T.C.

DOKUZ EYLUL UNIVERSITY

IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

THE ROLE OF CENP-A CONTAINING NUCLEOSOME IN THE ASSEMBLY OF THE INNER KINETOCHORE

ONUR ÖNDER

MOLECULAR BIOLOGY AND GENETICS

MASTER OF SCIENCE THESIS

IZMIR – 2023

Thesis code: DEU.IBG.MSc.2020850016

T.C.

DOKUZ EYLUL UNIVERSITY

IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

THE ROLE OF CENP-A CONTAINING NUCLEOSOME IN THE ASSEMBLY OF THE INNER KINETOCHORE

MOLECULAR BIOLOGY AND GENETICS

MASTER OF SCIENCE THESIS

ONUR ÖNDER

Supervisor: Asst. Prof. Athanasia Pavlopoulou

IZMIR – 2023

Thesis code: DEU.IBG.MSc.2020850016

YÜKSEK LİSANS TEZ SAVUNMA SINAVI TUTANAĞI

Dokuz Eylül Üniversitesi İzmir Uluslararası Biyotıp ve Genom Enstitüsü Genom Bilimleri ve Moleküler Biyoteknoloji Anabilim Dalı, Moleküler Biyoloji ve Genetik Yüksek Lisans programı öğrencisi Onur Önder **"CENP-A İÇEREN NÜKLEOZOMUN İÇ KİNETOKOR YAPISINDAKİ ROLÜ"** konulu Yüksek Lisans tezini 29/08/2023 tarihinde başarılı olarak tamamlamıştır.

BAŞKAN

Dr. Öğr. Üyesi Athanasia PAVLOPOULOU

Dokuz Eylül Üniversitesi

JÜRİ ÜYESİ

Dr. Öğr. Gör. Hani ALOTAİBİ

Dokuz Eylül Üniversitesi

JÜRİ ÜYESİ

Asst. Prof. Dr. Tuğba Önal-Süzek

Muğla Sıtkı Koçman Üniversitesi

Dokuz Eylül University Izmir International Biomedicine and Genome Institute Department of Genomics and Molecular Biotechnology, Molecular Biology and Genetics graduate program Master of Science student Onur Önder has successfully completed his Master of Science thesis titled **"THE ROLE OF CENP-A CONTAINING NUCLEOSOME IN THE ASSEMBLY OF THE INNER KINETOCHORE"** on the date of 29/08/2023.

HEAD OF THE COMMITTEE

Asst. Prof. Dr. Athanasia PAVLOPOULOU

Dokuz Eylül University

MEMBER

Asst. Prof. Dr. Hani ALOTAİBİ

Dokuz Eylül University

MEMBER

Asst. Prof. Dr. Tuğba Önal-Süzek

Muğla Sıtkı Koçman University

TABLE OF CONTENTS

TABLE OF CONTENTS II
INDEX OF TABLES IV
INDEX OF FIGURES V
ABBREVATIONS VI
ACKNOWLEDGEMENTS VII
ABSTRACT X
ÖZET XI
1. INTRODUCTION AND PURPOSE 1
1.1. Statement and Importance of the Problem 1
1.2. Aim of the Study 1
1.3. The Hypothesis of Study 2
2. **BACKGROUND AND SIGNIFICANCE 3**
2.1. Centromere-Specific Nucleosome 3
2.1.1. Cell division and variant of the canonical H3 histone 3
2.1.2. Structure of the CENP-A nucleosome 3
2.1.3. Electrostatic interactions DNA, histones and environment 4
2.1.4. Flux of ions changes in mitotic cell division 5
2.2. Centromere Formation and Neocentromere 6
2.2.1. The constitutive centromere associated network (CCAN) complex. 6
2.2.2. The role of the nucleosome acidic patch 7
2.2.3. Human neocentromere 7
2.2.4. Labile nucleosome 8
2.3. DNMT3A and DNA Methylation Inhibition 8
2.3.1. Insights into DNTM3A's structure 9
2.3.2. DNMT3A's implication in cancer 10
2.3.3. Procaine as a DNA methylation inhibitor ..... 10 **3.**
MATERIALS AND METHODS 12
3.1. Type of Research 12

3.2. Time and Place of the Research 12
3.3. Research Population, Sampling, and Experimental Groups 12
3.4. Research Materials 13
3.5. Research Variables 14
3.6. Data Collection Tools/Methods 16
3.6.1. Understanding DNA flexibility through molecular dynamics simulations and one pot assay 16
3.6.2. Molecular dynamics simulation of different histone variant constructs that aligned with CENP-C 19
3.6.3. Investigating how procaine affects DNMT3A through MD simulations to understand its mode of action 21
3.7. Research Plan And Timeline 22
3.8. Data Analysis 22
3.9. Limitations of the Research 23
3.10. Ethical Committee Approval ..... 24 **4.**
RESULTS 25
4.1. Detecting Allosteric Influences on Nucleosomes Containing CENP-A 25
4.1.1. Mitotic-like ionic conditions cause dna flexibility on both ends of the CENP-A nucleosome. 25
4.1.2. Core octamers respond differently depending on their variant-composition, DNA sequence, and the ionic environment. 27
4.1.3. Divalent cation modulates DNA-histone interactions by modulating the minor-groove binding of select lysine or arginine residues. 29
4.1.4. Measurement of the flexibility of the canonical H3- or the CENP-A containing nucleosome ends by using the one pot assay. 31
4.2. H2A.Z exerts an inhibitory impact on the process of centromere formation 33
4.2.1. Molecular dynamics simulations attribute the loss of CENP-C affinity to H2A.Z-specific substitutions. 33
4.2.2. Molecular dynamics simulations attribute the loss of CENP-C affinity to labile nucleosome. 36

4.3. Allosteric Inhibition of DNMT3A by Procaine..... 38 **5. DISCUSSION 41**

5.1. Allosteric Alterations in the Centromere-Specific Nucleosome Induced by Physical Factors 41

5.2. H2A.Z Suppresses Ectopic Kinetochore Formation by Inhibiting CENP-A Nucleosome Recognition through CENP-C 43

5.3. Understanding DNMT3A Inhibition by Procaine Using Molecular Dynamics Simulations ... 44

6. CONCLUSION AND FUTURE PROSPECTS ..... 45 7. REFERENCES ..... 46 8. APPENDICES 55

8.1. Ethics Committee Report 55

8.2. Curriculum Vitae 56

INDEX OF TABLES

Table 1: The primary sequences of the various inserts 18
Table 2: Research Plan 21
Table 5: CENP-C affinities and the corresponding dissociation constants K_D of three chimeric nucleosomes as inferred from MD simulations 35

INDEX OF FIGURES

Figure 1: Structures of CENP-A containing nucleosomes.. 3
Figure 2: DNA unwrapping can be facilitated even though the whole nucleosome still is highly stable. 5
Figure 3: Multiple sequence alignment of the H2AZ and H2A. 7
Figure 4: Allosteric modifications in the centromere-specific nucleosome result in the unfolding of DNA and its symmetrical opening. 27
Figure 5: Analysis of the Nucleosome Core's Variant-Specific Responses to Alterations in DNA and Ionic Environments. 28
Figure 6: Exploring the Interaction between Divalent Cations and DNA-Histone via Lysine and Arginine Residues. 31
Figure 7: One-pot restriction-accessibility assay of the CENP-A nucleosome. 33
Figure 8: H2A.Z influences CENP-C recognition, leading to its detachment from the nucleosome surface..... 36

Figure 9: H2A.Z/H3.1 (Labile) nucleosome reject binding of CENP-C.. 37
Figure 10: Molecular simulations of DNA-bound DNMT3A in procaine. 40

ABBREVATIONS

CENP-A: Centromere Protein A

CENP-C: Centromere Protein C

CENP-N: Centromere Protein N

Cryo-EM: Cryogenic electron microscopy

DNA: Deoxyribonucleic Acid

MD: Molecular Dynamics

DNMT3A: DNA methyltransferase 3 alpha

CCAN: Constitutive Centromere Associated Network

PBS: Phosphate Buffered Saline

LLC: Limited Liability Company

VMD: Visual Molecular Dynamics

OPC: Optimal Point Charges

PDB: Protein Data Bank

NVT: Number of atoms (N), and the volume (V), the absolute temperature (T)

NPT: Number of particles (N), system pressure (P) and temperature (T)

EDTA: Ethylenediamine Tetra acetic Acid

BSA: Bovine Serum Albumin

SDS: Sodium Dodecyl Sulfate

PAGE: Polyacrylamide Gel

Electrophoresis PME: Particle Mesh Ewald

ns: Nanosecond nM: Nanomolar

K_d: Dissociation Constant kcal/mol:

Kilocalorie per Mole

TRD: Target Recognition Domain

AML: Acute Myeloid Leukemia

SAH: S-adenosylhomocysteine

SCNT: Somatic Cell Nuclear Transfer

HMDB: The Human Metabolome Database

CHARMM: Chemistry at HARvard Molecular Mechanics

CGENFF: CHARMM General Force Field

SAM: S-adenosyl-L-methionine

ACKNOWLEDGEMENTS

First and foremost, my profound gratitude goes to Dr. Seyit Kale. From the outset, he presented me with this invaluable opportunity, consistently marveling at me with his expertise, knowledge, and inherent talents. His unwavering financial and intellectual support, alongside his unparalleled dedication to nurturing my ideas and projects, marks him as a truly distinctive mentor in my academic journey.

I am indebted to Asst. Prof. Dr. Athanasia Pavlopoulou. Her partnership in our joint project, her consistent support from day one at IBG, and her infectious enthusiasm have been invaluable. I cherish the privilege of knowing such a figure, and I'm endlessly grateful for her friendship, understanding, and encouragement throughout my master's degree journey.

This thesis, a facet of the France Epigenetic Consortium, is a product of collective effort. My deepest appreciation is extended to Prof. Dr. Stefan Dimitrov, whose guidance, invaluable suggestions, and insights were pivotal to my research. Equally, I'm honored to have worked

alongside Prof. Dr. Dimitar Angelov. Collaborating with such brilliant and thoughtful minds has been both a privilege and a source of immense learning. Special mention must go to our collaborative partners Dr. Ali Hamiche and Dr. Ramachandra Boopathi, without whom the depth and breadth of this research would have been unattainable.

My gratitude extends to my thesis advisory committee members, Asst. Prof. Dr. Hani Alotaibi and Asst. Prof. Dr. Tuğba Önal-Süzek. Their insightful feedback and dedication of time have been indispensable.

A special part for the community of the Kale Lab – both past and present members. I will always hold a special place in my heart for those people who supported me in whatever matter it was, made every day of ours enjoyable, and provided new insights, knowledge, and inspiration. It is with certainty that I believe our paths, in scientific realms and beyond, will intertwine again. Wishing each of you success and happiness.

On a more personal note, my deepest gratitude goes to my family, the pillars of my life. Their emotional and financial backing, ceaseless love, and sacrifices have been the wind beneath my wings. This work stands as a testament to their enduring faith and support. I'd also like to express my gratitude to my girlfriend, Ezgi Akpınar, who, despite the challenges, does everything to overcome them, constantly enhancing and loving me in every aspect of life.

Further, my journey was made brighter by friends like Bilal Çam, İbrahim Önal, Alper Başar, Aleyna Usta, Uğur Bora, Burak Çaylı, and Öykü Durak. Your support means the world.

Lastly, my sincere thanks to TUBITAK (The Scientific and Technological Research Council of Turkey) Scientist Support Office Presidency (BIDEB) for their financial assistance through the 2210-A project. My appreciation also extends to TUBITAK ULAKBIM (National Academic Network and Knowledge Center) for their computational resources and the alwayshelpful team at ULAKBIM.

THE ROLE OF CENP-A CONTAINING NUCLEOSOME IN THE ASSEMBLY OF THE INNER KINETOCHORE

Onur ÖNDER

Izmir International Biomedicine and Genome Institute,
Dokuz Eylül University, Health Campus, Balçova, 35340 İzmir/ TURKEY

ABSTRACT

This comprehensive research has demonstrated the histone variant effects on chromosomal interactions and the potential therapeutic implications of molecular dynamics simulations in cancer. In the first part of this study, we explored the behavior of CENP-A

nucleosomes, focusing on their flexibility compared to canonical H3 nucleosomes. We investigated how divalent cations, especially magnesium, influence these nucleosomes. Notably, our findings highlighted the potential of changing magnesium concentrations to impact nucleosome stability, shedding light on chromosome segregation processes during cell division. Motivated by this, we next explored the interactions between CENP-A nucleosomes and other kinetochore proteins, with a spotlight on CENP-C. Through molecular dynamics simulations, we identified the critical role of histone variants in these interactions. A novel finding was that H2A.Z could modulate the affinity between CENP-A and CENP-C, potentially inhibiting ectopic kinetochore assembly. Finally, we conducted a study to understand the interactions between DNMT3A, an important enzyme in DNA methylation, and procaine. Procaine showed promise as a promising DNA methylation inhibitor. Given the role of DNMT3A in numerous types of cancers, understanding its mode of action with procaine was important. Our molecular dynamics simulations provided a dynamic view of how procaine might inhibit DNMT3A, offering potential therapeutic pathways. Collectively, these interconnected studies offer a holistic understanding of chromosomal interactions, the role of molecular dynamics in these processes, and their implications in the broader context of gene regulation and cancer therapeutics.

Key words: CENP-A, CENP-C, Nucleosome, Ionic Environment, Molecular Dynamics

CENP-A İÇEREN NÜKLEOZOMUN İÇ KİNETOKOR YAPISINDAKİ ROLÜ

Onur ÖNDER

İzmir Uluslararası Biyotıp ve Genom Enstitüsü,
Dokuz Eylül Üniversitesi Sağlık Kampüsü, Balçova, 35340 İzmir/ TÜRKİYE

ÖZET

Bu kapsamlı araştırma, histon varyantının kromozomal etkileşimler üzerindeki etkilerini ve kanserde moleküler dinamik simülasyonlarının potansiyel terapötik etkilerini göstermiştir. Bu

çalışmanın ilk bölümünde CENP-A nükleozomlarının davranışını, kanonik H3 nükleozomlara kıyasla esnekliklerine odaklanarak araştırdık. İki değerlikli katyonların, özellikle magnezyumun bu nükleozomları nasıl etkilediğini gösterdik. Özellikle bulgularımız, magnezyum konsantrasyonlarını değiştirmenin nükleozom stabilitesini etkileme potansiyelini vurguladı ve hücre bölünmesi sırasında kromozom ayrışma süreçlerine ışık tuttu. Bundan motive olarak, daha sonra CENP-A nükleozomları ve diğer kinetekor proteinleri arasındaki etkileşimleri CENP-C'ye odaklanarak araştırdık. Moleküler dinamik simülasyonları aracılığıyla histon varyantlarının bu etkileşimlerdeki kritik rolünü belirledik. Yeni bir bulgu, H2A.Z'nin CENP-A ve CENP-C arasındaki afiniteyi modüle ederek potansiyel olarak ektopik kinetekor düzeneğini inhibe edebilmesiydi. Son olarak DNA metilasyonunda önemli bir enzim olan DNMT3A ile prokain arasındaki etkileşimi anlamak için bir çalışma gerçekleştirdik. Prokain umut verici bir DNA metilasyon inhibitörü olarak kullanılıyor. DNMT3A'nın çeşitli kanser türlerindeki rolü göz önüne alındığında, prokainle olan etki mekanizmasının anlaşılması önemliydi. Moleküler dinamik simülasyonlarımız, prokainin DNMT3A'yı nasıl engelleyebileceğine dair dinamik bir görünüm sunarak potansiyel terapötik yollar sundu. Toplu olarak, bu birbirine bağlı çalışmalar, kromozomal etkileşimlerin, bu süreçlerde moleküler dinamiğin rolünün ve bunların gen düzenlemesi ve kanser tedavilerinin daha geniş bağlamındaki etkilerinin bütünsel bir anlayışını sunmaktadır.

Anahtar Kelimeler: CENP-A, CENP-C, Nükleozom, İyonik Ortam, Moleküler Dinamik Simülasyonları

1. INTRODUCTION AND PURPOSE

1.1. Statement and Importance of the Problem

Epigenetics comprises the intricate interplay between chromatin structure, histone variants, and DNA modifications which influence gene expression and other DNA-templated processes. Among the essential players in epigenetics are the centromere-specific histone CENP-A, a variant of H3; histone H2A.Z, variant of H2A; and the DNA methyltransferase DNMT3A. Their roles in maintaining chromatin integrity, regulating gene accessibility, and epigenetic modifications are pivotal for cellular function. Investigating the effects of environmental ionicity, nucleosome dynamics, and interactions among these components posits a paramount significance in deciphering the mechanisms governing epigenetic regulation.

1.2. Aim of the Study

The aim of the initial research is to investigate the effect of environmental ionicity and changes in DNA sequences, which are different physical parameters, on the dynamic conformational behavior of the centromere-specific nucleosome CENP-A. This project has been studied through molecular dynamics simulations and biochemical wet-lab assays. The overarching aim of this project is to uncover the complex allosteric effects that occur in nucleosomal architecture in various physiological conditions, and thus to understand the role of the influence between physical factors and nucleosome dynamics in the inner kinetochore assembly during mitotic division.

The purpose of secondary research is to investigate the effect of histone variants governing the suppression of abnormal kinetochore formation. Central to this search is the histone variant H2A.Z, whose physiological role we argue goes beyond gene expression. In this project we demonstrate a mechanism whereby H2A.Z inhibits CENP-A nucleosome recognition by the kinetochore protein CENP-C, master regulator of the kinetochore attachment to chromatin. This project employs a sophisticated combination of *in silico* structural work and biochemical analyses towards unraveling the mechanisms underlying this suppression. In doing so, it contributes to the ever-evolving landscape of our understanding of centromeric functionality, rendering insights that transcend the boundaries of conventional molecular investigations.

The aim of tertiary research is a project in which we aim to reveal the binding mode while it is known that the DNMT3A enzyme is inhibited by the anesthetic drug procaine. To explain the inhibitory mechanism using molecular dynamics simulations and sensitive structural analysis. By understanding the interaction between procaine and DNMT3A, it aims to shed light on potential allosteric sites and modulation mechanisms that govern the complex enzymatic activity of DNMT3A.

These different scientific projects collectively underscore a collaborative effort to explain and extend the knowledge gained in the field of epigenetics. Each project demonstrates the complexity and importance of this field by focusing on the highlights of epigenetics. The molecular dynamics simulations we use are a tool that helps us better understand complex biological phenomena. In these projects, we emphasize once again the critical role of simulations in understanding epigenetic events.

1.3. The Hypothesis of Study

Our hypothesis of the first project postulates that changes in both environmental ionicity and DNA sequence modifications cause allosteric conformational changes within the centromere-specific nucleosome. These changes greatly affect things like how stable the nucleosome is, how flexible it is, and how it interacts with DNA. These shifts are thought to play a large role in guiding how chromatin moves and functions in the complex world of gene regulation, especially when faced with different physical changes.

In the context of the second project, we put forward the proposition that the histone variant H2A.Z exerts a regulatory effect through its interaction with CENP-C. This interaction has a restrictive role in the conformation and accessibility of CENP-A nucleosomes, thereby preventing the recognition of these nucleosomes by CENP-C and thus preventing the initiation of inappropriate neocentromere formation. This mechanism explains how structural modifications and kinetic changes affect cell division.

Our third research study suggests that procaine changes the structure of DNMT3A, which then affects its main functions, especially how it binds and operates as an enzyme. Our in-depth study aims to outline the relationship between procaine and DNMT3A. By understanding this, we hope to discover new directions for developing treatments in the field of gene regulation.

2. BACKGROUND AND SIGNIFICANCE

2.1. Centromere-Specific Nucleosome

2.1.1. Cell division and variant of the canonical H3 histone

Genetic inheritance relies on faithful segregation of sister chromatids. Proper association between spindle apparatus fibers and chromosomes is critical during cell division.[1] Kinetochore function creates load-bearing attachments of sister chromatids to spindle microtubules during cell division. Kinetochores assemble in specialized chromatin regions, known as centromeres.[2] Centromeric assemblies of the kinetochore are driven by the epigenetic factor CENP-A, a centromere-specific non-allelic variant of histone H3, one of the proteins that constitutes the core of nucleosomes, the smallest repeating unit of the chromatin fiber.[3] The DNA must be compacted to fit the nucleus. Nucleosomes are the basic structural units of DNA wrapping in eukaryotes.

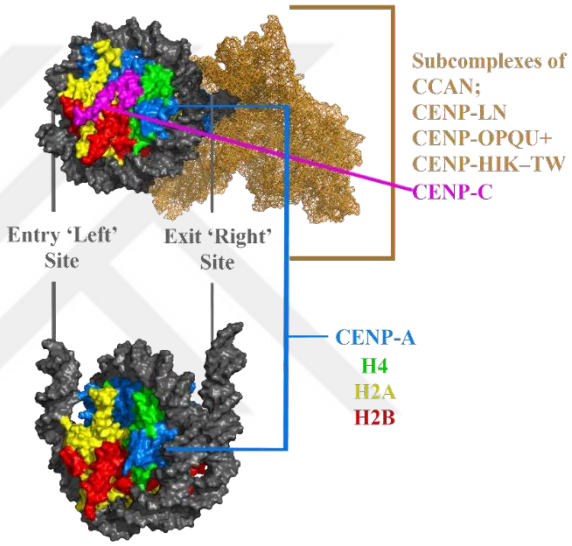


Figure 1: Structures of CENP-A containing the CENP-A nucleosome is similar to that of nucleosomes. Cartoon representation of CCAN which DNA terminals are not illustrated due to the higher flexibility. Molecular dynamics simulation output shows the exit 'right' site is farther than the entry 'left' site.

Each nucleosome holds a histone octamer, which is composed of a little less than two rounds of DNA wrapped around a collection of eight proteins called histones. The histone proteins H2A, H2B, H3, and H4 have two copies in each histone octamer. The crystal structure of the H3 nucleosome, consisting of histones H2A, H2B, CENP-A, and H4, as well as ~145 base pairs of DNA.

2.1.2. Structure of the CENP-A nucleosome

A unique genetic locus called the centromere interacts with the mitotic spindle to promote chromosomal segregation. This interaction occurs via the recognition of the CENP-A containing nucleosome.[4] The 16-subunit human inner kinetochore constitutive centromereassociated network (CCAN) complex is constitutively associated with centromeric

chromatin throughout the cell cycle according to CENP-A nucleosomes.[5] CENP-A structure consists of a disordered N-terminal tail, followed by a dynamic αN helix and three stable α-helices that constitute the histone fold. The αN helix of CENP-A is half a turn shorter than its H3 counterpart, and this "defect" promotes the flexibility of the DNA ends via looser DNA binding.[6] Molecular dynamics simulations demonstrated that the DNA termini of H3 nucleosomes are highly stable. This pronounced flexibility of the DNA termini of CENP-A nucleosome is caused by a lysine (Lys) residue located on the αN helix of CENP-A, which substitutes the DNA-binding R49 (Arg) of the canonical histone H3.[7] The global properties of CENP-A also support a weak association between DNA and the histone core in comparison to the H3-containing core octamer, and this difference could be implicated in the heightened accessibility of CENP-A nucleosomal DNA.[8] In an x-ray crystallographic analysis the DNA in the entry and exit regions could not be identified, suggesting that the nucleosomal DNA terminals of CENP-A are quite flexible. Furthermore, this flexibility is crucial for the CCAN complex to come together and maintain the fidelity of the mitotic pathway.[9] Recently, the cryoelectron microscopy structure suggests that a total of 20 bp equal amounts of DNA are unwrapped at each DNA terminal of CENP-A-containing nucleosomes, while the H3 histone maintains the canonical H3 nucleosome-wrapped DNA termini[10] (Figure 1). An international collaboration, including my lab, demonstrated that Widom 601, a synthetic DNA sequence commonly used for its high binding affinity toward the core histone octamer,[11] leads to the CENP-A nucleosome core particles open asymmetrically. Molecular dynamics simulations confirmed the cryo-EM results and revealed that the exit DNA 'right' has higher flexibility. (Figure 1) This suggest that it is possible to fine-tune the flexibility of DNA ends in a sequencedependent manner.[12] On the other hand, conditions for a doubly open CENP-A nucleosome were yet to be identified as of the beginning of this project.

2.1.3. Electrostatic interactions DNA, histones and environment

The strength of the nucleosomal DNA interactions with the histone core directly influences its accessibility.[13] Electrostatic attraction between the large and opposite charges of the globular histone core and DNA creates a stable complex between the nucleosomal DNA and the histone octamer.[14] On the other hand, the two charges do not fully cancel each other out, requiring solvent cations for the electrostatic counterbalance. This feature of the nucleosome causes its stability to be critically dependent on the solvent environment. Small ions in a solvent

can neutralize a large portion of the negative DNA charge to maintain the wrapped state of the DNA.[15] This sensitivity further pronounced in the CENP-A containing nucleosome because the CENP-A core is less positively charged than the canonical H3containing core, suggesting that the nucleosomal DNA is closer to unwrapping.(Figure 2) In this project, we have investigated and found that this critical balance can be influenced by the DNA sequence as well as the type of positively charged counterions present in solution.

2.1.4. Flux of ions changes in mitotic cell division

In higher eukaryotic cells, the nuclear envelop breaks down at the beginning of mitosis, exposing chromatin to the cytoplasmic environment.[16] Chromosomes counter other environments that contain different small ions. Free divalent cation Mg^{++} levels increase, while free divalent cation Ca^{++} levels remain the same in the cytoplasm during mitosis. This change in magnesium could affect nucleosomes.[17]

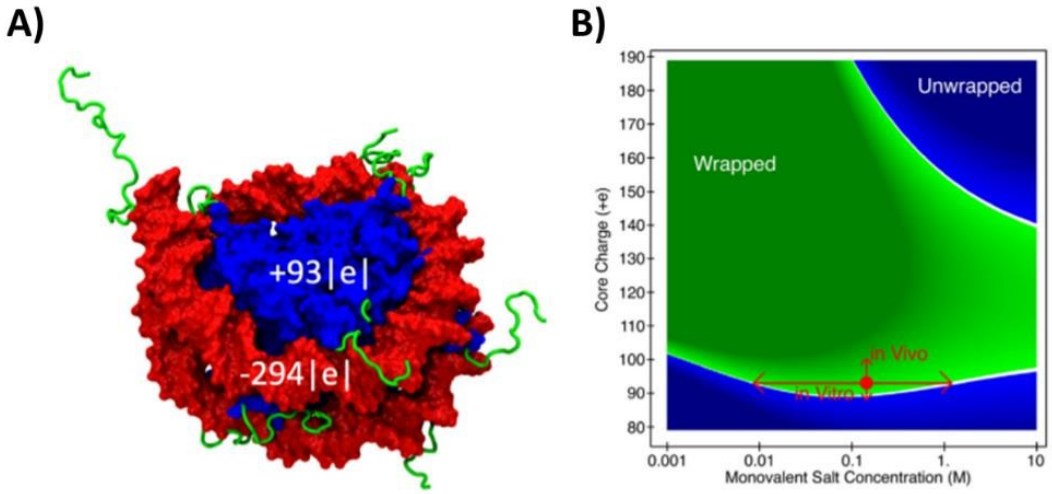


Figure 2: DNA unwrapping can be facilitated even though the whole nucleosome still is highly stable. A) Strong electrostatic attraction between the oppositely charged globular histone core (blue) and the DNA (red) provides high stability of the nucleosome. B) The state of the "canonical" nucleosome (red dot) is found in a "wrapped" state at physiological conditions. A small drop in the charge of the histone core alters nucleosome stability from wrapped to unwrapped so DNA becomes more accessible.[14]

To understand the sequence and solvent-dependent behaviors of CENP-A-containing nucleosomes, we performed molecular dynamics simulations of CENP-A containing Widom 601 and mutated Widom 601 in a highly divalent solvent that mimics mitotic conditions, and a physiological solvent that mimics phosphate-buffered saline (PBS). Here, we present computational and experimental evidence that DNA sequences and solvent environments can alter the structure and dynamics of CENP-A-containing nucleosomes. This evidence could

provide insights into the kinetic and epigenetic effects that control the formation of the internal kinetochore assembly.

2.2. Centromere Formation and Neocentromere

2.2.1. The constitutive centromere associated network (CCAN) complex.

The fundamental processes of gene expression and DNA-templated cellular activities are orchestrated within the dynamic world of the chromatin, the nucleoprotein complex located in the nucleus that protects and regulates the access to the genetic information encoded within DNA.[18] The nucleosome is as the smallest repeating unit of chromatin, with an approximately 145 bp of DNA wrapped around the histone octamer. These histones, along with non-histone proteins, not only provide structural support, but also intricately govern the regulation of gene expression through the binding of various proteins.[19]

All core histone proteins but H4 exist in variants that replace their canonical counterparts in a temporally regulated manner. Among the diverse families of histone variants, H2A.Z is a highly conserved evolutionary isoform of H2A that is typically associated with elevated gene expression among other physiological roles. Extensive studies have demonstrated its involvement in gene activation, chromosome segregation, heterochromatin silencing, and cell cycle progression.[20] Counterintuitively, however, H2A.Z has recently been demonstrated to be completely dispensable for cells that lack their ability to divide. (Stefan Dimitrov, personal communication) This suggests a potential role for H2A.Z in cell division.

Centromeres are critical regions that ensure proper chromosome segregation during cell division. In this context, CENP-A is a critical centromere-specific epigenetic marker that takes center stage as a specialized variant of the canonical H3 histone. CENP-A containing nucleosomes provide the foundation for the assembly of kinetochore proteins, which are essential for chromosome attachment to the mitotic spindle.[3,21] Notably, two key proteins of the constitutive centromere-associated network (CCAN), CENP-C and CENP-N, make specific contacts with CENP-A nucleosomes and play a pivotal role in centromere function.[2] CENP-C, a crucial player in the centromere-associated network, exhibits a unique interaction with CENPA nucleosomes through an "arginine anchor" mechanism, recognizing the acidic patch on the H2A-H2B dimer. This interaction not only stabilizes the nucleosome core but also extends its influence on the positioning of DNA gyres.[22,23]

2.2.2. The role of the nucleosome acidic patch

The study of chromatin dynamics has revealed a remarkable feature known as the "acidic patch," which is a specific region within the histone octamer that possesses a high concentration of negatively charged amino acid residues. This unique acidic "spot" on the nucleosome surface plays a crucial role in determining the higher-order organization of chromatin, influencing compaction, accessibility, nucleosome assembly, and the overall nucleosomal landscape.[24]

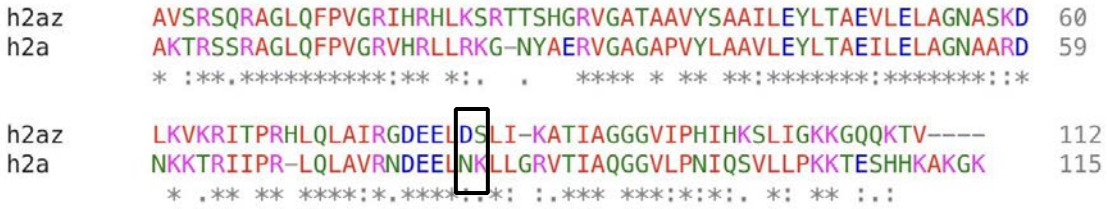


Figure 3: Multiple sequence alignment of the H2AZ and H2A. Part in the square shows an extended acidic patch.

Among histone variants, H2A.Z has emerged as a fascinating contender because of its extended acidic patch (Figure 3), offering additional possibilities for interactions with neighboring nucleosomes or non-histone proteins.[25]

2.2.3. Human neocentromere

Neocentromeres, a rare and novel class of centromeres with enigmatic structures that appear at chromosomal locations distinct from the original centromere, have emerged as a fascinating area of research. Despite lacking the typical centromere DNA, they demonstrate remarkable stability and form primary constrictions that are essential for proper chromosome inheritance.[26]

Proteins known as CENP-A, CENP-C, and CENP-E have been widely recognized as crucial markers of active centromere function and play essential roles in the proper functioning of conventional centromeres. Interestingly, recent research has revealed that these same proteins also localize to neocentromeres, challenging the conventional notion that neocentromeres lack the essential components necessary for functional centromeres.[27] Notwithstanding its beneficial functions, the creation of neocentromeres is commonly associated with cancer or genetic anomalies. Numerous constitutional human neocentromeres have been identified in individuals with congenital abnormalities, developmental delays, or intellectual disabilities.[28] Perturbations

in CENP-A expression and localization have been associated with genomic instability and have emerged as potential drivers of tumorigenesis.[29] A recent study explored the mechanism by which the CENP-A chaperone Scm3 is accommodated in non-centromeric regions. Researchers identified transient and immature neocentromeres in the fission yeast *Schizosaccharomyces* pombe, which exhibited reduced association with the CENP-A chaperone Scm3. This reduced association was attributed to the persistence of the histone H2A variant, H2A.Z. These findings offer novel insights into the epigenetic control of neocentromere establishment, mediated by histone variants.[30]

Here, we suggest that H2A.Z might play a significant role in preventing the formation of neonucleosomes, a phenomenon arising from aberrant nucleosome assembly triggered by CENP-C recognition of CENP-A. Modulating the binding of CENP-C to CENP-A/H2A. Zcontaining nucleosomes, H2A.Z, act as safeguards, ensuring the integrity of chromatin architecture and, ultimately, the faithful execution of genomic processes.

2.2.4. Labile nucleosome

Advanced tools are helping researchers understand deeper into nucleosomes, the DNA packaging units in our cells. These nucleosomes can undergo changes that influence gene activity. A recent focus has been on specific types of nucleosomes called histone variants. Notably, a study by Felsenfeld's team identified certain nucleosomes containing both H3.3 and H2A.Z as being particularly labile (unstable).[31] These labile nucleosomes are predominantly found at gene regulatory sites, suggesting they might play a key role in gene activation or silencing. This new knowledge gives scientists a clearer understanding of DNA packaging and its potential impact on our genes.[32]

2.3. DNMT3A and DNA Methylation Inhibition

In the field of molecular biology, DNA methyltransferases (DNMTs) stand out as key players in epigenetic regulation. By adding a methyl group to DNA, they directly influence gene expression patterns. The DNMT family consists of five main members: DNMT1, DNMT2, DNMT3A, DNMT3B, and DNMT3L. DNMT1 ensures the methylation patterns remain consistent after DNA replication, while DNMT2, distinctively, targets tRNAs instead of DNA. DNMT3A and DNMT3B are responsible for de novo methylation, setting up new methylation sites, with DNMT3L acting as a facilitator without direct methyltransferase activity. Notably, these DNMTs operate preferentially in specific genomic areas, with DNMT1, DNMT3A, and DNMT3B primarily found in the nucleus, playing essential roles in maintaining genomic integrity, supporting embryonic development, and guiding cell differentiation.[33]

2.3.1. Insights into DNTM3A's structure

DNMT3A is a protein with a molecular weight of 130 kDa. It is made up of a sequence encoded by 23 exons, located on the human chromosome 2p23.[34] DNMT3A is very similar across different mammals, which is quite surprising. In fact, the human version of this protein is almost the same (98% similarity) to the one found in mice. This enzyme is special because it adds methyl groups to DNA both at unique spots and at repeating sections.[35] DNMT3A has a specialized way of interacting with DNA, and this interaction is largely driven by three critical segments of the protein. First, there is a loop in the Target Recognition Domain (TRD), specifically the section spanning residues R831 to F848. Next, there is the catalytic loop, which covers residues G707 to K721. Lastly, there is a region designed for two DNMT3A proteins to come together, known as the homodimeric interface. When these three regions work in tandem, they create a continuous and well-defined surface for the DNA to attach. What is particularly interesting is the adaptability of these segments. When they come in contact with DNA, their structure changes in such a way that they can snugly fit and interact with the wider parts of the DNA spiral, known as the major groove. One residue, R882, plays a special role in strengthening this interaction by forming a hydrogen bond, which is a type of strong attraction between molecules. The crystal structure of the DNMT3A combined with DNMT3L, when

paired with a short segment of DNA, provides a detailed snapshot. This view allows scientists to see in high resolution how DNMT3A, DNMT3L, and DNA come together and how their structures influence their functions.[36]

2.3.2. DNMT3A's implication in cancer

Cancer research has spent a lot of time looking at how DNA methylation, or the addition of methyl groups to DNA, goes wrong. Early studies mainly looked at how there was too much methylation (hypermethylation) at the starting points (promoters) of genes. This extra methylation can turn off genes that normally stop tumors from growing, which is not good for our bodies.[37] At the same time, research has also looked at the opposite problem: not enough methylation (hypomethylation) throughout the whole DNA. This lack of methylation can make the DNA unstable, which is not good either. One important finding in this area is that DNMT3A often has mutations in certain blood cancers, especially one called acute myeloid leukemia (AML). When this happens, it can affect how the disease progresses and its outcome.[38] These mutations disrupt the intricate balance of cell cycle control and the established DNA methylation patterns, resulting in aberrant gene expression. Particularly, DNMT3A mutations play a significant role in this context, impacting the regulation of key cellular processes. These perturbations culminate in the dysregulation of critical genes, which is a hallmark of cancer development. Intriguingly, DNMT3A has also been observed to be excessively active (overexpressed) in a variety of human cancers. This heightened activity corresponds to the excessive methylation of specific genes known to be involved in suppressing tumor formation, thereby contributing to the progression of cancer.[39]

2.3.3. Procaine as a DNA methylation inhibitor

In the field of cancer treatment, researchers are trying to turn back on genes that normally help stop tumor growth. They do this by stopping the excessive methylation that happens at the start of these genes (promoters), which can silence them. Interestingly, a substance called procaine, which was originally used as a local anesthetic, has a unique role in this effort. It works by blocking the process of DNA methylation, effectively preventing those genes from being silenced. This makes procaine a potential inhibitor of DNA methylation in the context of

cancer therapy.[40] This ability of procaine is especially promising for certain types of cancers, pointing to a promising direction for treatments. By inhibiting DNA methylation, procaine might offer a new way to intervene and potentially reverse some of the damaging effects of cancer.[41] Recent investigations indicate that when procaine or Sadenosylhomocysteine (SAH) are administered individually or in combination to cell culture media, they exhibit no detrimental impact on cellular proliferation or morphological characteristics. Significantly, procaine has demonstrated its efficacy by diminishing tumor dimensions while preserving cellular viability in specific experimental conditions. This therapeutic approach, which involves the utilization of procaine, SAH, or a synergistic combination of both, presents considerable potential for protocols such as Somatic Cell Nuclear Transfer (SCNT). These findings suggest a viable methodology to regulate DNA methylation levels without adversely affecting cellular proliferation or maintaining structural integrity.[42]

Considering these findings, we suggest a detailed investigation into the potential mechanisms governing the interactions among S-adenosylmethionine (SAM), procaine, and DNMT3A. By utilizing techniques such as docking and molecular dynamics simulations, our research aims to uncover the specifics of how procaine interacts with various regions of DNMT3A. This exploration will further our understanding of the interactions between these molecules, potentially paving the way for optimized therapeutic strategies.

3. MATERIALS AND METHODS

3.1. Type of Research

In the investigation of allosteric changes involves the collaborative work of biochemical wet lab experiments and computational research methods.

In the study of suppression of neocentromere formation contains computational research methods validated by biochemical affinity measurement assays (the latter not conducted at the host institute of the thesis author).

In the study regarding the inhibition of DNMT3A contains computational research methods.

3.2. Time and Place of the Research

These projects were carried out with the financial aid provided by the EMBO installation grant IG 5056 awarded to Seyit Kale as well as internal funds of the Izmir Biomedicine and Genome Center (IBG). Experimental work of project 01 has been carried out at Université Grenoble Alpes and Université de Lyon. Computational work of these projects has been carried out by the thesis author under the supervision of Dr. Seyit Kale, IBG group leader of the Computational Biophysics Laboratory. Computational work of these has been carried out between July 2021 and August 2023.

3.3. Research Population, Sampling, and Experimental Groups

CENP-A containing nucleosome structure obtained from Protein Data Bank (PDB). The experimental material used in this research study is the pEX-A128- 601 mutants containing Widom 601 nucleosome positioning sequence expressed in *Escherichia coli* DH5α cells.

CENP-A, CENP-C, H2A-Z containing nucleosome structures obtained from the Protein Data Bank (PDB).

DNMT3A structure obtained from protein data bank and procaine structure is obtained from human metabolome database.[43]

3.4. Research Materials

In the context of environmental changes of CENP-A containing nucleosome, an effort was made to elucidate the nuanced interactions within nucleosomes and their adaptations to various environments. A cornerstone of our investigation is the Widom 601 DNA sequence. This sequence is fundamental for determining nucleosome placement, and we introduced specific alterations in order to align it with the one-pot assay D6L structure. In our examination, histones, especially the variants CENP-A and H3, were pivotal. Their inclusion was essential to reconstruct nucleosomes, offering insights into the distinct effects various histones can exert on nucleosome dynamics.

The main aim of this work is mimicking the cell environment in the simplest way. By introducing Na^{+} and Cl^{-} ions into our models, we approximated physiological conditions. Similarly, Introduction of Mg^{++} ions are aimed at understanding its function in cell division. We created salt solutions to appropriately reflect certain cellular conditions, such as cell division.

The critical enzyme of the study was the restriction enzyme HaeIII, an enzyme used to digest nucleosomal DNA. This enzymatic assay facilitated the assessment of DNA accessibility within nucleosomes. Vectors and cloning materials, notably the pEX-A128 vector, were utilized to house and manipulate DNA fragments containing Widom 601 nucleosome positioning sequences.

The research was facilitated by sophisticated laboratory equipment and software tools. Gromacs software has been used to collect the molecular dynamics simulations' trajectories, while PyMOL (The PyMOL Molecular Graphics System, Version 2.5.0 Schrödinger, LLC) and VMD[44] were instrumental for the visualization and analysis of molecular structures and trajectories. The CHARMM36m force field shaped the molecular dynamics simulations, while the four-domain OPC water model defined the water representation in the simulated environment. Gel electrophoresis equipment enabled the separation and visualization of DNA fragments, pivotal to the one-pot assay. The application of a phosphor imager, alongside ImageQuant TL and Typhoon FLA 9500 software, permitted the exposure and meticulous analysis of gels.

In the study of inhibition of ectopic formation, four primary atomistic structural models were constructed for molecular dynamics simulations: the CENP-A and H2A.1 containing

nucleosome, the CENP-A and H2A.Z containing nucleosome, the H3.1 and H2A.Z containing nucleosome, and the H3.1- and H2A.1-containing nucleosomes. These models were based on atomic coordinates sourced from various Protein Data Bank (PDB) structures, specifically PDB ID:6TEM[12], PDB ID:7PII[45], PDB ID:3LZ0[46], PDB ID:5NL0[47], and PDB ID4X23[48]. The PyMOL Molecular Graphics System (version 2.5.4, Schrödinger LLC) was employed to extend the missing regions according to the Histone Variants Database. The simulations also involved modeling with the extended Widom 601 DNA sequence and incorporated two atomic models of CENP-C into each construct. Each of these models was solved in cubic water boxes, accompanied by Na^+ and Cl^- ions to mimic experimental CENP-C binding assay conditions. The modified CHARMM36m force field, the four-site OPC water model, and updated Na^+ and Cl^- ion parameters were utilized along with the Cufix correction and PME for defining chemical interactions. The Gromacs software (version 2019.4) was the backbone for the molecular dynamics' trajectories, while the analysis was primarily conducted using Visual Molecular Dynamics (VMD, version 1.9.4), its Python wrapper library (VMD-Python, version 3.0.1), and the Prodigy algorithm. The Python libraries and PyMOL were crucial for visualization tasks.

Our study in the respect of understanding procaine's mode of action, the primary resources included the protein structures of an enzyme, which is associated with the Protein Data Bank ID: 6F57.[49] Additionally, we procured the atomic details for procaine from the human metabolome database with the HMDB ID: HMDB0014859. Our study was made possible through an array of software tools and databases. The CGENFF[50] tool was vital for obtaining force field parameters for both procaine and SAH. The core simulations were conducted with the assistance of Gromacs version 2019.4 software. Post-simulation, we analyzed and visualized our data using several tools, including Visual Molecular Dynamics (VMD) version 1.9.4, VMD Python version 3.0.1, and the PyMOL Molecular Graphics System version 2.5.4 by Schrodinger LLC.

3.5. Research Variables

In the respect of allosteric factors effects on CENP-A containing nucleosome, independent variables include DNA sequence, histone variant, ion concentration, and Mg^{++} concentration. The DNA sequence is modified specifically within CENP-A containing nucleosomes to reveal its effect on nucleosome stability and interactions. Selection between

CENP-A and canonical H3 histone variants serves as another independent variable, enabling exploration of how different histone variants affect nucleosome behavior. The ion concentration, including Na^+ and Cl^- ions, is varied to distinguish how different salt conditions affect DNA-histone interactions. The concentration of Mg^{++} ions is adjusted independently to investigate its effect on interactions between DNA and lysine/arginine residues.

Dependent variables correspond to the results of simulation, analysis, and experiments. These likely include changes in DNA-histone interactions, nucleosome stability, and chromatin structure dynamics.

Relating to initial centromere formation and inhibition, four atomistic structural models served as key independent variables: CENP-A paired with H2A.1, CENP-A with H2A.Z, H3.1 combined with H2A.1, and an H2A.Z and H3.1 duo. These models were constructed using atomic coordinates derived from specific Protein Data Bank (PDB) IDs, including PDB ID:6TEM, PDB ID:7PII, PDB ID3LZ0, PDB ID:5NL0[47], and PDB ID:4X23. Another set of independent variables are the simulation environment factors, such as the use of the extended Widom 601 DNA sequence and the ion concentration with the presence of Na^+ and Cl^- ions, all designed to mimic CENP-C binding assay conditions.

On the flip side, our dependent variables aligned to the outcomes of the simulations and subsequent analyses. These outcomes included the molecular dynamics trajectories, changes in nucleosome structure, DNA-histone interactions, and the affinity of the incorporated atomic representations of CENP-C to the nucleosome constructs.

Our research method for inhibition of DNMT3A carefully considered both dependent and independent variables. For dependent variables, we looked at things like the path's molecules took over time (molecular dynamics trajectories), how procaine was arranged and behaved inside the SAH pocket, and how the enzyme acted when different amounts of SAH were present. On the other hand, our independent variables were defined by whether SAH was present or not and by the specific concentration of procaine used in the simulation, which we set at 0.07 M. This setup allowed us to systematically study the interactions and changes under different conditions.

3.6. Data Collection Tools/Methods *3.6.1. Understanding DNA flexibility through molecular dynamics simulations and one pot assay*

3.6.1.1. Preparation and molecular dynamics of CENP-A and H3-containing nucleosomes.

We prepared 149 base pairs (Widom 601) of the structures of CENP-A-containing nucleosomes (Carlo Petusa, personal communication)[51] and H3 nucleosome-containing nucleosomes (PDB ID:3LZ0[52]). These were used as the initial atomic coordinates. We used Web 3DNA 2.0[53] to mutate the Widow 601 DNA sequence of CENP-A-containing nucleosomes, which was compatible with the one-pot assay D6L structure. Mutated DNA sites were determined by aligning the DNA sequence on Clustal Omega.[54] All systems were solvated in a cubic water box that was sufficiently large to provide a minimum buffer zone of 12 Å between the biological material and system boundaries. For the PBS mimicking boxes, Na^+ and Cl^- ions were randomly distributed to neutralize the systems electrostatically at a salt concentration of 0.1615 M. For the mitosis-mimicking boxes, Mg^{++}, Na^+, and Cl^- were randomly distributed to neutralize the systems electrostatically at a salt concentration of 0,035 M. CHARMM36m force field[55,56], four-site OPC water model[57], and adjusted ion parameters for Mg^{++}, Na^+, and Cl^- were chosen to represent the interactions between particles.[58] Energy minimization of the system was carried out using a combination of conjugate gradient and steepest descent methods, then equilibrated in the NVT ensemble at 100 K for 1 ns using a 1 fs timestep, then at 310 K for 1 ns, but using a time step of 2 fs. In the NPT ensemble at 310 K and atmospheric pressure (1atm), the production trajectories were collected using an integration time stage of 2 fs for a total of 500 ns for the nucleosomal system. The velocity-rescaling thermostat[59] and Parrinello-Rahman barostat[60] were used to maintain constant temperature and pressure, respectively. Atomic coordinates were saved every 100 ps. We performed the analysis using in-house Python v3.8 scripts and Visual Molecular Dynamics (VMD), version 1.9.4[44], and its Python library, VMD Python (version 3.0.1). Structural alignment and generation of molecular images were performed using PyMOL (The PyMOL Molecular Graphics System, Version 2.5.0 Schrödinger, LLC).

3.6.1.2. Setup preparation and molecular dynamics of CENP-A and H3 alpha N helix

To investigate the influence of Mg^{++} concentration on the specific interaction between DNA and histone lysine or arginine residues that bind DNA on minor grooves, a detailed *in silico* analysis was conducted. Two Widom 601 sequence regions, corresponding to the αN region between the 47th and 55th residues of CENP-A and the 44th and 55th residues of canonical H3, were prepared using PyMOL. The molecular structures were energy-minimized, and the systems were simulated independently in a cubic OPC water box for a duration of 1 μs using the CHARMM36 force field and Gromacs software. To maintain the physiological conditions, Na^+ and Cl^- ions were added to the system to a concentration of 165 mM. Additional Na^+ ions are introduced to balance the excess negative system charge.

The same set of simulations were replicated under conditions of excess Mg^{++} concentration (40 mM) to examine the potential impact of this divalent cation known to have an impact on the secondary structure of chromatin.[61] Trajectories from these simulations were obtained, and the effect of magnesium concentration on the interaction between DNA and lysine/arginine residues within the histones was analyzed using VMD and Python. This comprehensive analysis provides valuable insights into the dynamic behavior and stability of DNA-histone complexes under varying Magnesium ion concentrations. The visualization and computational tools allowed for a detailed examination of the specific interactions and their changes in response to the Magnesium ion concentration. Through this approach, this study aimed to unravel the intricate interplay between divalent cations, DNA, and histones, shedding light on the mechanisms governing the structural organization and the stability of chromatin.

3.6.1.3. Preparation of DNA fragments and one-pot assay

A set of 7 pEX-A128- 601 mutants containing Widom 601 nucleosome positioning sequence were utilized for "one pot assay." Briefly, 240bp 601 sequences were cloned in the pEX-A128 vector between the EcoRI and AflII restriction sites and expressed in E. coli DH5α cells. Each construct harbors HaeIII sites at different superhelical locations, as previously described before Wu C, Travers A. A 'one-pot' assay for the accessibility of DNA in a nucleosome core particle.[62]

Inserts were excised from the vector using the restriction enzymes EcoRI and AflII and purified using a 1% crystal violet agarose gel. Fragments were labeled using Klenow enzyme with either α32P-dATP (top strand) or α32P-dTTP (bottom strand).

Nucleosome reconstitution was performed using the salt dialysis method. Briefly, 8µg of chicken carrier DNA, 1µg of 207bp 601, and an equimolar mixture of the seven different α32P -labeled 240bp 601 DNA fragments (100ng) were mixed with an equimolar amount of CENP-A histone octamer in HFB buffer (2M NaCl, 10mMTris pH-7.4, 1mM EDTA pH-8 and 10mM β-mercaptoethanol). For canonical nucleosome reconstitution, CENP-A was replaced by an equimolar amount of canonical H3 in the histone octamer, and reconstitution was performed as described above. After reconstitution approximately 70ng of each nucleosome was digested in 10µl reaction buffer (25mM Potassium Acetate, 10mM Trisacetate (pH-7.9), 5mM Magnesium Acetate, 100µg/ml BSA and 0.025% NP40) with 3.75U/µl

Hae III at 30°C. The reactions were arrested at different time points by adding 100µl stop buffer containing 0.2% SDS and 40 mM EDTA. The digested DNA was purified by phenolchloroform extraction, and ethanol precipitation was performed using 8% denaturing PAGE.[63]The screens were scanned on a Typhoon FLA 9500 (GE Healthcare) and analyzed using the ImageQuant TL (GE Healthcare) software.

Table 1: The primary sequences of the various inserts.

D6L	ATCAGAATCGGCCTGCCGAGCCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCGAT
D5L	ATCAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCGAT
D4L	ATCAGAATCCCGGTGCCGAGCCCGCTCAAGGCCTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCGAT

D0	ATCAGAATCCCGGTGCCGAGCCCGCTCAATTGGTCGTAGACAGCTCTA GCACCGCTTAAACGCACGTACGGGCCGTCCCCCGCGTTTTAACCGCCA AGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCGA T
D4R	ATCAGAATCCCGGTGCCGAGCCCGCTCAATTGGTCGTAGACAGCTCTA GCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCA AGGGGATTACTCCCTAGGCCCCAGGCACGTGTCAGATATATACATCG AT
D5R	ATCAGAATCCCGGTGCCGAGCCCGCTCAATTGGTCGTAGACAGCTCTA GCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCA AGGGGATTACTCCCTAGTCTCCAGGCGGCCGTCAGATATATACATCGA
D6R	ATCAGAATCCCGGTGCCGAGCCCGCTCAATTGGTCGTAGACAGCTCTA GCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCA AGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGAGGCCTACATCG AT

3.6.2. Molecular dynamics simulation of different histone variant constructs that aligned with CENP-C

3.6.2.1. Setup preparation and molecular dynamics of CENP-A, H3, H2A.Z H2A.1 and CENP-C containing nucleosomes.

We constructed three atomistic structural models for the molecular dynamics simulations: 1) CENP-A and H2A.1 containing nucleosome, 2) CENP-A and H2A.Z containing nucleosomes, and 3) H3.1- and H2A.1-containing nucleosomes. For the nucleosome part of construct #1, we used the atomic coordinates of the structure with PDB ID:6TEM[12] and incorporated the missing C-terminal region of the two copies of CENP-A proteins using the corresponding segment from the structure with PDB ID:7PII.[45] For the nucleosome part of construct #2, we used our own cryo-EM model, onto which we incorporated the same Cterminus extension for CENP-A as we have done for construct #1. For the nucleosome part of construct #3, we used the structure with PDB ID:3LZ0.[46] For all three constructs, we extended the missing C-terminal regions of H2A.Z or H2A.1 using PyMOL Molecular Graphics System, version

2.5.4, Schrödinger LLC. We used the human histone isoforms available in the Histone Variants Database for all sequences of experimentally unresolved peptide segments[64]. To avoid potential DNA end artifacts, we modeled an additional 15 bp on all DNA termini using PyMOL. We built the DNA ends based on the extended Widom 601 sequence, as used in the structure with PDB ID:5NL0.[47] For all three constructs, we incorporated two copies of the atomic models of CENP-C from the structure with the PDB ID4X23[48] by aligning the core octamers of each construct with the core octamers of the 4X23 structure. We solved all three constructs in cubic water boxes sufficiently large to provide a buffer zone of at least 12 Å between the biological material and container boundaries. To mimic the experimental CENP-C binding assay conditions, we randomly incorporated Na^+ and Cl^- ions at a salt concentration of 0.050 M throughout the container box. In addition, we introduced sufficient Na^+ ions to ensure the overall charge neutrality. We employed the modified CHARMM36m[55,56] force field, four-site OPC water model[8], and updated Na^+ and Cl^- ion parameters[58] together with the Cufix correction[65] to describe the chemical interactions between the particles. We used PME[66] to account for long-range electrostatics.

Molecular dynamics trajectories were collected using Gromacs version 2019.4.[67] To ensure stable simulations, we employed a combination of conjugate gradient and steepest descent techniques to minimize the potential energy of each system box. We then thermalized and equilibrated each system gradually, first to 100 K for 1 ns, and then to 310 K for another 1 ns, both in the NVT ensemble and using an integration time step of 1 fs. Following these steps, we collected our production trajectories in the NPT ensemble at 310 K and 1 atm, using an integration time step of 2 fs. Each simulation was performed for 500 ns. To maintain a constant temperature and pressure, we utilized the velocity-rescaling thermostat[59] and ParrinelloRahman barostat[60], respectively. The atomic coordinates were recorded every 100 ps. We analyzed molecular trajectories using Visual Molecular Dynamics (VMD), version 1.9.4[44] ,and its Python wrapper library, VMD-Python, version 3.0.1. Python libraries and PyMOL were used for graphical and molecular visualization, respectively. To estimate CENP-C affinities, we used the Prodigy[68–70] algorithm, whose standalone code was modified to go through all molecular trajectory frames.

3.6.3. Investigating how procaine affects DNMT3A through MD simulations to understand its mode of action

3.6.3.1. Molecular Dynamics Simulations of DNMT3A in Procaine Environment

We prepared two atomistic simulations:1) SAH-free apoenzyme bound to procaine at its SAH-pocket, and 2) SAH-bound enzyme in excess (0.07 M) procaine. We used the atomic coordinates of the enzyme obtained from the Protein Data Bank (PDB ID:6F57).[71] For the atomic coordinates of procaine, the human metabolome database (HMDB ID: HMDB0014859) was used. We used CGENFF[72] to obtain the force field parameters for procaine and SAH. We solvated three of the structures in a cubic water box with a minimum 12 Å buffer zone between the limits of the biological material and the system. We randomly positioned Na^+ and Cl^- ions into each water box to mimic a physiological salt concentration of 0.161 M. The system was electrostatically neutralized by adding Na^+ or Cl^- ions, depending on the net excess charge. For simulations that contained procaine, we introduced randomly distributed procaine molecules to reach a concentration of 0.07 M. We used the CHARMM36m force field[55,56], the four-site OPC water model[57], and updated the Na^+ and Cl^- ion parameters to describe the interactions between atoms and chemical groups.[58] We collected molecular dynamics trajectories using Gromacs version 2019.4.[67] We energy-minimized each system box using a combination of the conjugate gradient and steepest descent algorithms, followed by equilibration in the NVT ensemble for 1 ns at 100 K, and 1 ns at 300 K, and at the NVT ensemble using an integration time step of 1 fs for each. The production trajectories for each of the four enzyme systems were gathered in the NPT ensemble at 300 K and atmospheric pressure (1 atm) using an integration time step of 2 fs for a total of 500 ns. We used the velocity-rescaling thermostat[59] and the Parrinello-Rahman barostat[60] to maintain the temperature and pressure constant, respectively. The atomic coordinates were recorded every 100 ps. We used Visual Molecular Dynamics (VMD), version 1.9.4[44] and its Python wrapper library, VMD Python version 3.0.1. Molecular images were generated using PyMOL Molecular Graphics System, version 2.5.4, Schrodinger LLC.

3.7. Research Plan And Timeline

Table 2: Research Plan

Project 01	Months								
	1-4	5-8	9-12	13-16	17-21	22-26	26-29	29-32	32-36
Literature Review	**X**	**X**	**X**						
Data Collection	**X**	**X**	**X**						
Model Preparation		**X**	**X**	**X**					
Basic MD Simulations and analysis			**X**	**X**	**X**				
Enhanced Molecular Dynamics Simulations				**X**	**X**	**X**	**X**		
In-depth Analysis					**X**	**X**	**X**	**X**	
Experimental Validation					**X**	**X**	**X**	**X**	
Thesis Writing							**X**	**X**	**X**

3.8. Data Analysis

The collected data regarding allosteric effects on CENP-A containing nucleosomes are important points for understanding nucleosome interactions. The first steps include analysis aimed at improving the data towards understanding the structure and setting up the simulation system. Depending on what the data reveals, more customized approaches are used. Molecular dynamics simulations' trajectories and certain statistical approaches shed light on the behavior of nucleosomes. We try to understand the trajectories we have obtained by visualizing using VMD and PyMOL and by writing home-based python codes to get a mathematical result. On the biochemical wet-lab research side, the screens were scanned on Typhoon FLA 9500 (GE Healthcare) and analyzed using the ImageQuant TL (GE Healthcare) software.

For the analysis of the molecular dynamics' trajectories with respect to the inhibition of CENP-A recognition by CENP-C, Gromacs version 2019.4 was employed. Production trajectories were collected in the NPT ensemble at 310 K and 1 atm, integrating every 2 fs for a duration of 500 ns. The atomic coordinates from these simulations were recorded at intervals of 100 ps. These trajectories were analyzed using Visual Molecular Dynamics (VMD), version 1.9.4, and its Python extension, VMD-Python version 3.0.1. Both Python libraries and PyMOL facilitated detailed graphical and molecular visualizations. Furthermore, the Prodigy algorithm was adapted and utilized to estimate CENP-C affinities across the entirety of the molecular trajectory frames.

The production phase of DNTM3A with procaine, we monitored the system under constant conditions of 300 K temperature and regular atmospheric pressure. This was done for a total time of 500 ns. Throughout this period, we kept a record of where each atom was every 100 ps. We used Visual Molecular Dynamics (VMD), version 1.9.418 and its Python wrapper library, VMD Python version 3.0.1. Molecular images were generated using PyMOL Molecular Graphics System, version 2.5.4, Schrodinger LLC.

3.9. Limitations of the Research

Studying ionic environment and DNA variant on CENP-A containing nucleosome offers essential insights into how DNA, histones, and certain ions work together inside nucleosomes. However, it is important to remember that there are some parts we might have missed or oversimplified. When we use molecular dynamics simulations, like in this research, we are making some educated guesses and approximations. How accurate our guesses are depending a lot on things called force fields and some computer settings. Even though we ran these simulations for a long time, they might not show everything that can happen, especially events that do not occur often. We also made decisions about how to represent ions and liquids in our simulations. Different choices might give different results. Real cell conditions are complex, and our computer models might not capture all those details. We also focused on a particular kind of DNA sequence for our study. This means that our findings might not apply to all types of DNA sequences. In our study, we mainly focused on two types of histones,

CENP-A and the usual H3. There are other histones like H2A.Z and H2A.X that also play important roles, but we did not include them. This could affect our results. Therefore, when reading our findings, it is crucial to remember these limitations.

Considering how histone variants affect centromere formation, we took data from various sources (PDB IDs) to make our four main models. This means our models could be a bit different from each other because they come from different places. Also, for one model, we used data that was made in our own lab, so there might be some small mistakes or biases. Using PyMOL for Missing Parts: Some parts of our models were missing, so we used a software called PyMOL to fill in the gaps. This could cause some differences when comparing with the real, natural sequences. CENP-C selection: We added two copies of a model called CENP-C motif to our study. This was a choice we made because we trusted the structure of CENP-C. Adding Ions (Na^+ and Cl^-): We added some ions to try and mimic certain conditions. These ions are like cell environments, but our setup might not act the same way as it does in a real cell.

While our results about explaining mode of action procaine are encouraging, there are certain limitations to our study. One of the main challenges in molecular dynamics simulations is the accurate representation of molecular interactions, and the choice and evolution of force fields play a crucial role in this. As force fields have evolved over time, they've been refined to provide better predictions for different types of molecules and interactions. However, they still contain approximations and may not be perfectly suited for every system. Therefore, our simulations, reliant on these force fields, necessitate validation through experimental studies. Moreover, the detailed interactions between DNMT3A and procaine, and the possible ways they influence each other's activity (known as allosteric mechanisms), still need more in-depth research for a full understanding.

3.10. Ethical Committee Approval

The study with protocol number 2021-028 underwent evaluation by the ethics committee on July 26, 2021, and was determined to be ethically appropriate.

4. RESULTS

4.1. Detecting Allosteric Influences on Nucleosomes Containing CENP-A

Herein, we explored the effects of environmental ionicity and DNA sequence alteration on one of the important histone variants, CENP-A, on nucleosome dynamics. Based on the crystal structures of the octameric nucleosomes, including CENP-A (PDB ID: 6TEM) and H3 (PDB: ID 3LZ0), we examined the structural variations and main modes of motion of the CENP-A nucleosome in comparison to the canonical H3 containing nucleosome under different ionicities using all-atom molecular dynamics (MD) simulations. Five systems were studied: (1) Widom 601 DNA sequence containing canonical H3 nucleosomes in PBS; (2) Widom 601 DNA sequence containing CENP-A nucleosome in PBS; (3) Widom 601 DNA sequence containing CENP-A nucleosome in excess Mg^{++} ;(4) D6L DNA sequence containing CENP-A nucleosome in PBS; and (5) D6L DNA sequence containing CENP-A nucleosome in excess Mg^{++}.

4.1.1. Mitotic-like ionic conditions cause dna flexibility on both ends of the CENP-A nucleosome.

We computed distances to the tetramer core to understand how ionicity and DNA sequences affect nucleosomal DNA flexibility. (Figure 4A) Widom 601 containing CENP-A nucleosome (Figure 4B) exhibited ionicity changes that increased the flexibility of the DNA arms. D6L containing CENP-A nucleosomes (Figure 4C) indicated that the DNA sequence and ionicity increased the flexibility of the DNA arms even more when compared with the same environments of Widom 601. Interestingly, when the concentration of divalent atoms in the solution decreased, undesirable interactions could occur in a pre-mitotic environment. Assembly visualization of D6L and Widom 601 DNA containing nucleosome under varying conditions clearly show the DNA alteration with pre-mitotic condition increase flexibility of both DNA ends. (Figure 4D-E)

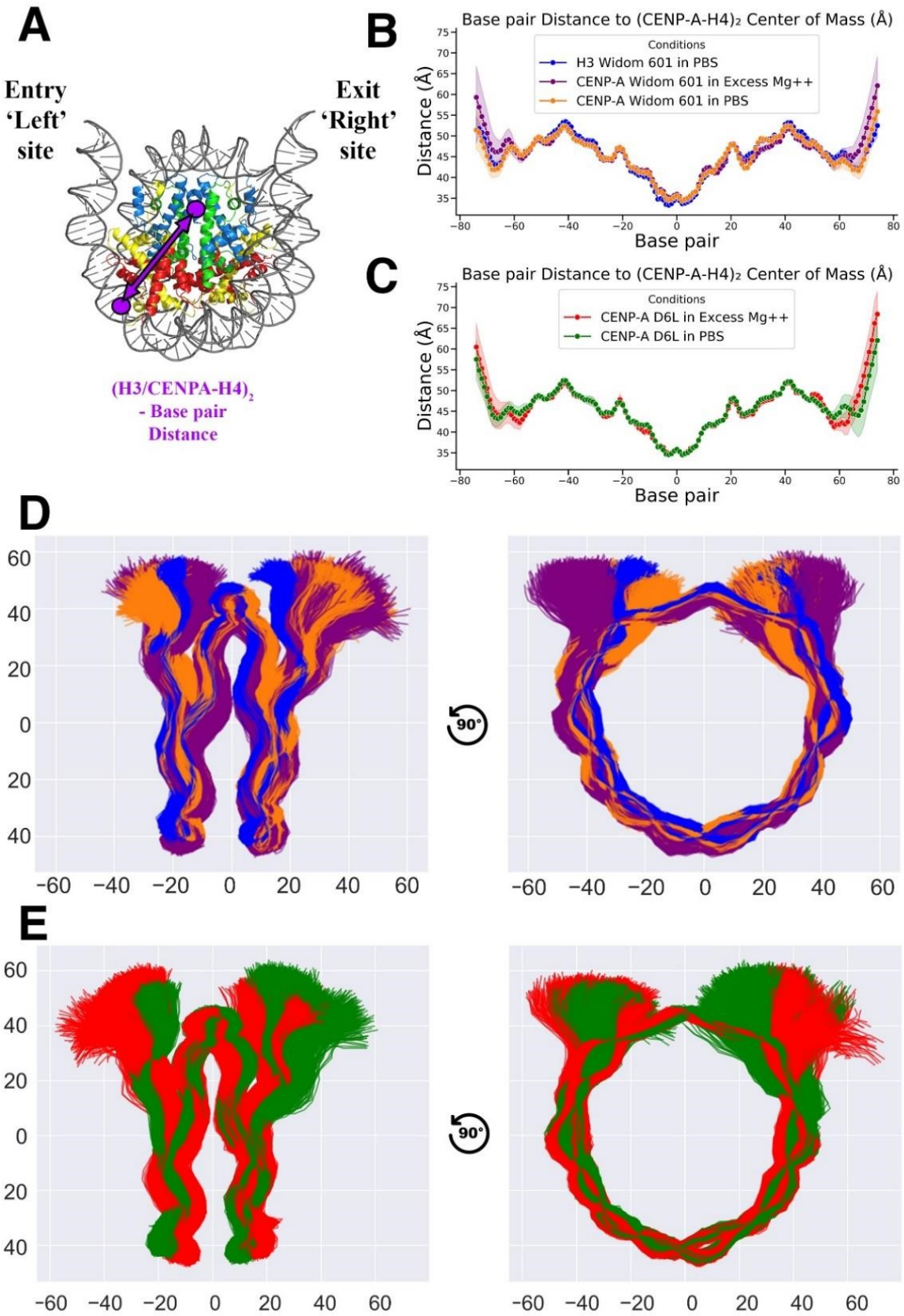


Figure 4: Allosteric modifications in the centromere-specific nucleosome result in the unfolding of DNA and its symmetrical opening. A) Illustration of the CENP-A-containing nucleosome structure. BC)

Measurement of distances between each base pair of nucleosomal DNA and the CENP-A/H4 tetramer. D-E) Visualization of the assembly of D6L DNA and Widom 601 DNA under different environmental conditions. The conditions represented are as follows: Widom 601 CENP-A nucleosome in an environment with excess Mg^{++} (depicted in purple), Widom 601 CENP-A nucleosome in PBS (orange), Widom 601 H3 nucleosome in PBS (also in orange), D6L CENP-A nucleosome with excess Mg^{++} (shown in red), and D6L CENP-A nucleosome in PBS (green).

4.1.2. Core octamers respond differently depending on their variant-composition, DNA sequence, and the ionic environment.

Beyond individual residues, structural variants and the environment generate structural fluctuations in nucleosomes at multiple spatial scales. Previous study demostrated that the core octamers of both H3- and CENP-A-containing nucleosomes undergo subtle but distinct structural changes. These alterations occur whether or not there is antibody binding.[73]Since CENP-A containing nucleosome is more delicate than the canonical H3 containing nucleosome, we want to understand this structural changes. We computed distance between : H3 (or CENPA)/H4 (Figure 5A), H2A/H2B pairs (Figure 5B), and between H3 (or CENP-A)/H4 and H2A/H2B(Figure 5C). When the Widom 601 with CENP-A histone replaces H3, it extends along the tetramer axis by roughly half an Ångström, regardless of the environmental context (whether in PBS or pre-mitotic conditions). There are also significant changes along the dimer axis of Widom 601 containing CENP-A nucleosome, with the mitotic environment further amplifying the expansion. (Figure 5D-E) In a pre-mitotic condition, the Widom 601 containing CENP-A shows no discernible changes along the dimer-tetramer axis compared to the H3containing nucleosome. However, when compared to the Widom 601 with CENP-A in a PBS condition, there is a compression observed. (Figure 5F) What was surprising to us was our initial expectation that DNA alterations would also lead to increased expansion. Contrary to this, when D6L contains CENP-A in pre-mitotic conditions, it compresses along both the tetramer axis and the dimer-tetramer axis by approximately half an Ångström. In a PBS environment, the Widom 601 with CENP-A does not display any notable changes along either the dimer-tetramer axis or the dimer axis when compared to the D6L in PBS conditions. Yet, the D6L with CENP-A in pre-mitotic conditions expands along the dimer axis but compresses on the dimer-tetramer axis (Figure 5G-J).

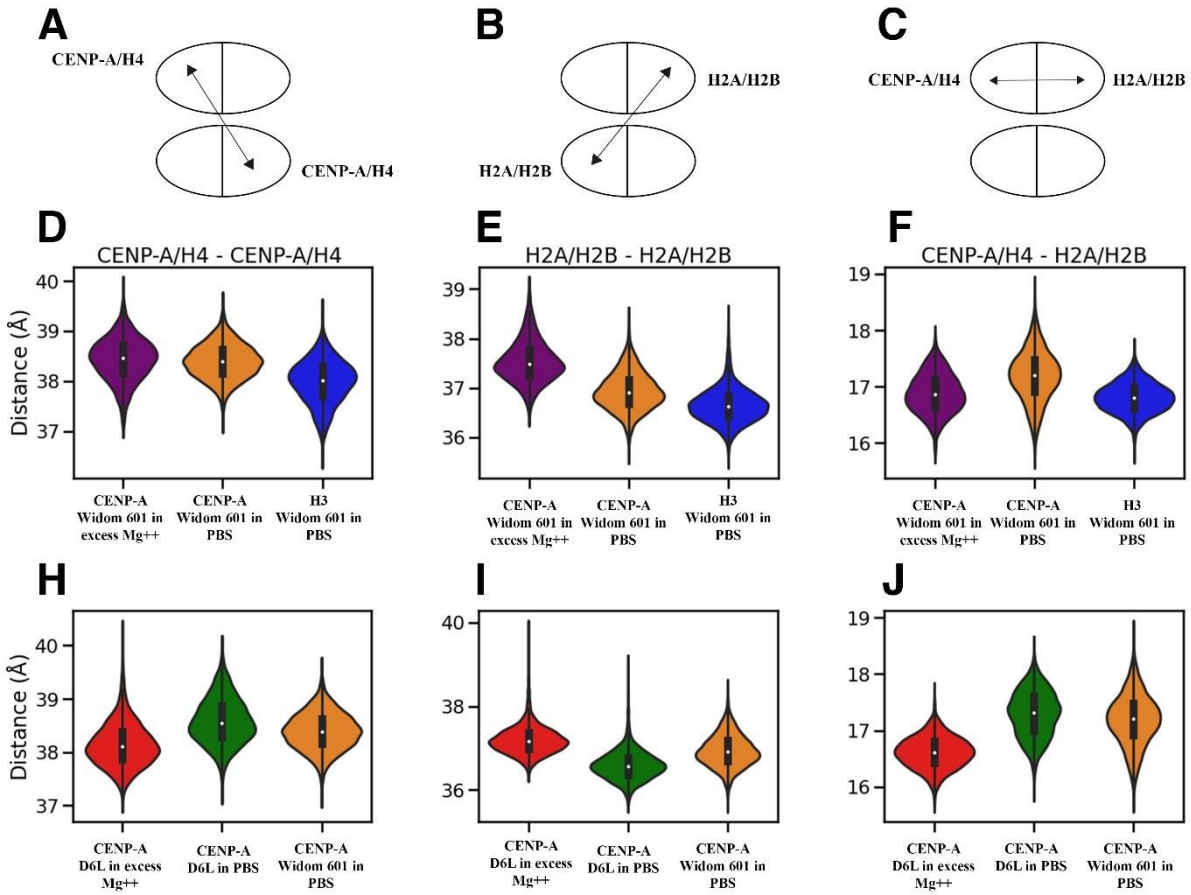


Figure 5: Analysis of the Nucleosome Core's Variant-Specific Responses to Alterations in DNA and Ionic Environments. Panels A-C depict the calculated distances between pairs: H3 (or CENP-A)/H4 (shown in A), H2A/H2B pairs (in B), and between H3 (or CENP-A)/H4 and H2A/H2B (in C). The distance distributions for these pairs are showcased in panels E-J. Specifically, the following conditions are represented: Widom 601 CENP-A nucleosome in excess Mg^{++} (purple), Widom 601 CENP-A nucleosome in PBS (orange), Widom 601 H3 nucleosome in PBS (also in orange), D6L CENP-A nucleosome in excess Mg^{++} (red), and D6L CENP-A nucleosome in PBS (green). The average values for these distributions are marked by white dots.

4.1.3. Divalent cation modulates DNA-histone interactions by modulating the minor-groove binding of select lysine or arginine residues.

To gain a deeper understanding of the interactions between divalent cations, DNA, and histones, comprehensive simulations were undertaken. These simulations focused on two distinct histone parts called αN region using the same DNA sequence. In our observations, a particular sequence engaged with the K49 of CENP-A, a crucial substitution also this identical sequence also interacted with R49 within the corresponding region of the canonical H3. At this point, these constructs were also tested in pre-mitotic and PBS conditions. (Figure 6A-D)

After analyzing the simulation results, an interesting pattern emerged. The separation of K49, which is the substitution contained in CENPA, from DNA was only seen when exposed to an environment where magnesium is abundant. This separation underscores the potentially reduced stability of CENP-A compared to H3, considering interactions with lysine and arginine residues. It was observed that the heightened concentration of magnesium only makes CENPA more susceptible to separating from the DNA than its H3 counterpart. (Figure 6E-F)

Computational analysis shows pivotal role of magnesium in determining the complex interplay between lysine, arginine residues, DNA, and histones. For a comprehensive analysis, we determined the distance between the center of mass coordinates for both αN region and DNA over the course of the simulation. Specifically, increased magnesium concentrations resulted in a more notable separation of the K49 of αN region CENP-A from DNA compared to the R49 of H3. Furthermore, the shielding effect of magnesium[74] on DNA did not significantly influence the R49 of H3 αN region.

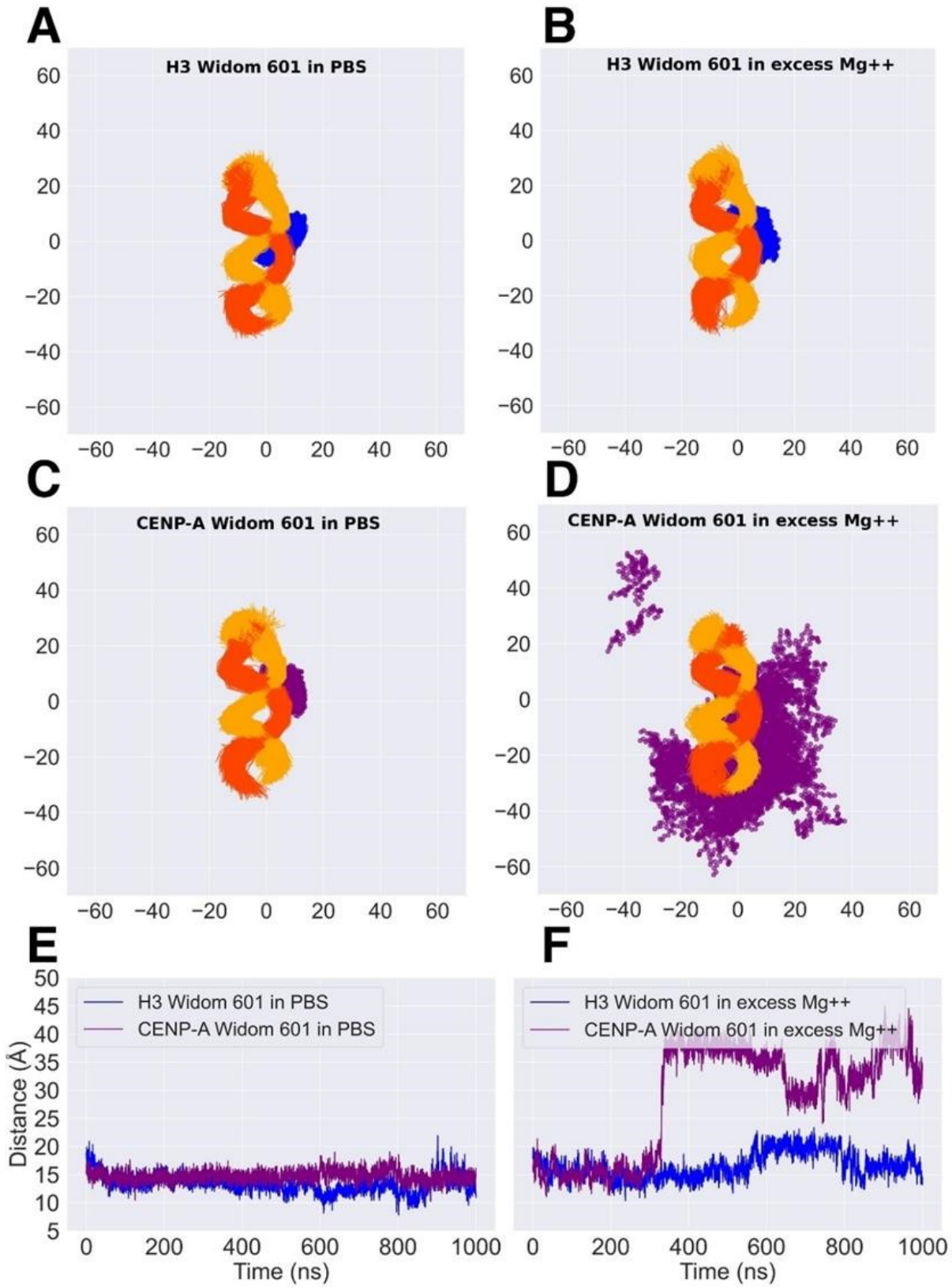


Figure 6: Exploring the Interaction between Divalent Cations and DNA-Histone via Lysine and Arginine Residues. A-D) The panels provide visual representation of the assembly of Widom 601 DNA

with lysine (from CENP-A) and arginine (from H3) in varied environmental contexts. Specifically, the interactions are illustrated between Widom 601 and the lysine residue of the alpha-N region of the CENP-A nucleosome (displayed in purple), and between Widom 601 and the arginine residue of the alpha-N region of the H3 nucleosome in PBS (represented in blue). E-F) These panels chart the changes in distances between the center of mass of the alpha-N region and the center of mass of the DNA during the simulations, specifically focusing on the initial (0th) and final (1000th) nanoseconds.

4.1.4. Measurement of the flexibility of the canonical H3- or the CENP-A containing nucleosome ends by using the one pot assay.

To analyze the above hypothesis, we studied the accessibility of the CENP-A nucleosomal ends by using a one-pot assay[62,75], using seven different 32P-end labeled (either at the bottom or at the top strand) mutated 255 bp 601.2 DNA sequences to reconstitute centrally positioned conventional and CENP-A nucleosomes, each sequence bearing a unique HaeIII restriction site. The dyad of the nucleosomes was designated dyad-0 (d0). The three sites inserted in the left side were designated as dyad 4 L (d4L), dyad 5 L (d5L), and dyad 6 L (d6L), while those on the right side were designated as dyad 4R (d4R), dyad 5R (d5R), and dyad 6R (d6R), where the number indicates the number of helical turns from the dyad (Figure 7A). An equimolar mixture of the seven CENP-A centrally positioned nucleosomes was incubated with HaeIII for different times ranging from 1 min to 32 min, and DNA was isolated from the HaeIIIdigested particles and separated by 8% denaturing PAGE (Figure 7B). The same experiment was performed using the conventional H3 nucleosomes. The gel was then dried and exposed, and the cleavage bands were quantified and expressed as a percentage of the cut fraction (Figure 7C). The data clearly show that: (i) the accessibility to HaeIII of both the right (R) and the left (L) ends of the CENP-A nucleosome is much higher (close to tenfold) than that for both ends of the conventional H3 nucleosome; (ii) no difference in the HaeIII accessibility of the right and the left end of the CENP-A nucleosome as well between the ends of the conventional nucleosome was observed.

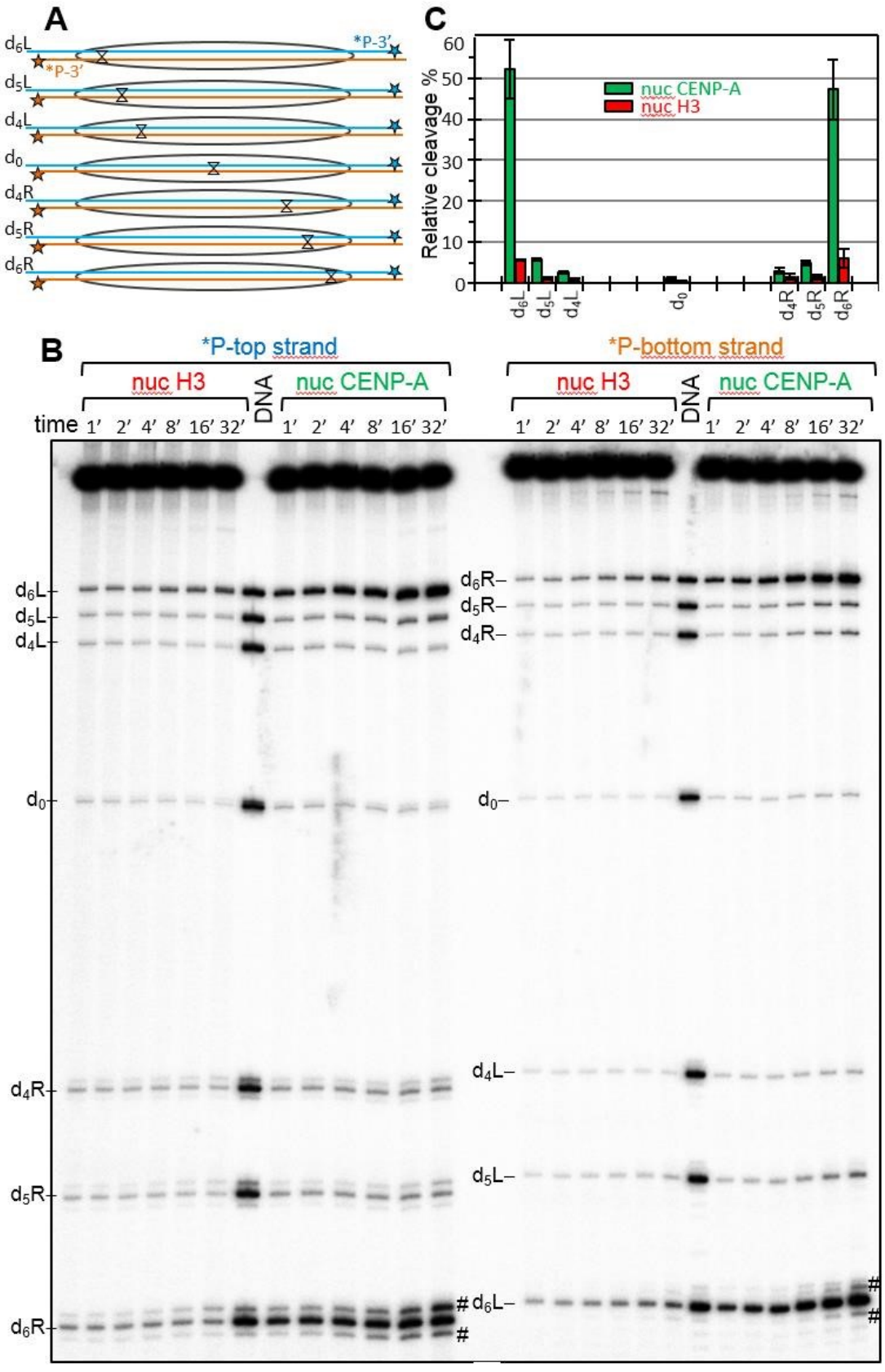

A
d6L
*P-3′
d5L
*P-3′
d4L
d0
d4R
d5R
d6R
C
Relative cleavage %
60
50
40
30
20
10
0
nuc CENP-A
nuc H3
d6L
d5L
d4L
d0
d4R
d5R
d6R
B
*P-top strand
*P-bottom strand
nuc H3
DNA
nuc CENP-A
time 1′ 2′ 4′ 8′ 16′ 32′
d6L–
d5L–
d4L–
d0–
d4R–
d5R–
d6R–
d6L–
#
#

Figure 7: One-pot restriction-accessibility assay of the CENP-A nucleosome. a. Schematics of the nucleosome used for the one-pot assay. b. Time course of HaeIII DNA digestion pattern of the CENPA and conventional H3 nucleosomes. Both samples were incubated under identical conditions with Hae III for the indicated times, DNA was then isolated and run-on denaturing PAGE. Either the top or bottom DNA stand was ^{32}P-labeled; #, bands associated with the exonuclease activity of the Klenow enzyme. c. Quantification of the data presented in b. The data are the averages of two independent experiments using two sets of different reconstituted nucleosomes. Two independent measurements of each dyad corresponding to the ^{32}P-labeled top and bottom strands are used for the calculation of the average accessibility. The bands associated with the Klenow exonuclease activity are also considered in the calculation, bar, and standard deviation.

4.2. H2A.Z exerts an inhibitory impact on the process of centromere formation

4.2.1. Molecular dynamics simulations attribute the loss of CENP-C affinity to H2A.Z-specific substitutions.

We performed sub-microsecond atomistic simulations of three chimeric nucleosome core particles: 1) CENP-A and H2A.1 containing nucleosome, 2) CENP-A and H2A.Z containing nucleosome, and 3) H3.1 and H2A.1 containing nucleosome. All nucleosomes are initially bound to two copies of the 25-residue long CENP-C motif of Rattus norvegicus CENP-C on their octamer façades (Figure 8A, also see Materials and Methods). The starting atomistic configuration of CENP-C follows the crystal structure of the nucleosome containing an H3/CENP-A hybrid construct[48] where CENP-C exhibits favorable ionic contacts with the core acidic patch through residues R717 and R719 (Figure 8B-C). These residues present a common chromatin-binding motif known as the "Arginine anchor"[76] whose crystallographic bound configuration is best preserved in the MD simulation of the CENP-A/H2A.1 containing nucleosome (construct #1, Figure 8D). In all three simulations, CENP-C exhibits a general structural heterogeneity which is significantly more pronounced in the presence of H2A.Z or H3 (Figure 8E-F). Canonical H3's C terminal is 3 amino acids shorter than the CENP-A tail. The conserved residues Y725 and W726 in the CENP-C motif are responsible for recognizing the residues I133 and L137 in the IEGGLG tail of CENP-A through hydrophobic interactions. However, this interaction is absent in the H3 tail. Instead, only the R129 residue interacts with Y725 and W726, leading to a destabilization effect (Figure 8G-H). [77] In the CENP-A/H2A.Z containing nucleosome (construct #2), the C-terminal half of CENP-C detaches from the nucleosome facilitated by favorable interactions with several H2A.Z docking domain residues. (Figure 8I). This type of interception is absent in the H2A.1 containing constructs (#1 and #3)

where the docking domain dynamically and non-specifically anchors to linker DNA via H123 H124, K125, K127 and K129, two residues do not present in H2A.Z and only K121, K122 and K126 anchors to DNA. (Figure 8J-K). In the CENP-A/H2A.Z-containing nucleosome (construct #2), the detachment of CENP-C is more severe compared to other scenarios. This detachment is mainly due to the reduction in hydrophobic contacts, such as those between CENP-C's D732 residues and the hydrophobic C-terminal tail residues L134 and L139 of CENP-A. Instead, strong contacts are formed between the C-terminal residues I119 and K126 of H2A.Z and the residues R730 and D732 of CENP-C, respectively. These changes in contacts contribute to the milder form of detachment observed in this nucleosome construct. These changes lead to a tougher form of detachment observed in this nucleosome configuration (Figure 8L).

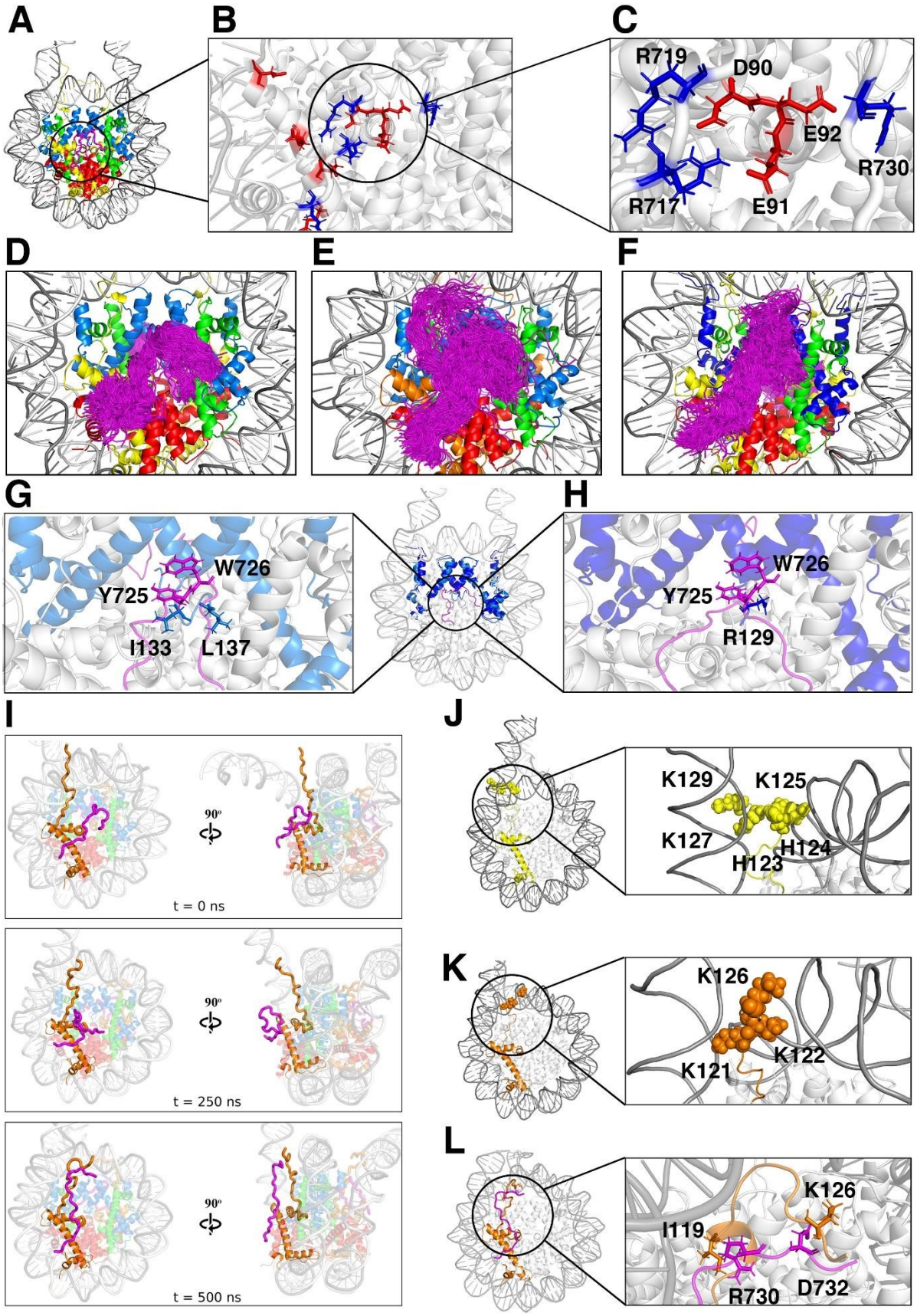
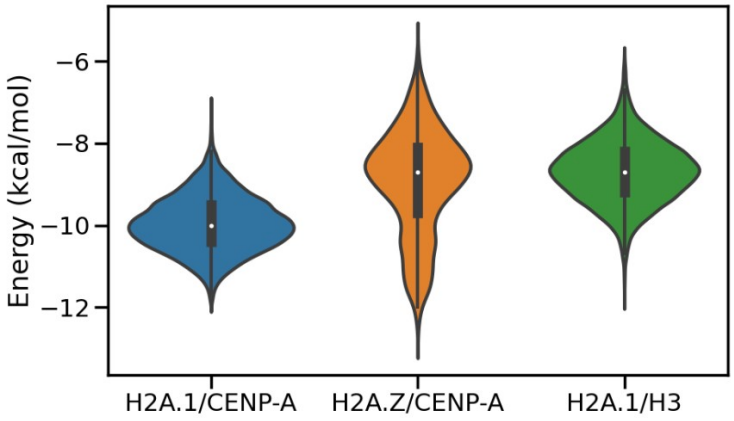


Figure 8: H2A.Z influences CENP-C recognition, leading to its detachment from the nucleosome surface. A-C, CENP-C arginine anchors interact with the nucleosome core via the acidic patches of H2A

and H2B. D-F, CENP-C assembly in different nucleosome configurations, including H2A.1/CENP-A, H2A.Z/CENP-A, and H2A.1/H3. G-H: The crucial role of the conserved residues Y725 and W726 in the CENP-C motif is to recognize CENP-A's, H2A.Z's C-terminal tail. I, CENP-C detachment from H2A.Z/CENP-A containing in time intervals. J, Stabilization of the C-terminal tail by histidine residues of H2A.1 through anchoring to DNA. K, H2A.Z lacks histidine residues, resulting in release of the Cterminal tail. The release of H2A.Z's C-terminal tail leads to CENP-C detachment, mediated by electrostatic forces between K126 of H2A.Z and D732 of CENP-C.

Table 5: CENP-C affinities and the corresponding dissociation constants KD of three chimeric nucleosomes as inferred from MD simulations. Violin plots show the distributions obtained from all MD frames recorded. White dots indicate the mean values.

Nucleosome	Affinity (kcal/mol)	K_D (nM)
H2A.1/CENP-A	–10.0	85.5
H2A.Z/CENP-A	–8.60	630
H2A.1/H3	–8.79	858



4.2.2. Molecular dynamics simulations attribute the loss of CENP-C affinity to labile nucleosome.

We also performed sub-microsecond atomistic simulation for chimeric nucleosome core particles: H3.1 and H2A.Z containing nucleosome. As we expected CENP-C exhibits a general structural heterogeneity more pronounced in the presence of H2A.Z and H3 (Figure 9). In the context of H3.1/H2A.1, our findings showed that the R129 residue exclusively interacts with Y725 and W726, resulting in a destabilizing influence. The added interactions between residues I119 and K126 of H2A.Z with residues R730 and D732 of CENP-C amplify this detachment effect. We also calculated the affinity and K_D as -8.54 kcal/mol and 956 nM, respectively.

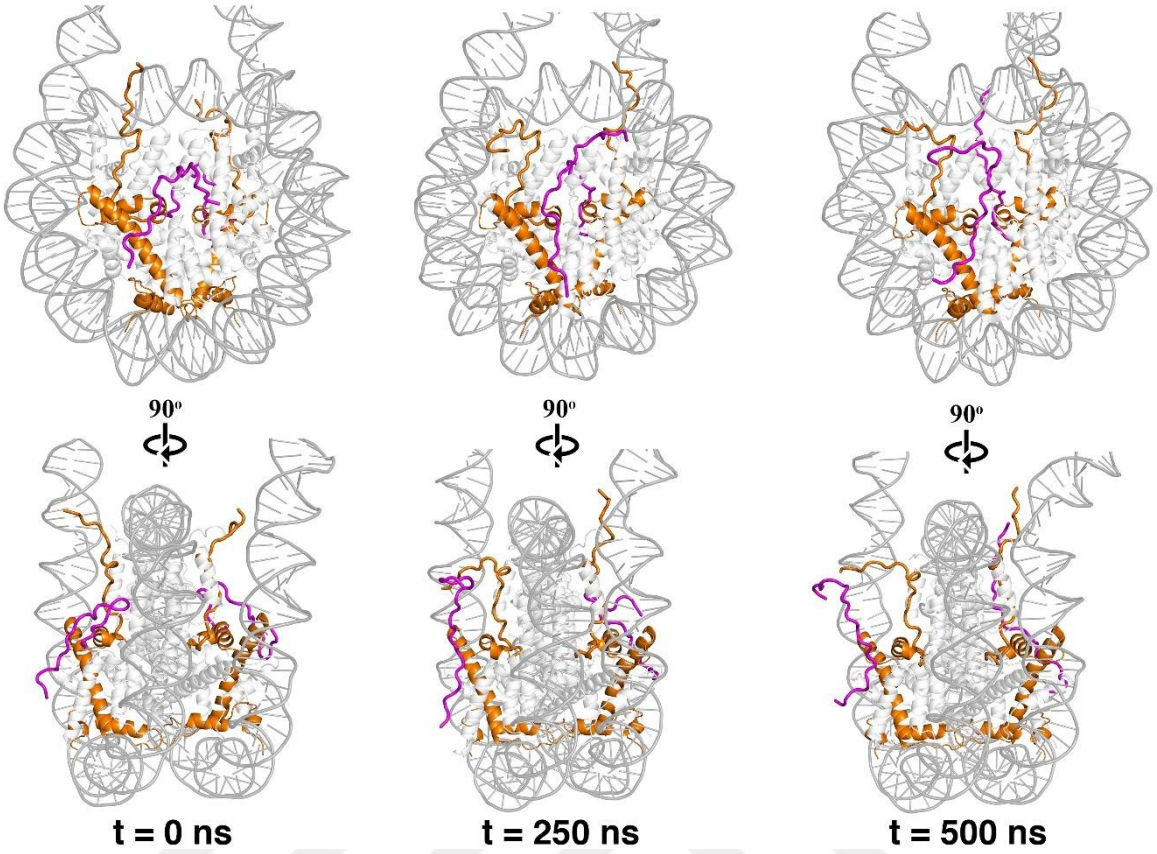


Figure 9: H2A.Z/H3.1 (Labile) nucleosome reject binding of CENP-C. Time interval from 0 ns to 500 ns, The release of H2A.Z's C-terminal tail and decreased interaction of H3.1 C terminal leads to CENPC detachment.

4.3. Allosteric Inhibition of DNMT3A by Procaine

Several DNA methyltransferases (MTases), including DNMT3A, utilize S-adenosyl-Lmethionine (SAM) as the methyl group donor, resulting in the conversion of SAM into Sadenosyl-L-homocysteine (SAH) following the transfer of the methyl group onto the DNA.[78] SAH is present in various cellular compartments, including the nucleus, cytoplasm, and the extracellular environment. An increase in SAH levels promotes the binding of DNA methyltransferases to active sites, thereby reducing DNA methylation levels (Figure 10A top). Despite identifying several non-nucleoside inhibitors of DNA methyltransferases, including the anti-arrhythmia drug procainamide, the search for novel demethylating agents remains ongoing. Therefore, we investigated the potential use of procaine, an anesthetic drug closely related to procainamide, as a chemical agent that promotes demethylation.[79] This investigation was motivated by the fact that procaine exhibits binding affinity towards CpG-enriched DNA and inhibits DNA methyltransferases. Notably, a recent study demonstrated that DNA methylation levels are reduced by SAH alone and in combination with procaine.[80] Our simulations indicate that procaine strongly interacts with DNMT3A, albeit at sites distal from its catalytic site, suggesting a potential allosteric mechanism. Furthermore, we found that procaine does not favorably bind to the pocket otherwise occupied by SAH, suggesting that competitive binding between SAH and procaine is an unlikely mode of inhibitory action.

To gain molecular-level understanding of the inhibition mechanism of procaine, we conducted a comprehensive investigation utilizing ligand docking and atomistic molecular dynamics (MD) simulations. The precise mechanism by which procaine interacts with DNMT3A for inhibition remains unknown, necessitating a thorough analysis to elucidate the binding dynamics between procaine and specific regions of DNMT3A (Figure 10A bottom). We employed unbiased docking using SwissDock[72,81] to identify the most favorable binding pocket of procaine on the DNMT3A surface. Interestingly, Swissdock identified the active site of DNMT3A, near which the essential cofactor SAH is known to reside, as the preferred binding pocket (Figure 10B). This observation suggests a potential interaction between procaine and the crucial catalytic site of the enzyme, which was subsequently tested using MD.

Starting with the active site of DNMT3A occupied by procaine instead of SAH, we performed a sub-microsecond MD simulation (Figure 10C left, see *Methods*). Contrary to the docking prediction, procaine abandoned the SAH pocket (Figure 10C right) and relocated to an

alternative region on the outer enzyme façade (Figure 10D). This unexpected observation led to the hypothesis that procaine may exert allosteric effects on DNMT3A by modulating its activity through interactions with distal sites.

To investigate the presence of additional allosteric sites on DNMT3A, we performed a second sub-microsecond simulation, in which we placed SAH back into its binding pocket and introduced excess procaine into the environment. Procaine molecules in the solvent were randomly distributed in the environment. (Figure 10E). Throughout this simulation, SAH remained bound to the active site of DNMT3A, emphasizing its critical role in enzymatic function. Procaine molecules mapped out a small number of distinct loci on the enzyme surface, all distant from the active site (Figure 10F). Considering the inhibitory role of procaine in DNMT3A activity, these sites may be implicated in allosteric modulation of the enzyme.

In conclusion, our combined approach of docking and MD simulations provides valuable insights into the mode of action through which procaine inhibits DNMT3A. We found evidence that procaine-induced inhibition could occur through the allosteric regions of the enzyme instead of chemical competition with the cofactor SAH. This hypothesis can be experimentally tested using a double-titration assay including procaine and SAH. Finally, we noted the differences between the docking- and MD-predicted binding modalities of the ligand. This highlights the importance of energetic contributions due to the inherent flexibility of the ligand and the enzyme, as well as the delicate chemical interactions in an explicitly represented solvent environment.

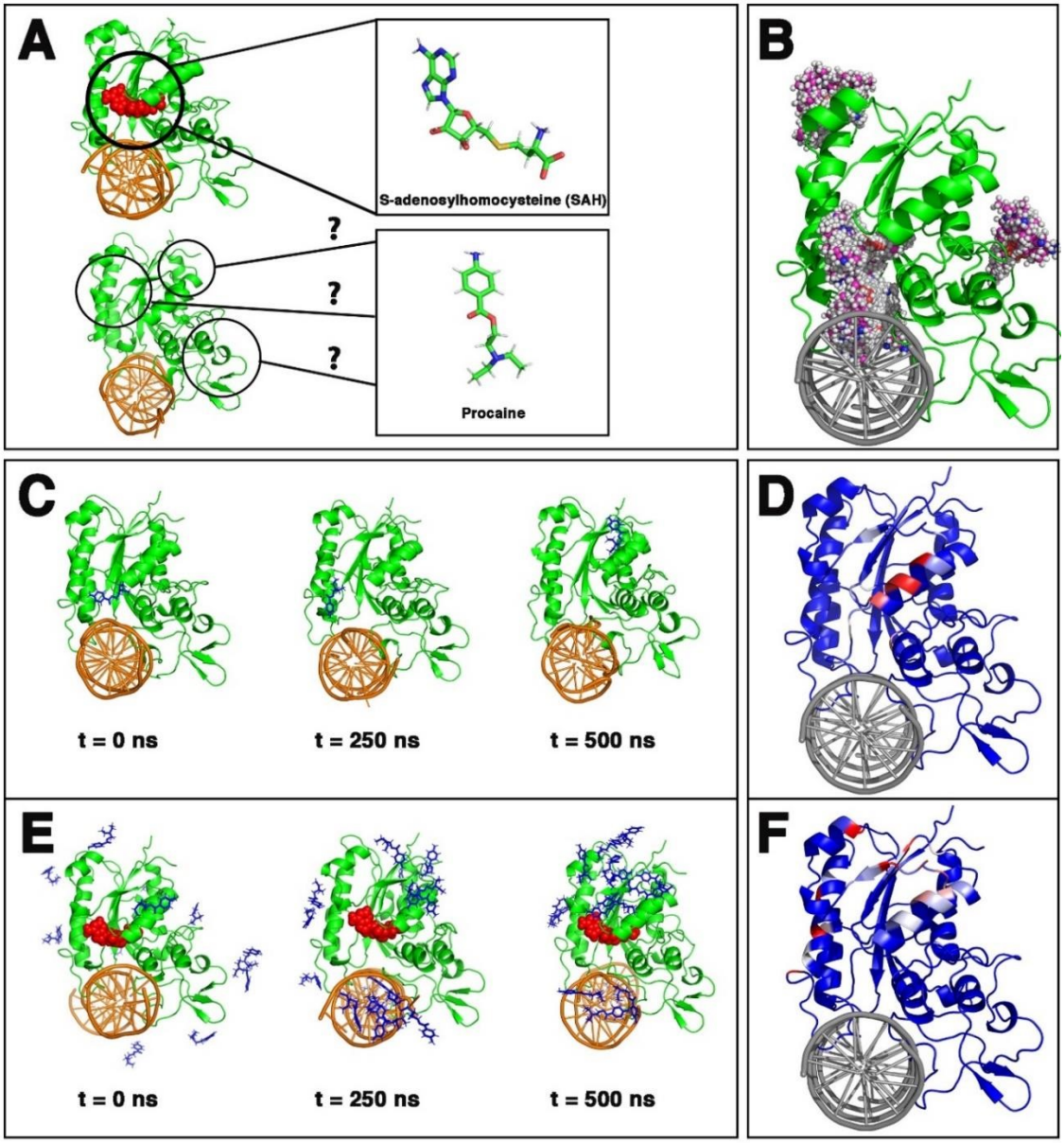


Figure 10: Molecular simulations of DNA-bound DNMT3A in procaine. A) DNMT3A (green) bound to DNA (orange). The cofactor SAH is in red spheres in top model. Bottom model illustrates the possibilities by which procaine could be interacting with DNMT3A. B) Ensemble of the molecular conformations of procaine as predicted by SwissDock to bind DNMT3A most favorably. C) Snapshots of the initial (time=0), mid-point (time=250 ns), and final (time=500 ns) states from the apoenzyme simulation. The cofactor SAH is replaced by one copy of procaine (blue) which departs this position at the end of the simulation. D) Residues that make strong contacts with procaine are indicated in shades of red (red refers to the strongest interaction). E) Snapshots of the initial (time=0), mid-point (time=250 ns), and final (time=500 ns) states from the SAH-bound enzyme in 0.07 M procaine. F) Following panel D, strong procaine contacts are indicated.

5. DISCUSSION 5.1. Allosteric Alterations in the Centromere-Specific Nucleosome Induced by Physical Factors

Segregation of sister chromatids during cell division is a fundamental process in genetic inheritance. Proper association between spindle apparatus fibers and chromosomes is crucial for ensuring accurate chromosome segregation. The kinetochore, which assembles specialized chromatin regions called centromeres, plays a key role in establishing load-bearing attachments between the sister chromatids and spindle microtubules. Central to kinetochore formation is the epigenetic factor CENP-A, which is a centromere-specific variant of histone H3. In this study, we investigated the effect of divalent cation concentration on the structure and dynamics of CENP-A-containing nucleosomes. We performed molecular dynamics simulations and experimental assays to gain insight into the behavior of nucleosomes under different conditions. Our findings revealed several important aspects of the CENP-A nucleosome. First, molecular dynamics simulations demonstrated that the DNA termini of CENP-A-containing nucleosomes exhibit higher flexibility than canonical H3 nucleosomes. The αN helix of CENP-A, which contains lysine residues, contributes to increased flexibility of the DNA ends. This flexibility is believed to be crucial for the assembly of the CCAN complex and the maintenance of the fidelity of the mitotic pathway. This structural difference suggests that CENP-A nucleosomes have a weaker association between DNA and the histone core, potentially leading to increased accessibility of the nucleosomal DNA. These observations highlight the unique properties of CENP-A nucleosomes that distinguish them from the canonical H3 nucleosomes.

Additionally, we investigated the influence of divalent cations, specifically magnesium (Mg^{++}), on CENP-A nucleosomes. During mitosis, the nuclear membrane ruptures, exposing the chromosomes to a cytoplasmic environment characterized by altered levels of divalent cations. An important factor is the increase in free Mg^{++} levels, while Ca^{++} levels remain constant, which could affect the stability of nucleosomes. The electrostatic attraction between the oppositely charged globular histone core and the DNA plays a crucial role in maintaining nucleosome stability. However, an unbalanced charge distribution between DNA and histones can lead to DNA unfolding. In this context, the reduced positive charge of the CENP-A core compared to that of the canonical H3 core suggests that CENP-A nucleosomal DNA is closer to the unwrapping state. This implies that changes in divalent cation concentrations, specifically

Mg^{++}, in the cytoplasmic environment during mitosis can potentially affect CENP-A nucleosome stability and structure.

Our *in silico* results validated by the one-pot assay further demonstrated that DNA sequence alterations in combination with changes in ionicity influence the dynamics of CENPA nucleosomes. Molecular dynamics simulations indicate that the Widom 601 DNA sequence, known for its strong affinity for histone octamers, exhibited increased flexibility of DNA arms when combined with the CENP-A nucleosome. Moreover, the D6L DNA sequence showed induced flexibility in both low ionic and low magnesium environments. These findings suggest that specific DNA sequences and ionic conditions can modulate the structural dynamics of the CENP-A nucleosome.

Furthermore, we explored the interplay between divalent cations and DNA-histone interactions, specifically focusing on lysine and arginine residues within the αN region of CENP-A. Our simulations revealed that an excess concentration of divalent cations, specifically Mg^{++}, leads to a reduction in the interaction strength between lysine and DNA. This disruption weakens the stability of CENP-A nucleosomes and increases the mobility of the αN helix.

Overall, our study provides valuable insights into the structure and dynamics of CENPA nucleosomes and their sensitivity to divalent cation concentration. We demonstrated that the unique properties of CENP-A nucleosomes, such as increased flexibility at DNA termini and unwrapping of DNA, contribute to their distinct behavior compared with canonical H3 nucleosomes. Additionally, we highlighted the role of divalent cations, particularly Mg^{++}, in modulating CENP-A nucleosome stability and structure. These findings contribute to our understanding of the molecular mechanisms underlying accurate chromosome segregation during cell division and may have implications for the development of therapeutic strategies targeting aberrant chromatin dynamics in diseases, such as cancer.

5.2. H2A.Z Suppresses Ectopic Kinetochore Formation by Inhibiting CENP-A Nucleosome Recognition through CENP-C

Further research in this field should focus on investigating the interplay between CENP-A nucleosomes and other components of the kinetochore complex. Understanding how CENPA nucleosomes interact with proteins such as CENP-C and CENP-N as well as the impact of these interactions on chromosome segregation would provide a more comprehensive understanding of the kinetochore assembly process. Additionally, exploring the effects of other divalent cations and their interplay with DNA-histone interactions could uncover further nuances in CENP-A nucleosome behavior.

To elucidate the molecular mechanisms underlying specific recognition of CENP-C, we performed molecular dynamics simulations of CENP-C in complex with chimeric nucleosomes containing H2A.1, H2A.Z, CENP-A, and/or H3.1. Our simulations revealed that in the presence of H2A.Z, the conformational ensembles of nucleosome-bound CENP-C exhibit significant structural heterogeneity, reminiscent of weak encounters. In addition, CENP-C deviated greatly from its initial configuration, as resolved in the crystal structure bound to the nucleosome core particle. Our MD simulations indicate that the primary driver of this effect is the docking domain of H2A.Z, which binds to the C-terminal half of the CENP-C motif and mobilizes it away from the octamer façade. In contrast, the C-terminal tail of the canonical H2A.1 is rich in lysine and histidine residues, which interact extensively with DNA. As a result, the docking domain of H2A.1 does not interfere with CENP-C binding. These findings suggest that H2A.Z could downregulate CENP-A affinity for CENP-C, as could be operational in, for example, inhibiting the assembly of ectopic kinetochores.

5.3. Understanding DNMT3A Inhibition by Procaine Using Molecular Dynamics Simulations

This study provides valuable insights into the complex relationship between DNMT3A and procaine, revealing potential mechanisms behind DNMT3A inhibition. By merging structural knowledge, implications in cancer, and molecular dynamics simulations, we've deepened our grasp of these vital interactions and paved the path for future research and therapeutic advancements. The fact that DNMT3A is so well-preserved across different mammals highlights its essential function in DNA methylation. Understanding the structure and role of specific sections, like the TRD loop and catalytic loop, is crucial because they play a significant part in attaching to DNA, which is essential for methylation. Realizing how these parts work helps us better understand DNMT3A's role in controlling gene activity. What stands out is DNMT3A's role in cancer. Its frequent mutations, especially in blood-related cancers like AML, highlight its importance in predicting disease outcomes. These mutations mess with the usual patterns of DNA methylation and gene activity, which can lead to cancer growth. Seeing DNMT3A overactive in many cancers further points to its potential as a target for treatments. The use of procaine, a drug typically known as an anesthetic, as a DNA methylation blocker is particularly fascinating. Its ability to shrink tumors without killing cells makes it a promising treatment. By delving into how procaine removes methyl groups from DNA, we're uncovering new treatment routes, especially in cancers where turning on defense genes is key. Using computer simulations to look at how DNMT3A and procaine interact was a vital part of our study. These simulations give us a moving picture of how procaine might stop DNMT3A from working. While these findings need to be tested in real-life settings, they offer a solid starting point for more research on how procaine might change DNMT3A's behavior.

To sum up, our comprehensive look into DNMT3A, its role in cancer, and procaine's unique abilities offers a significant contribution to molecular biology and cancer research. The use of computer simulations added depth to our understanding of DNMT3A and procaine's relationship, setting the stage for potential treatments. As we continue to unravel the complex world of gene regulation, our study's findings promise to enhance our understanding of diseases and lead to better treatments.

6. CONCLUSION AND FUTURE PROSPECTS

Our research focused on CENP-A nucleosomes and their sensitivity to environmental changes, particularly salt levels, and DNA sequences. We discovered that magnesium ions significantly affect histone-DNA interactions, with CENP-A being more sensitive than H3. This has implications for nucleosome stability and cellular processes. We also found that CENP-A nucleosomes are more accessible to enzymes than H3 nucleosomes, suggesting potential impacts on cellular functions. Future research will explore post-translational modifications and the interactions of CENP-A nucleosomes with the CCAN complex, shedding light on cell division mechanisms. Studying various histone types will provide a comprehensive understanding of nucleosome behavior, and CENP-A's unique behavior in specific environments could have practical applications in drug development and disease treatment.

Our research on four distinct chimeric nucleosome core particles with relevance to centromere formation has revealed insights into how CENP-C motifs interact with diverse nucleosome structures. CENP-C exhibited varying structural behaviors, particularly in conjunction with H2A.Z or H3, unveiling intricate nucleosome dynamics. Our study emphasized the distinct roles of CENP-A, H3, H2A.1, and H2A.Z in nucleosome stability and their collaborative function, hinting at potential structural and functional variations among these combined nucleosomes. These findings lay the foundation for further investigations into their impact on centromere formation. Our study highlights the importance of molecular dynamics simulations for understanding functional effects, particularly in epigenetic regulation. It raises questions about the potential influence of these simulations on cellular processes, including DNA damage recognition and repair, involving H2A.Z and CENP-A.

Our study used ligand docking and computer simulations to investigate how procaine interacts with the DNMT3A enzyme. Surprisingly, we found allosteric modulation instead of competitive inhibition, suggesting procaine influences the enzyme via non-active sites. This opens doors for novel drug development targeting DNMT3A. To delve deeper, we plan cellbased experiments to study procaine's impact on DNA methylation and gene activity. Our findings also hint at other compounds with similar interactions, which advanced computational techniques can help identify. While our initial simulations were promising, longer and more sophisticated simulations may reveal further insights. This research offers an exciting start towards therapies based on gene regulation understanding.

7. REFERENCES

1. Allu PK, Dawicki-McKenna JM, Van Eeuwen T, et al. Structure of the Human Core Centromeric Nucleosome Complex. *Current Biology*. 2019;29(16):26252639.e5. doi:10.1016/j.cub.2019.06.062

2. Musacchio A, Desai A. A molecular view of kinetochore assembly and function. *Biology (Basel)*. 2017;6(1). doi:10.3390/biology6010005

3. Boopathi R, Danev R, Khoshouei M, et al. Phase-plate cryo-EM structure of the Widom 601 CENP-A nucleosome core particle reveals differential flexibility of the DNA ends. *Nucleic Acids Res*. 2020;48(10):5735-5748. doi:10.1093/nar/gkaa246

4. Sharma AB, Dimitrov S, Hamiche A, Van Dyck E. SURVEY AND SUMMARY Centromeric and ectopic assembly of CENP-A chromatin in health and cancer: old marks and new tracks. *Nucleic Acids Res*. 2019;47(3):1051-1069. doi:10.1093/nar/gky1298

5. Yatskevich S, Muir KW, Bellini D, et al. Structure of the human inner kinetochore bound to a centromeric CENP-A nucleosome. *Science (1979)*. 2022;376(6595):844-852. doi:10.1126/science.abn3810

6. Ali-Ahmad A, Sekulić N. CENP-A nucleosome-a chromatin-embedded pedestal for the centromere: lessons learned from structural biology. *Essays Biochem*. 2020;64(2):205-221. doi:10.1042/EBC20190074

7. Kono H, Shirayama K, Arimura Y, Tachiwana H, Kurumizaka H. Two arginine residues suppress the flexibility of nucleosomal dna in the canonical nucleosome core. *PLoS One*. 2015;10(3):1-15. doi:10.1371/journal.pone.0120635

8. Bernstein BE, Liu CL, Humphrey EL, Perlstein EO, Schreiber SL. Global nucleosome occupancy in yeast. *Genome Biol*. 2004;5(9). doi:10.1186/gb-20045-9-r62

9. Roulland Y, Ouararhni K, Naidenov M, et al. The Flexible Ends of CENP-A Nucleosome Are Required for Mitotic Fidelity. *Mol Cell*. 2016;63(4):674-685.

doi:10.1016/j.molcel.2016.06.023

10. Yan K, Yang J, Zhang Z, et al. Structure of the inner kinetochore CCAN complex assembled onto a centromeric nucleosome. *Nature*. 2019;574(7777):278-282. doi:10.1038/s41586-019-1609-1

11. Vasudevan D, Chua EYD, Davey CA. Crystal Structures of Nucleosome Core Particles Containing the "601" Strong Positioning Sequence. *J Mol Biol*. Published online 2010. doi:10.1016/j.jmb.2010.08.039

12. Boopathi R, Danev R, Khoshouei M, et al. Phase-plate cryo-EM structure of the Widom 601 CENP-A nucleosome core particle reveals differential flexibility of the DNA ends. *Nucleic Acids Res*. 2021;48(10):5735-5748. doi:10.1093/NAR/GKAA246

13. Fenley AT, Anandakrishnan R, Kidane YH, Onufriev A v. Modulation of nucleosomal DNA accessibility via charge-altering post-translational modifications in histone core. *Epigenetics Chromatin*. 2018;11(1). doi:10.1186/s13072-018-0181-5

14. Onufriev A V., Schiessel H. The nucleosome: from structure to function through physics. *Curr Opin Struct Biol*. 2019;56:119-130. doi:10.1016/j.sbi.2018.11.003

15. Iwaki T, Saito T, Yoshikawa K. How are small ions involved in the compaction of DNA molecules? *Colloids Surf B Biointerfaces*. 2007;56(1-2):126-133. doi:10.1016/j.colsurfb.2006.10.040

16. Aoki K, Hayashi H, Furuya K, et al. Breakage of the nuclear envelope by an extending mitotic nucleus occurs during anaphase in Schizosaccharomyces japonicus. *Genes to Cells*. 2011;16(9):911-926. doi:10.1111/j.13652443.2011.01540.x

17. Maeshima K, Matsuda T, Shindo Y, et al. A Transient Rise in Free Mg2+ Ions Released from ATP-Mg Hydrolysis Contributes to Mitotic Chromosome Condensation. *Current Biology*. 2018;28(3):444-451.e6. doi:10.1016/j.cub.2017.12.035

18. Zhou K, Gaullier G, Luger K. Nucleosome structure and dynamics are coming of age. *Nat Struct Mol Biol*. 2019;26(1):3-13. doi:10.1038/s41594-018-0166-x

19. Mariño-Ramírez L, Kann MG, Shoemaker BA, Landsman D. *Histone Structure and Nucleosome Stability NIH Public Access*.; 2007.

20. Giaimo BD, Ferrante F, Herchenröther A, Hake SB, Borggrefe T. The histone variant H2A.Z in gene regulation. *Epigenetics Chromatin*. 2019;12(1). doi:10.1186/s13072-019-0274-9

21. Allu PK, Dawicki-McKenna JM, Van Eeuwen T, et al. Structure of the Human Core Centromeric Nucleosome Complex. *Current Biology*. 2019;29(16):2625-2639.e5. doi:10.1016/j.cub.2019.06.062

22. Kato H, Jiang J, Zhou BR, et al. A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science (1979)*. 2013;340(6136):1110-1113. doi:10.1126/science.1235532

23. Guo LY, Allu PK, Zandarashvili L, et al. Centromeres are maintained by fastening CENP-A to DNA and directing an arginine anchor-dependent nucleosome transition. *Nat Commun*. 2017;8:15775. doi:10.1038/ncomms15775

24. Paul S. Histone "acidic patch": a hotspot in chromatin biology. *Nucleus (India)*. 2021;64(3):271-275. doi:10.1007/s13237-021-00357-0

25. Subramanian V, Mazumder A, Surface LE, et al. H2A.Z Acidic Patch Couples Chromatin Dynamics to Regulation of Gene Expression Programs during ESC Differentiation. *PLoS Genet*. 2013;9(8). doi:10.1371/journal.pgen.1003725

26. Amor DJ, Choo KH. Neocentromeres: role in human disease, evolution, and centromere study. Am J Hum Genet. 2002 Oct;71(4):695-714. doi: 10.1086/342730. Epub 2002 Aug 26. PMID: 12196915; PMCID: PMC378529.

27. Saffery R, Irvine D V, Griffiths B, Kalitsis P, Wordeman L, Choo KHA. *Human Centromeres and Neocentromeres Show Identical Distribution Patterns of >20 Functionally Important Kinetochore-Associated Proteins*. Vol 9.; 2000.

28. Barra V, Fachinetti D. The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA. *Nat Commun*. 2018;9(1). doi:10.1038/s41467-018-06545-y

29. Sharma AB, Dimitrov S, Hamiche A, Van Dyck E. SURVEY AND SUMMARY Centromeric and ectopic assembly of CENP-A chromatin in health and cancer: old marks and new tracks. *Nucleic Acids Res*. 2019;47(3):1051-1069. doi:10.1093/nar/gky1298

30. Ogiyama Y, Ohno Y, Kubota Y, Ishii K. Epigenetically induced paucity of histone H2A.Z stabilizes fission-yeast ectopic centromeres. *Nat Struct Mol Biol*. 2013;20(12):1397-1406. doi:10.1038/nsmb.2697

31. Jin C, Felsenfeld G. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev*. 2007;21(12):1519-1529. doi:10.1101/gad.1547707

32. Henikoff S. Labile H3.3+H2A.Z nucleosomes mark "nucleosome-free regions." *Nat Genet*. 2009;41(8):865-866. doi:10.1038/ng0809-865

33. de Macedo MP, Glanzner WG, Gutierrez K, Bordignon V. Chromatin role in early programming of embryos. *Animal Frontiers*. 2021;11(6):57-63. doi:10.1093/af/vfab054

34. Robertson KD, Uzvolgyi E, Liang G, et al. *The Human DNA Methyltransferases (DNMTs) 1, 3a and 3b: Coordinate MRNA Expression in Normal Tissues and Overexpression in Tumors*. Vol 27.; 1999.

35. Xie S, Wang Z, Okano M, et al. *Cloning, Expression and Chromosome Locations of the Human DNMT3 Gene Family k*. Vol 236.; 1999. www.elsevier.com/locate/gene

36. Zhang ZM, Lu R, Wang P, et al. Structural basis for DNMT3A-mediated de novo DNA methylation. *Nature*. 2018;554(7692):387-391. doi:10.1038/nature25477

37. Zhang J, Yang C, Wu C, Cui W, Wang L. DNA methyltransferases in cancer: Biology, paradox, aberrations, and targeted therapy. *Cancers (Basel)*. 2020;12(8):1-22. doi:10.3390/cancers12082123

38. Yang L, Rau R, Goodell MA. DNMT3A in haematological malignancies. *Nat Rev Cancer*. 2015;15(3):152-165. doi:10.1038/nrc3895

39. Spencer DH, Russler-Germain DA, Ketkar S, et al. CpG Island Hypermethylation Mediated by DNMT3A Is a Consequence of AML Progression. *Cell*. 2017;168(5):801-816.e13. doi:10.1016/j.cell.2017.01.021

40. Villar-Garea A, Fraga MF, Espada J, Esteller M. *Procaine Is a DNADemethylating Agent with Growth-Inhibitory Effects in Human Cancer Cells 1*.
Vol 63.; 2003. http://aacrjournals.org/cancerres/article-pdf/63/16/4984/2506524/ch1603004984.pdf

41. Li YC, Wang Y, Li DD, Zhang Y, Zhao TC, Li CF. Procaine is a specific DNA methylation inhibitor with anti-tumor effect for human gastric cancer. *J Cell Biochem*. 2018;119(2):2440-2449. doi:10.1002/jcb.26407

42. Schumann NAB, Mendonça AS, Silveira MM, et al. Procaine and S-Adenosyl-lHomocysteine Affect the Expression of Genes Related to the Epigenetic Machinery and Change the DNA Methylation Status of in Vitro Cultured Bovine Skin Fibroblasts. *DNA Cell Biol*. 2020;39(1):37-49. doi:10.1089/dna.2019.4934

43. Wishart DS, Tzur D, Knox C, et al. HMDB: The human metabolome database. *Nucleic Acids Res*. 2007;35(SUPPL. 1). doi:10.1093/nar/gkl923

44. Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD—Visual Molecular Dynamics. Journal of Molecular Graphics, 14, 33-38. https://doi.org/10.1016/0263-7855(96)00018-5

45. Yatskevich S, Muir KW, Bellini D, et al. Structure of the human inner kinetochore bound to a centromeric CENP-A nucleosome. *Science (1979)*. 2022;376(6595):844-852. doi:10.1126/science.abn3810

46. Vasudevan D, Chua EYD, Davey CA. Crystal Structures of Nucleosome Core Particles Containing the “601” Strong Positioning Sequence. *J Mol Biol*. 2010;403(1):1-10. doi:10.1016/j.jmb.2010.08.039

47. Bednar J, Garcia-Saez I, Boopathi R, et al. Structure and Dynamics of a 197 bp

Nucleosome in Complex with Linker Histone H1. *Mol Cell*. 2017;66(3):384-397.e8. doi:10.1016/j.molcel.2017.04.012

48. Kato H, Jiang J, Zhou BR, et al. A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science (1979)*. 2013;340(6136):1110-1113. doi:10.1126/science.1235532

49. Zhang ZM, Lu R, Wang P, et al. Structural basis for DNMT3A-mediated de novo DNA methylation. *Nature*. 2018;554(7692):387-391. doi:10.1038/nature25477

50. Vanommeslaeghe K, Hatcher E, Acharya C, et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem*. 2010;31(4):671-690. doi:10.1002/jcc.21367

51. R. B, R. D, M. K, et al. Phase-plate cryo-EM structure of the Widom 601 CENPA nucleosome core particle reveals differential flexibility of the DNA ends. *Nucleic Acids Res*. Published online 2020.

52. Vasudevan D, Chua EYD, Davey CA. Crystal Structures of Nucleosome Core Particles Containing the "601" Strong Positioning Sequence. *J Mol Biol*. Published online 2010. doi:10.1016/j.jmb.2010.08.039

53. Li S, Olson WK, Lu XJ. Web 3DNA 2.0 for the analysis, visualization, and modeling of 3D nucleic acid structures. *Nucleic Acids Res*. 2019;47(W1):W26W34. doi:10.1093/nar/gkz394

54. Madeira F, Pearce M, Tivey ARN, et al. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res*. 2022;50(W1):W276W279. doi:10.1093/nar/gkac240

55. Best RB, Zhu X, Shim J, et al. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and sidechain χ1 and χ2 Dihedral Angles. *J Chem Theory Comput*. 2012;8(9):3257-3273. doi:10.1021/ct300400x

56. Huang J, Rauscher S, Nawrocki G, et al. CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat Methods*. 2016;14(1):71-73.

doi:10.1038/nmeth.4067

57. Izadi S, Anandakrishnan R, Onufriev AV. Building Water Models: A Different Approach. J Phys Chem Lett. 2014 Nov 6;5(21):3863-3871. doi: 10.1021/jz501780a. Epub 2014 Oct 16. PMID: 25400877; PMCID: PMC4226301.

58. Luo Y, Roux B. Simulation of osmotic pressure in concentrated aqueous salt solutions. *Journal of Physical Chemistry Letters*. 2010;1(1):183-189. doi:10.1021/jz900079w

59. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *Journal of Chemical Physics*. 2007;126(1). doi:10.1063/1.2408420

60. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J Appl Phys*. 1981;52(12):7182-7190. doi:10.1063/1.328693

61. Garcia-Saez I, Menoni H, Boopathi R, et al. Structure of an H1-Bound 6Nucleosome Array Reveals an Untwisted Two-Start Chromatin Fiber Conformation. *Mol Cell*. 2018;72(5):902-915.e7. doi:10.1016/j.molcel.2018.09.027

62. Shukla MS, Syed SH, Montel F, et al. Remosomes: RSC generated non-mobilized particles with approximately 180 bp DNA loosely associated with the histone octamer. *Proc Natl Acad Sci U S A*. 2010;107(5):1936-1941. doi:10.1073/pnas.0904497107

63. Lone IN, Shukla MS, Charles Richard JL, Peshev ZY, Dimitrov S, Angelov D. Binding of NF-κB to Nucleosomes: Effect of Translational Positioning, Nucleosome Remodeling and Linker Histone H1. *PLoS Genet*. 2013;9(9). doi:10.1371/journal.pgen.1003830

64. Draizen EJ, Shaytan AK, Mari~ No-Ram Irez L, Talbert PB, Landsman D, Panchenko AR. HistoneDB 2.0: a histone database with variants-an integrated resource to explore histones and their variants. *Database*. 2016;2016:14. doi:10.1093/data

65. Yoo J, Aksimentiev A. New tricks for old dogs: Improving the accuracy of biomolecular force fields by pair-specific corrections to non-bonded interactions. *Physical Chemistry Chemical Physics*. 2018;20(13):8432-8449. doi:10.1039/c7cp08185e

66. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. *J Chem Phys*. 1995;103(19):8577-8593. doi:10.1063/1.470117

67. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: Fast, flexible, and free. *J Comput Chem*. 2005;26(16):1701-1718. doi:10.1002/jcc.20291

68. Vangone A, Mjj Bonvin A. Contacts-based prediction of binding affinity in protein-protein complexes. doi:10.7554/eLife.07454.001

69. Kastritis PL, Rodrigues JPGLM, Folkers GE, Boelens R, Bonvin AMJJ. Proteins feel more than they see: Fine-tuning of binding affinity by properties of the noninteracting surface. *J Mol Biol*. 2014;426(14):2632-2652. doi:10.1016/j.jmb.2014.04.017

70. Xue LC, Rodrigues JP, Kastritis PL, Bonvin AM, Vangone A. PRODIGY: A web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics*. 2016;32(23):3676-3678. doi:10.1093/bioinformatics/btw514

71. Zhang ZM, Lu R, Wang P, et al. Structural basis for DNMT3A-mediated de novo DNA methylation. *Nature*. 2018;554(7692):387-391. doi:10.1038/nature25477

72. Vanommeslaeghe K, Hatcher E, Acharya C, et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem*. 2010;31(4):671-690. doi:10.1002/jcc.21367

73. Doğan D, Arslan M, Uluçay T, Kalyoncu S, Dimitrov S, Kale S. CENP-A Nucleosome is a Sensitive Allosteric Scaffold for DNA and Chromatin Factors. *J Mol Biol*. 2021;433(6). doi:10.1016/j.jmb.2020.166789

74. Špringer T, Šípová H, Vaisocherová H, Štepánek J, Homola J. Shielding effect of monovalent and divalent cations on solid-phase DNA hybridization: Surface plasmon resonance biosensor study. *Nucleic Acids Res*. 2010;38(20):7343-7351. doi:10.1093/nar/gkq577

75. Wu C, Travers A. A "one-pot" assay for the accessibility of DNA in a nucleosome core particle. *Nucleic Acids Res*. 2004;32(15). doi:10.1093/nar/gnh121

76. McGinty RK, Tan S. Recognition of the nucleosome by chromatin factors and enzymes. *Curr Opin Struct Biol*. 2016;37:54-61. doi:10.1016/j.sbi.2015.11.014

77. Yatskevich S, Muir KW, Bellini D, et al. Structure of the human inner kinetochore bound to a centromeric CENP-A nucleosome. *Science (1979)*. 2022;376(6595):844-852. doi:10.1126/science.abn3810

78. Rudenko AY, Mariasina SS, Sergiev P V., Polshakov VI. Analogs of S-AdenosylL-Methionine in Studies of Methyltransferases. *Mol Biol*. 2022;56(2):229-250. doi:10.1134/S002689332202011X

79. Villar-Garea A, Fraga MF, Espada J, Esteller M. *Procaine Is a DNADemethylating Agent with Growth-Inhibitory Effects in Human Cancer Cells 1*.
Vol 63.; 2003. http://aacrjournals.org/cancerres/article-pdf/63/16/4984/2506524/ch1603004984.pdf

80. Schumann NAB, Mendonça AS, Silveira MM, et al. Procaine and S-Adenosyl-lHomocysteine Affect the Expression of Genes Related to the Epigenetic Machinery and Change the DNA Methylation Status of in Vitro Cultured Bovine Skin Fibroblasts. *DNA Cell Biol*. 2020;39(1):37-49. doi:10.1089/dna.2019.4934

81. Grosdidier A, Zoete V, Michielin O. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res*. 2011;39(SUPPL. 2). doi:10.1093/nar/gkr366

8. APPENDICES

8.1. Ethics Committee Report

İZMİR BİYOTIP VE GENOM MERKEZİ
GİRİŞİMSEL OLMAYAN ARAŞTIRMALAR ETİK KURULU (İBG-GOEK)
KARARI

Toplantı Tarihi	: 26/07/2021	**Toplantı Günü**	: Pazartesi
Toplantı Sayısı	: 8	**Toplantı Saati**	: 10:30

Sayın Dr Seyit KALE,

2021-028 Protokol No'lu; sorumlusu olduğunuz "CENP-A içeren nükleozom yapısının iç kinetokor aparatının kurulumdaki rolü" başlıklı araştırmanın uygulanmasında etik açıdan sakınca olmadığına oy birliği ile karar verilmiştir.

Bilgilerinizi ve gereğini rica ederiz.

Prof. Dr. H. Alper BAĞRIYANIK **Başkan** (Katılmadı)	**Dr. Serap ERKEK ÖZHAN** **Başkan Yardımcısı**
Prof. Dr. Sedef AKGÜNGÖR **Üye**	**Prof. Dr. İnci ALACACIOĞLU** **Üye**
Prof. Dr. Gülgün OKTAY **Üye** (Katılmadı)	**Prof. Dr. Hilmi ORHAN** **Üye**
Prof. Dr. Şermin GENÇ **Üye** (Katılmadı)	**Doç. Dr. Canan Aslı YILDIRIM** **Üye** (Katılmadı)
Doç. Dr. Can KÜÇÜK **Üye**	**Dr. Öğr. Üyesi Yavuz OKTAY** **Üye**

8.2. Curriculum Vitae

Onur Önder

MSc(c) Molecular Biology and Genetics

EDUCATION

2023/06 – 2020/09 MASTER OF SCIENCE : MOLECULAR BIOLOGY AND GENETICS - IN PROGRESS
Dokuz Eylul University /
Izmir Biomedicine and Genome Center

2020/06 – 2016/09 BACHELOR OF SCIENCE : MOLECULAR BIOLOGY AND GENETICS -
Istanbul Technical University

EXPERIENCE

Current – 2020/09 COMPUTATIONAL BIOPHYSICS MSC STUDENT
Izmir Biomedicine and Genome Center
Molecular dynamics simulations of biological systems that we model and analyzing data of trajectories with using in-house Python and Linux scripts, Visual Molecular Dynamics (VMD), and PyMOL.

2018/08 – 2018/06 MOLECULAR BIOLOGY INTERN
Managed day to day activities of lab and provided both assitance and supervision to new interns.
DNA and RNA isolation, PCR (PCR and Q-PCR), and NGS techniques for the research of cancer and FMF.

VOLUNTEER EXPERIENCE

2021/03 – 2020/12 DIGITAL CONTENT MANAGER
Askıda Eğitim
Content creating for marketing of blog that helps students to get what they want to learn.

2019/04 – 2018/09 DIGITAL MARKETING MANAGER/TEAM LEADER
AIESEC Istanbul
Directed a digital marketing team to promote international non-governmental organization that offers social responsibility and professional internship opportunities for students.

CONFERENCE

2022/10 THE INTERNATIONAL SYMPOSIUM ON HEALTH INFORMATICS AND BIOINFORMATICS
Selected Talk - *'Allosteric changes in the centromere-specific nucleosome in response to physical factors'*

2022/07 8TH INTERNATIONAL CONGRESS OF THE MOLECULAR BIOLOGY ASSOCIATION OF TURKEY
Poster Presentation - *'Allosteric changes in the centromere-specific nucleosome in response to physical factors'*

2022/12 8TH INTERNATIONAL BAHÇEŞEHIR UNIVERSITY (BAU) DRUG DESIGN CONGRESS
Autodock Vina for drug repurposing

2023/04 III. STRUCTURAL BIOLOGY WORKSHOP
X-Ray crystallography applications
Data-Driven modelling with HADDOCK
Cryo-EM data processing
SAXS data analysis

2023/07 MUNICH-PARIS EPIGENETICS WORKSHOP

AWARDS

TUBITAK BIDEB
2210-A Scholarship / 2020

EUROPEAN MOLECULAR BIOLOGY ORGANIZATION
Scholarship / 2022

PAPERS

ALLOSTERIC CHANGES IN THE CENTROMERE SPECIFIC NUCLEOSOME IN RESPONSE TO PHYSICAL FACTORS - IN PROGRESS
Onur Önder, Ramachandran Boopathi, Stefan DIMITROV, Seyit KALE

SKILLS

// RESEARCH SKILLS

Python, R, Linux Programming Languages
Machine Learning, Data Analysis
Biological System modelling and Docking
Molecular Dynamics and Metadynamics
Simulations VMD and PyMOL molecular visualization Droplet PCR, Q-PCR
Good Labotary Practice(GLP)

// OTHER SKILLS

Problem Solving, Handing Pressure
Time Management, Teamwork

LANGUAGES

TURKISH	Native Speaker
ENGLISH	Excellent
FRENCH	Beginner

CONTACTS

+90 537 259 43 42
ond .com
onu /
Izmi