Studies using the 601 repeats have provided structural insights for oligonucleosomes with different linker lengths in presence of divalent cations or H1. However it is important to note than even in presence of such elaborated nucleosomal templates the controversy on the organization of the 30 nm fiber is still continuing.

A very compelling experiment was performed on compacted nucleosome arrays stabilized by disulfide cross-links. Indeed, the known interaction between the H4 tail and the H2A acidic patch was used to create disulfide cross-links and stabilize the interactions between the neighboring nucleosomes. Single mutations were performed introducing a cysteine residue in the acidic patches of either H2A or H2B, and in the histone H4 tail. The combined mutations H4V21C and H2AE64C yielded the most cross-linkable species. Divalent cation Mg^{2+} or H1 were used to induce compaction, and after histone-histone cross-linking between nucleosomes, the arrays were digested to a mono-nucleosome state and subjected to EM analysis. The images clearly showed two parallel stacks of nucleosomes, which is in accordance with the two start-model (Figure 20A) [25]. These combined biochemical and microscopy analyses, were confirmed by the X-ray structure of a tetranucleosome at 9 resolution. The tetranucleosome used was made up of four copies of 147 bp of the 601 sequence with 20 bp DNA linker length and *Xenopus laevis* histone octamer. The results are consistent with a with two-start (zigzag) helix organization rather than with a one-start (solenoid) helix (Figure 20B). A 12 nucleosome array compacted model was built by stacking tetranucleosomes one on another [214]. However, the weak resolution and the lack of linker histone H1 (chromatin compaction was induced by Mg^{2+}) make these X-ray results questionable. Another cryo-EM study using the '601 sequence' nucleosome array in the presence of H1 suggested a compact interdigitated solenoid structure of the 30 nm fiber. Fitting models of the EM images of a compacted chromatin fiber of 33 nm-diameter presented a left-handed one-start helix with interdigitation of nucleosomes from successive gyres leading to a high packing ratio (Figure 20C) [215]. However, the electron microscopy resolution was the limiting point in this experiment coupled with computational fitting where the parameter differs from one program to another.

More recently, a 12 3D cryo-EM images of reconstituted and H1-folded 12 tandems repeats of '601 sequence' (12x177 bp and 12x187 bp) crosslinked with formaldehyde were compatible with a two-start model. These mathematically derived mass-density cryo-EM maps showed a two-start helix twisted by tetranucleosomal structural units that was not affected by the 10 bp increase of NRL. The 12–mer fiber can be separated into three tetranucleosomal units, with unit 1 being for nucleosome N1-N4, unit 2 for N5-N8 and unit 3 for N9-N12. Within a tetranucleosome unit, nucleosome cores are connected with a straight linker and stacked head-to-head, whereas the linker DNA between two tetranucleosome units is twisted (Figure20D) [216]. It should be noted that these observations were made by docking the crystal structure of the Richmond's 4x167 bp tetranucleosome without the linker histone H1 [214] into cryo-EM density maps obtained with 177 and 187 bp [216].

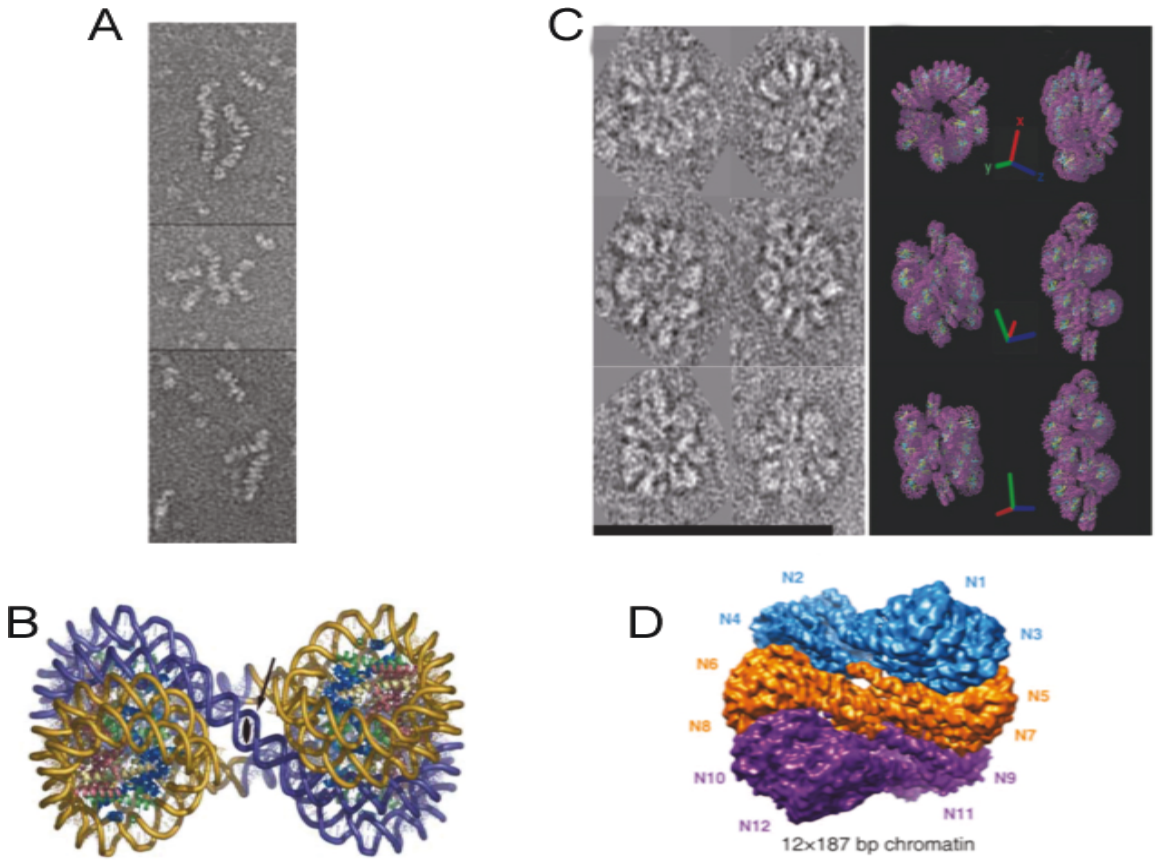


Figure 20: The 30 nm fiber structure 3. (A) Electron micrographs showing the two-start organization of the nucleosome arrays. A 48mer nucleosome array crosslinked and cleaved in the linker DNA shows (three separate examples) two stacked column suggesting zigzag arrangement. (B) A view down of the crystallographic images of the tetranucleosome with a bent linker DNA segment. (C) Comparison of models to raw EM images of a folded 22-mer array of 177 bp. (D) 3D cryo-EM map of the 30-nm chromatin fibers reconstituted on 12x187bp DNA with the three tetranucleosomal structural units highlighted by different colors. Figures adapted from [25], [214], [215] and [216] respectively.

3.2.3 Atomic force microscopy

Atomic force microscopy is a technique that can directly image single molecules in solution. It provides a powerful tool for obtaining insights into the basics of biological materials. In both, objects are bound (immobilized) on a surface. In EM, analysis is made by electron beam (microscopy), while in AFM analysis is made by a vibrating nano-tip. EM has in principle a higher resolution, while AFM is more flexible and easier to use.

The first AFM images coupled with geometrical modeling showed the chromatin fiber as having an irregular organization without adding salt or H1 but once salt was added, the chromatin tended to organize in zigzag form (Figure 21) [217]. The irregular organization of the chromatin beads-on-chain was challenged when EM studies coupled with computer modeling suggested that at low salt the chromatin is organized into a zigzag model with loose entry/exit angles. This angle decreased once an increasing concentration of salt was added. In summary, the 30 nm fiber zigzag organization is predisposed by the organization of the unfolded chromatin [104].

The accessibility of sites within folded and unfolded chromatin was studied by combining biochemical analysis to AFM. The data revealed that folded nucleosome arrays display a less important accessibility to the sites, but the array remains intrinsically dynamic allowing for sites to be slightly accessible [218].

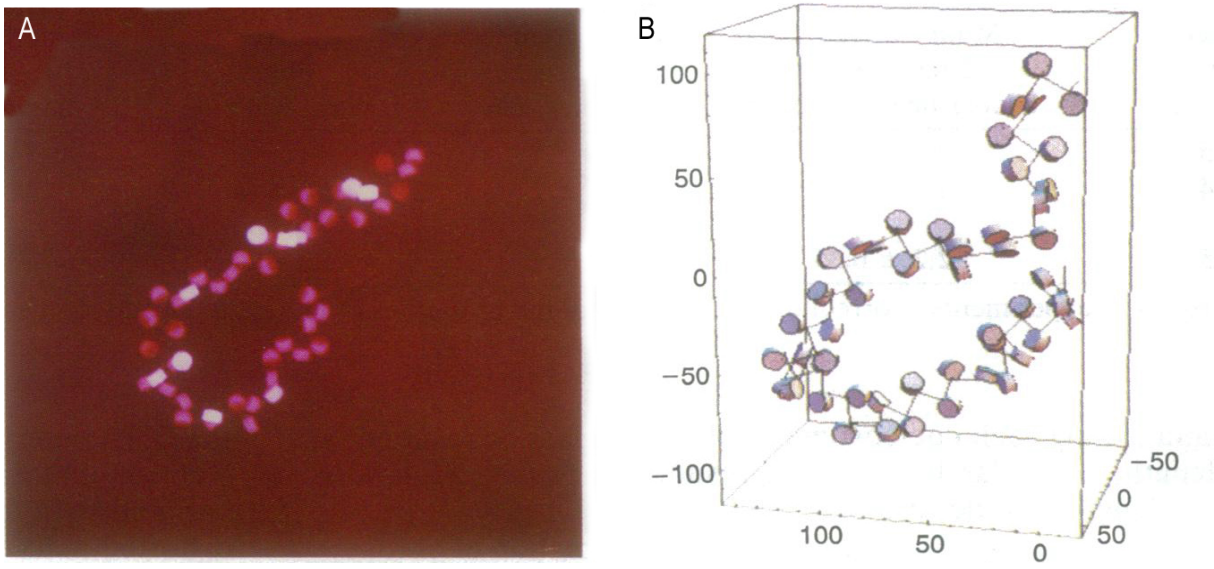


Figure 21: The 30 nm fiber structure 4. (A) Simulated AFM images of chromatin after partial flattening with random linker length between 60 and 64 bp. (B) Model used for the chromatin simulation in A. Figure adapted from [217]

3.2.4 Optical tweezers

For a further understanding of the chromatin structure in presence or absence of salt, the inter-nucleosome interactions were measured using optical tweezers followed by computer modeling. Optical tweezers is a single-molecule technique that can exert forces up to 100 pN on particles ranging from nanometers to micrometers. Focusing laser creates the optical trap; in this case the laser beam traps a polystyrene bead with one chromatin fiber stuck to it. Stretching was applied with a glass micropipette at different salt concentrations. The stretching forces should act on the entry/exit angle of the DNA, which is fixed by the interactions with the linker histone H1. The stretch and release experiment revealed that at physiological ionic strength the fibers present a dynamic structure that interconvert between open and closed states suggesting an irregular zigzag (Figure 22A) [219].

Single-molecule force spectroscopy (SMFS) is a unique new way to study the interactions involved in chromatin folding. This technique does not expose the chromatin fiber to staining, drying, extreme buffer conditions or surfaces. Using optical tweezers, this method provides a quantitative measure of the force needed for stretching the chromatin and predicting the inter-nucleosome interactions that modulate it structure. Surprisingly, it was reported that the linker histone H1 does not affect the stiffness of the fiber but only stabilizes it. Interestingly, these results are in accordance with a solenoid, one-start helix model (Figure 22B) [220].

Whatever the real structure of the chromatin within the cell nuclei id, the *in vitro* structure of the 30 nm fiber remains elusive and controversial, mainly due to the lack of direct data because of insufficient resolution requiring extensive mathematical treatment, modeling and fitting.

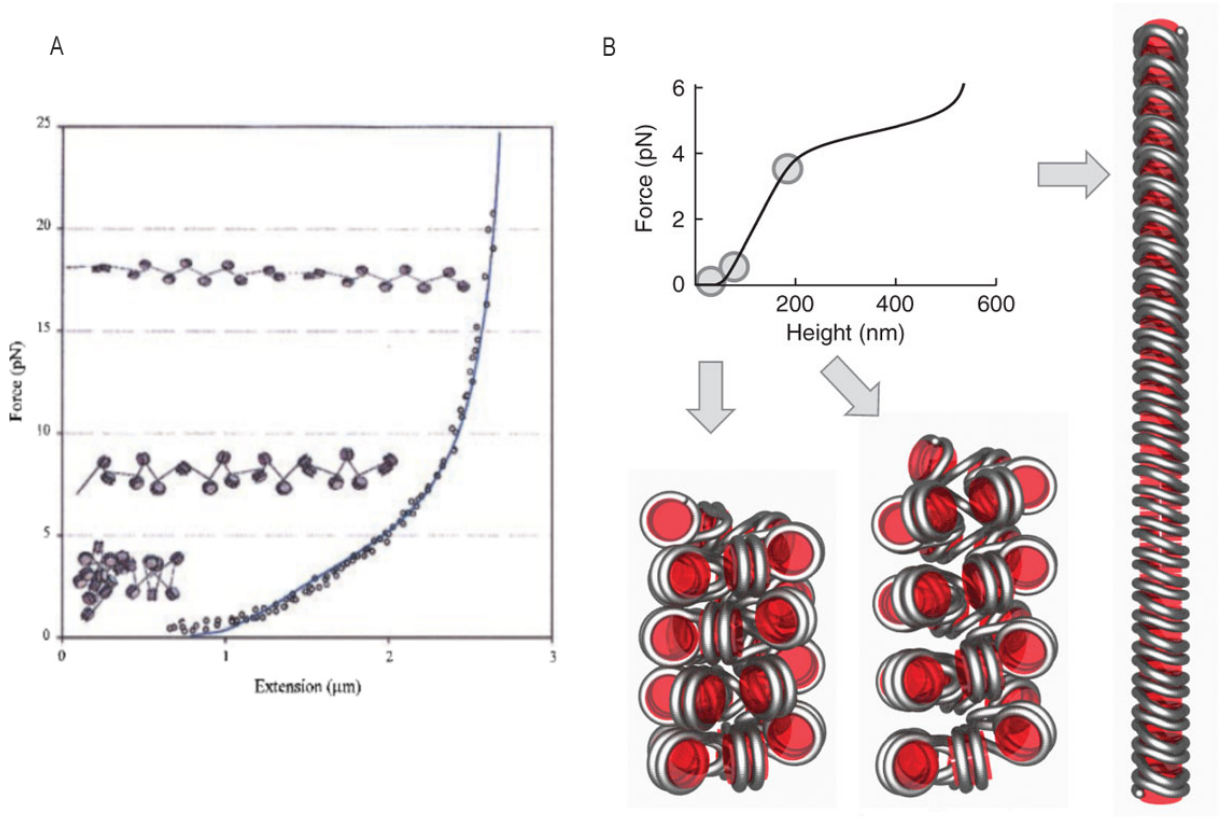


Figure 22: The 30 nm fiber structure 5. (A) Fit of the release of the force-extension curve of a chromatin fiber at low ionic strength by using an extensible worm-like chain model. Schematic drawing of the chromatin qualitatively represents the continuous deformation of the chromatin fiber as it is subjected to increasing tension at low ionic strengths. (B) Schematic representations of chromatin fiber subjected to different forces. Figure adapted from [219] and [220], respectively.

3.3 The nucleosome repeat length effect on the chromatin structure

Chromatin fibers of various species and tissues are characterized by different Nucleosome Repeat Length (NRL) of linker DNA. For example, single cellular organisms have short NRLs (160-189 bp), whereas mature cells usually have longer NRLs (190-220bp). Recent studies showed a strong positively-correlated linear relationship between the number of H1/nucleosome and the NRL, with increasing the NRLs showing a higher H1/nucleosome ratio [221]. Furthermore, EM imaging combined with sedimentation coefficient measurement suggests that the folding of the chromatin is both NRL- and linker histone-dependent. Indeed, it has been shown that only 197 bp NRLs can fold into 30 nm fiber whereas 167 bp have a limited compaction resulting in a thinner fiber (diameter of 21 nm only) with a two-start helix (zigzag) organization [222]. In addition, it was suggested that the medium-sized NRLs (177-207 bp) have a constant diameter of 33 nm and that longer linkers (217-237 bp) form a 44 nm-

diameter fiber, while both fold into interdigitated one-start helices (solenoid) [215]. These same EM data were analyzed independently by another group while taking in consideration the linker histone H1. The analysis showed that NRLs ranging from 155 to 211 bp reveal a compacted fiber with possible one-, two- and three- start structures [223].

Computational modeling studies using an already established mesoscale model suggest that short to medium NRLs (173 to 209 bp) with linker histone condense into irregular zigzag structures. However, for longer NRLs, the chromatin folds into solenoid structures [224]. Models taking into consideration the basics physics of chromatin such as electrostatics, DNA and nucleosome mechanics, histone tail flexibility, hydrogen bonding fluctuations, showed that the linker histone H1 has very little effect on the short NRLs but an important one on medium NRLs [224]. Moreover, Monte Carlo simulations of oligonucleosome model showed that non-uniform NRL fiber containing one short linker forms a bent ladder rather than a compact 30 nm fiber. But for chromatin fibers with medium and long NRLs, the preferred organization would be a regular zigzag conformation that has been described in uniform fibers. However, long NRLs are described to compact into irregular fiber, with linker DNA having a bent or interdigitated conformation. These findings led to the conclusion that the polymorphism of the chromatin structure is triggered by the NRL variations.

3.4 The nucleosome acidic patch effect on higher-order structure

The nucleosome acidic patch is negatively charged region formed by a cluster of eight acidic residues: Glu56, Glu61, Glu64, Glu90, Glu91, Glu92 of H2A and Glu102, Glu110 of H2B. The crystal structure of the nucleosome first presented this cluster as a surface with uneven charge distribution [65]. This cluster forms a curved and asymmetrically charged surface called the acidic patch. As shown in Figure 18, the acidic patch has the form of a narrow pocket where the bottom is formed by the α2-helix and the C-terminal extension of H2A and the edges are formed by the α1- and αC-helices of H2B (Figure 23). Moreover, the presence of the H2A residues Y50, V54 and Y57 at the bottom of the pocket makes this pocket hydrophobic due to the non-polar character of these residues. The mechanistic details of the H4 tail-acidic patch interaction remain unresolved but proven to be necessary to the chromatin folding. However, molecular dynamics simulation of the H4 tail acidic patch showed that residues V21 of H4 and E64 of H2A

are close in space when chromatin is folded. This model also suggested that residues R3, K16, R19 and K20 of the histone H4 face the acidic patch groove and form non-covalent interactions with the acidic patch residues [225]. Even though the acidic patch plays a role in chromatin folding, its absence was not shown to inhibit the compaction of the 30 nm fiber. The acidic patch is also found to bind to non-histone proteins, which creates a competition between these proteins and the H4 tail domain, thereby triggering remodeling of the higher-order structure of the chromatin [226].

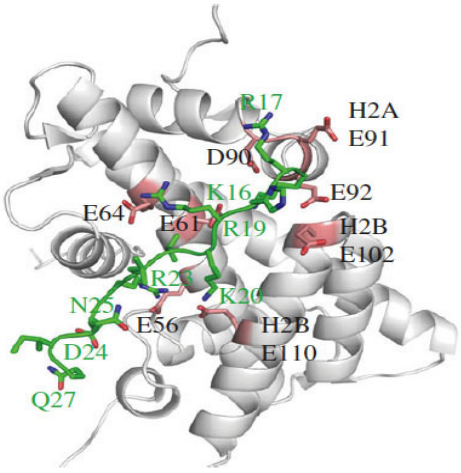


Figure 23: Close-up View of the H4 tail domain–acidic patch interactions observed in crystal. The H4 tail is shown in lime green. Histones H2A are in grey. Acidic patch residues are in shaded pink. Figure adapted from [226]

3.5 Role of histone variants in chromatin structure

The biological roles of histone variants have been largely discussed and many questions still need to be answered. The structural organization of chromatin containing histone variants is one of these remaining questions. X-ray structures, cryo-EM and AFM images of nucleosomes containing histone variants such as H2A.Z, H2A.Bbd or CENP-A helped to shed some light on this issue. However, much less is known about the 30 nm fiber containing these histone variants.

3.5.1 Role of H2A histone variant H2A.Z

3.5.1.1 Intra-nucleosomal interactions

The H2A.Z nucleosome crystal structure [164] showed an extended acidic patch compared to the H2A nucleosome, and Asn and Lys residues were replaced by Asp and

Ser respectively. The main question remains about the involvement of this extended acidic patch in the folding of the chromatin.

In order to provide an answer a mutated H2A containing the additional acidic patch residues was expressed and purified. Sedimentation studies compared H2A.Z and the mutated H2A to H2A arrays and the results showed that the H2A.Z and mutated H2A arrays showed similar folding patterns and were considered more compact compared to H2A arrays under the same divalent cation concentration, confirming that due to their extended acidic patch, H2A.Z arrays promote the formation of the 30 nm. [227, 228]

3.5.1.2 Inter-nucleosomal interactions

Increasing the concentrationsof divalent cations above the intramolecular folding range induces a cooperative and reversible oligomerization. During this oligomerization, the arrays self-associate into large nucleoprotein complexes resembling a chromosomal fiber during interphase. In order to study the H2A.Z implication in this self-association, sedimentation experiments were carried out on H2A.Z-containing nucleosome arrays. The results showed that with 2 mM $MgCl_2$ the H2A.Z fiber presented a much slower sedimentation than the H2A array [228]. Moreover, microfuge-based assay provides further evidence that H2A.Z inhibits oligomerization and requires a much greater amount of divalent cation in order to achieve it [227]. Briefly, H2A.Z array seem to promote the 30 nm fiber formation but at the same time inhibits the oligomerization of this array.

3.5.2 Role of H2A histone variants: H2A.Bbd

H2A.Bbd has different characteristics from H2A, for example the absence of a C-terminal tail, no lysine residues in its N-terminal tail and the lack of residues that contribute to the acidic patch in the nucleosome core particle in mammals [178]. The H2A.Bbd nucleosome was reported to have a more relaxed conformation than the H2A nucleosome [178]. Furthermore, the H2A.Bbd nucleosome displays lower stability *in vitro* and the dimer H2A.Bbd-H2B is rapidly exchanged *in vivo* [229] [230]. All these indications of an unfolded chromatin, together with the H2A.Bbd enrichment in euchromatin regions of the genome strongly suggests that it might play a role in up-regulation of chromatin transcription [177, 179].

3.5.3 Role of H3 histone variant: CENP-A

For the past 20 years, the structural and mechanical features of the CENP-A nucleosomes have been widely studied. Even though a recent crystal structure of the CENP-A-containing nucleosome helped uncover some of its important features, the controversy about CENP-A chromatin structure and dynamics is still lingering.

The variant histone CENP-A was first described to be present in nucleosome-like structures following micrococcal nuclease digestions and that it co-purified with nucleosome particles [187]. CENP-A has been a marker for centromere location because of its dynamic nucleosomes, its histone stoichiometry and the nature of the wrapping of the DNA, from left to right (while in canonical nucleosome it is from right to left). Even though the central role of CENP-A arrays has been identified, the inter-nucleosomal interactions occurring are still completely unexplored.

The histone stoichiometry within the CENP-A-containing nucleosome remains the subject of many speculations. AFM studies on an *ex vivo* CENP-A nucleosome arrays showed a size reduced by half relative to canonical H3 nucleosomes. These results led to believe that the CENP-A nucleosomes are tetrameric (containing one copy of each histone). Moreover, EM, like AFM, revealed a small particle separated by long linkers that resisted ionic condensation [231]. However, these results were controversial, because mammalian centromeric nucleosomes are described as octameric, whereas budding yeast nucleosome can go through two forms, i.e. octameric or hexameric complexes. These different models are due to the various eukaryotic species used to isolate the CENP-A nucleosomes.

The crystal structure of the CENP-A nucleosome gave a larger idea of the intra-nucleosome interaction, describing the CENP-A nucleosome to have loose terminal DNA contacts, resulting in DNA wrapping [190]. The loose terminal DNA contact was also confirmed by micrococcal nuclease assay, where CENP-A nucleosome was shown to cover 20 bp less than the canonical nucleosome.

To further illustrate the CENP-A nucleosome controversy, it was recently published by *Miell et al.* [232] that the CENP-A nucleosome does indeed have a reduction of height in octameric nucleosomes. CENP-A and H3 were purified from two distant organisms, human and *S. pombe*, reconstituted into an octameric nucleosome and crosslinked before conducting AFM experiments [232]. These results were indeed surprising, because they contradicted other AFM studies done on octameric CENP-A nucleosomes

as well as the CENP-A nucleosome structure published a couple of years earlier. To solve this matter, another team repeated the same AFM experiment with human CENP-A, yeast CENP-A and canonical H3 that they obtained from four independent laboratories including the one used by Miell *et al*. As expected, no changes in height or diameter were noticed with both human and yeast CENP-A compared to H3. However, Miell et al. replied by collecting all AFM data that have been published on CENP-A/H3 difference of height (Figure 24). They stated that inconsistencies for AFM measurements collected from different groups are not unusual and might be due to imaging conditions such as the force and frequencies of the AFM tip, the adsorption of salt on the surface, humidity among other factors. The limitations of 2D imaging techniques and the indirect approaches have been the source of most contradictory informations collected about the chromatin structure and its dynamics.

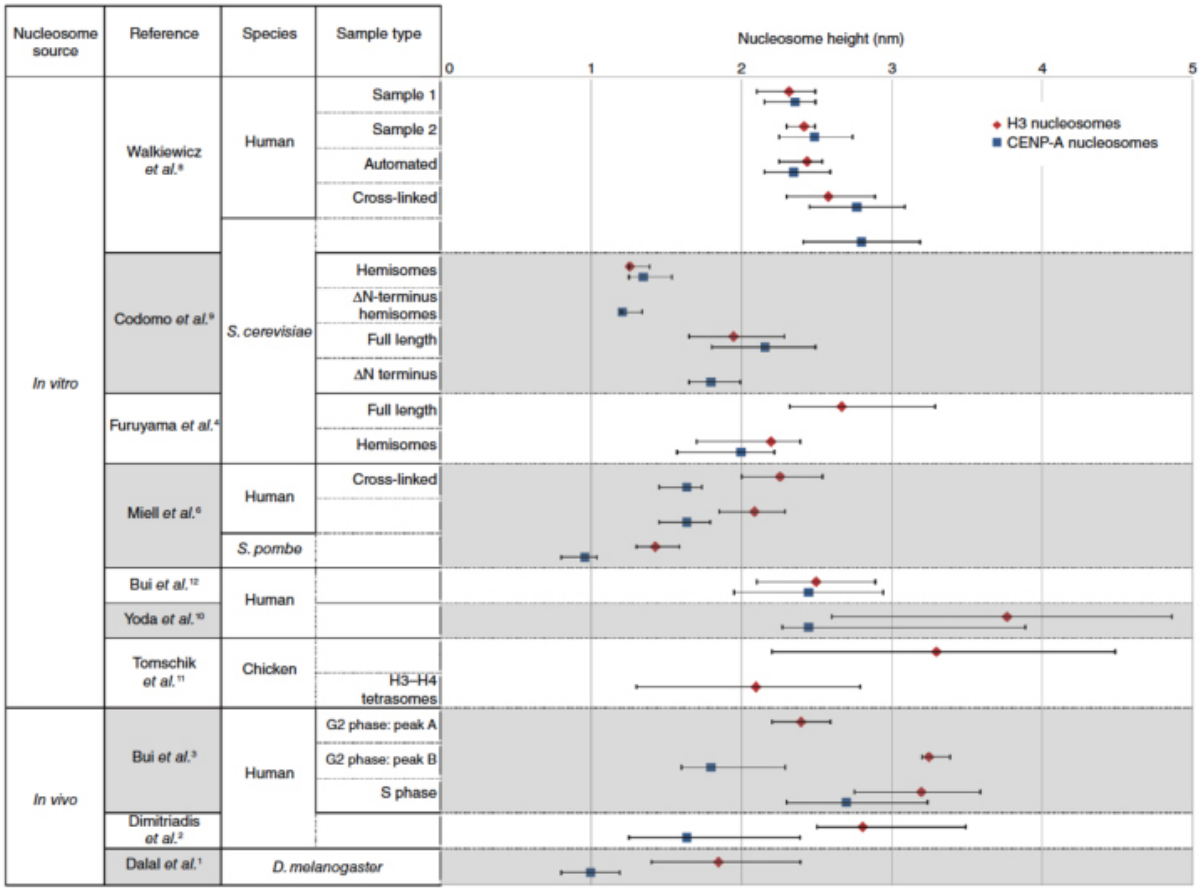


Figure 24: Variability in the distribution of nucleosome heights measured with AFM. Points represent the median height of H3 (red diamonds) and CENP-A (blue squares) nucleosomes measured by AFM from several publications. Figure adapted from [233].

3.6 Beyond the 30 nm fiber

3.6.1 Does the 30 nm exist *in vivo*?

The controversy regarding the structure of 30 nm fibers has been substantiated by few reports because of inability to observe the 30 nm fiber in *in vivo* cryo-EM vitreous sections visualization (CEMOVIS). This was first done 20 years ago on mammalian mitotic cells and only an aggregation of 11 nm nucleosome fiber was observed [234]. However, it was argued that the cryo-EM images had been corrupted by the CTF (contrast transfer function), but once CTF were corrected, the spectrum still showed no evidence of 30 nm peaks. Furthermore, 3D structure of *Xenopus* chromosomes assembled *in vitro* using electron tomography detected no fiber-like structures [235]. It is important to note that the original concept of the 30 nm fiber started with conventional transmission or scanning EM showing fibers of 30 nm diameters in isolated chromatin, nuclei and mitotic chromosomes [236]. However, if the 30 nm fiber can be visualized by EM only, this might be explained by environment changes triggered by sample preparation. In addition, it has been noted that nuclei of specific cell types such as starfish spermatozoids of chicken erythrocytes are almost full of 30 nm fiber. These observations have led to the hypothesis that in vivo the chromatin does not fold into a 30 nm fiber. Therefore, much more experiments are needed to exclude artifacts and to eventually confirm these hypotheses.

3.6.2 Tertiary structure

As for proteins the concepts of primary, secondary, tertiary and quaternary structures can be applied to chromatin structure hierarchies (Table 4).

Even though no definitive structure has been appointed to the secondary structure of the chromatin, there is no doubt about the higher-order organization, which is obvious in the metaphase chromosome.

Level of chromatin structure	Example of global structure	Example of local structure
Primary- beads on string	Nucleosome repeat length	Preferred locations of nucleosomes and features such as DH sites on a specific DNA sequence.
Secondary- formed by interactions of nucleosomes	The 30 nm chromatin fiber	3D architecture of nucleosomes and regulatory proteins on a specific DNA sequence.
Tertiary- formed by interactions between secondary structures	The thicker fiber seen in nuclei and composed of 30 nm fibers	Long distance contacts possibly involving locus control regions, enhancers and promoters.

Table 4: Level of chromatin structure and their possible global and local interactions.

In order to study the tertiary level of chromatin structure, experiments were performed using optical microscopy, which is limited to a ~300 nm resolution. The images revealed that the tertiary structure in natural chromosome is beaded. The number of adjacent beads can go from one to six beads with diameters going from 0.4 to 0.8 μm. Similar results were obtained with tandem DNA. Moreover, these beaded structures can be adapted to chromonema models and radial-loop/scaffold models (Figure 25). The chromonema models propose a fiber that yields different fiber thicknesses. It starts with a thinner fiber that yields into a thicker fiber that appears as beads under the light of the microscope. This structure is stabilized by fiber-fiber interactions. The thickness of beads depends on the thickness of the underlying fiber. For example, if the underlying fiber measures around 30 nm, the beads would appear to be 500 nm in diameter. In contrast, if the underlying fiber is thicker (around 100 nm fiber), the beads would only be twice as big, around 200 nm. On the other side, it has been described that chromatin fibers form loops attached to a protein scaffold. Under the light of the microscope, the cluster of loops would form a bead with a diameter around 0.3 μm (which is close to the average diameter of the beads) [237].

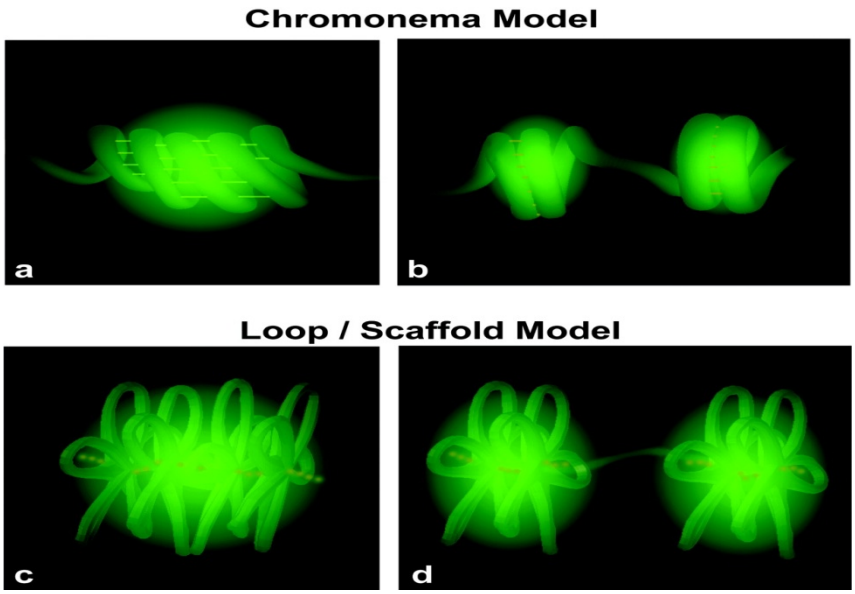


Figure 25 Chromonema and radial-loop/protein scaffold models. (a) Chromonema models propose a hierarchically folded chromatin fiber that yields a series fiber of different thicknesses. (b) Local disruption of fiber-fiber interactions would lead to the unraveling of the fibers. (c) Radial-loop models propose that the chromatin fiber forms loops attached to a protein scaffold. (d) The structure shown in panel c could decondense into two beads by local detachment of a loop from the scaffold. Figure adapted from [237].

3.6.3 The metaphase chromosome

The metaphase chromosome is part of the chromatin higher-order structure, where DNA is compacted 10,000 to 20,000 fold. To this day, there is no accepted model because of the several features that must be accounted for. One of these features for example is the diameter of the chromosome arm, which is different from one species to another. This chromosome has been described as DNA loops surrounding a core structure called scaffold, which is at the base of the chromosome architecture. Studies of some of the component isolated from the scaffolds revealed a family of structural maintenance of chromosomes (SMC) proteins. Also, two proteins complexes called Cohesin and Condensing play a role in sister chromatin adhesion and chromosome condensation [238]. Experiments showing the extension of the chromosome revealed that chromosomes exhibit a remarkable degree of elasticity and repeatedly return to their original state after stretching [239]. In contrast, these results do not apply to hierarchical models of chromosome structure, which predicts that stretching the chromosome would lead to unfolding and reduction of the chromosome diameter. The metaphase chromosomes appear to be more complicated than they look like, hence understanding their composition, formation, maintenance and architecture remains to be achieved.

3.6.4 Chromosome territories

Different functional regions in eukaryotic chromosomes are determined by a large variety of chromatin biochemistry. A complex system of accessory proteins modifies, binds and reorganizes histone complexes. In the last decade, decoding the chromatin "languages" such as DNA methylation [240], histone modifications [134], and chromatin remodeling events for gene regulation [241, 242], has made an impressive progress. But it is obvious that these advances do not suffice to understand the different epigenomes present in various cell types [243]. Indeed, epigenomes and their functional implications depend on differences in higher-order chromatin organization and nuclear architecture. Beyond the fine scale arrangement of chromatin, what is the higher-order structure of chromosomes. By combining multicolor technologies with 3D imaging tools, it has been possible to visualize all 46-chromosome territories in an intact human cell as can be seen in figure 26 [244]. The current view is that chromosomes are compartmentalized and occupy distinct, non-overlapping, sub-nuclear regions named chromosome territories. The location of a gene within the chromosome territory seems to influence its access to DNA template machineries.

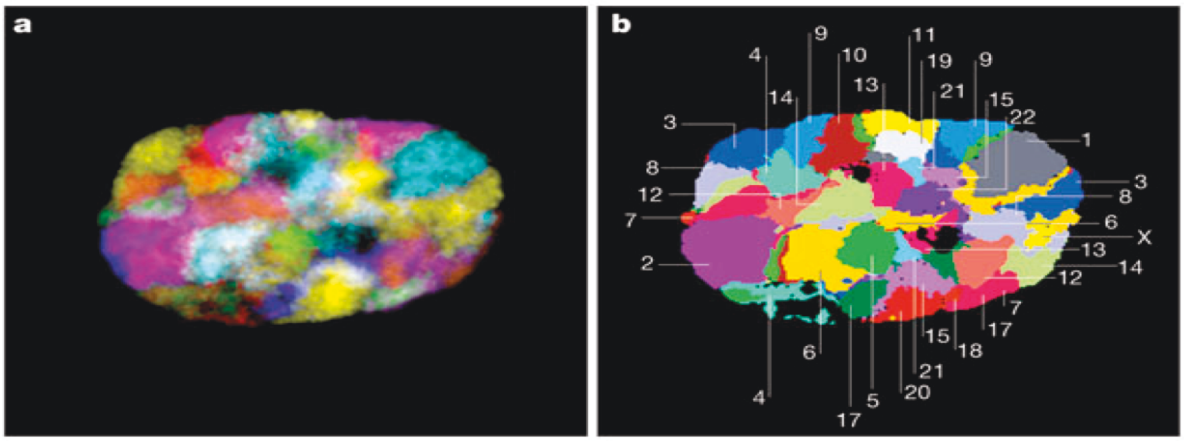


Figure 26 : Chromosome territories. All the chromosome territories that make up the human genome can be visualized simultaneously in intact interphase nuclei, each represented by different color. (a) A red, green and blue image of the 24 labeled chromosomes (1–22, X and Y). (b) Scheme adapted from (a) showing the chromosome territories of all 23 chromosomes. Figure adapted from [245]

In the last decade, the analysis of nuclear organization has known a breakthrough due to the development of chromosome conformation capture (3C) technology and its subsequent genomic variants. The 3C technology is a biochemical strategy that analyzes contact frequencies between selected genomic sites in a cell population [246]. 3C-based methods can put observations made on single genes in selected cells in the context of

genomic behavior in cell populations, thus enabling DNA topology studies at a higher resolution. These techniques rely on the remarkably simple idea that digestion and re-ligation of fixed chromatin in cells, followed by quantification of the ligation junctions, allows the determination of the DNA contact frequencies. In the original study that describing the 3C techniques, the average 3D conformation of yeast chromosome III was determined, showing that it forms a contorted ring [246]. The method was then adapted for mammalian systems and used to prove the existence of chromatin loops *in vivo* between regulatory DNA elements and their target genes [247]. Other 3C-based methods have seen the light since then including 4C (Chromosome Conformation Capture-on-Chip), 5C (Chromosome Conformation Capture Carbon Copy), HiC and ChIA-PET. However, the 3C-derived methods fail to detect cell-to-cell variation and cannot assess the dynamics in the system. Integration of the 3C-based methods with high resolution live-cell imaging will lead to a better understanding of the dynamics underlying chromosomal interactions and their cell-to-cell variation.

II. Objectives

The compaction of genomic DNA into chromatin is involved in the regulation of key biological processes such as transcription, replication and DNA repair. Nucleosomes, which form the repeating units of chromatin, vary in their histone composition.

Linear arrays of nucleosomes are folded and compacted into 3D assemblies of higher-order structures. Secondary chromatin structures are defined as the structures emerging from the folding of an individual array (strings-on-beads) to produce a particular fiber, the 30 nm fiber. Given the fundamental role of the 30 nm fiber structures in regulating all DNA- and chromosome-dependent processes, its precise molecular organization and 3D arrangement have been extensively debated. However, progress has been slow and result interpretation controversial due to the limitation of current microscopy approaches and to the complexity of the question. Initially, structural studies on fibers assembled on natural DNA sequences were hindered by variation of the length of the linker DNA. However, significant advances have come from the construction of regularly spaced tandem DNA repeats for precise nucleosome positioning [72]. Under physiological salt conditions, and or/presence of H1 an array of nucleosomes folds into a secondary chromatin structure with the shape of a 30 nm fiber. Despite decades of effort, the structure of the 30 nm fiber has not been resolved yet due to its highly compacted nature, which prevents the path of the DNA from being visualized by electron microscopy. Two models have been proposed on the basis of *in vitro* data, the solenoid and zigzag arrangements, which were both supported by different studies. Crosslinking studies and the crystallization of a tetranucleosome array have provided strong evidence that the 30 nm fiber adopts a zigzag form. However, the short repeat length and the absence of linker histone H1 made these results questionable [25, 214]. On the contrary, cryo-EM on long regular arrays with constant repeat length and in the presence of linker histone H1 concluded to a multiple-start interdigitated solenoid model [248]. Finally, a recent study using cryo-EM suggested that there is not one uniform type of organization but rather heterogeneity of nucleosome interactions. In another study, the fiber showed a predominantly two-start organization that was interspersed with bent DNA typical of 'one-start' solenoids [26]. Of note, the nucleosome repeat lengths have been reported to affect the chromatin folding as well. Indeed, some results suggest that short to medium NRLs (173-209 bp) favored a zigzag structure, whereas longer NRLs (218 bp and above) favored solenoid structures [224].

Chromatin, at all levels of organization, is not static but very dynamic. This dynamicity and plasticity is crucial to ensure proper functioning of the cell. Modification of chromatin structure is prime step in regulation of all the DNA-templated processes like transcription, replication, repair and recombination. These processes require quick changes in the chromatin organization and structure. The dynamic control of genome accessibility is governed by contributions from DNA sequence, ATP-dependent chromatin remodeling, post-translational modification of histones and histone variants incorporation [249]. However, controversial results were reported on the folding of the chromatin and its organization upon replacement of conventional histones and histones variants.

The first aim of this work was to develop a novel robust in vitro biochemical approach, we termed ICNN (Identification of Closest Neighbor Nucleosomes), which allows the direct and unambiguous determination of the 3D nucleosome arrangement within the 30 nm chromatin fiber. The method is based on direct identification of the closest interacting neighbors of each nucleosome, by using disulfide crosslinking between nucleosomes within a compacted chromatin, and affinity pull-dawn and qPCR.

To clarify the chromatin folding structure, the ICNN approach is applied in 3 complementary work axes:

1) The characterization of the 30 nm fiber organization (zigzag or solenoid) in presence of H1 on 12-mer nucleosome arrays with different NRLs to investigate the contribution of H1 and of the NRL length on the resulting chromatin structure.

2) The identification of the role played by the histone variant H2A.Z (H2A variant) on the chromatin folding structure.

3) The analysis of the CENP-A (H3 variant)-containing chromatin folding behavior in presence of linker histone H1, and more precisely the specific role of the αN domain of CENP-A in the process of folding.

III. Results

Manuscript 1:

The novel ICCN approach reveals a 3D non-consecutive, independent of both H2A.Z and nucleosome repeat length, arrangement of nucleosomes in compact chromatin

The novel ICCN approach reveals a 3D non-consecutive, independent of both H2A.Z and nucleosome repeat length, arrangement of nucleosomes in compact chromatin.

Soueidan,L ; Shukla,M ; Ben Simon,E; Tonchev,O; Dimitrov,S; Angelov,A.

Abstract

The 3D organization of the compact chromatin fiber, remains, despite the numerous efforts, still not well defined and a matter of debates. Here, we report a novel approach for analyzing the 3D arrangement of nucleosomes within the fiber and its application for studying chromatin samples with different repeat length. The approach, termed Identification of Closest Neighbor Nucleosomes (ICNN), is based on the crosslinking of the nucleosomes within the compact chromatin fiber. ICNN measures with very high accuracy the probability of a close contact of a selected nucleosome (N) with other nucleosomes in the fiber. The data show that the 3D organization of the nucleosomes in the fiber is of the type N±2 for all three nucleosomal arrays studied with nucleosome repeat length of 177 bp, 197 bp and 227 bp, respectively. In addition, the chromatin fiber, reconstituted with the histone variant H2A.Z, shows the same N±2 3D nucleosome arrangement. The data are compatible with a two-start helix structure of the chromatin fiber. ICCN has the potential for a broad range of applications in *In vitro* and *In vivo* chromatin studies.

Introduction

Chromatin exhibits repeating structure. The nucleosomes, the fundamental units of chromatin, are connected with linker DNA with different length (varying between ~10 and ~80 bp) and form the 10 nm chromatin filament [1]. Upon binding of the linker histone H1 the chromatin filament is folded into the 30 nm chromatin fiber [2-8]. Although, the structures of the conventional [9] and histone variant nucleosomal core particles [10, 11] were solved at very high resolution, the 3D organization of the chromatin fiber is still poorly understood. The available data suggests that the structure of the fiber, which contains ~6 nucleosomes/11 nm [2, 12, 13] could be described either as one start helix with consecutive arrangement of nucleosomes and bent linker (the solenoid model, [2] and **Figure 1a**) or as two start structure with non-consecutive

arrangement of nucleosomes and straight linker ([6, 14-18] and **Figure1b**). The two-start structure was described, in turn, either as a helical ribbon arrangement of nucleosomes or as two intertwined helical stacks of nucleosomes. Recently a polymorphic structure of the chromatin fiber based on biochemistry experiments and EM imaging was proposed [19]. In addition, the folding of the fiber might depend on the nucleosome repeat length [20].

Each model for a regular 3D arrangement of the nucleosomes within the chromatin fiber supposes well-defined neighboring nucleosomes. For example, in the case of the solenoid model the neighbors of the N nucleosome are N±1 nucleosomes (**Figure 1a**). In contrast, in the two-start structure model the N nucleosome has for neighbors the N±2 nucleosomes (see **Figure 1b**). Therefore, unambiguous determination of the spatial arrangement of the nucleosomes in the compact chromatin fiber will help to discriminate between the different models and will allow shedding light on the chromatin condensation process.

Histone variants are key epigenetic players, which are usually coded by two distinct genes. Each cell is expressing histone variants. The current view is that the incorporation of histone variants confers novel structural and functional properties of the nucleosome [21]. H2A.Z is a universal histone variant present from yeast to men, which is implicated in several vital cell processes, including transcription, repair and mitosis [22-26]. Several studies were focused on both the structure of the H2A.Z nucleosome and the H2A.Z chromatin fiber [27-32]. The H2A.Z chromatin organization was, however, investigated with mainly low resolution physical-chemistry methods and, in addition, the available data show some controversy [29-33].

Description of the novel Identification of Closest Neighbor Nucleosomes (ICNN) approach

In this work we are describing a novel robust approach, termed ICNN (Identification of Closest Neighbor Nucleosomes), which allows the determination of the 3D nucleosome arrangement within the compact fiber with very high precision (**Figure 1c-h**). The method is based on the ability to crosslink nucleosomes containing both single amino-acid cysteine substituted both histone H4 (H4-V21C) and H2A (H2A-E64C) within

compact chromatin [34]. ICNN uses strongly positioned 601 DNA nucleosomal arrays reconstituted with the histone octamer consisting of H4-V21C and H2A-E64C mutated histones and both wild type H3 and H2B. Linker histone H1 is deposited in *In vivo*-like manner by using its chaperone Nap1. Each individual 601 repeat within the array contains 2 sets of 2 distinct mutated nucleotides (each individual set of the 2 mutated nucleotides is in proximity to the respective nucleosomal end, **Figure 1e**). The use of specific primers, which 3'-end is "finishing" at these sets of mutated nucleotides, allows the specific amplification of the repeat of interest, and thus, the "visualization" of the nucleosome of interest. The repeats are separated by a blunt ScaI restriction site. In addition, each individual repeat contains an inserted cleavage site for a distinct restriction enzyme (**Figure 1e**). The presence of these last distinct restriction cleavage sites allows the specific replacement of a selected repeat with the same repeat, but containing biotin, a procedure necessary for the usage of ICNN (see below).

The ICNN approach consists of the following steps (see schematics in **Figure 1c-h**): (i) induction of disulfide crosslinks (H4-H4 and H2A-H4) between the neighboring nucleosomes in the compact arrays in solutions containing 1:1 ratio of reduced and oxidized glutathione as described [34]; (ii) Digestion the chromatin samples to completion with the ScaI restriction enzyme; (iii) Checking the completion of cleavage at the level of DNA, isolated from the cleavage reaction products; [34]. Note that upon complete cleavage only individual 601 repeats DNA and mono-nucleosomes have to be observed; (v) Demonstration of the efficiency of crosslinking; (vi) Native streptavidin chromatin pull down and, (vii) Quantitative PCR of the DNA isolated from the pull down samples.

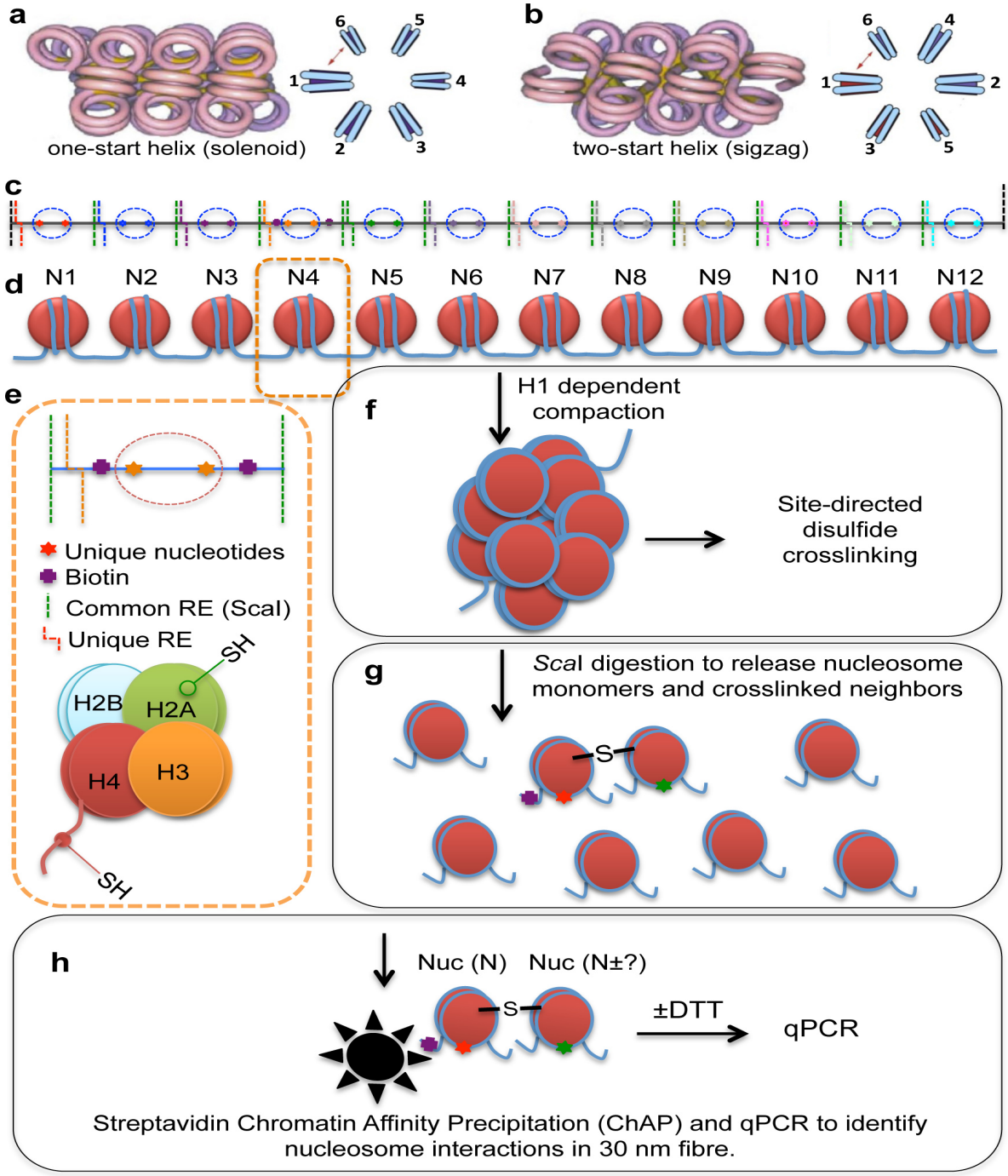


Figure 1: Schematic presentation of the Identification of Closest Neighbor Nucleosomes (ICNN) approach. (a) The one-start helix (solenoid) model with N±1 nucleosome arrangement (b) The N±2 "zigzag" model of two intertwined helical stacks with linkers crossing the fiber (c) Schematics of the 12 repeats of 601 DNA used for nucleosomal array reconstitution; the repeats are separated by a common blunt ScaI site (vertical green discontinuous line) and contains in close vicinity the cleavage site (indicated in orange as I_) for specific for this repeat « overhang » restriction enzyme (see also panel 2 in (d). (d) Linear alignment of the 601 nucleosomal array, containing 12 (from N1 to N12) distinct nucleosomes. (e) Upper part, schematics for an individual nucleosomal 601 repeat used for streptavidin pull down; the end of the repeats are defined by the blunt end (in green) ScaI site and it contains an unique site for an « overhang » end restriction enzyme(orange dotted line); two nucleotides (orange stars) in vicinity of each the nucleosome DNA (indicate as discontinuous oval) are mutated, which allows their unique "visualization" by QPCR, incorporated biotin

(purple cross) ; lower part, the core histone octamer (used for reconstitution and crosslinking) containing the cysteine mutated H2A (H2A-E64C) and H4 (H4-V21)C, respectively. The other nucleosomes, not used for the streptavidin pull down in a given experiment, do not contain inserted biotin. (f-h) Schematical illustration of the different consecutive steps of ICNN.

Analysis of the 3D nucleosome arrangement within compact chromatin with varying repeat lengths

We have initially applied the ICNN approach for analyzing the neighbors of nucleosome #5 within compact chromatin, reconstituted by using 12 tandem 197 bp 601 repeats (see schematics of **Figure 1**). We have first characterized the efficiency of reconstitution (**Figures 2a, b**). In the absence of H1, the reconstituted arrays migrate as an up-shifted band compared to the band of free DNA (**Figure 2a, lanes 2 and 3**). Addition of Nap1 does not affect the migration of the reconstituted chromatin (**Figure 2a, lane 3**), suggesting, as expected, no Nap1 association with chromatin. The addition of the Nap1-H1 complex results, however, in increase of the migration rate of the arrays (**Figure 2, lane 4**), which reflects the Nap1 assisted deposition and binding of H1 to the arrays. The bisulfide crosslinking does not affect the mobility of the samples indicating no changes in their structure upon crosslinking (**Figure S1, lane2-5**). Digestion of chromatin with AvaI-ScaI, a restriction enzyme having a single cleavage site within the linker region of each nucleosomal repeat, generates only mono-nucleosome particles, with no free DNA detected (**Figure 2b, lane 3**). Although, chromatin DNA was inaccessible for digestion with the restriction enzyme HhaI, which has a single cleavage site within in the interior of each individual nucleosome of the arrays (**Figure 2b, lane 4**). In contrast, incubation of free DNA tandem repeats with either one of the enzymes leads to their complete cleavage resulting in the presence of monomeric repeat DNA only (**Figure 2b, lanes 5, 6**). Taken as a whole, the above data demonstrate the high quality of our reconstituted 197 bp arrays.

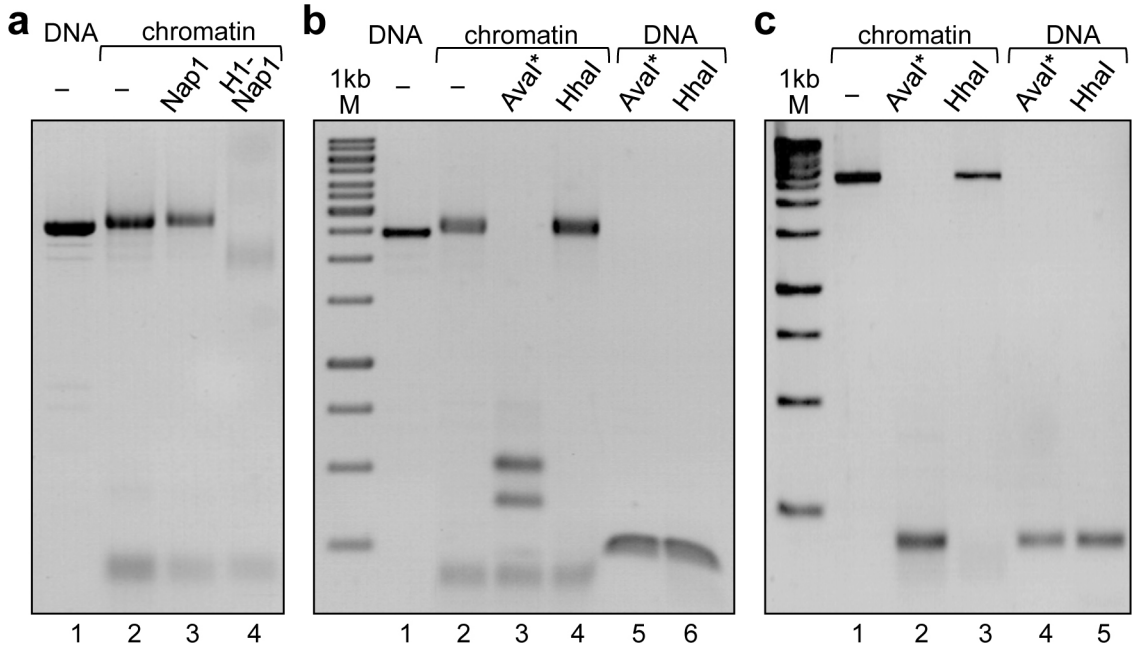


Figure 2: Characterization of the reconstituted chromatin used for ICCN analysis. (a) Agarose gel image of reconstituted 12x197 bp nucleosome arrays without (lane 3) and with H1 (lane 4). The histone chaperone Nap1 was used for the deposition of H1. Lane 3, only N added to the reconstituted chromatin. (b) Agarose gel image of 12x197 bp nucleosome arrays digested with either AvaI-ScaI (lane 3; AvaI-ScaI unique cleavage sites for 601 DNA is in the linker of the carrier and the biotinylated DNA substrates respectively) or HhaI (lane 4), having an unique cleavage site in the nucleosomal DNA. Lanes 5,6, digested free DNA tandem arrays with either AvaI-ScaI (lane 5) or with HhaI (lane 6) (c) Agarose gel of isolated DNA of 12x197 nucleosome arrays digested by AvaI-ScaI (lane 2) and HhaI (lane 4) after phenol chloroform. Lane 4, 5 digested free DNA tandem arrays with AvaI-ScaI (lane 4) and HhaI (lane 5).

We next condensed the arrays by using H1-Nap1 complewe in 50 mM monovalent ions buffer and carried out the bisulfide crosslinking. The efficiency of crosslinking was tested by SDS PAGE. Two additional bands, corresponding to H4-H4 and H2A-H4 adducts, were observed in the H1 bound arrays (**Figure 3a, lane 2**), a result in agreement with [34]. The composition of these bands was further confirmed by mass-spectrometry analysis (**Figure S2**). No such bands were, however, observed in the arrays, incubated with Nap1 only (**Figure 3a, lane 1**). We attribute this absence of nucleosome crosslinking in the arrays without H1 to reflect the very weak compaction of the arrays under these salt conditions [34]. The efficiency of nucleosome crosslinking within the H1-bound arrays was further revealed upon digestion with the restriction enzyme AvaI-ScaI (**Figure 3b**). Extensive digestion of the H1-folded crosslinked arrays results in the generation of several species, showing slower migration rates compared to this of the mono-nucleosomes. Note that under our conditions of crosslinking up to 3-4

additional bands with lower electrophoretic mobility were detected (**Figure 3b, left panel, lane3**). Treatment with 100 mM DTT leads to complete disappearance of these slower migrating bands and increasing the intensity of the mono-nucleosome band (**Figure 3b, left panel, lane 4**). We attributed these bands, in agreement with the available data [34], to reflect the presence of crosslinked monomer species (containing up to 3-4 monomers), which integrity is maintained thanks to the disulfide histone-histone crosslinking. Accordingly, DTT treatment leads to de-crosslinking and the presence of only monomeric particles in the treated sample. This is further supported by the electrophoretic analysis of the DNA isolated from the non-treated and DTT-treated AvaI-ScaI crosslinked arrays (**Figure 3b, right panel**). Indeed, in both cases only monomeric repeat 601 DNA was observed. This demonstrates that the digestion is achieved to completion and that the detected slower migrating bands in the nucleoprotein gel originated from the crosslinked monomer species.

We have next carried out a streptavidin chromatin pull-down. After isolation of DNA from the pull-down samples, we have used qPCR to identify the neighboring nucleosomes, which are crosslinked with nucleosome #5 within the compact array. By using specific set of primers, we have amplified (in 12 separate reactions) and thus measured the amount of DNA corresponding to each nucleosome (repeat) in the sample and normalized it relative to the amount of amplified DNA, corresponding to nucleosome #5. As seen (**Figure 3c**), the signals for nucleosome # 1, 3, 7, 9, and 11 compared to this of nucleosome #5, are 8%, 22% 42%, 24% and 13 %, respectively. No signals were detected for the any one of the "even" nucleosomes in the array. In addition, in the DTT treated samples, only a signal for nucleosome #5 was observed, showing that the qPCR amplified DNA reflects the crosslinking of neighboring nucleosomes via the disulfide histone H4-H4 and H2A-H4 bridges (**Figure 3c**). Accordingly, in the absence of histone H1, again only a signal for nucleosome #5 is detected (**Figure 3c**). All this indicates that within the 197 bp compact fiber, nucleosome #5 is in contact with both nucleosomes #3 and #7. As for the signal for nucleosome 1, 9 and 11, we attributed it to the pull-down of crosslinked "trimeric" and "tetrameric" particles, containing different combinations of crosslinked "odd" nucleosomes. Thus, the 197 bp chromatin fiber should exhibit a "N±2 " 3D nucleosome arrangement (see **Figure 1b**). To further confirm this, we have carried out the same

type of experiments but with biotin-labeled either nucleosome #4, or #8 or #9 (**Figure 4b**). The obtained results are in complete agreement with this for nucleosome #5. We conclude that within the 197 bp compact chromatin the nucleosomes exhibit a N±2 type of non-consecutive arrangement, fully compatible with the recent cryo-EM data [18].

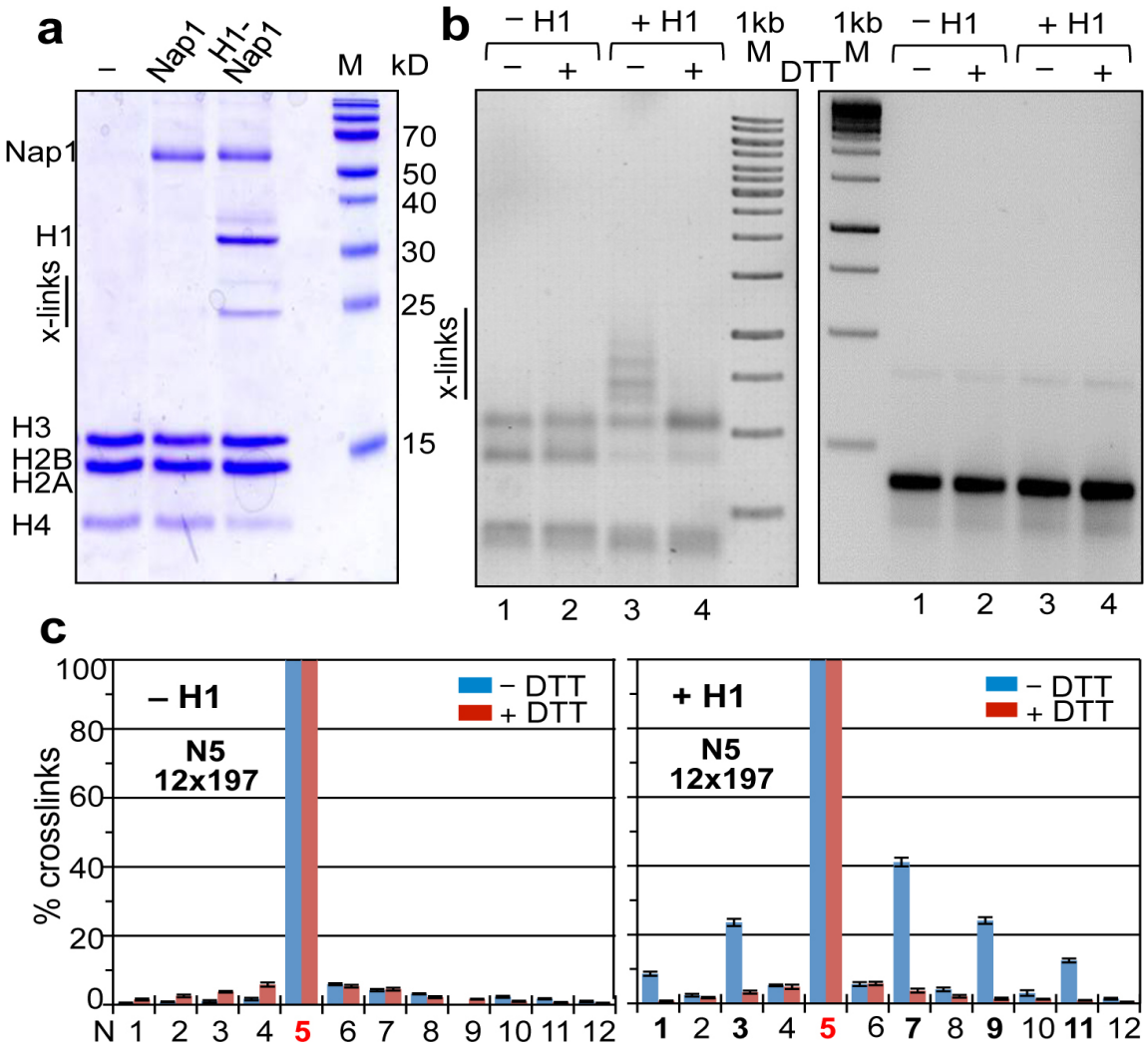


Figure 3: 3D organization of nucleosomes in reconstituted 12x197 bp nucleosomal arrays. (a) SDS PAGE image of internucleosomal disulfide crosslinking of histone H4-H4 and H2A-H4 in compacted nucleosomal arrays in the absence (lanes 1, 2) and presence of histone H1 (lane 3). Lane 2, only Nap1 was added to the nucleosome arrays. (b) Left panel, treatment with DTT results in complete reversal of the disulfide crosslinking. The crosslinked arrays were digested to completion with AvaI and then run on agarose gel before (lane 3) or after treatment with 100 mM DTT (lane 4). Right panel, DNA isolated from the samples run on lanes 1-4, respectively. (c) Right panel, N±2 3D nucleosome arrangement in histone H1 bound 12x197 bp nucleosomal arrays. The measured probability of crosslinking between the different nucleosomes is presented. Nucleosome #5 (N5) contained inserted biotin. The amount of streptavidin pull down N5 DNA was taken as 100% and the amount of the co-precipitated DNA from the other nucleosomes was normalized to it; Left panel, same as upper panel, but for 12x197 bp nucleosomal arrays without H1.

To test if the same type of 3D organization is also common for fibers with different repeat length, we have performed the same experiments, but by using compact nucleosome arrays with 177 and 227 bp repeat lengths. These arrays show an electrophoretic behavior similar to this of the 197 bp ones (**Figure S3, S4**). In contrast to the arrays without histone H1, the histone H1-bound both 177 bp and 227bp arrays exhibits efficient nucleosome-nucleosome crosslinking (**Figures S3, S4**), a result in agreement with the data for the 197 bp arrays. The qPCR data, summarized in (**Figure 4a,c)**, clearly show a N±2 spatial nucleosome arrangement for both samples. Therefore, this type of arrangement is general and does not depend on the length of the linker DNA.

Of note, the efficiency of crosslinking of a given nucleosome to its neighbor, located towards the center of the fiber, is ~1.5-2-fold higher compared to this of nucleosome located to the end of the fiber (**Figures 3d, 4a-c**). For example, the relative crosslinking for nucleosome #8 to nucleosome #6 is ~40%, while this for nucleosome #10 is ~ 20%. We attribute this to reflect the higher dynamics ("breathing") of the end located nucleosomes. Furthermore, the qPCR data without H1 (summarized in figure S7) shows clearly no chromatin compaction, in accordance with electrophoresis results.

Our results are in agreement with the structure in figure 4d, as well as a more refined version presented in figure 4e. However, our results do not confirm the irregular space between nucleosomes. This model resulted from fitting of cryo-EM images with the crystal structure of tetramer without H1 [37]. This irregular spacing might be an artifact of the fitting process. We suggest a two start model similar to figure 4e but with regular alignment of the tetranucleosome.

It is important to note that the ICNN approach showed reproducible results. Two distinct experiments were performed on both arrays 12x197 #9 and 12x177 #4 presented same type of nucleosomal arrangements (respectively **Figure S5a, b and S5c, d**).

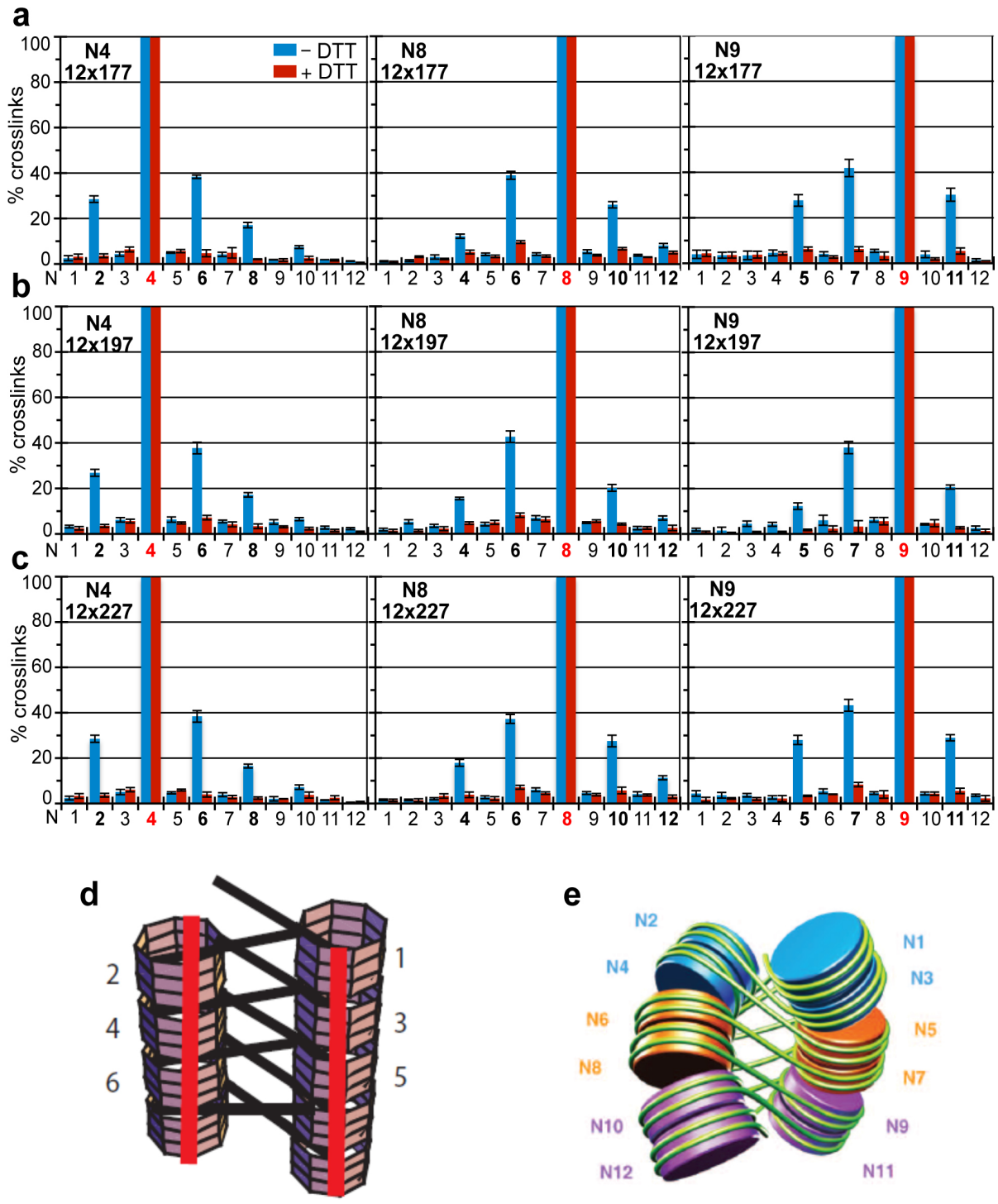


Figure 4: The N±2 3D nucleosomal arrangement in condensed arrays is independent of DNA linker length. (a) Measured probability of crosslinking of nucleosomes within the condensed 177 bp repeats array. Biotin was inserted in either nucleosome #4 (N4, left), or #8 (N8, middle) or #9 (N9, right), respectively. (b, c) Same as (a) but for condensed arrays with repeat length of 197 (b) and 227 bp (c), respectively. d) Zigzag model proposed by [35] e) zigzag model proposed by [36].

The H2A.Z chromatin fiber exhibits N±2 nucleosome spatial organization

H2A.Z is an essential histone variant and its genome wide localization and structure of the H2A.Z containing chromatin is believed, as noted above, to be of crucial importance for the regulation of transcription as well as other important events in the cell (recently reviewed in [38]). With this in mind, we next studied, by using ICNN, the 3D arrangement of compact 197 bp H2A.Z nucleosome arrays with nucleosome #8 being biotin labeled. The reconstituted under our conditions chromatin was of excellent quality (**Figure S6**). The crosslinking at 50 mM NaCl was as efficient as for conventional H2A 197 bp nucleosomal arrays (**Figure 3 and Figure S6**). As expected, either the absence of H1 or the treatment with DTT abolished completely the crosslinking in the H2A.Z compact chromatin (**Figure S6**). The data for the 3D arrangement of nucleosome #8 in H2A.Z arrays are identical to these for the nucleosome arrangement in conventional H2A compact arrays, i.e. H2A.Z chromatin exhibits N±2 3D nucleosome arrangement. We conclude that the presence of H2A.Z does not affect the 3D arrangement of nucleosomes in compact chromatin. This suggests that the implication of H2A.Z in numerous vital for the cell phenomena, should not be achieved via the H2A.Z fiber structure, but instead by its mere presence at specific genomic locations. In this context, the role of H2A.Z chaperones in selective deposition/removal of the H2A.Z [39] would be of crucial importance, since this allows modulating at a "wish" the accessibility of the DNA binding factors cognate sequences buried in the H2A.Z nucleosomes.

In summary, we have described the application of the ICNN approach for identification of the *In vitro* 3D nucleosome organization in the compact both conventional and H2A.Z nucleosomal arrays. However, ICNN could have a broader application in chromatin studies, including analysis of the alterations in chromatin fiber conformations and their spreading along the fiber upon H1 loss, binding of transcription factors or chromatin remodeling machines. Of particular interest would be to study the conformational changes along the fiber upon active transcription and replication.

The ICNN approach uses tandem arrays, which contain a biotinylated repeat inserted at specific site (see Materials and Methods). The insertion of the biotinylated repeat is performed by its ligation with the remaining two DNA fragments of the tandem repeat. Then arrays are reconstituted by salt dialysis and Nap1 is used to deposit H1.

Importantly, a further development of the approach could be envisaged, which could use already reconstituted chromatin templates and the ligation could be carried out at the level of these templates. This would allow the insertion of one or several nucleosomes at a selected site of the arrays. This nucleosome could contain epigenetic modifications such as acetylation or methylation and thus, the local effect on the structure of the fiber due to these histone modifications could be analyzed by the ICNN approach. Some recent, mainly cryo-Electron Microscopy data, defied the longtime accepted existence of genome-wide ordered 30 nm chromatin fiber *In vivo* (for recent review, see [40]. The potential of application of ICNN for *In vivo* studies would shed light on this question.

References

1. van Holde, K., Chromatin. Springer-Verlag KG, Berlin, Germany. 1988.
2. Thoma, F., T. Koller, and A. Klug, Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. J Cell Biol, 1979. 83(2 Pt 1): p. 403-27.
3. Makarov, V.L., S.I. Dimitrov, and P.T. Petrov, Salt-induced conformational transitions in chromatin. A flow linear dichroism study. Eur J Biochem, 1983. 133(3): p. 491-7.
4. Felsenfeld, G. and J.D. McGhee, Structure of the 30 nm chromatin fiber. Cell, 1986. 44(3): p. 375-7.
5. Bednar, J., et al., Chromatin conformation and salt-induced compaction: three-dimensional structural information from cryoelectron microscopy. J Cell Biol, 1995. 131(6 Pt 1): p. 1365-76.
6. Bednar, J., et al., Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. Proc Natl Acad Sci U S A, 1998. 95(24): p. 14173-8.
7. Tremethick, D.J., Higher-order structures of chromatin: the elusive 30 nm fiber. Cell, 2007. 128(4): p. 651-4.
8. Perez-Montero, S., et al., The embryonic linker histone H1 variant of Drosophila, dBigH1, regulates zygotic genome activation. Dev Cell, 2013. 26(6): p. 578-90.
9. Luger, K., et al., Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature, 1997. 389: p. 251-260.
10. Tachiwana, H., et al., Crystal structure of the human centromeric nucleosome containing CENP-A. Nature, 2011. 476(7359): p. 232-5.
11. Tachiwana, H., et al., Structures of human nucleosomes containing major histone H3 variants. Acta Crystallogr D Biol Crystallogr, 2011. 67(Pt 6): p. 578-83.

12. Gerchman, S.E. and V. Ramakrishnan, Chromatin higher-order structure studied by neutron scattering and scanning transmission electron microscopy. Proc Natl Acad Sci U S A, 1987. 84(22): p. 7802-6.
13. Ghirlando, R. and G. Felsenfeld, Hydrodynamic studies on defined heterochromatin fragments support a 30-nm fiber having six nucleosomes per turn. J Mol Biol, 2008. 376(5): p. 1417-25.
14. Williams, S.P., et al., Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length. Biophys J, 1986. 49(1): p. 233-48.
15. Bordas, J., et al., The superstructure of chromatin and its condensation mechanism. II. Theoretical analysis of the X-ray scattering patterns and model calculations. Eur Biophys J, 1986. 13(3): p. 175-85.
16. Bordas, J., et al., The superstructure of chromatin and its condensation mechanism. I. Synchrotron radiation X-ray scattering results. Eur Biophys J, 1986. 13(3): p. 157-73.
17. Makarov, V., et al., A triple helix model for the structure of chromatin fiber. FEBS Lett, 1985. 181(2): p. 357-61.
18. Song, F., et al., Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. Science, 2014. 344(6182): p. 376-80.
19. Grigoryev, S.A., et al., Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. Proc Natl Acad Sci U S A, 2009. 106(32): p. 13317-22.
20. Correll, S.J., M.H. Schubert, and S.A. Grigoryev, Short nucleosome repeats impose rotational modulations on chromatin fibre folding. EMBO J, 2012. 31(10): p. 2416-26.
21. Boulard, M., et al., Histone variant nucleosomes: structure, function and implication in disease. Subcell Biochem, 2007. 41: p. 71-89.
22. Nekrasov, M., et al., Histone H2A.Z inheritance during the cell cycle and its impact on promoter organization and dynamics. Nat Struct Mol Biol, 2012. 19(11): p. 1076-83.
23. Greaves, I.K., et al., H2A.Z contributes to the unique 3D structure of the centromere. Proc Natl Acad Sci U S A, 2007. 104(2): p. 525-30.
24. Xu, Y., et al., Histone H2A.Z controls a critical chromatin remodeling step required for DNA double-strand break repair. Mol Cell, 2012. 48(5): p. 723-33.
25. Weber, C.M., J.G. Henikoff, and S. Henikoff, H2A.Z nucleosomes enriched over active genes are homotypic. Nature structural & molecular biology, 2010. 17(12): p. 1500-7.
26. Weber, C.M., S. Ramachandran, and S. Henikoff, Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. Mol Cell, 2014. 53(5): p. 819-30.
27. Jin, C., et al., H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. Nat Genet, 2009. 41(8): p. 941-5.

28. Ishibashi, T., et al., Acetylation of vertebrate H2A.Z and its effect on the structure of the nucleosome. Biochemistry, 2009. 48(22): p. 5007-17.
29. Placek, B.J., et al., The H2A.Z/H2B dimer is unstable compared to the dimer containing the major H2A isoform. Protein Sci, 2005. 14(2): p. 514-22.
30. Fan, J.Y., et al., The essential histone variant H2A.Z regulates the equilibrium between different chromatin conformational states. Nat Struct Biol, 2002. 9(3): p. 172-6.
31. Fan, J.Y., et al., H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding. Mol Cell, 2004. 16(4): p. 655-61.
32. Abbott, D.W., et al., Characterization of the stability and folding of H2A.Z chromatin particles. Implications for transcriptional activation. J. Biol. Chem., 2001. 276: p. 41945-41949.
33. Suto, R.K., et al., Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. Nat Struct Biol, 2000. 7(12): p. 1121-4.
34. Dorigo, B., et al., Nucleosome arrays reveal the two-start organization of the chromatin fiber. Science, 2004. 306(5701): p. 1571-3.
35. Staynov, D.Z., The controversial 30 nm chromatin fibre. Bioessays, 2008. 30(10): p. 1003-9.
36. Song, F., et al., Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. Science. 344(6182): p. 376-80.
37. Schalch, T., et al., X-ray structure of a tetranucleosome and its implications for the chromatin fibre. Nature, 2005. 436(7047): p. 138-41.
38. Talbert, P.B. and S. Henikoff, Environmental responses mediated by histone variants. Trends Cell Biol, 2014.
39. Obri, A., et al., ANP32E is a histone chaperone that removes H2A.Z from chromatin. Nature, 2014. 505(7485): p. 648-53.
40. Ausio, J., The shades of gray of the chromatin fiber: Recent literature provides new insights into the structure of chromatin. Bioessays, 2014.

Supplementary Results:

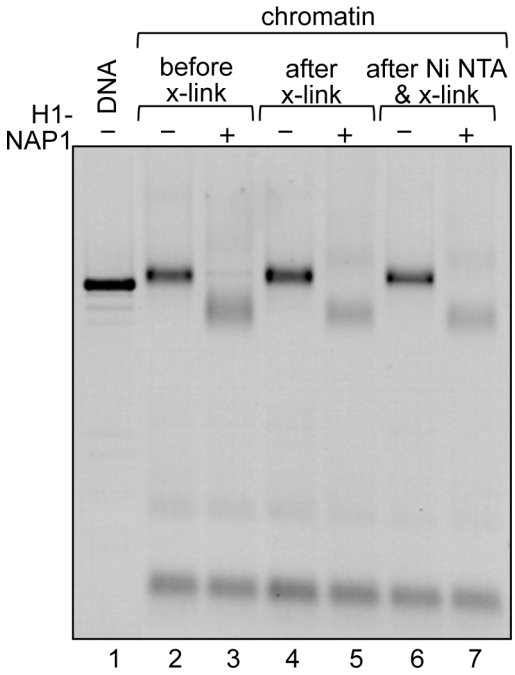
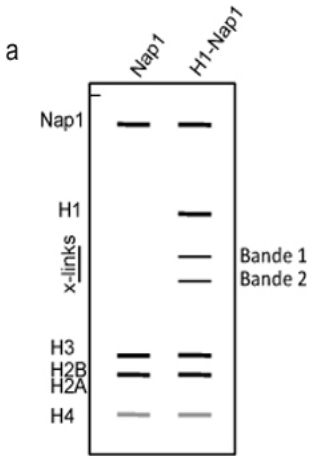


Figure S1: The bisulfide crosslinking does not affect the mobility of the nucleosomal arrays, agarose gel analysis. Free DNA tandem 12x227 bp (lane 1). Nucleosome arrays 12X227 bp without histone H1 (lanes 2, 4, 6) and with histone H1 (lanes 3, 5, 7). The histone chaperone Nap1 was used for the deposition of H1. Removal of Nap1 was performed by Ni-NTA beads (lane 6 and 7). As seen Ni-NTA beads shows no effect on the chromatin crosslinking.

a



b

Réf. CCMP	Réf. Client	Analyse LC-MSMS	Base de données			
11_705	Bande 1	705a.wiff	Sprot_2011_07_XENLA			
N(1)	**Total (2)**	**% Cov (3)**	**Accession #**	**Name**	**Species**	**Peptide (95%)(4)**
1	13,65	37,7	sp|P06897|H2A1_XENLA	Histone H2A type 1 OS= Xenopuis laevis PE=1 SV=2	XENLA	6
2	12,05	78,6	sp|P62799|H4_XENLA	Histone H4 OS= Xenopus laevis PE=1 SV=2	XENLA	5
11_706	Bande 2	706a.wiff	Sprot_2011_07_XENLA			
N(1)	**Total (2)**	**% Cov (3)**	**Accession #**	**Name**	**Species**	**Peptide (95%)(4)**
1	32,16	80,6	sp|P62799|H4_XENLA	Histone H4 OS= Xenopus laevis PE=1 SV=2	XENLA	16

Figure S2: Mass spectroscopy analysis of the crosslinked bands. (a) Schematic of the SDS gel used for the mass spectroscopy analysis. (b) Mass stectroscpy results showing that 'bande 1' contains histone H2A and histone H4, whereas 'bande 2' contains only histone H4 confirming that these band corresponds to H2A-H4 and H4-H4 crosslinks. These results are consistent with crosslinking results published in [37].

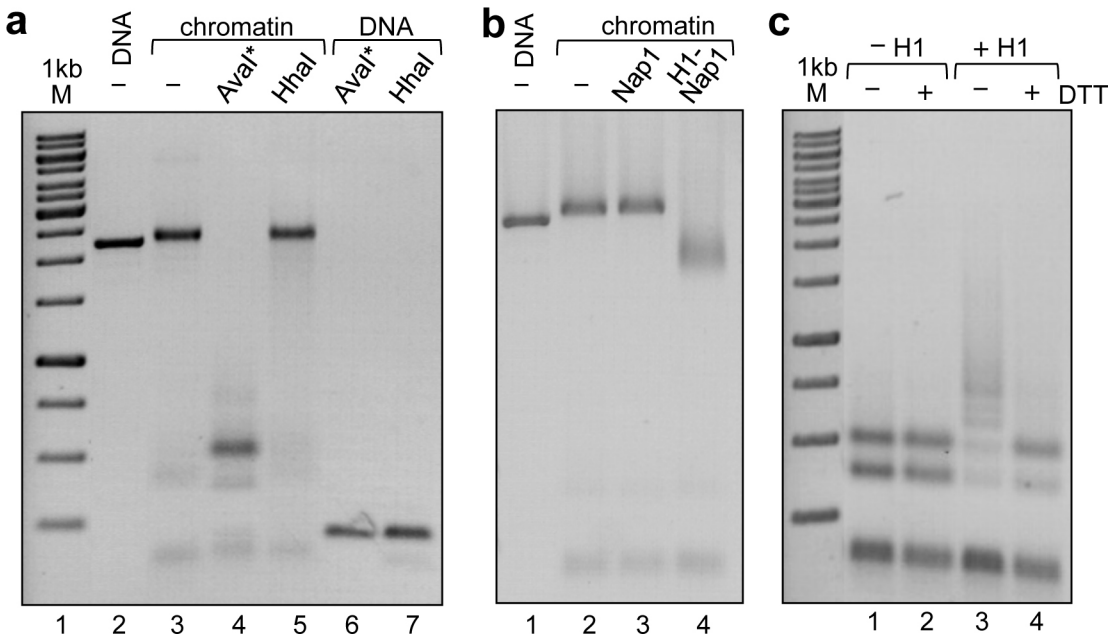


Figure S3: Agarose gel characterization of the reconstituted 12x177 bp nucleosome arrays by restriction enzymes treatment. Free DNA and nucleosome array without (a) and with H1 (b). . The histone chaperone Nap1 was used for the deposition of H1. (c) The crosslinked arrays were digested to completion with AvaI/*Sca*I before (lane 1, 3) or after treatment with 100 mM DTT (lane 2, 4) Treatment with DTT results in complete reversal of the disulfide crosslinking in 12x177 bp arrays (compare lanes 3 and 4).

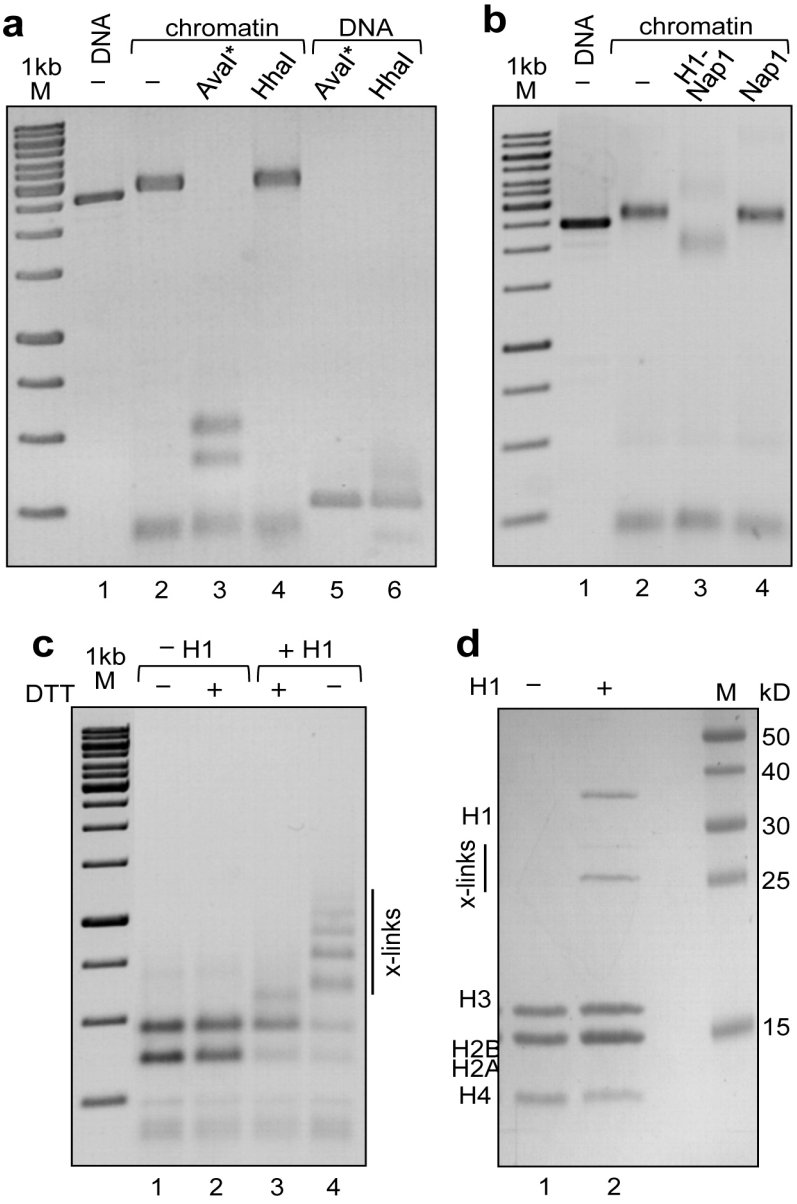


Figure S4: Efficient disulfide nucleosome crosslinking in compact 12x227 bp nucleosome arrays. Characterization of the reconstituted 12x227bp nucleosome arrays by restriction enzymes cleavage and agarose gels (a), (b), (c) and by SDS PAGE (d). The histone chaperone Nap1 was used for the deposition of H1. (b) Nucleosome arrays without (lanes 2, 4), with H1 (lane 3), and with only Nap1 added (lane 4). (c) Treatment with DTT results in complete reversal of the disulfide crosslinking in 12x227 bp arrays. The crosslinked arrays were digested to completion with AvaI/ScaI and then run on agarose gel before (lane 1, 4) or after treatment with 100 mM DTT (lane 2, 3). (d) SDS PAGE image of internucleosomal disulfide crosslinking of histone H4-H4 and H2A-H4 in compacted nucleosomal arrays in the absence (lanes 1) and presence of histone H1 (lane 2). Lane 1, only Nap1 was added to the nucleosome arrays. Nap1 were removed by Ni-NTA beads but it did not affect the crosslinked chromatin (Figure S2, Lane 6, 7).

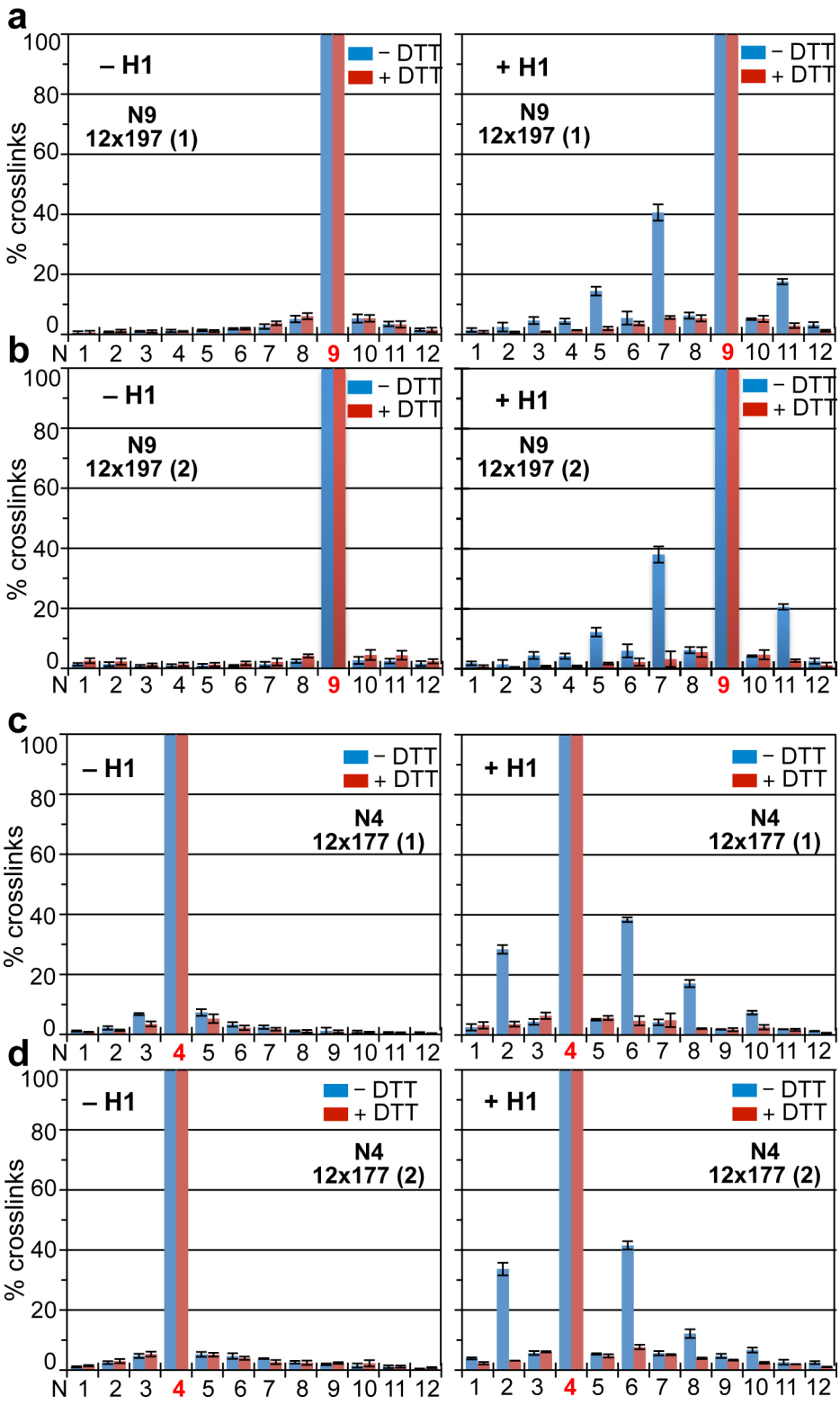


Figure S5: QPCR analysis for two fully independent experiment with arrays 12x197 #9 (a, b) and 12x177 #4 (c, d) demonstrating the high reproducibility of the results.

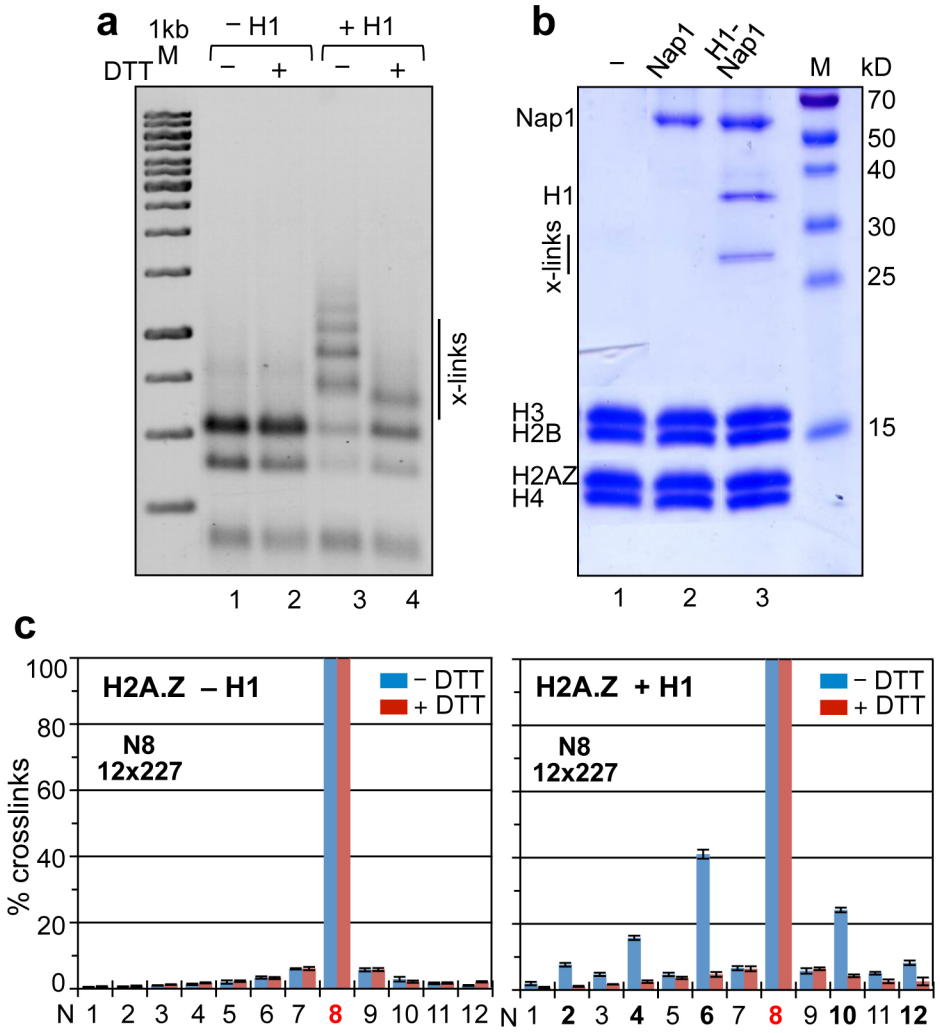


Figure S6: 3D organization of nucleosomes in reconstituted 12x197 bp nucleosomal arrays containing H2A.Z instead of H2A. (a) Treatment with DTT results in complete reversal of the disulfide crosslinking. The crosslinked arrays were digested to completion with AvaI/ScaI and then run on agarose gel before (lane 1, 3) or after treatment with 100 mM DTT (lane 2, 4). (b) SDS PAGE image of internucleosomal disulfide crosslinking of histone H4-H4 and H2A.Z-H4 in compacted nucleosomal arrays in the absence (lanes 1, 2) and presence of histone H1 (lane 3). Lane 2, only Nap1 was added to the nucleosome arrays. (c) qPCR analysis clearly showing a similar N±2 3D arrangement as conventional H2A.

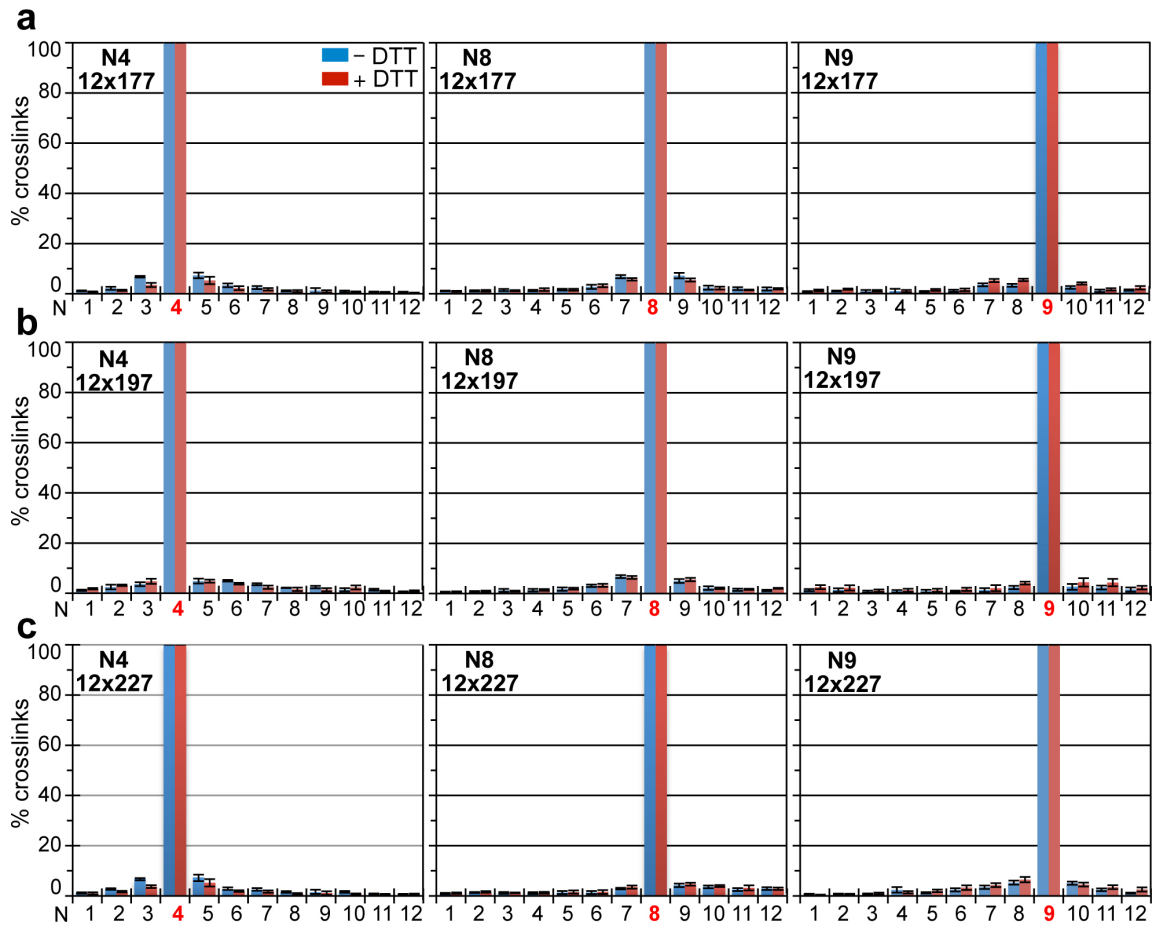


Figure S7: qPCR analysis of streptavidin pull down DNA isolated from AvaI-ScaI-digested 177 bp (a), 197 (b) and 227 bp nucleosome arrays (c) with biotin labeled nucleosomes #4, #8 and #9 (from left to right) without H1. Note the absence of crosslinking for all arrays demonstration the essential role of H1 in chromatin folding at low monovalent and absence of divalent ions.

Manuscript 2:

CENP-A centromeric chromatin exhibits non-canonical 30 nm fiber structure

CENP-A centromeric chromatin exibits non-canonical 30 nm fiber structure

Soueidan,L; Dimitrov,S; Angelov,D.

Background

All the conventional histones, except H4, have histone variants. Histone variants are non-allelic forms of the conventional histones (1). The incorporation of histone variants confers novel structural and functional properties to the nucleosomes (3). A spate of recent studies have implicated the incorporation of histone variants into the nucleosome in the regulation of transcription, repair, senescence, cell division, meiosis, epigenomics events, etc (4).

Histone variant CENP-A

CENP-A replaces histone H3 at the pericentromeric chromatin. It is a key player of centromere organization and contributes to centromere identity. It has a crucial role in mitosis, as it is required for the proper recruitment of kinetochore components which drive the alignment and segregation of chromosomes (5). CENP-A comprises a well-conserved histone-fold domain and a highly divergent amino-terminal tail. Very recently the crystal structure of the CENP-A nucleosome was solved (2). Intriguingly, in contrast to the canonical nucleosome (where 147 bp of DNA are wrapped around the histone octamer), only the central 121 bp were visible (Figure 1). The thirteen base pairs from both ends of the DNA are invisible in the crystal structure, and the αN helix of CENP-A is shorter than that of H3, which is known to be important for the orientation of the DNA ends in the canonical H3 nucleosome (6).

Onto centromeric CENP-A chromatin is assembled the so-called constitutive centromere associated network (CCAN) of 16 proteins (termed generally as CENPs) distributed in several functional groups as follows: CENP-C, CENP-H/CENP-I/CENP-K/, CENP-L/CENP-M/CENP-N, CENP-O/CENP-P/CENP-Q/CENP-R/CENP- U(50), CENP-T/CENP-W, and CENP-S/CENP-X (7). Importantly, CENP-C and CENP-T function to direct kinetochore formation (8).

Structure-function relationship of the CENP-A nucleosome

Association of histone H1 with nucleosomal arrays leads to compaction and inaccessibility of protein factors to the underlying DNA sequences. CENP-A chromatin has to be, however, accessible to the CENPC's from the constitutive centromere associated network, in particular to CENP-B, CENP-C, CENP-T and CENP-N. How this could be achieved? We hypothesized that the specific structure of the CENP-A nucleosome interferes with the binding of histone H1 and in turns, a compact 30 CENP-A fiber with nucleosomes in close contacts cannot be formed. This basic property of centromeric chromatin would be essential for the *In vivo* assembly of active centromeres. We have recently mapped in the lab at one base pair resolution the interaction of H1 with the nucleosome (9). It was shown that H1 binding requires a specific orientation of the exit and entry angle of the nucleosomal DNA ends (10). These requirements appear, however, not to be met in the CENP-A nucleosome (13 bps from each DNA end of the NCP are quite flexible and thus, no rigid orientation of the linker DNA should be expected, see **Figure 1**).

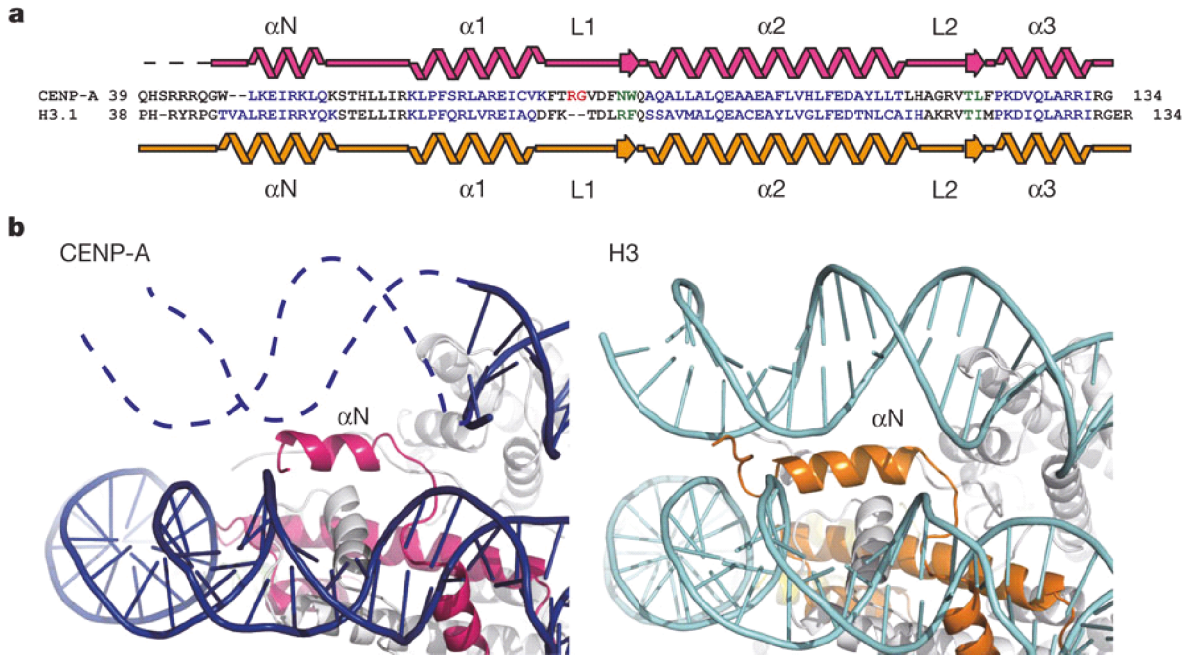


Figure 1: Structure of the DNA entrance and exit of the human CENP-A nucleosome (2). (a) Secondary structure of CENP-A in the nucleosome. The sequences of human CENP-A and H3 are aligned with the secondary structure. (b) Close-up views of the αN helices and the DNA edge regions of the CENP-A (left panel) and H3 (right panel) nucleosomes. The dashed line in the left panel shows the DNA region that is not visible in the crystal structure. The CENP-A and H3 molecules are shown in magenta and orange, respectively.

3D nucleosome arrangement in the CENP-A chromatin

Very recently, in our laboratory it was shown that H1 is indeed unable to properly bind to the CENP-A nucleosome (Imtiaz Lone, personal communication). This suggests that the CENP-A chromatin would not be able to condense in a structure of the "N±2" type, characteristic for the conventional 30 nm fiber. To test this we have used the ICNN approach described in the attached paper. Briefly, we have reconstituted 12x197 bp nucleosome arrays by using nucleosome #8 (N8) with inserted biotin and histone octamer, consisting of H4-V21C and H2A-E64C mutated histones and both wild type H3 and H2B (see Materials and Methods section of the attached manuscript for details). Linker histone H1 was deposited in *In vivo*-like manner by using its chaperone NAP1. Digestion with either *Ava*I/ScaI (*Ava*I/scaI unique cleavage site for 601 DNA is in the linker DNA) or *Hha*I (*Hha*I has an unique cleavage site in the nucleosomal DNA) shows the full nucleosome occupancy of the arrays. Of note, upon addition of the NAP1-H1 complex, the band corresponding to the 12x197 bp CENP-A nucleosome arrays, is up-shifted compared to these of both conventional H3 nucleosomal arrays with and without H1 (**Figure 2a)**. Note that the conventional H1-bound arrays exhibits higher mobility relative to the mobility of arrays without H1. This suggests that histone H1 is not properly associated with the CENP-A arrays. If this, as indicated also by other data in the lab is true, we may expect a weaker and distinct condensation of the arrays and no crosslinking of the CENP-A nucleosomes.

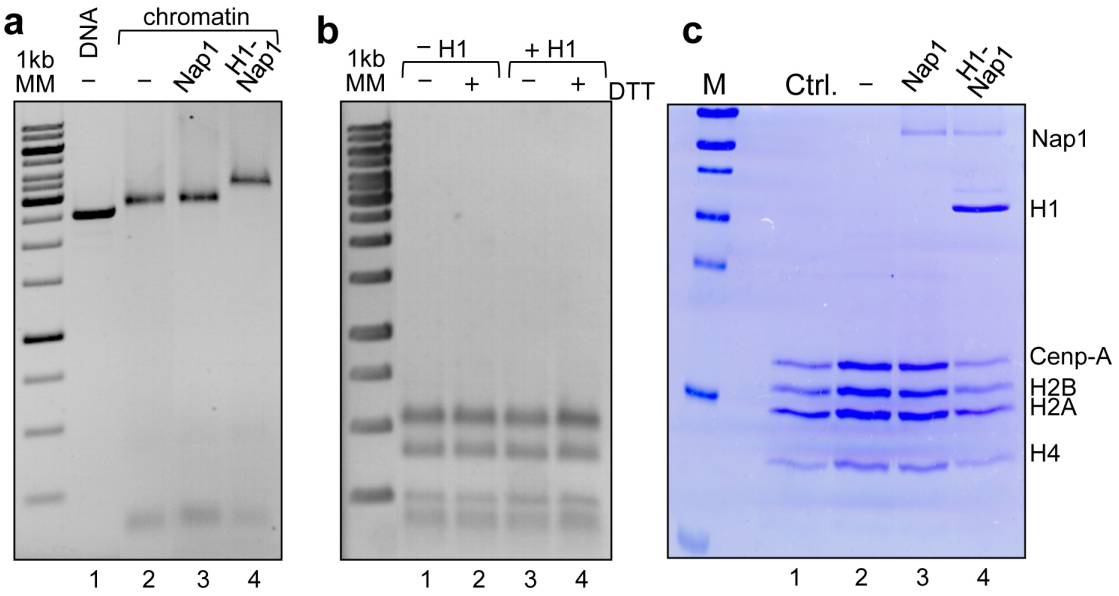


Figure 2: Characterization of the reconstituted CENP-A 12x197 bp nucleosomal arrays used for ICCN analysis. (a) Agarose gel of reconstituted 12x197 bp CENP-A nucleosome arrays without (lane 3) and with H1 (lane 4). The histone chaperone NAP1 was used for the deposition of H1. For the sample in lane 3, only NAP1 added to the reconstituted chromatin; lane 1, free DNA 12x197 bp tandem arrays (b) Arrays with (lane1, 2) or without H1 (lane3, 4) were treated with 100 mM DTT. The crosslinked arrays were digested to completion with ScaI and then run on agarose gel before (lane1, 3) or after (lane2, 4) treatment with 100 mM DTT. Note the absence of crosslinking (no lower migrating bands are present). (c) SDS PAGE analysis of internucleosomal disulfide crosslinking in CENP-A 12x197 bp nucleosomal arrays.

The crosslinking was tested first by SDS electrophoresis (**Figure 2c**). As seen, in contrast to conventional arrays, no H4-H4 or H2A-H4 crosslinking was observed in the H1 bound samples. This was further supported by the ScaI-AvaI digestion of both conventional and CENP-A nucleosomal arrays (**Figure 2b**). The appearance of oligomers with higher molecular masses, reflecting the crosslinked monomeric units, was detected in the ScaI digested conventional samples. Of note, treatment with DTT led to disappearance of the oligomers. No such bands were, however, detected in both non-DTT treated and treated digested with AvaI/ScaI CENP-A arrays (incubated with either NAP1-H1 or with H1 only). This reveals that the 3D organization of the nucleosomes in the CENP-A arrays should be distinct from the N±2 organization of the conventional chromatin fiber, which is further demonstrated by the qPCR results presented on **Figure 3**.

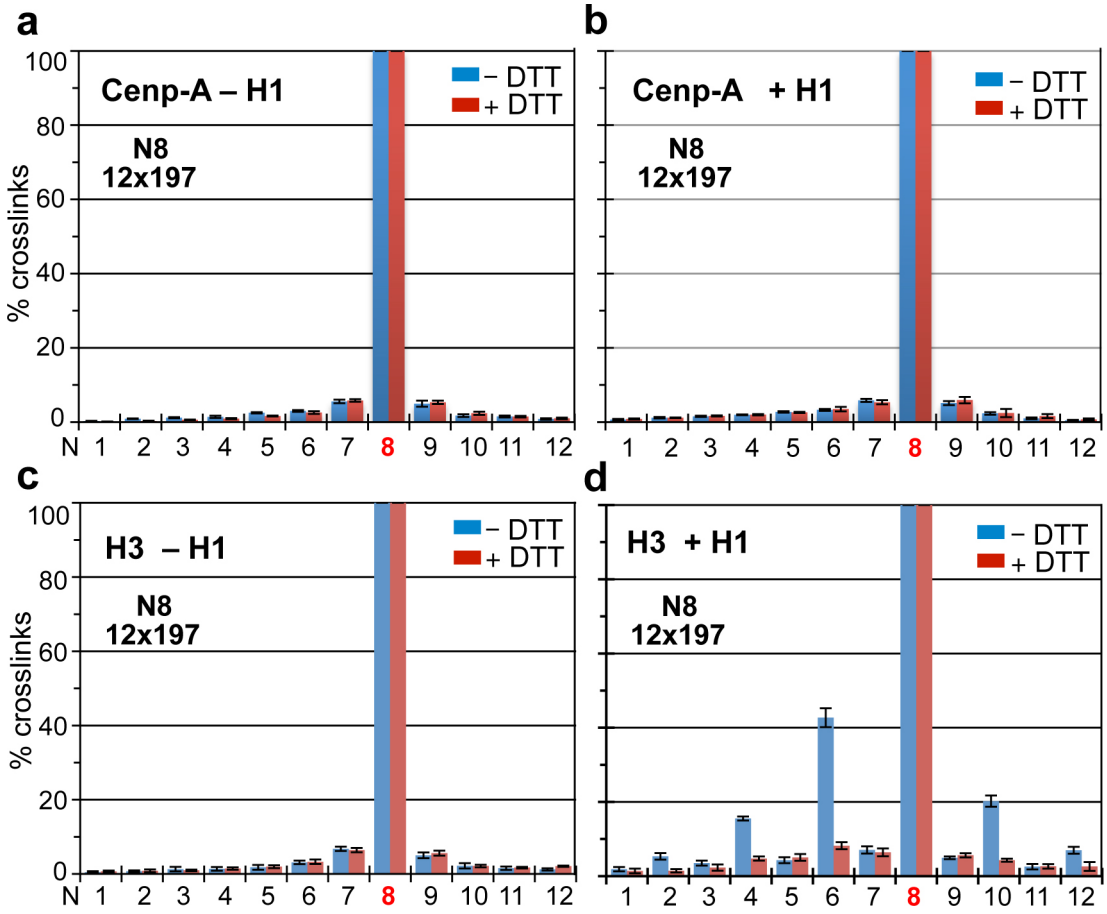


Figure 3: Measured probability of crosslinking of nucleosomes within the condensed 197 bp repeats arrays containing CENP-A octamer or conventional one. Biotin was inserted in either nucleosome 8. To note, the absence of the N+2 3D structure in CENP-A array in presence or absence of H1 (a,b) in comparision with conventional chromatin (c,d).

"Rigidifying" the DNA ends of the CENP-A-A nucleosome allows the assembly of fiber with 3D "N±2"-type arrangement of nucleosomes.

The flexibility of the DNA ends of CENP-A nucleosome reflects the peculiar structure of CENP-A amino-terminal region within the nucleosome (2). Indeed, the αN helix of CENP-A is one helical turn shorter than this of conventional H3 and the preceding region, in contrast to this of H3, is completely disordered ((2), see also **Figure 1)**. However, both the length of αN helix and the loop segment preceding the αN helix (this loop interacts directly with the DNA ends in H3 nucleosome (6)) are required for maintaining the DNA orientation at the entrance and exit of H3 nucleosomes. The "defects" in the organization of the amino-terminal region of CENP-A appear, thus, to be responsible for inherent flexibility of the DNA ends of the CENP-A nucleosome (2) and the inability of histone H1 to bind to it. This would, in turn, determine the inability of the CENP-A nucleosome arrays to adopt 3D "N±2" - type nucleosome arrangement (see above).

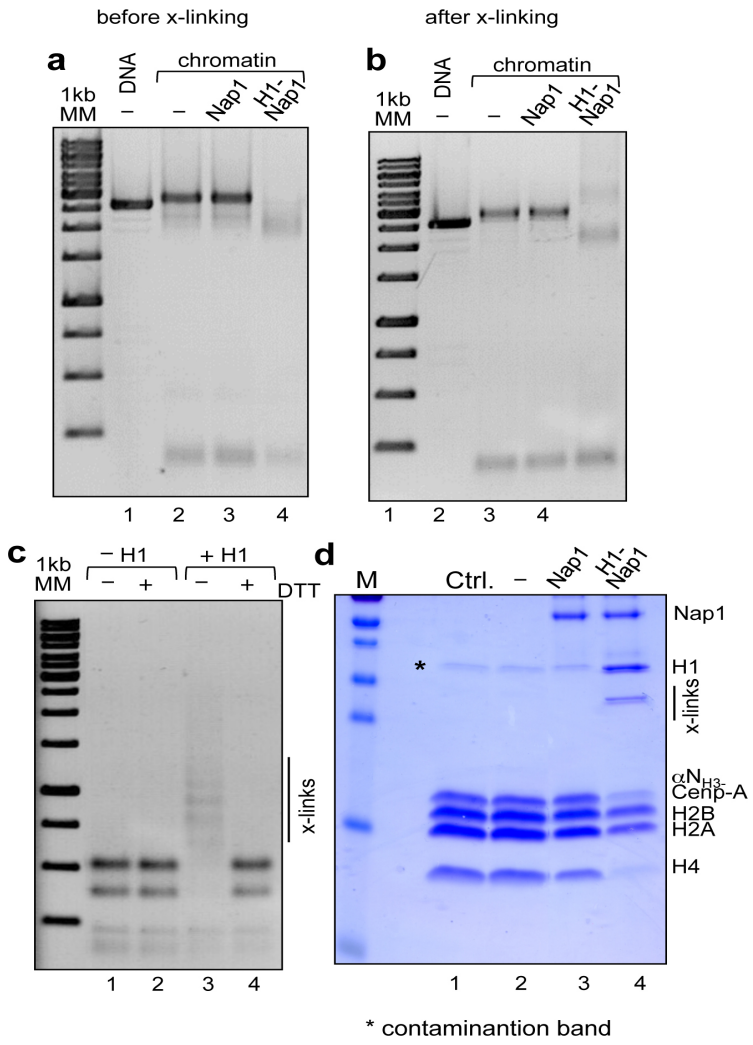


Figure 4: Characterization of the reconstituted αN_{H3} CENP-A 12x197 bp nucleosomal arrays used for ICCN analysis. (a) Agarose gel of reconstituted 12x197 bp CENP-A nucleosome arrays without (lane 3) and with H1 (lane 4) before crtosslinking. (b) Same as (a) but after crosslinking. (c) Arrays with (lane1, 2) or without H1 (lane3, 4) were treated with 100 mM DTT. The crosslinked arrays were digested to completion with *Sca*I and then run on agarose gel before (lane1, 3) or after (lane2, 4) treatment with 100 mM DTT.(d) SDS PAGE analysis of internucleosomal disulfide crosslinking in αN_{H3} CENP-A 12x197 bp nucleosomal arrays in the absence (lanes 1, 2, 3) and presence of histone H1 (lane 4). Lane 3, only Nap1 was added to the 12x197 bp nucleosome arrays. The samples were run on 18% SDS-PAGE. Note the H4-H4 crosslinks in lane 4.

With this in mind, we hypothesized that swapping of the CENP-A αN helix and the segment preceding it with these of conventional H3 would rigidify the ends of nucleosomal CENP-A DNA and would allow binding of H1. The nucleosome-reconstituted arrays with a histone octamer comprising this mutant of CENP-A would then be able to properly bind H1 and thus, to adopt a 3D "N±2"- type nucleosome arrangement. To test this we have constructed a swapped CENP-A mutant (αN_{H3}- CENP-A) containing the αN helix and the preceding loop region of H3, expressed it in bacteria

and used for reconstitution of αN_{H3}- CENP-A 12x197 bp nucleosomal arrays. The deposition of H1 was performed by using the NAP1/H1 complex and was not affected by the crosslinking (**Figure 4a,b**). As seen (**Figure 4**), the αN_{H3}- CENP-A 12x197 bp H1-bound nucleosomal arrays, show identical electrophoretic behavior as the conventional ones. Importantly, both SDS PAGE and ScaI digestion unambiguously reveal nucleosome crosslinking in the αN_{H3}- CENP-A arrays with the same efficiency as this between the nucleosomes in the conventional condensed nucleosome arrays (**Figure 4c,d**). The qPCR analysis shows that the αN_{H3}- CENP-A arrays have identical to the conventional H1-bound 12x197 bp arrays a "N±2"-type spatial arrangement (**Figure 5a,b**). We conclude, that the flexible DNA ends of the CENP-A nucleosomes are determining the distinct fiber organization of CENP-A chromatin.

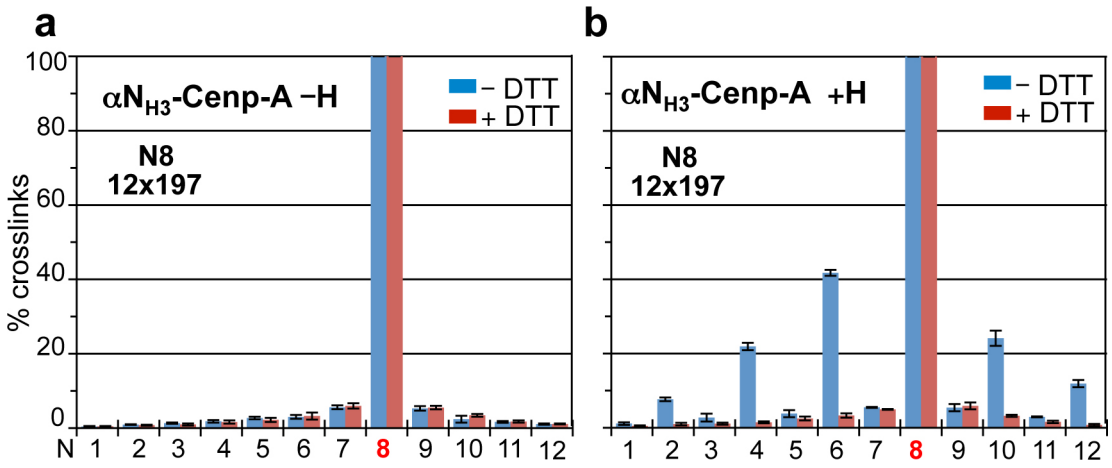


Figure 5: Measured probability of crosslinking of nucleosomes within the condensed 197 bp repeats arrays containing αN_{H3} CENP-A octamer. Biotin was inserted in either nucleosome 8. To note, the N+2 3D structure in αN_{H3} CENP-A array in presence of H1.

References

1. van Holde K (1988) Chromatin. Springer-Verlag KG, Berlin, Germany.
2. Tachiwana H, et al. (2011) Crystal structure of the human centromeric nucleosome containing CENP-A. Nature 476(7359):232-235.
3. Boulard M, Bouvet P, Kundu TK, & Dimitrov S (2007) Histone variant nucleosomes: structure, function and implication in disease. Subcell Biochem 41:71-89.
4. McKittrick E, Gafken PR, Ahmad K, & Henikoff S (2004) Histone H3.3 is enriched in covalent modifications associated with active chromatin. Proc Natl Acad Sci U S A 101(6):1525-1530.
5. Buscaino A, Allshire R, & Pidoux A (2010) Building centromeres: home sweet home or a nomadic existence? Curr Opin Genet Dev 20(2):118-126.

6. Luger K, Mäder AW, Richmond RK, Sargent DF, & Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389:251-260.
7. Perpelescu M & Fukagawa T (2011) The ABCs of CENPs. Chromosoma.
8. Gascoigne KE, et al. (2011) Induced ectopic kinetochore assembly bypasses the requirement for CENP-A nucleosomes. Cell 145(3):410-422.
9. Syed SH, et al. (2010) Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome. Proc Natl Acad Sci U S A 107(21):9620-9625.
10. Shukla MS, et al. (2011) The docking domain of histone H2A is required for H1 binding and RSC-mediated nucleosome remodeling. Nucleic Acids Res 39(7):2559-2570.

IV. Discussion

The structure of the chromatin plays a key role in epigenetic regulation of gene expression and therefore solving the 3D structure of the chromatin might lead to better understanding of the chromatin function and dynamics. The fundamental unit of the chromatin arrays is the nucleosome, which repeats every 160 to 240 bp across the genome [287]. The nucleosome has three essential functions: first, it is the first level of genomic compaction. Second, the nucleosome acts as signaling hub for DNA-templated processes by providing a platform for the binding of transcription factors or other chromatin-binding enzyme and by displaying post-translational histone modifications or histone variants. Third, the nucleosome can self-assemble into a 30 nm chromatin fiber, which is the first level of higher-order structure compaction.

The basic structure of the nucleosome core, the subunit of the chromatin, is known at a resolution of 1.9 [65] [59]. However, despite 30 years of research efforts, the structure of the 30 nm chromatin fiber remains unsolved. The 30 nm fiber is indeed too compact to allow the visualization of individual nucleosomes and of the DNA linking each of them. Several models have been created based on experimental data using cryo-EM and X-ray crystallography. The first model proposed by Klug and colleagues is the solenoid model (reviewed by [248]), in which consecutive nucleosomes are located next to each other in the fiber (nucleosome N interacts with N±1). This model is described to fold into a simple one-start helix (Figure 28, middle panel). In the second model, nucleosomes are arranged as a zigzag such that alternating nucleosomes become interacting partners (N interacts with N±2). This latter model forms a two-start helix described as two rows of nucleosomes attached by straight or crisscrossed linker DNA [104]. The crystal structure of a tetranucleosome array at a resolution of 9 was solved by Richmond and colleagues [214] (Figure 28, left panel). Even though the resolution was low, it was possible to visualize the positions of the linker DNA and the relative nucleosome positioning. However, the lack of the linker histone H1 made these results questionable. Electron microscopy data of long and regular chromatin fibers with incorporated linker histone H5 (chicken subtypes of H1) leaned towards an interdigitated solenoid [215]. Indeed, the role of the linker histone H1 in the chromatin fiber structure remains to be determined. It is important to note that the proposed models assume that the conformation of the nucleosome does not change upon compaction. More recently, 11 resolution cryo-EM structures of a reconstituted 12

tandem repeats of 601 DNA sequence in presence of linker histone H1 showed a zigzag structure with straight linker DNA. Moreover, the structure described the chromatin as a left-handed twist of repeating tetranucleosome structural units [216] (Figure 28, right panel).

The crystal structure of the nucleosome core revealed that the surface of the nucleosome has an uneven distribution of charges [59, 65]. The most striking feature is a cluster of 7 acidic AA contributed from histone H2A, later referred to as acidic patch. It was reported that the N-terminal tail of histone H4 (residues 16-25) originating from adjacent nucleosome interacts with histone H2A through the acidic patch leading to chromatin compaction [288]. Crosslinking studies demonstrated that disulfide bridges can be generated between the N-terminal tail of histone H4 and the H2A acidic patch in condensed chromatin only [25].

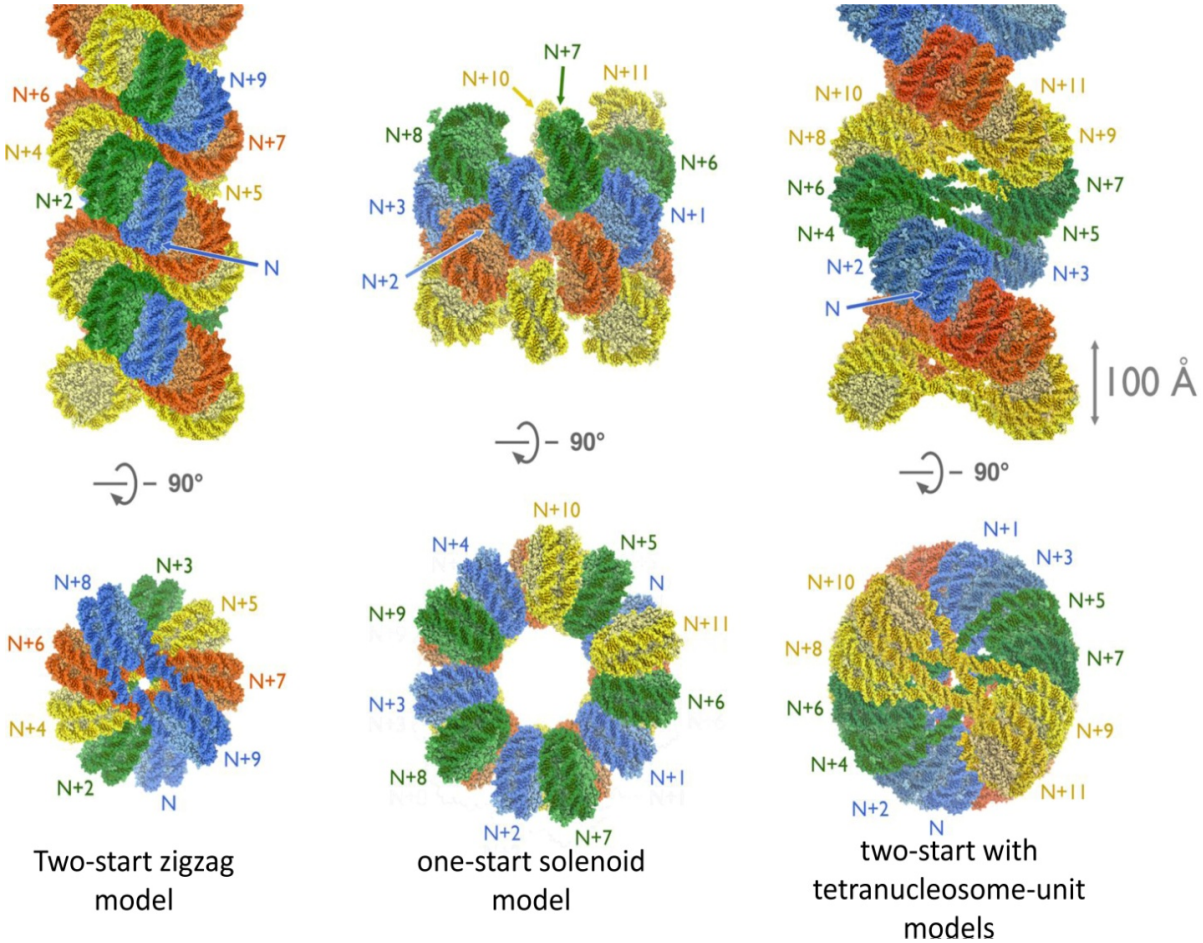


Figure 27: Models of the 30 nm fiber. Orthogonal views perpendicular to the 30 nm fiber axis (top) and down the axis (bottom) of the two-start zigzag model (left), one-start solenoid model (center) and two-start tetranucleosome-unit model (right). In the two-start model, each sequential pair of nucleosomes across the fiber is colored similarly. For the one-start model, all nucleosomes in the same turn of the solenoid are colored similarly. In the two-start tetranucleosome model, each tetranucleosome repeating unit is colored similarly

In this work, we developed a new approach called the ICNN (identification of the closest neighbor nucleosome) allowing unambiguous determination the 3D arrangement of nucleosomes within a linker histone H1-dependent compacted fiber. The ability of disulfide cross-linking between neighboring nucleosomes due to interactions between the acidic patch of H2A and the H4 tail is the main feature of this approach. 601 DNA nucleosomal arrays were reconstituted with histone octamers containing mutated H2A (H2A-E64C) and H4 (H4-V21C). Linker histone H1 was deposited using histone chaperone NAP-1 in order to induce compaction.

The ICNN approach was first used to investigate the 3D structure a 12x197 bp nucleosome array containing biotin at the nucleosome N5. After chromatin compaction in presence of linker histone H1 and crosslinking of adjacent nucleosome, the fiber was cleaved to monomers and the biotinylated nucleosome N5 was pulled down by streptavidin Chromatin Affinity Precipitation and the DNA analysed by qPCR. The results showed signal for nucleosomes N # 1, 3, 7, 9 and 11 with respectively 8%, 22%, 42%, 24% and 13% of crosslinking. However, once treated with DTT these signal disappeared proving that the disulfide bonds between neighboring nucleosomes are responsible for the signal detected. Nucleosome N5 interacts with its N±2 neighbors, which corresponds to a zigzag structure.

It is clear that the linker histone H1 binds to the nucleosome core particle and promotes compaction of chromatin array into 30 nm fiber. The linker histone has 2 structural domains, the globular domain (GH) and the C-terminal domain. The binding mechanism of the linker histone H1 has been debated for years, as has its effect on the 30 nm structure. To investigate the direct effect of the linker histone H1 on the structure of the 30 nm fiber, we conducted a similar set of experiments where we substituted the H1-dependent compaction by Mg^{2+}-dependent compaction. Indeed, the analysis of the 12x197 bp nucleosome array containing N5-biotin nucleosomes in presence or absence of 1mM divalent cation Mg^{2+} by the ICNN approach showed a similar behavior of the compacted chromatin (Figure 29). Of note, in presence of Mg^{2+} the nucleosome N5 interacts with nucleosomes # 1, 3, 7, 9 and 11 with respectively 6%, 15%, 35%, 20% and 10% of crosslinking. As expected, without Mg^{2+} the nucleosome arrays show no cross-linking.

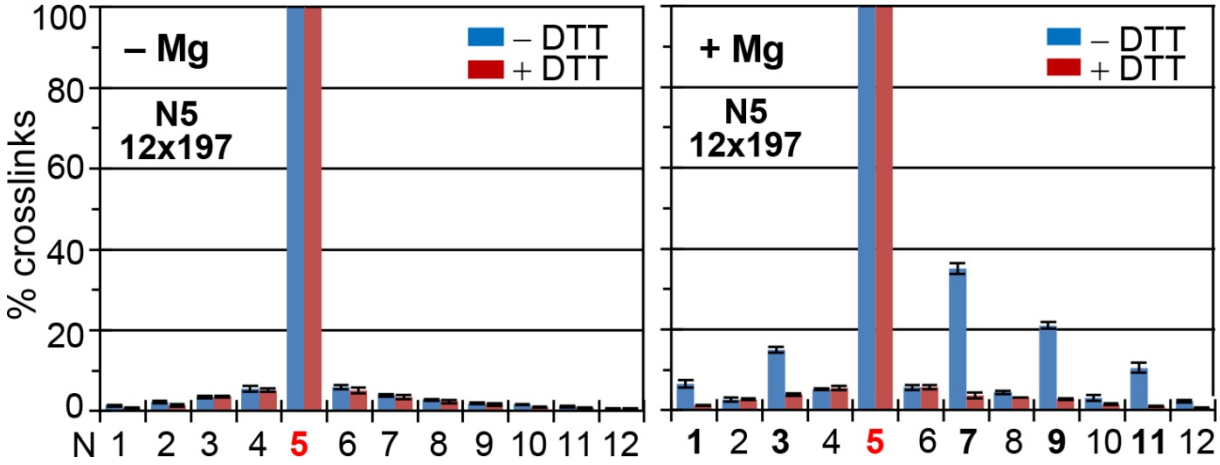


Figure 28: Measured probability of crosslinking of nucleosomes within the condensed 197 bp repeats arrays containing cannonical octamer. Biotin was inserted in nucleosome 8. Left panel, the chromatin display no folding in absence of divalent cation Mg^{2+}. Right panel, the N+2 3D structure of the folded array in presence of divalent cation Mg^{2+}

Furthermore, the effect of the position of the pulled down nucleosome on the 30 nm structure was evaluated by applying the ICNN approach. Three additional 12x197 bp arrays were produced with biotin-labeled either nucleosomes N4, N8 and N9. This demonstrates that in a uniform chromatin, no structural polymorphism is present. The data for all 3 constructs confirmed that the nucleosome N interacts with the nucleosomes N±2 independently of the position of N. These results led us to conclude that in presence of H1 the 12x197 bp nucleosome array folds into a uniform zigzag conformation. Of note, our data did not confirm the two-start with tetranucleosomal units repeat models suggested after cryo-EM analysis [216]. Indeed, if the chromatin was in a tetranucleosomal unit repeat conformation, the interactions between nucleosomes N5 and N7 should be stronger than the interactions between N4 and N6, because the former falls within the same tetranucleosome, while the latter happens within the limits of two tetranucleosomes. However, we did not notice a significant difference of the percentage of crosslinking between N4-labeled arrays and N5-labeled arrays and their respective interacting N±2 partners.

A long-time persisting paradigm is that nucleosome repeat length variation might play a role in both regulating fiber compaction and selective DNA exposure. NRLs are not uniform within a single fiber, but how this variation affects fiber compaction is still unclear. Indeed, it was suggested that short 167 bp NRLs have a limited compaction, which results in a thin fiber [289], and an interdigitated one-start helix for medium (177-207 bp) and long (217-237 bp) NRLs but with different diameters of 33 and 44 nm

respectively [215]. However, chromatin modeling presented a zigzag structure of small and medium NRLs and a solenoid structure for longer NRLs [224].

To investigate this dilemma, we completed our study on NRLs effect by applying the ICNN approach to two additional fibers with different NRLs, 177 and 227 bp, that correspond respectively to the range of short and long NRLs.

As seen, both fibers exhibit the same N±2 (zigzag) conformation as the 197 bp fiber upon H1 compaction. We concluded that the NRL does not affect the 3D organization of the 30 nm fiber, which remains a two-start helix for all repeat lengths tested. Even though these results go against some data claiming different 3D conformations depending on the NRL, it does not give a definitive answer to the main question we are facing. Recent Monte-Carlo simulation data of coarse-grained oligonucleosome models described the chromatin as a polymorphic form that might include solenoid, zigzag, and irregular zigzag within the same non-uniform fiber [290]. However, this simulation takes in consideration data confirming that the small and long NRLs fold into zigzag and solenoid organization respectively, which makes their data bias and their polymophism theory unfounded.

Moreover, our biochemical approach does not allow us to decipher the DNA linker position (crosslinked, helical ribbon) within the zigzag model. Electron microscopy or X-ray scattering in crystal studies are needed for further comprehension of the 3D organization of the 30 nm fiber.

The efficiency and reproducibility of the ICNN techniques allowed us to extend our scope of research to new horizons. In recent years, countless structural studies have been done on histone variants revealing some variant-specific roles in the stability and exposure of the NCP [291]. Much like PTMs, histone variants can engender changes in the interactions with nucleosomal DNA and solvent accessible nucleosome surface. The most studied histone variants in the past few years are H2A.Z, H2A.Bbd (H2A variants) and CENP-A (H3 variant). Of note, information on the 3D structure of NCP organization and on the compaction levels of histone variants-containing arrays is scarce.

The ICNN approach can be used as a powerful biochemical technique that describes the degree of compaction of a given chromatin. With that in mind, we decided to investigate the 3D organization of chromatin fibers containing either H2A.Z or CENP-A. H2A.Z was largely studied for its implication in many important cellular processes such as transcription, DNA repair and mitosis. The crystal structure solved by Luger and

colleagues [164] described the nucleosome core containing H2A.Z to be similar to the canonical with the exception of the extended acidic patch, which might interfere with chromatin folding. However, contradictory data comparing H2A.Z- and H2A-containing arrays described the H2A.Z arrays to be more compact than H2A-containing arrays in one case and less compact in the other.

We used the ICNN technique on reconstituted 12x197 bp arrays with biotin-labeled nucleosome N8 with H1-dependent compaction in the same conditions as canonical arrays. The data showed similar behavior for the H2A.Z containing-array displaying a zigzag organization where nucleosome N interacts with N±2. We did not notice any significant difference leading to believe that the extended acidic patch helps the chromatin to be more or less compact. The signals reported for nucleosomes N8 (# 2, 4, 6, 10, 12) were 8%, 15%, 40%, 24%, 8% of crosslinking respectively for H2A.Z chromatin and 6%, 15%, 42%, 20%, 7% of crosslinking for H2A chromatin. We then concluded that the H2A.Z chromatin displays a zigzag compaction with similar intensity as the H2A chromatin.

In most structural studies concerning array condensation, the chosen DNA template, histone origins and production systems, and the presence of linker histone H1 play a crucial role in determining the structural organization. For example, a biochemical analysis of the folding using 5S DNA template and divalent cations for inducing compaction, shows an accentuated folding into higher-order structure [228]. In our case, the 601 DNA template and linker histone H1-dependent folding were used to perform the experiments, resulting in equal folding levels of H2A.Z- and H2A- containing arrays.

Finally, we focused on the structural changes occasioned by the presence of CENP-A in a nucleosome array. The centromere-specific CENP-A is a H3 histone variant, which is important for recruitment of the kinetochore necessary to the alignment and segregation of the chromosomes. CENP-A proteins are conserved and essential components of the centromere, and thus are the best candidates for epigenetically marking the location of the centromere. In 2011, the crystal structure of NCP containing CENP-A revealed that CENP-A octamers only occupy 121 bp of nucleosomal DNA with the remaining ±13 bp of DNA invisible within the structure. This major structural difference was mainly attributed to a shorter α-N_{CENP-A} helix [190].

In our lab, early studies on CENP-A nucleosome using •OH footprinting and EMSA methods showed that the linker histone H1 do not bind specifically to the CENP-A

nucleosome (data not shown). To investigate the specific role played by the "defective" α-N_{CENP-A} helix in linker H1 binding and linker histone H1-dependent folding, a mutated CENP-A (αN_{H3}-CENP-A) protein containing the α-N_{H3} helix and the preceding loop region of H3 was purified. The one base pair resolution •OH footprinting was used to evaluate the H1 binding to αN_{H3}-CENP-A-containing nucleosome. The data showed a clear recovery of H1 binding onto the nucleosome (not shown). We thus hypothesized that the αN_{H3} helix and the loop 1 are responsible for rigidifying the end of the nucleosomal DNA, which allows binding of the linker histone H1 as seen with canonical nucleosome. These results suggested that in presence of linker histone H1, CENP-A nucleosomal arrays will not fold into higher-order structures. However, H1-dependent folding might be recovered within replacing CENP-A with mutated αN_{H3}- CENP-A.

To investigate these hypotheses, we first used the ICNN approach on reconstituted 12x197 bp arrays with CENP-A octamer containing the H4V21C and H2AE64C mutated histones, allowing intra-fiber crosslinking. The biotin-labeled arrays showed no compaction in presence of linker histone H1. The absence of folding of the CENP-A chromatin is in agreement with our early observations, in which the linker histone H1 does not bind to the CENP-A nucleosome. We speculated that the absence of binding of the linker histone H1 to the CENP-A nucleosomal array leads to an unfolded chromatin. Indeed, the recovery of linker histone binding ability in the mutated αN_{H3}-CENP-A nucleosome led, as expected, to a proper H1 binding and folding of the αN_{H3}-CENP-A array. As seen, the N±2 interactions between nucleosomes were detected similarly to the conventional chromatin. These combined results confirm that the αN helix "defect" of CENP-A is responsible to the absence of H1-induced folding of CENP-A, since H1 cannot stably bind to the nucleosome to induce the linker DNA stem necessary for compaction. However, a large number of questions surrounding the 3D organization remain unanswered. Even though the ICNN technique allowed an unambiguous recognition of the two-start model of the 30 nm fiber, it did not uncover linker DNA arrangement and linker histone H1 binding site. Advancement in cryo-EM and X-ray crystallography in the future will aid the visualization of the 30 nm fiber and the characterization of nucleosome binding enzymes. Despite the extraordinary progress in the past years, it is clear that we only uncovered a small part of the role of nucleosome in coordinating chromatin-templated processes. Additional work will create a new and

heightened understanding of the nucleosome recognition and the role of 30 nm fiber in the organization of eukaryotic genome.

V. Conclusion

This thesis described the development of a new biochemical technique, called the ICNN, which was specifically designed for obtaining structural information. The ICNN approach allowed us by a relatively simple crosslinking experiment to recognize the nucleosome positions within a compacted fiber. Our *in vitro* provide direct unambiguous evidence that uniform 601-arrays, containing properly bound linker histoneH1, fold into a two-start zigzag conformation with a regular nucleosome spacing, thus closing a 40 years old debates and controversy. Interestingly, the regular two-start conformation is invariant in respect to the nucleosome linker length variation within 177 and 227 bp NRL. This finding is important with respect to existing speculations on chromatin structural polymorphisms and its origin. In addition, we established with certitude the implication of histone variants in the organization of the chromatin structure. While H2A.Z fiber did not exhibit any particular changes in the folding conformation compared to H2A fibers, CENP-A fiber displayed a high level of folding inhibition, This finding is in accordance with the observation that linker histone H1 does not stably binds to CENP-A mononucleosomes. This "open" conformation of CENP-A chromatin is likely to be a means for allowing the recruitment of CENPs proteins and the formation of the consecutive centromere associated network.

VI. Future perspective

Understanding chromatin structure is fundamental for understanding its functional role in cellular regulation.

Nowadays, and in near future, special attention will be focused on the following problems:

- The data presented in this thesis point out to a clear edge-dependent asymmetry, meaning that we noticed a more efficient N±2 crosslinking from the longer side of the array. This asymmetry occurs for fibers with different linker lengths and different biotin-labeled nucleosomes. We tentatively attributed it to the finite length of our arrays and the more dynamic conformation at the edges. To directly address the origin of these observations, we will extend the length of our N1-N12 601 arrays by adding at each end one flanking 6x601 repeat. The ICNN detection technique will only "visualize" the nucleosomes N1-N12 located at the central part of the 24-mer array. Since the chromatin is a never-ending array *in vivo*, the physiological relevance of this question is obvious.

- The 3D organization of chromatin containing heterogeneous NRLs. We will create heterogeneous arrays containing a single or multiple nucleosomes with different NRLs and by means of the ICNN technique we will address the fiber organization. The resulting structural disordering, sometimes called polymorphic structure, has only been discussed by mesoscale modeling but has never been assessed by direct experimental approaches, which make these experiments relevant in addition to being closer to the *in vivo* conditions.

- Altering linker length is not the only possibility to create heterogeneity. Incorporation of a single CENP-A nucleosome within a conventional array will also most probably create local structural disturbances. We are wondering whether and how the lack of H1 binding on the CENP-A nucleosome might affect its binding to neighboring conventional nucleosomes and what structural perturbations might be induced. Will the disturbance be localized to the CENP-A nucleosome position or will it affect the whole fiber? These questions will be addressed by a combination of our ICNN approach and cryo-EM imaging. The evidence of mixed H3/CENP-A nucleosomes at centromeres *in vivo* makes this investigation physiologically relevant.

- A combination of our biochemical ICNN method with X-ray diffusion in crystals and cryo-EM imaging appears to be the best choice to gain more structural information on

the chromatin fiber. Efforts in this direction and the necessary collaborations have been already undertaken.

- The most compelling question is the *in vivo* organization of the chromatin fiber. It would be a remarkable achievement to extend the ICNN approach *in vivo*, in order to deepen our understanding of the local chromatin organization in its physiological environment. Besides, such an approach would provide direct experimental answers at the molecular level to many speculations on the role of the chromatin structure in genetic and epigenetic regulation. Efforts in this direction are already undertaken in our lab.

VII. Materials and methods

Chapter 1: DNA Production and purification

1.1 Multiple length array production

1.1.1 Carrier arrays: 601-12X

The 601 Widom DNA sequence (strong positioning DNA sequence) was used as a matrix to produce the 601x12X carrier arrays. AvaI restriction site was introduced at the end of the linker DNA by PCR. Placing the AvaI at different linker lengths allowed the production of arrays with different NRLs. 601x12X with AvaI restriction sites on both sides of the linker allowed the construction of the 12 repeat arrays by ligation.

Ligation was performed for 1h at 25°C By DNA T4 ligase. The agarose gel pattern displayed a ladder-like migration. Each band corresponding to number of repeats (2,3,4, etc.) of the 601 fragments ligated together. The resulting DNA band migrating at around 2400 bp was eluted from the gel and purified with the PCR clean-up kit (Promega).

The 601x12X fragment was then cloned into a pGEMT-easy plasmid (promega). 100 ng of vector and 300 ng of 601x12X repeats were mixed together with a DNA T4 ligase for 2 h at 25°C, transformed into DH-5α *E.coli* bacteria, plated on ampicillin-containing agar plates, and incubated overnight at 37°C. Single colonies were picked and incubated overnight in LB medium. Plasmids were extracted by miniprep and the presence of multimers was verified by an EcoRI digestion that releases the entire array. 3 different linker length were chosen for the 12 repeat arrays: 30, 50, and 80 bp (respectively 177, 197, and 227 bp NRLs).

1.1.2 N1-N12 array design.

1.1.2.1 Designing experiment.

Based on the 601 Widom sequence of the 147 bp corresponding to core DNA, 12 clones named N1 to N12 were designed, with 3 unique mutations on each end of the core DNA (illustrated in figure 30). These mutations will allow specific amplification of the repeat of interest (produced by MWG (Eurofins) and delivered in the pMAT vector). Each DNA repeat is separated by a blunt ScaI cleavage site. Each individual repeat has an inserted distinct restriction enzyme (other than ScaI). This distinct restriction site allows the specific replacement of a selected repeat with a biotin-labeled repeat.

1.1.2.2 Assembling the N1-N12 array

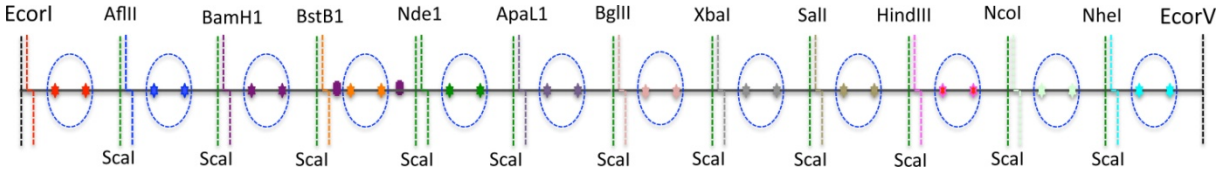


Figure 29: Schematic representation of the restriction enzyme sites in the N1-N12 repeats.

Production PCRs: PCR reactions were performed 10 times in a total reaction volume of 100 µl, containing 1x Pfu polymerase buffer, 250 µM dNTPs, 20 µM primers and 2.5 units Pfu DNA polymerase. The following program was used on the Uno cycler (VWR): 95°C for 3 min, 34 cycles of : 94°C for 40 sec, Primer melting temperature (Tm) 65 °C for 40 sec, and 72°C for 45 sec, followed by a final extension step at 72°C for 5 min. PCR products were analyzed on 1.2% agarose gels. DNA was extracted by PCR clean-up kit (Promega) and approximately 60 µg were digested for 3 h in a reaction volume of 200 µl with 150 units restriction enzyme (NEB) and reaction buffer. The reaction was stopped by phenol/chloroform/isoamyl alcohol (25 :24 :1), and precipitated with ethanol and resuspended in TE (100 mM Tris pH 8.0, 0.1 mM EDTA). 3 M sodium acetate pH 5.2 was added to reach a 300 mM salt concentration. The reaction was mixed with 3 volumes of absolute ethanol to a total concentration of 70% ethanol. The pellet was collected by centrifugation for 20 min at 21.000 g at 4°C. The supernatant was immediately removed and the pellet resuspended in 70% ethanol for washing. The DNA was pelleted again for 10 min at room temperature. The supernatant was removed and the pellet resuspended in desired buffer after air-drying.

Ligation reactions: 50 µg of each fragment were mixed with 15 µl of 10 ligase buffer and 7 µl of ligase (25 weiss units) in a total volume of 150 µl, and incubated overnight at 4°c. The reaction was stopped by adding an equivalent amount of phenol/chloroforom/isoamyl alcohol (25 :24 :1) and precipitated with ethanol.

N1-N12 repeats Amplification : 4 L of LB were inoculated with colonies and incubated at 37°C under stirring at 250 rpm overnight. The bacteria were pelleted at 4000 rpm for 20 min at 4°C. The plasmid was extracted by Giga prep nucleobond PC10000 columns (Macherey-Nagel).

N1-N12 repeats purification: 5ml of N1-N12 plasmid containing around 10 mg of DNA were mixed with 1ml of 10x restriction buffer (NEB Buffer 4), 4000 U of EcoRI-HF, 4000 U of EcoRV-HF, 4000 U of DraI and 4000 U of HaeII in 10ml total volume was incubated overnight at 37°C. The desired fragement was purified from aragorese gel

1.1.3 N1-N12 REPEATS with one repeat biotin labeled

4 biotin-labeled arrays were prepared, each one on a different fragment (N4, N5, N8 and N9).

# Of the modified fragment		Primers sequence
N4	Forward	5' CATCAGTACTAGGTCTTCGAACAATACATGCXCAGGATGTA 3'
	Reverse	5' GTGCATGTATTGACATATGACCTAGTACTGAXGGACCCTATACG 3'
N5	Forward	5' CTAGGTCATATGTCAATACATGCXCAGGATG 3'
	Reverse	5' TATTGAACGTGCACCTAGTACTGAXGGACCCTATACGC 3'
N8	Forward	5' CATCAGTACTAGGTTCTAGATCAATACATGCXCAGGATGTA 3'
	Reverse	5' GTGCATGTATTGAGTCGACACCTAGTACTGAXGGACCCTATACG 3'
N9	Forward	5' CATCAGTACTAGGTGTCGACTCAATACATGCXCAGGATGTA 3'
	Reverse	5' GTGCATGTATTGAAGCTTGACCTAGTACTGAXGGACCCTATACG 3'

Table 1: PCR primers used for inserting a biotin label. **X** : biotin position.

PCR using biotin-labeled primers were performed in the same conditions described earlier.

Chapter 2: Protein production

2.1 *Xenopus* core histones

The histone mutants H4V21C, H3C111A, H2AE64C and the wild-type H2B were produced using a modified version of the Luger procedure. 200 ng of plasmid (for each histone) were used to transform BL21 (DE3) pLYsS bacteria. Bacteria was induced with 0.2 mM IPTG (isopropyl-beta-D-thiogalactopyranoside) at 37°C for 3-4 hours at 200 rpm. Each liter of bacteria was pelleted at 5000g for 20 min at room temperature. 20 ml of wash buffer containing 50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1mM Na-EDTA pH8, 1mM benzamidine and 10 mM ß-mercaptoethanol was added per liter of pelleted bacteria. Once the pellet were resuspended, the mixture was sonicated at 70%

amplitude for 20 min and then pelleted for 30 min at maximum speed at 4°C. The supernatant was discarded and the pellet resuspended with wash buffer supplemented with 1% Triton X-100.

The mixture was pelleted again for 30 min at maximum speed at 4°C. This step was repeated twice and then 3 additional times without Triton X-100. After careful washing, the pellet was resuspended in 10 ml guanidinium buffer containing 7 M guanidinium-HCl, 20 mM Tris-HCl pH 7.4 and 10 mM DTT, incubated for 1 h at 4°C with gentle rotation (5-7 rpm) and then centrifuged for 30 min at maximum speed at room temperature. The supernatant was recovered gently and kept aside at room temperature. This procedure was repeated once. The two supernatants recovered were pooled and dialyzed overnight against SAU100 buffer (7 M Urea, 20 mM sodium acetate pH 5.2, 100 mM NaCl, 5mM β-mercaptoethanol and 1 mM Na-EDTA). The next day, samples were spun at 8000 g for 5 min before loading the supernatant onto an SP sepharose cation exchange column (GE Healthcare). The elution was done by SAU buffer with increasing NaCl concentration, ranging from 100 mM to 1500 mM. After checking on 18 % SDS-PAGE gel, the fractions containing the purified protein (500 to 700 mM NaCl) were mixed together. Before loading the purified protein to the ion exchange column Resource S 6 ml (GE Healthcare), the NaCl concentration was reduced to around 70-100mM by buffer exchange. The elution of the histone was done with a linear gradient ranging from 0.1 to 1 M NaCl in SAU buffer. After checking purity on SDS gel, these pure fractions were mixed and stored at -80°C.

2.2 Variant histones

CENP-A C75A, αN_{H3}CENP-A C75A and H2A.Z E64C were all sub-cloned into pET28 plasmid, which allows the production of 6x His-tagged histones (Tanaka Y et al 2004). BL21 (DE3) pLYsS we amplified in the same conditions as for conventional core histone.

Chapter 3: Chromatin reconstitution and chromatin check

3.1 Histone octamer

The histone cores were all mixed in an equimolar ratio, and dialyzed overnight in HFB buffer (2M NaCl, 10 mM Tris pH7.4, 1 mM EDTA pH 8 and 10 mM β-mercaptoethanol).

H3C111A, 12 μl H2AE65C, 15 μl of H2B and 46 μl of H4V21C were mixed with 50 μl of 8 M urea and 20 μl of 1 M DTT to have a total volume of 160 μl. After dialysis, 40 μl of NaCl saturated glycerol was added to the mixture, adding up the volume to 200 μl, which makes the octamer concentration 0.5 μg/μl.

3.2 Chromatin reconstitution

For 20 μg reconstituted chromatin, 20 μg of the 601 wt arrays and 1 μg of biotinilated arrays were mixed with 4 μg of chicken carrier DNA (20 % of total amount of DNA), ~20 μg octamer, 2 M NaCl, 10 mM Tris pH7.4, 1 mM EDTA and 100 mM DTT. The mix was transferred into dialysis tubing and the reconstitution was done by dialysis against a slowly decreasing salt buffer. The NaCl concentration starts at 2 M and decreases slowly up to 500 mM NaCl. Indeed, with the help of a peristaltic pump, low salt buffer is added to the high salt buffer beaker at the rate of 1.5 ml/min for 18 h. Once finished, the dialysis bags were transferred to a 300 mM NaCl buffer and left for buffer exchange for 2 h, which was followed by a final dialysis in 10 mM NaCl buffer overnight. All NaCl buffers for reconstitution include 10 mM Tris pH 7.4, 0.25 mM EDTA, 10 mM β-mercaptoethanol and the desired amounts of NaCl.

3.3 Chromatin check by restriction enzymes

300 ng of chromatin were mixed with 0.5 mM $MgCl_2$ and 10 units of AvaI or HhaI and incubated at 37°C for 2 h. AvaI restriction site is present in the linker DNA and HhaI restriction site is present close to the dyad site in the core histone DNA. In case of undersaturation, the histone octamer is not sufficient to wrap the entire DNA amount, which gives access for HhaI digestion. In case of oversaturation, the excess of histone binds non-specifically to the linker DNA blocking the access for AvaI. Exact nucleosome saturation is achieved (i.e. exact equilibrium between the DNA and the octamer) when the array is fully digested to mononucleosomes by Ava I, while no any digestion is observed with HhaI.

3.4 Chromatin crosslinking and affinity precipitation

25 μg of nucleosomal array was first treated with 100 mM DTT in order to break the pre-existing disulfide bonds on ice for 2hrs.

3.4.1 Chromatin compaction by H1 deposition

Histone H1 mixed with histone chaperone Nap1 at 1:2 molar ratio was incubated for 30 min at 30°C in the following buffer: 20 mM Tris–HCl pH 7.5, 0.5 mM EDTA, 100 mM NaCl, 1mM DTT, 10% glycerol, 0.1mM PMSF, added to the chromatin (at a concentration of 40 ng/μl) at a 1.25: 1 ratio and incubated for 1h at 30°C in a total volume of 100 μl .

3.4.2 Crosslinking with glutathion

Chromatin was first dialyzed against crosslinking buffer (10 mM Tris pH 9, 50 mM NaCl, 250 μM EDTA) for 3 h at room temperature using the cellusepT2 dialysis tubing 10 mm width. The purpose of this dialysis is to eliminate the DTT present in the chromatin mixture, whereas its presence will block the crosslinking. After dialysis, the chromatin (30 ng/μl) was adjusted to 25 μM of oxidized glutathione, 50 mM NaCl, 10 mM Tris pH 9 and was incubated at 37°C overnight for histone-histone crosslinking.

3.4.3 Digestion and DTT treatment

Before ScaI-AvaI digestion, the crosslinked chromatin was incubated with 2.5 mM iodoacetamide at room temperature in the dark for 1 h. Later on, the mixture was dialyzed against NT buffer containing 50 mM NaCl, 10mM Tris pH 7.4 for 5 h at 4°C. After dialysis, 5μl of highly concentrated NapI (9.2 μg/μl) was added in order to remove the linker Histone H1 from the chromatin. Chromatin was digested with 250 units of AvaI, 100 units of ScaI and 25 units of XbaI (methylation site) overnight at 37°C in 200 μl of buffer containing 10 mM Tris pH7.4, 50 mM NaCl, 1.25mM $MgCl_2$. The next day, each tube was separated in two 100 μl each. In one of them, 10 μl of 1M DTT was added and in the other one 10 μl of NT buffer (50 mM NaCl, 10 mM Tris pH7.4). Reactions were incubated at room temperature for 2 h. The digestion and efficiency of DTT treatment was checked on 1% agarose gel with 0.5X TBE before proceeding to affinity precipitation.

3.4.4 Affinity precipitation

To each tube containing 2 μg of chromatin (-/+DTT), 1.5 ng of IC nucleosomes (estimated amount of the N1-N12 present in each tube) and 1 μl of 1% NP40 were

added. 5% of the total volume was taken for input. In this case, 5 µl was kept at 4°C for later use. The rest was mixed to blocked streptavidin magnetic beads.

3.4.4.2 Blocking streptavidine beads

5µl Dynabeads M-280 Streptavidin (Life technology) were added to 1.5 ml siliconized Eppendorf tubes and mixed with 200 µl of blocking buffer containing 10 mM Tris pH 7.4, 1 mg/ml BSA, 50 mM NaCl, 0.01% NP40. Magnetic beads were shook at 23 rpm to prevent them from settling, which would limit the blocking process. Beads were separated from supernatant on a Magnetic Particle concentrator (Dynabeads MCP - Life Technologies). The blocking step was repeated twice for 20 minute before proceeding to AP. Once the beads were ready to use, chromatin was added for binding overnight at 4°C under 20 rpm shaking. Supernatant was discarded, and beads were washed 3 times for 30 min with a buffer containing 10 mM Tris pH7.4, 50 mM NaCl, 100 µM EDTA pH 8, 100 µg/ml BSA, 0.01% NP40 and 10 ng/µl erythrocyte chicken nucleosome. The beads were then subjected to proteinase K treatment in order to recover the attached DNA. Indeed, 200 µl of STOP buffer were added to the beads alongside 1 µl of proteinase K (934 U/ml - Fermentas) and incubated at 55°C for 2 h in an hybridization oven for its capacity to provide a rotation of the tubes and prevent the beads from settling down. Alongside the tubes containing the beads, the input tubes were subjected to the same treatment (stop buffer and proteinase K for 2 hours at 55°C). The supernatant was collected and DNA extracted by phenol/chloroform and precipitated by ethanol.

3.4.5 DNA quantification and qPCR.

After DNA precipitation, The total DNA was quantified using the Quant-it Picogreen dsDNA assay kit (Invitrogen). The concentration of each sample (usually between 1.5 and 2 ng/µl) was determined with the help of a standard calibration curve. The DNA was diluted up to 16.6 pg/µl with ultrapure Millipore water and 3 µl (50 pg) were added to each well of qPCR 196-well plates (Eurogentec), together with 1x Sybr Green Mix (Fast start Universal SYBR Green (Roche)) and 10 pmol/µl detection primers (Figure30) and completed to 15 µl with ultrapure water. The amplification was done on a StepOnePlus real-time PCR system (Applied Biosystem).

VIII. Bibliography

1. Flemming, W., Leipzig, 1882.
2. Hughes, A., A history of cytology. 1959, London: Albelard-schuman.
3. Miescher-R√ºsch, F., Ueber die chemische Zusammensetzung der Eiterzellen. 1871.
4. Kossel, A., Ueber die chemische Beschaffenheit des Zellkerns. Munchen Med. Wochenschrift 1911. 58: p. 65-69.
5. Mazia, D., enzyme studies of chromosomes, in cold spring harb. symp. quant. biol. 1941: cold spring. p. 40-46.
6. schultz, j., The evidence of nucleoprotein nature of the gene. Cold Spring Harb Symp Quant Biol, 1941. IX: p. 55-65.
7. Pauling, L. and R.B. Corey, The structure of synthetic polypeptides. Proc Natl Acad Sci U S A, 1951. 37(5): p. 241-50.
8. Watson, J.D. and F.H. Crick, Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature, 1953. 171(4356): p. 737-8.
9. Wilkins, M.H., A.R. Stokes, and H.R. Wilson, Molecular structure of deoxypentose nucleic acids. Nature, 1953. 171(4356): p. 738-40.
10. Franklin, R.E. and R.G. Gosling, Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. Nature, 1953. 172(4369): p. 156-7.
11. Zubay, G. and M.R. Watson, The absence of histone in the bacterium Escherichia coli. I. Preparation and analysis of nucleoprotein extract. J Biophys Biochem Cytol, 1959. 5(1): p. 51-4.
12. Gall, J.G., Kinetics of deoxyribonuclease action on chromosomes. Nature, 1963. 198: p. 36-8.
13. Johns, E.W. and T.A. Hoare, Histones and gene control. Nature, 1970. 226(5246): p. 650-1.
14. Allfrey, V.G., R. Faulkner, and A.E. Mirsky, Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. Proc Natl Acad Sci U S A, 1964. 51: p. 786-94.
15. Olins, A.L. and D.E. Olins, Spheroid chromatin units (v bodies). Science, 1974. 183(4122): p. 330-2.
16. woodcock, C.L.F., Ultrastrucutre of inactive chromatin. J Cell Biol, 1973. 59: p. A368.
17. Kornberg, R.D., Chromatin structure: a repeating unit of histones and DNA. Science, 1974. 184(4139): p. 868-71.
18. Oudet, P., M. Gross-Bellard, and P. Chambon, Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell, 1975. 4(4): p. 281-300.
19. Olins, D.E. and A.L. Olins, Chromatin history: our view from the bridge. Nat Rev Mol Cell Biol, 2003. 4(10): p. 809-14.
20. Luger, K., et al., Characterization of nucleosome core particles containing histone proteins made in bacteria. J Mol Biol, 1997. 272(3): p. 301-11.
21. Martins, R.P., et al., Mechanical regulation of nuclear structure and function. Annu Rev Biomed Eng. 14: p. 431-55.
22. Woodcock, C.L., L.L. Frado, and J.B. Rattner, The higher-order structure of chromatin: evidence for a helical ribbon arrangement. J Cell Biol, 1984. 99(1 Pt 1): p. 42-52.

23. Finch, J.T. and A. Klug, Solenoidal model for superstructure in chromatin. Proc Natl Acad Sci U S A, 1976. 73(6): p. 1897-901.
24. Widom, J. and A. Klug, Structure of the 300A chromatin filament: X-ray diffraction from oriented samples. Cell, 1985. 43(1): p. 207-13.
25. Dorigo, B., et al., Nucleosome arrays reveal the two-start organization of the chromatin fiber. Science, 2004. 306(5701): p. 1571-3.
26. Grigoryev, S.A., et al., Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. Proc Natl Acad Sci U S A, 2009. 106(32): p. 13317-22.
27. Kireeva, M.L., et al., Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription. Mol Cell, 2002. 9(3): p. 541-52.
28. Hu, Y., et al., Large-scale chromatin structure of inducible genes: transcription on a condensed, linear template. J Cell Biol, 2009. 185(1): p. 87-100.
29. Wiki, M. From DNA to metaphase chromosome. 2014; Available from: http://www.mechanobio.info/figure/figure/1389944285515.jpg.html.
30. Frenster, J.H., V.G. Allfrey, and A.E. Mirsky, Repressed and Active Chromatin Isolated from Interphase Lymphocytes. Proc Natl Acad Sci U S A, 1963. 50: p. 1026-32.
31. Grewal, S.I. and S.C. Elgin, Heterochromatin: new possibilities for the inheritance of structure. Curr Opin Genet Dev, 2002. 12(2): p. 178-87.
32. Sun, F.L., M.H. Cuaycong, and S.C. Elgin, Long-range nucleosome ordering is associated with gene silencing in Drosophila melanogaster pericentric heterochromatin. Mol Cell Biol, 2001. 21(8): p. 2867-79.
33. Sugiyama, T., et al., SHREC, an effector complex for heterochromatic transcriptional silencing. Cell, 2007. 128(3): p. 491-504.
34. Grewal, S.I. and S. Jia, Heterochromatin revisited. Nat Rev Genet, 2007. 8(1): p. 35-46.
35. Lu, B.Y., et al., Heterochromatin protein 1 is required for the normal expression of two heterochromatin genes in Drosophila. Genetics, 2000. 155(2): p. 699-708.
36. Ris, H. and J. Korenberg, Chromosome structure and levels of chromosome organization. cell biol, 1979. 2: p. 267-361.
37. Fahrner, J.A. and S.B. Baylin, Heterochromatin: stable and unstable invasions at home and abroad. Genes Dev, 2003. 17(15): p. 1805-12.
38. Grewal, S.I. and D. Moazed, Heterochromatin and epigenetic control of gene expression. Science, 2003. 301(5634): p. 798-802.
39. Craig, J.M., Heterochromatin--many flavours, common themes. Bioessays, 2005. 27(1): p. 17-28.
40. Shi, Y., Histone lysine demethylases: emerging roles in development, physiology and disease. Nat Rev Genet, 2007. 8(11): p. 829-33.
41. Shilatifard, A., Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annu Rev Biochem, 2006. 75: p. 243-69.
42. Grewal, S.I. and S.C. Elgin, Transcription and RNA interference in the formation of heterochromatin. Nature, 2007. 447(7143): p. 399-406.
43. Nakatsu, S.L., et al., Activity of DNA templates during cell division and cell differentiation. Nature, 1974. 248(446): p. 334-5.
44. Feldman, N., et al., G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nat Cell Biol, 2006. 8(2): p. 188-94.

45. Heard, E., Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome. Curr Opin Genet Dev, 2005. 15(5): p. 482-9.
46. Skok, J.A., et al., Reversible contraction by looping of the Tcra and Tcrb loci in rearranging thymocytes. Nat Immunol, 2007. 8(4): p. 378-87.
47. Talbert, P.B. and S. Henikoff, Spreading of silent chromatin: inaction at a distance. Nat Rev Genet, 2006. 7(10): p. 793-803.
48. Henikoff, S., Heterochromatin function in complex genomes. Biochim Biophys Acta, 2000. 1470(1): p. O1-8.
49. Rusche, L.N., A.L. Kirchmaier, and J. Rine, The establishment, inheritance, and function of silenced chromatin in Saccharomyces cerevisiae. Annu Rev Biochem, 2003. 72: p. 481-516.
50. Eichler, E.E., Repetitive conundrums of centromere structure and function. Hum Mol Genet, 1999. 8(2): p. 151-5.
51. Earnshaw, W., et al., Three human chromosomal autoantigens are recognized by sera from patients with anti-centromere antibodies. J Clin Invest, 1986. 77(2): p. 426-30.
52. Earnshaw, W.C. and C.A. Cooke, Proteins of the inner and outer centromere of mitotic chromosomes. Genome, 1989. 31(2): p. 541-52.
53. Palmer, D.K., et al., Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. Proc Natl Acad Sci U S A, 1991. 88(9): p. 3734-8.
54. Pierce, b.A., Genetics: A Conceptual Approach. 2nd edition ed. 2012: W.H. Freeman. 745.
55. Greider, C.W., Telomere length regulation. Annu Rev Biochem, 1996. 65: p. 337-65.
56. Blackburn, E.H., Telomeres and telomerase: their mechanisms of action and the effects of altering their functions. FEBS Lett, 2005. 579(4): p. 859-62.
57. Lingner, J., et al., Reverse transcriptase motifs in the catalytic subunit of telomerase. Science, 1997. 276(5312): p. 561-7.
58. Richmond, T.J. and C.A. Davey, The structure of DNA in the nucleosome core. Nature, 2003. 423(6936): p. 145-50.
59. Davey, C.A., et al., Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. J Mol Biol, 2002. 319(5): p. 1097-113.
60. Hayashi, K., T. Hofstaetter, and N. Yakuwa, Asymmetry of chromatin subunits probed with histone H1 in an H1-DNA complex. Biochemistry, 1978. 17(10): p. 1880-3.
61. Stuart, D., The Mechanisms of DNA Replication: InTech.
62. Arents, G., et al., The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. Proc Natl Acad Sci U S A, 1991. 88(22): p. 10148-52.
63. Alberts, B., Molecular Biology of the Cell: Reference edition. 2008: Garland Science.
64. Xie, X., et al., Structural similarity between TAFs and the heterotetrameric core of the histone octamer. Nature, 1996. 380(6572): p. 316-22.
65. Luger, K., et al., Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature, 1997. 389(6648): p. 251-60.

66. Ransom, M., B.K. Dennehey, and J.K. Tyler, Chaperoning histones during DNA replication and repair. Cell. 140(2): p. 183-95.
67. Angelov, D., et al., Preferential interaction of the core histone tail domains with linker DNA. Proc Natl Acad Sci U S A, 2001. 98(12): p. 6599-604.
68. Yuan, G.C., et al., Genome-scale identification of nucleosome positions in S. cerevisiae. Science, 2005. 309(5734): p. 626-30.
69. Valouev, A., et al., Determinants of nucleosome organization in primary human cells. Nature. 474(7352): p. 516-20.
70. Zaret, K.S. and J.S. Carroll, Pioneer transcription factors: establishing competence for gene expression. Genes Dev. 25(21): p. 2227-41.
71. Drew, H.R. and A.A. Travers, DNA bending and its relation to nucleosome positioning. J Mol Biol, 1985. 186(4): p. 773-90.
72. Lowary, P.T. and J. Widom, New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J Mol Biol, 1998. 276(1): p. 19-42.
73. Hansen, J.C., Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. Annu Rev Biophys Biomol Struct, 2002. 31: p. 361-92.
74. Cairns, B.R., The logic of chromatin architecture and remodelling at promoters. Nature, 2009. 461(7261): p. 193-8.
75. Johns, E.W. and S. Forrester, Studies on nuclear proteins. The binding of extra acidic proteins to deoxyribonucleoprotein during the preparation of nuclear proteins. Eur J Biochem, 1969. 8(4): p. 547-51.
76. Bonne, C., M. Duguet, and A.M. de Recondo, Single-strand DNA binding protein from rat liver: interactions with supercoiled DNA. Nucleic Acids Res, 1980. 8(21): p. 4955-68.
77. Hartman, P.G., et al., Studies on the role and mode of operation of the very-lysine-rich histone H1 in eukaryote chromatin. The three structural regions of the histone H1 molecule. Eur J Biochem, 1977. 77(1): p. 45-51.
78. Cole, R.D., Microheterogeneity in H1 histones and its consequences. Int J Pept Protein Res, 1987. 30(4): p. 433-49.
79. Ramakrishnan, V., et al., Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. Nature, 1993. 362(6417): p. 219-23.
80. Simpson, R.T., Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. Biochemistry, 1978. 17(25): p. 5524-31.
81. Thomas, J.O., Histone H1: location and role. Curr Opin Cell Biol, 1999. 11(3): p. 312-7.
82. Staynov, D.Z. and C. Crane-Robinson, Footprinting of linker histones H5 and H1 on the nucleosome. EMBO J, 1988. 7(12): p. 3685-91.
83. Syed, S.H., et al., Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome. Proc Natl Acad Sci U S A. 107(21): p. 9620-5.
84. An, W., et al., Linker histone protects linker DNA on only one side of the core particle and in a sequence-dependent manner. Proc Natl Acad Sci U S A, 1998. 95(7): p. 3396-401.
85. Brown, D.T., T. Izard, and T. Misteli, Mapping the interaction surface of linker histone H1(0) with the nucleosome of native chromatin in vivo. Nat Struct Mol Biol, 2006. 13(3): p. 250-5.

86. Shintomi, K., et al., Nucleosome assembly protein-1 is a linker histone chaperone in Xenopus eggs. Proc Natl Acad Sci U S A, 2005. 102(23): p. 8210-5.
87. Hansen, J.C., et al., Intrinsic protein disorder, amino acid composition, and histone terminal domains. J Biol Chem, 2006. 281(4): p. 1853-6.
88. Subirana, J.A., Analysis of the charge distribution in the C-terminal region of histone H1 as related to its interaction with DNA. Biopolymers, 1990. 29(10-11): p. 1351-7.
89. Allan, J., et al., Roles of H1 domains in determining higher order chromatin structure and H1 location. J Mol Biol, 1986. 187(4): p. 591-601.
90. Hendzel, M.J., et al., The C-terminal domain is the primary determinant of histone H1 binding to chromatin in vivo. J Biol Chem, 2004. 279(19): p. 20028-34.
91. Liao, L.W. and R.D. Cole, Condensation of dinucleosomes by individual subfractions of H1 histone. J Biol Chem, 1981. 256(19): p. 10124-8.
92. Talasz, H., et al., In vitro binding of H1 histone subtypes to nucleosomal organized mouse mammary tumor virus long terminal repeat promotor. J Biol Chem, 1998. 273(48): p. 32236-43.
93. Verdaguer, N., et al., Helical structure of basic proteins from spermatozoa. Comparison with model peptides. Eur J Biochem, 1993. 214(3): p. 879-87.
94. Caterino, T.L. and J.J. Hayes, Structure of the H1 C-terminal domain and function in chromatin condensation. Biochem Cell Biol. 89(1): p. 35-44.
95. Shimamura, A., et al., Histone H1 represses transcription from minichromosomes assembled in vitro. Mol Cell Biol, 1989. 9(12): p. 5573-84.
96. Brown, D.T., B.T. Alexander, and D.B. Sittman, Differential effect of H1 variant overexpression on cell cycle progression and gene expression. Nucleic Acids Res, 1996. 24(3): p. 486-93.
97. Woodcock, C.L., A.I. Skoultchi, and Y. Fan, Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. Chromosome Res, 2006. 14(1): p. 17-25.
98. Th'ng, J.P., et al., H1 family histones in the nucleus. Control of binding and localization by the C-terminal domain. J Biol Chem, 2005. 280(30): p. 27809-14.
99. Misteli, T., et al., Dynamic binding of histone H1 to chromatin in living cells. Nature, 2000. 408(6814): p. 877-81.
100. Phair, R.D., et al., Global nature of dynamic protein-chromatin interactions in vivo: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. Mol Cell Biol, 2004. 24(14): p. 6393-402.
101. Raghuram, N., et al., Molecular dynamics of histone H1. Biochem Cell Biol, 2009. 87(1): p. 189-206.
102. Lever, M.A., et al., Rapid exchange of histone H1.1 on chromatin in living human cells. Nature, 2000. 408(6814): p. 873-6.
103. Thoma, F., T. Koller, and A. Klug, Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. J Cell Biol, 1979. 83(2 Pt 1): p. 403-27.
104. Bednar, J., et al., Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. Proc Natl Acad Sci U S A, 1998. 95(24): p. 14173-8.
105. Clark, D.J. and T. Kimura, Electrostatic mechanism of chromatin folding. J Mol Biol, 1990. 211(4): p. 883-96.
106. Shen, X., et al., Linker histones are not essential and affect chromatin condensation in vivo. Cell, 1995. 82(1): p. 47-56.

107. Bates, D.L. and J.O. Thomas, Histones H1 and H5: one or two molecules per nucleosome? Nucleic Acids Res, 1981. 9(22): p. 5883-94.
108. Fan, Y., et al., Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. Cell, 2005. 123(7): p. 1199-212.
109. Carozzi, N., et al., Clustering of human H1 and core histone genes. Science, 1984. 224(4653): p. 1115-7.
110. Eick, S., et al., Human H1 histones: conserved and varied sequence elements in two H1 subtype genes. Eur J Cell Biol, 1989. 49(1): p. 110-5.
111. Albig, W., et al., Isolation and characterization of two human H1 histone genes within clusters of core histone genes. Genomics, 1991. 10(4): p. 940-8.
112. Albig, W., T. Meergans, and D. Doenecke, Characterization of the H1.5 gene completes the set of human H1 subtype genes. Gene, 1997. 184(2): p. 141-8.
113. Drabent, B., et al., Isolation of two murine H1 histone genes and chromosomal mapping of the H1 gene complement. Mamm Genome, 1995. 6(8): p. 505-11.
114. Tanaka, H., et al., Expression profiles and single-nucleotide polymorphism analysis of human HANP1/H1T2 encoding a histone H1-like protein. Int J Androl, 2006. 29(2): p. 353-9.
115. Tanaka, M., et al., A mammalian oocyte-specific linker histone gene H1oo: homology with the genes for the oocyte-specific cleavage stage histone (cs-H1) of sea urchin and the B4/H1M histone of the frog. Development, 2001. 128(5): p. 655-64.
116. Yan, W., et al., HILS1 is a spermatid-specific linker histone H1-like protein implicated in chromatin remodeling during mammalian spermiogenesis. Proc Natl Acad Sci U S A, 2003. 100(18): p. 10546-51.
117. Happel, N., E. Schulze, and D. Doenecke, Characterisation of human histone H1x. Biol Chem, 2005. 386(6): p. 541-51.
118. Doenecke, D. and R. Tonjes, Differential distribution of lysine and arginine residues in the closely related histones H1 and H5. Analysis of a human H1 gene. J Mol Biol, 1986. 187(3): p. 461-4.
119. Happel, N. and D. Doenecke, Histone H1 and its isoforms: contribution to chromatin structure and function. Gene, 2009. 431(1-2): p. 1-12.
120. Marzluff, W.F., Metazoan replication-dependent histone mRNAs: a distinct set of RNA polymerase II transcripts. Curr Opin Cell Biol, 2005. 17(3): p. 274-80.
121. Franke, K., B. Drabent, and D. Doenecke, Expression of murine H1 histone genes during postnatal development. Biochim Biophys Acta, 1998. 1398(3): p. 232-42.
122. Kim, K., et al., Isolation and characterization of a novel H1.2 complex that acts as a repressor of p53-mediated transcription. J Biol Chem, 2008. 283(14): p. 9113-26.
123. Konishi, A., et al., Involvement of histone H1.2 in apoptosis induced by DNA double-strand breaks. Cell, 2003. 114(6): p. 673-88.
124. Meergans, T., W. Albig, and D. Doenecke, Varied expression patterns of human H1 histone genes in different cell lines. DNA Cell Biol, 1997. 16(9): p. 1041-9.
125. De Lucia, F., et al., Histone-induced condensation of rat testis chromatin: testis-specific H1t versus somatic H1 variants. Biochem Biophys Res Commun, 1994. 198(1): p. 32-9.
126. Tanaka, M., et al., H1oo: a pre-embryonic H1 linker histone in search of a function. Mol Cell Endocrinol, 2003. 202(1-2): p. 5-9.
127. Gabrielli, F., et al., Histone complements of human tissues, carcinomas, and carcinoma-derived cell lines. Mol Cell Biochem, 1984. 65(1): p. 57-66.

128. Zlatanova, J. and D. Doenecke, Histone H1 zero: a major player in cell differentiation? FASEB J, 1994. 8(15): p. 1260-8.
129. Stoldt, S., et al., G1 phase-dependent nucleolar accumulation of human histone H1x. Biol Cell, 2007. 99(10): p. 541-52.
130. Bohm, L. and C. Crane-Robinson, Proteases as structural probes for chromatin: the domain structure of histones. Biosci Rep, 1984. 4(5): p. 365-86.
131. Allan, J., et al., Participation of core histone "tails" in the stabilization of the chromatin solenoid. J Cell Biol, 1982. 93(2): p. 285-97.
132. Schwarz, P.M., et al., Reversible oligonucleosome self-association: dependence on divalent cations and core histone tail domains. Biochemistry, 1996. 35(13): p. 4009-15.
133. Gordon, F., K. Luger, and J.C. Hansen, The core histone N-terminal tail domains function independently and additively during salt-dependent oligomerization of nucleosomal arrays. J Biol Chem, 2005. 280(40): p. 33701-6.
134. Kouzarides, T., Chromatin modifications and their function. Cell, 2007. 128(4): p. 693-705.
135. Mutskov, V., et al., Persistent interactions of core histone tails with nucleosomal DNA following acetylation and transcription factor binding. Mol Cell Biol, 1998. 18(11): p. 6293-304.
136. Usachenko, S.I., et al., Rearrangement of the histone H2A C-terminal domain in the nucleosome. Proc Natl Acad Sci U S A, 1994. 91(15): p. 6845-9.
137. Lee, K.M. and J.J. Hayes, Linker DNA and H1-dependent reorganization of histone-DNA interactions within the nucleosome. Biochemistry, 1998. 37(24): p. 8622-8.
138. Lee, K.M. and J.J. Hayes, The N-terminal tail of histone H2A binds to two distinct sites within the nucleosome core. Proc Natl Acad Sci U S A, 1997. 94(17): p. 8959-64.
139. Ebralidse, K.K., S.A. Grachev, and A.D. Mirzabekov, A highly basic histone H4 domain bound to the sharply bent region of nucleosomal DNA. Nature, 1988. 331(6154): p. 365-7.
140. Kan, P.Y. and J.J. Hayes, Detection of interactions between nucleosome arrays mediated by specific core histone tail domains. Methods, 2007. 41(3): p. 278-85.
141. Portela, A. and M. Esteller, Epigenetic modifications and human disease. Nat Biotechnol. 28(10): p. 1057-68.
142. Tse, C., et al., Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. Mol Cell Biol, 1998. 18(8): p. 4629-38.
143. Wang, X. and J.J. Hayes, Acetylation mimics within individual core histone tail domains indicate distinct roles in regulating the stability of higher-order chromatin structure. Mol Cell Biol, 2008. 28(1): p. 227-36.
144. Roth, S.Y., J.M. Denu, and C.D. Allis, Histone acetyltransferases. Annu Rev Biochem, 2001. 70: p. 81-120.
145. Brownell, J.E. and C.D. Allis, Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. Curr Opin Genet Dev, 1996. 6(2): p. 176-84.
146. Lee, J.S. and A. Shilatifard, A site to remember: H3K36 methylation a mark for histone deacetylation. Mutat Res, 2007. 618(1-2): p. 130-4.
147. Shogren-Knaak, M., et al., Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science, 2006. 311(5762): p. 844-7.

148. Sun, Y., et al., A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM. Proc Natl Acad Sci U S A, 2005. 102(37): p. 13182-7.
149. Haldar, D. and R.T. Kamakaka, Schizosaccharomyces pombe Hst4 functions in DNA damage response by regulating histone H3 K56 acetylation. Eukaryot Cell, 2008. 7(5): p. 800-13.
150. Hohmann, P., R.A. Tobey, and L.R. Gurley, Phosphorylation of distinct regions of f1 histone. Relationship to the cell cycle. J Biol Chem, 1976. 251(12): p. 3685-92.
151. Wolffe, A.P., et al., Transcription factor access to DNA in the nucleosome. Cold Spring Harb Symp Quant Biol, 1993. 58: p. 225-35.
152. Sauve, D.M., et al., Phosphorylation-induced rearrangement of the histone H3 NH2-terminal domain during mitotic chromosome condensation. J Cell Biol, 1999. 145(2): p. 225-35.
153. Zeitlin, S.G., R.D. Shelby, and K.F. Sullivan, CENP-A is phosphorylated by Aurora B kinase and plays an unexpected role in completion of cytokinesis. J Cell Biol, 2001. 155(7): p. 1147-57.
154. Downs, J.A., N.F. Lowndes, and S.P. Jackson, A role for Saccharomyces cerevisiae histone H2A in DNA repair. Nature, 2000. 408(6815): p. 1001-4.
155. Tan, E., et al., Histone H4 histidine kinase displays the expression pattern of a liver oncodevelopmental marker. Carcinogenesis, 2004. 25(11): p. 2083-8.
156. Rice, J.C., et al., Mitotic-specific methylation of histone H4 Lys 20 follows increased PR-Set7 expression and its localization to mitotic chromosomes. Genes Dev, 2002. 16(17): p. 2225-30.
157. Jenuwein, T. and C.D. Allis, Translating the histone code. Science, 2001. 293(5532): p. 1074-80.
158. Goldknopf, I.L., et al., Isolation and characterization of protein A24, a "histone-like" non-histone chromosomal protein. The Journal of biological chemistry, 1975. 250(18): p. 7182-7.
159. Wang, H., et al., Histone H3 and H4 ubiquitylation by the CUL4-DDB-ROC1 ubiquitin ligase facilitates cellular response to DNA damage. Mol Cell, 2006. 22(3): p. 383-94.
160. Zhao, Y., et al., A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing. Mol Cell, 2008. 29(1): p. 92-101.
161. Zhu, B., et al., Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. Mol Cell, 2005. 20(4): p. 601-11.
162. Kamakaka, R.T. and S. Biggins, Histone variants: deviants? Genes Dev, 2005. 19(3): p. 295-310.
163. Zlatanova, J. and A. Thakar, H2A.Z: view from the top. Structure, 2008. 16(2): p. 166-79.
164. Suto, R.K., et al., Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. Nat Struct Biol, 2000. 7(12): p. 1121-4.
165. Marques, M., et al., Reconciling the positive and negative roles of histone H2A.Z in gene transcription. Epigenetics. 5(4): p. 267-72.
166. Mavrich, T.N., et al., Nucleosome organization in the Drosophila genome. Nature, 2008. 453(7193): p. 358-62.
167. Mizuguchi, G., et al., ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science, 2004. 303(5656): p. 343-8.
168. Obri, A., et al., ANP32E is a histone chaperone that removes H2A.Z from chromatin. Nature. 505(7485): p. 648-53.

169. Rogakou, E.P., et al., Megabase chromatin domains involved in DNA double-strand breaks in vivo. J Cell Biol, 1999. 146(5): p. 905-16.
170. Celeste, A., et al., Genomic instability in mice lacking histone H2AX. Science, 2002. 296(5569): p. 922-7.
171. Stucki, M., et al., MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks. Cell, 2005. 123(7): p. 1213-26.
172. Chakravarthy, S., et al., Structural characterization of the histone variant macroH2A. Mol Cell Biol, 2005. 25(17): p. 7616-24.
173. Chakravarthy, S. and K. Luger, The histone variant macro-H2A preferentially forms "hybrid nucleosomes". J Biol Chem, 2006. 281(35): p. 25522-31.
174. Angelov, D., et al., The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. Mol Cell, 2003. 11(4): p. 1033-41.
175. Doyen, C.M., et al., Mechanism of polymerase II transcription repression by the histone variant macroH2A. Mol Cell Biol, 2006. 26(3): p. 1156-64.
176. Chadwick, B.P. and H.F. Willard, A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. J Cell Biol, 2001. 152(2): p. 375-84.
177. Gautier, T., et al., Histone variant H2ABbd confers lower stability to the nucleosome. EMBO Rep, 2004. 5(7): p. 715-20.
178. Bao, Y., et al., Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. EMBO J, 2004. 23(16): p. 3314-24.
179. Angelov, D., et al., SWI/SNF remodeling and p300-dependent transcription of histone variant H2ABbd nucleosomal arrays. EMBO J, 2004. 23(19): p. 3815-24.
180. Sansoni, V., et al., The histone variant H2A.Bbd is enriched at sites of DNA synthesis. Nucleic Acids Res. 42(10): p. 6405-20.
181. Ishibashi, T., et al., H2A.Bbd: an X-chromosome-encoded histone involved in mammalian spermiogenesis. Nucleic Acids Res. 38(6): p. 1780-9.
182. Green, G.R., et al., Histone phosphorylation during sea urchin development. Semin Cell Biol, 1995. 6(4): p. 219-27.
183. Hake, S.B. and C.D. Allis, Histone H3 variants and their potential role in indexing mammalian genomes: the "H3 barcode hypothesis". Proc Natl Acad Sci U S A, 2006. 103(17): p. 6428-35.
184. Chen, P., et al., H3.3 actively marks enhancers and primes gene transcription via opening higher-ordered chromatin. Genes Dev. 27(19): p. 2109-24.
185. Schwartz, B.E. and K. Ahmad, Transcriptional activation triggers deposition and removal of the histone variant H3.3. Genes Dev, 2005. 19(7): p. 804-14.
186. Elsaesser, S.J. and C.D. Allis, HIRA and Daxx constitute two independent histone H3.3-containing predeposition complexes. Cold Spring Harb Symp Quant Biol. 75: p. 27-34.
187. Palmer, D.K., et al., A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. J Cell Biol, 1987. 104(4): p. 805-15.
188. Van Hooser, A.A., et al., Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. J Cell Sci, 2001. 114(Pt 19): p. 3529-42.
189. Sekulic, N., et al., The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. Nature. 467(7313): p. 347-51.
190. Tachiwana, H., et al., Crystal structure of the human centromeric nucleosome containing CENP-A. Nature. 476(7359): p. 232-5.

191. Bailey, A.O., et al., Posttranslational modification of CENP-A influences the conformation of centromeric chromatin. Proc Natl Acad Sci U S A. 110(29): p. 11827-32.
192. Mosammaparast, N., C.S. Ewart, and L.F. Pemberton, A role for nucleosome assembly protein 1 in the nuclear transport of histones H2A and H2B. EMBO J, 2002. 21(23): p. 6527-38.
193. Cook, A.J., et al., A specific function for the histone chaperone NASP to fine-tune a reservoir of soluble H3-H4 in the histone supply chain. Mol Cell. 44(6): p. 918-27.
194. Parthun, M.R., J. Widom, and D.E. Gottschling, The major cytoplasmic histone acetyltransferase in yeast: links to chromatin replication and histone metabolism. Cell, 1996. 87(1): p. 85-94.
195. Tagami, H., et al., Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell, 2004. 116(1): p. 51-61.
196. Burgess, R.J. and Z. Zhang, Histone chaperones in nucleosome assembly and human disease. Nat Struct Mol Biol. 20(1): p. 14-22.
197. Andrews, A.J., et al., The histone chaperone Nap1 promotes nucleosome assembly by eliminating nonnucleosomal histone DNA interactions. Mol Cell. 37(6): p. 834-42.
198. Saeki, H., et al., Linker histone variants control chromatin dynamics during early embryogenesis. Proc Natl Acad Sci U S A, 2005. 102(16): p. 5697-702.
199. Park, Y.J. and K. Luger, The structure of nucleosome assembly protein 1. Proc Natl Acad Sci U S A, 2006. 103(5): p. 1248-53.
200. Kellogg, D.R. and A.W. Murray, NAP1 acts with Clb1 to perform mitotic functions and to suppress polar bud growth in budding yeast. J Cell Biol, 1995. 130(3): p. 675-85.
201. Clapier, C.R. and B.R. Cairns, The biology of chromatin remodeling complexes. Annu Rev Biochem, 2009. 78: p. 273-304.
202. Narlikar, G.J., H.Y. Fan, and R.E. Kingston, Cooperation between complexes that regulate chromatin structure and transcription. Cell, 2002. 108(4): p. 475-87.
203. Langst, G., et al., Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. Cell, 1999. 97(7): p. 843-52.
204. Corona, D.F., et al., Modulation of ISWI function by site-specific histone acetylation. EMBO Rep, 2002. 3(3): p. 242-7.
205. Murawska, M. and A. Brehm, CHD chromatin remodelers and the transcription cycle. Transcription. 2(6): p. 244-53.
206. Schnetz, M.P., et al., Genomic distribution of CHD7 on chromatin tracks H3K4 methylation patterns. Genome Res, 2009. 19(4): p. 590-601.
207. Worcel, A., S. Strogatz, and D. Riley, Structure of chromatin and the linking number of DNA. Proc Natl Acad Sci U S A, 1981. 78(3): p. 1461-5.
208. Eggleston, A.K., Unraveling chromatin organization. Nat Struct Mol Biol, 2005. 12(1): p. 6.
209. Williams, S.P., et al., Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length. Biophys J, 1986. 49(1): p. 233-48.
210. Horowitz, R.A., et al., The three-dimensional architecture of chromatin in situ: electron tomography reveals fibers composed of a continuously variable zig-zag nucleosomal ribbon. J Cell Biol, 1994. 125(1): p. 1-10.

211. Woodcock, C.L., et al., A chromatin folding model that incorporates linker variability generates fibers resembling the native structures. Proc Natl Acad Sci U S A, 1993. 90(19): p. 9021-5.
212. Beard, D.A. and T. Schlick, Computational modeling predicts the structure and dynamics of chromatin fiber. Structure, 2001. 9(2): p. 105-14.
213. Arya, G. and T. Schlick, Role of histone tails in chromatin folding revealed by a mesoscopic oligonucleosome model. Proc Natl Acad Sci U S A, 2006. 103(44): p. 16236-41.
214. Schalch, T., et al., X-ray structure of a tetranucleosome and its implications for the chromatin fibre. Nature, 2005. 436(7047): p. 138-41.
215. Robinson, P.J., et al., EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. Proc Natl Acad Sci U S A, 2006. 103(17): p. 6506-11.
216. Song, F., et al., Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. Science. 344(6182): p. 376-80.
217. Leuba, S.H., et al., Three-dimensional structure of extended chromatin fibers as revealed by tapping-mode scanning force microscopy. Proc Natl Acad Sci U S A, 1994. 91(24): p. 11621-5.
218. Poirier, M.G., et al., Spontaneous access to DNA target sites in folded chromatin fibers. J Mol Biol, 2008. 379(4): p. 772-86.
219. Cui, Y. and C. Bustamante, Pulling a single chromatin fiber reveals the forces that maintain its higher-order structure. Proc Natl Acad Sci U S A, 2000. 97(1): p. 127-32.
220. Kruithof, M., et al., Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30-nm chromatin fiber. Nat Struct Mol Biol, 2009. 16(5): p. 534-40.
221. Fan, Y., et al., H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo. Mol Cell Biol, 2003. 23(13): p. 4559-72.
222. Robinson, P.J., et al., 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. J Mol Biol, 2008. 381(4): p. 816-25.
223. Wong, H., J.M. Victor, and J. Mozziconacci, An all-atom model of the chromatin fiber containing linker histones reveals a versatile structure tuned by the nucleosomal repeat length. PLoS One, 2007. 2(9): p. e877.
224. Perisic, O., R. Collepardo-Guevara, and T. Schlick, Modeling studies of chromatin fiber structure as a function of DNA linker length. J Mol Biol. 403(5): p. 777-802.
225. Yang, D. and G. Arya, Structure and binding of the H4 histone tail and the effects of lysine 16 acetylation. Phys Chem Chem Phys. 13(7): p. 2911-21.
226. Kalashnikova, A.A., et al., The role of the nucleosome acidic patch in modulating higher order chromatin structure. J R Soc Interface. 10(82): p. 20121022.
227. Fan, J.Y., et al., The essential histone variant H2A.Z regulates the equilibrium between different chromatin conformational states. Nat Struct Biol, 2002. 9(3): p. 172-6.
228. Fan, J.Y., et al., H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding. Mol Cell, 2004. 16(4): p. 655-61.
229. Doyen, C.M., et al., Dissection of the unusual structural and functional properties of the variant H2A.Bbd nucleosome. EMBO J, 2006. 25(18): p. 4234-44.
230. Shukla, M.S., et al., The docking domain of histone H2A is required for H1 binding and RSC-mediated nucleosome remodeling. Nucleic Acids Res. 39(7): p. 2559-70.

231. Dalal, Y., et al., Tetrameric structure of centromeric nucleosomes in interphase Drosophila cells. PLoS Biol, 2007. 5(8): p. e218.
232. Miell, M.D., et al., CENP-A confers a reduction in height on octameric nucleosomes. Nat Struct Mol Biol. 20(6): p. 763-5.
233. Miell, M.D., A.F. Straight, and R.C. Allshire, Reply to "CENP-A octamers do not confer a reduction in nucleosome height by AFM". Nat Struct Mol Biol. 21(1): p. 5-8.
234. Dubochet, J., et al., Cryo-electron microscopy of vitrified specimens. Q Rev Biophys, 1988. 21(2): p. 129-228.
235. Konig, P., et al., The three-dimensional structure of in vitro reconstituted Xenopus laevis chromosomes by EM tomography. Chromosoma, 2007. 116(4): p. 349-72.
236. Marsden, M.P. and U.K. Laemmli, Metaphase chromosome structure: evidence for a radial loop model. Cell, 1979. 17(4): p. 849-58.
237. Muller, W.G., et al., Generic features of tertiary chromatin structure as detected in natural chromosomes. Mol Cell Biol, 2004. 24(21): p. 9359-70.
238. Hagstrom, K.A. and B.J. Meyer, Condensin and cohesin: more than chromosome compactor and glue. Nat Rev Genet, 2003. 4(7): p. 520-34.
239. Marko, J.F., Micromechanical studies of mitotic chromosomes. Chromosome Res, 2008. 16(3): p. 469-97.
240. Fuks, F., DNA methylation and histone modifications: teaming up to silence genes. Curr Opin Genet Dev, 2005. 15(5): p. 490-5.
241. Maier, V.K., M. Chioda, and P.B. Becker, ATP-dependent chromatosome remodeling. Biol Chem, 2008. 389(4): p. 345-52.
242. Jiang, C. and B.F. Pugh, Nucleosome positioning and gene regulation: advances through genomics. Nat Rev Genet, 2009. 10(3): p. 161-72.
243. Cremer, T. and M. Cremer, Chromosome territories. Cold Spring Harb Perspect Biol. 2(3): p. a003889.
244. Bolzer, A., et al., Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol, 2005. 3(5): p. e157.
245. Speicher, M.R. and N.P. Carter, The new cytogenetics: blurring the boundaries with molecular biology. Nat Rev Genet, 2005. 6(10): p. 782-92.
246. Dekker, J., et al., Capturing chromosome conformation. Science, 2002. 295(5558): p. 1306-11.
247. Tolhuis, B., et al., Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell, 2002. 10(6): p. 1453-65.
248. Robinson, P.J. and D. Rhodes, Structure of the '30 nm' chromatin fibre: a key role for the linker histone. Curr Opin Struct Biol, 2006. 16(3): p. 336-43.
249. Bell, O., et al., Determinants and dynamics of genome accessibility. Nat Rev Genet. 12(8): p. 554-64.
250. van Holde, K., Chromatin. Springer-Verlag KG, Berlin, Germany. 1988.
251. Makarov, V.L., S.I. Dimitrov, and P.T. Petrov, Salt-induced conformational transitions in chromatin. A flow linear dichroism study. Eur J Biochem, 1983. 133(3): p. 491-7.
252. Felsenfeld, G. and J.D. McGhee, Structure of the 30 nm chromatin fiber. Cell, 1986. 44(3): p. 375-7.
253. Bednar, J., et al., Chromatin conformation and salt-induced compaction: three-dimensional structural information from cryoelectron microscopy. J Cell Biol, 1995. 131(6 Pt 1): p. 1365-76.

254. Tremethick, D.J., Higher-order structures of chromatin: the elusive 30 nm fiber. Cell, 2007. 128(4): p. 651-4.
255. Perez-Montero, S., et al., The embryonic linker histone H1 variant of Drosophila, dBigH1, regulates zygotic genome activation. Dev Cell, 2013. 26(6): p. 578-90.
256. Luger, K., et al., Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature, 1997. 389: p. 251-260.
257. Tachiwana, H., et al., Crystal structure of the human centromeric nucleosome containing CENP-A. Nature, 2011. 476(7359): p. 232-5.
258. Tachiwana, H., et al., Structures of human nucleosomes containing major histone H3 variants. Acta Crystallogr D Biol Crystallogr, 2011. 67(Pt 6): p. 578-83.
259. Gerchman, S.E. and V. Ramakrishnan, Chromatin higher-order structure studied by neutron scattering and scanning transmission electron microscopy. Proc Natl Acad Sci U S A, 1987. 84(22): p. 7802-6.
260. Ghirlando, R. and G. Felsenfeld, Hydrodynamic studies on defined heterochromatin fragments support a 30-nm fiber having six nucleosomes per turn. J Mol Biol, 2008. 376(5): p. 1417-25.
261. Bordas, J., et al., The superstructure of chromatin and its condensation mechanism. II. Theoretical analysis of the X-ray scattering patterns and model calculations. Eur Biophys J, 1986. 13(3): p. 175-85.
262. Bordas, J., et al., The superstructure of chromatin and its condensation mechanism. I. Synchrotron radiation X-ray scattering results. Eur Biophys J, 1986. 13(3): p. 157-73.
263. Makarov, V., et al., A triple helix model for the structure of chromatin fiber. FEBS Lett, 1985. 181(2): p. 357-61.
264. Song, F., et al., Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. Science, 2014. 344(6182): p. 376-80.
265. Correll, S.J., M.H. Schubert, and S.A. Grigoryev, Short nucleosome repeats impose rotational modulations on chromatin fibre folding. EMBO J, 2012. 31(10): p. 2416-26.
266. Boulard, M., et al., Histone variant nucleosomes: structure, function and implication in disease. Subcell Biochem, 2007. 41: p. 71-89.
267. Nekrasov, M., et al., Histone H2A.Z inheritance during the cell cycle and its impact on promoter organization and dynamics. Nat Struct Mol Biol, 2012. 19(11): p. 1076-83.
268. Greaves, I.K., et al., H2A.Z contributes to the unique 3D structure of the centromere. Proc Natl Acad Sci U S A, 2007. 104(2): p. 525-30.
269. Xu, Y., et al., Histone H2A.Z controls a critical chromatin remodeling step required for DNA double-strand break repair. Mol Cell, 2012. 48(5): p. 723-33.
270. Weber, C.M., J.G. Henikoff, and S. Henikoff, H2A.Z nucleosomes enriched over active genes are homotypic. Nature structural & molecular biology, 2010. 17(12): p. 1500-7.
271. Weber, C.M., S. Ramachandran, and S. Henikoff, Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. Mol Cell, 2014. 53(5): p. 819-30.
272. Jin, C., et al., H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. Nat Genet, 2009. 41(8): p. 941-5.
273. Ishibashi, T., et al., Acetylation of vertebrate H2A.Z and its effect on the structure of the nucleosome. Biochemistry, 2009. 48(22): p. 5007-17.

274. Placek, B.J., et al., The H2A.Z/H2B dimer is unstable compared to the dimer containing the major H2A isoform. Protein Sci, 2005. 14(2): p. 514-22.
275. Abbott, D.W., et al., Characterization of the stability and folding of H2A.Z chromatin particles. Implications for transcriptional activation. J. Biol. Chem., 2001. 276: p. 41945-41949.
276. Suto, R.K., et al., Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. Nat Struct Biol, 2000. 7(12): p. 1121-4.
277. Staynov, D.Z., The controversial 30 nm chromatin fibre. Bioessays, 2008. 30(10): p. 1003-9.
278. Talbert, P.B. and S. Henikoff, Environmental responses mediated by histone variants. Trends Cell Biol, 2014.
279. Obri, A., et al., ANP32E is a histone chaperone that removes H2A.Z from chromatin. Nature, 2014. 505(7485): p. 648-53.
280. Ausio, J., The shades of gray of the chromatin fiber: Recent literature provides new insights into the structure of chromatin. Bioessays, 2014.
281. McKittrick, E., et al., Histone H3.3 is enriched in covalent modifications associated with active chromatin. Proc Natl Acad Sci U S A, 2004. 101(6): p. 1525-30.
282. Buscaino, A., R. Allshire, and A. Pidoux, Building centromeres: home sweet home or a nomadic existence? Curr Opin Genet Dev, 2010. 20(2): p. 118-26.
283. Perpelescu, M. and T. Fukagawa, The ABCs of CENPs. Chromosoma, 2011.
284. Gascoigne, K.E., et al., Induced ectopic kinetochore assembly bypasses the requirement for CENP-A nucleosomes. Cell, 2011. 145(3): p. 410-22.
285. Syed, S.H., et al., Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome. Proc Natl Acad Sci U S A, 2010. 107(21): p. 9620-5.
286. Shukla, M.S., et al., The docking domain of histone H2A is required for H1 binding and RSC-mediated nucleosome remodeling. Nucleic Acids Res, 2011. 39(7): p. 2559-2570.
287. McGhee, J.D., G. Felsenfeld, and H. Eisenberg, Nucleosome structure and conformational changes. Biophys J, 1980. 32(1): p. 261-70.
288. Dorigo, B., et al., Chromatin fiber folding: requirement for the histone H4 N-terminal tail. J Mol Biol, 2003. 327(1): p. 85-96.
289. Routh, A., S. Sandin, and D. Rhodes, Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. Proc Natl Acad Sci U S A, 2008. 105(26): p. 8872-7.
290. Collepardo-Guevara, R. and T. Schlick, Chromatin fiber polymorphism triggered by variations of DNA linker lengths. Proc Natl Acad Sci U S A, 2014. 111(22): p. 8061-6.
291. Thambirajah, A.A., et al., H2A.Z stabilizes chromatin in a way that is dependent on core histone acetylation. J Biol Chem, 2006. 281(29): p. 20036-44.