Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering

and Natural Sciences

of the

Ruprecht – Karls – University

Heidelberg

Presented by

M.Sc. Harsh Ranawat

Born in: Mumbai, India

Oral examination: 30.01.2026

# Structural analysis of Asgard archaeal chromatin

Referees:

Prof. Dr. Robert Russell

Dr. Sebastian Eustermann

# Abstract

The origin of eukaryotic chromatin represents a pivotal question in understanding the evolution of cellular complexity. Chromatin organization underpins genome accessibility, transcriptional regulation, and cellular differentiation - features that distinguish eukaryotes from their prokaryotic ancestors. Recent phylogenomic analyses identify the Asgard superphylum of archaea, including Lokiarchaeota, Heimdallarchaeota, and Hodarchaeota, among others, as the closest relatives of eukaryotes. These lineages therefore provide a unique opportunity to explore how primitive DNA-binding proteins may have evolved into the dynamic chromatin systems characteristic of eukaryotic nuclei.

This doctoral research presents the first high-resolution structural and mechanistic analysis of Asgard archaeal chromatin, focusing on the core histone protein HHoB from a representative Hodarchaeal lineage. Using cryogenic electron microscopy (cryo-EM) complemented by advanced biochemical and biophysical methods, the study reveals that Asgard chromatin does not exist as a single static entity, but rather as a dynamic ensemble capable of adopting two fundamentally distinct conformations. The first, a novel open chromatin assembly, exhibits a loosely packed, filamentous architecture in which the DNA remains partially exposed, allowing potential access by transcriptional or repair machinery. The second, a closed chromatin assembly, forms a compact, highly ordered fiber stabilized by extensive inter-histone interactions.

The open conformation displays structural parallels to the eukaryotic (H3–H4) octasome, and is possibly an Asgard-specific innovation toward greater chromatin flexibility. In contrast, the closed state recapitulates structural features seen in other archaeal chromatin systems, indicating evolutionary continuity across archaeal lineages. Biochemical and structural assays further reveal that divalent cations such as $Mg^{2+}$ modulate the equilibrium between these states, with elevated $Mg^{2+}$ concentrations favoring chromatin compaction - highlighting a possible ionic regulatory mechanism. This study provides the first structure-based model of Asgard chromatin organization, expanding our understanding of chromatin architecture in an evolutionary context.

# Zusammenfassung

Der Ursprung des eukaryotischen Chromatins stellt eine zentrale Fragestellung für das Verständnis der Evolution zellulärer Komplexität dar. Die Organisation des Chromatins bildet die Grundlage für die Zugänglichkeit des Genoms, die Transkriptionsregulation und die zelluläre Differenzierung – Merkmale, die Eukaryoten deutlich von ihren prokaryotischen Vorfahren unterscheiden. Jüngste phylogenomische Analysen identifizieren das Asgard-Superphylum der Archaeen, einschließlich Lokiarchaeota, Heimdallarchaeota und Hodarchaeota, als die nächsten bekannten Verwandten der Eukaryoten. Diese Linien bieten somit eine einzigartige Möglichkeit, zu erforschen, wie sich primitive DNA-bindende Proteine zu den dynamischen Chromatinsystemen der eukaryotischen Zellkerne entwickelt haben könnten.

Diese Promotionsarbeit präsentiert die erste hochauflösende strukturelle und mechanistische Analyse des Asgard-archaealen Chromatins, mit Schwerpunkt auf dem histonähnlichen Kernprotein HHoB aus einer repräsentativen Hodarchaeen-Linie. Mithilfe der Kryo-Elektronenmikroskopie (Kryo-EM), ergänzt durch fortgeschrittene biochemische und biophysikalische Methoden, zeigt die Studie, dass Asgard-Chromatin keine statische Struktur darstellt, sondern als dynamisches Ensemble existiert, das zwei grundlegend unterschiedliche Konformationen einnehmen kann. Die erste, eine neuartige „offene“ Chromatinanordnung, weist eine locker gepackte, filamentöse Architektur auf, bei der die DNA teilweise exponiert bleibt und somit potenziell für Transkriptions- oder Reparaturmechanismen zugänglich ist. Die zweite, „geschlossene“ Chromatinanordnung, bildet dagegen eine kompakte, hochgeordnete Faser, die durch ausgeprägte intermolekulare Hisone-Interaktionen stabilisiert wird.

Die offene Konformation zeigt strukturelle Parallelen zum eukaryotischen (H3–H4)-Oktasom und stellt möglicherweise eine Asgard-spezifische Innovation dar, die eine größere Chromatinflexibilität ermöglicht. Im Gegensatz dazu spiegelt die geschlossene Konformation strukturelle Merkmale wider, die auch in anderen archaealen Chromatinsystemen beobachtet werden, und verdeutlicht damit eine evolutionäre Kontinuität innerhalb der Archaeenlinien. Biochemische und strukturelle Analysen zeigen ferner, dass zweiwertige Kationen wie $Mg^{2+}$ das Gleichgewicht zwischen diesen Zuständen modulieren, wobei erhöhte $Mg^{2+}$ Konzentrationen eine Chromatinkondensation begünstigen – ein Hinweis auf einen möglichen ionenabhängigen Regulationsmechanismus. Diese Arbeit liefert das erste strukturbasierte Modell der Asgard-Chromatinorganisation und erweitert damit das Verständnis der Chromatinarchitektur im evolutionären Kontext.

## Publications

**Ranawat H.**, Cajili M.K., Lopez-Barbosa N., Quail T., Dame R.T., Dodonova S.O. Cryo-EM reveals open and closed Asgard chromatin assemblies. *Molecular Cell* Volume 85, 1- 14 (2025). DOI : 10.1016/j.molcel.2025.10.001.

## Acknowledgments

This PhD would not have been possible without the help and support of a large group of people. I would like to take the opportunity to express my gratitude. I am grateful to Svetlana for accepting me into her lab and for setting up the lab together with her and Tina. I have learnt a lot about scientific thinking and supervision from you, and I carry that with me into my next steps. Thank you for your enthusiasm for new findings and for trusting me with new experiments. A big thank you to everyone in the Dodonova lab (Svetlana Dodonova, Tina Bohstedt, Gabriele Ubartaite, Megha Patro, Tiana Behr, Fredrika Rajer, Maximillian Dreimann, Anna Lina Bula, Quentin Durieux-Trouilleton, Enrico Salvatore Atorino, and Pei-Tzu Chen) for the amazing time and experience of working together. From tolerating my playlist to going for lab pool games, I will always be grateful for your support in helping me cope when times were stressful. I deeply appreciate the enthusiasm for science in the lab and the open space we shared for discussions and troubleshooting.

I express my gratitude to the members of the Quail lab (Thomas Quail, Natalia Lopez and Shilpi Dhar) for training and supporting me in using the optical tweezers for force spectroscopy and for the fun lab environment. I owe deep gratitude to the support staff at EMBL - Karin Lapouge for helping me with the biochemical characterisation by CD, nanoDSF, mass photometry, and MST experiments. You are a shining light and always instantly lift my mood with your energy. Thank you to Joseph Bartho and Sebastian Unger for keeping the microscopes in good condition, for training, and for helping me acquire and troubleshoot data. I also very much appreciate the new initiatives in the EM seminars and the updates to the scripts and the EM Wiki. I express my thanks to Thomas Hoffmann for always staying on top of all the IT infrastructure inquiries and for being patient with all my support requests. Thanks to Jonas Wiederhausen, Tina Bohstedt, and Fredrika Rajer for keeping the ÄKTA machines in order. I would like to thank my TAC members Dr. Sebastian Eustermann, Dr. Gautam Dey, and Dr. Robert Russell for their valuable input and discussions during the TAC meetings over the years.

I would like to thank the whole EMBL community and the 2021 predoc cohort for the amazing friendships and time at EMBL. From the innumerable coffee breaks, seminars, the predoc course, and shared existential crises, you all made the PhD experience richer. Special gratitude to Ana Cobos, Sarkis Tafnakaji, Bishoy Wadie, Ana Barrios and Shang-Chin Kao for being the best of friends, empathising with struggles, and celebrating the everyday victories. Lastly, I would like to thank my family—Meena Ranawat, Mahendra Ranawat, Pranav Ranawat, and Krishi Ranawat—for believing in me and trusting me, for checking up on me, and for sending care packages all the way from India. Life only makes sense in context, and for a big part of my person, you are the context. Thank you to my grandpa, the late Vagtawarmal Ranawat, for being my north star in life and for teaching me to strive to be humble and to appreciate the everyday.

## Abbreviations

| | |
|---|---|
| ATP | Adenosine Triphosphate |
| BSA | Bovine Serum Albumin |
| bp | Base Pairs |
| CTF | Contrast Transfer Function |
| cryo-EM | Cryogenic Electron Microscopy |
| cryoET | Cryo-Electron Tomography |
| DNA | Deoxyribonucleic Acid |
| DTT | Dithiothreitol (reducing agent) |
| EMSA | Electrophoretic Mobility Shift Assay |
| ESP | Eukaryotic Signature Protein |
| FAM | Fluorescein Amidite (fluorescent DNA label) |
| FSC | Fourier Shell Correlation (map resolution estimation) |
| GST | Glutathione S-Transferase |
| HEPES | 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic Acid (buffer) |
| HHoA/B/F/G | Histone variants from *Hodarchaeon LC_3* (Asgard archaea) |
| HMfB | Histone from *Methanothermus fervidus* |
| HTkA/B | Histones from *Thermococcus kodakarensis* |
| HU | Histone-like protein from bacteria |
| H3, H4, H2A, H2B | Canonical eukaryotic core histones |
| IHF | Integration Host Factor (bacterial NAP) |
| IPTG | Isopropyl $\beta$-D-1-thiogalactopyranoside (inducer) |
| LECA | Last Eukaryotic Common Ancestor |
| LIC | Ligation-Independent Cloning |
| MAFFT | Multiple Alignment using Fast Fourier Transform (alignment software) |
| NAP | Nucleoid-Associated Protein |
| PAGE | Polyacrylamide Gel Electrophoresis |
| PTM | Post-Translational Modification |
| RMSD | Root-Mean-Square Deviation |
| SMFS | Single-Molecule Force Spectroscopy |
| TBE | Tris/Borate/EDTA buffer |
| 3DVA | Three-Dimensional Variability Analysis (cryoSPARC algorithm) |

| | |
|---|---|
| 3DFSC | Three-Dimensional Fourier Shell Correlation (anisotropy metric) |
| AlphaFold2 | AI-based protein structure prediction (DeepMind) |
| AMBER | Assisted Model Building with Energy Refinement (molecular dynamics) |
| ChimeraX | UCSF ChimeraX molecular visualization and modeling software |
| COOT | Crystallographic Object-Oriented Toolkit (model building) |
| cryoDRGN | Neural-network framework for continuous cryo-EM heterogeneity analysis |
| FIJI | “Fiji Is Just ImageJ” – image processing software |
| GTDB | Genome Taxonomy Database |
| ISOLDE | Interactive molecular dynamics refinement tool in ChimeraX |
| JASP | Statistical analysis software |
| JalView | Multiple sequence alignment and visualization tool |
| LUMICKS Pylake | Python library for optical tweezers data analysis |
| MAFFT | Multiple Alignment using Fast Fourier Transform |
| MO.AffinityAnalysis | Microscale Thermophoresis analysis software (NanoTemper) |
| ModelAngelo | AI-based cryo-EM model building and fitting software |
| Phenix | Comprehensive structure refinement and validation suite |
| Python | Programming language for data analysis and visualization |
| RELION | REgularized LIkelihood OptimizatioN (cryo-EM refinement software) |
| SerialEM | Electron microscopy data acquisition software |
| Topaz | Neural-network particle picker for cryo-EM |
| WebLogo | Sequence logo generation tool |
| HistoneDB | Database of histone sequences and variants |
| UniProt | Universal Protein Resource (protein sequence database) |

# Contents

# Chapter I

# Introduction

## 1 Genome packaging in the tree of life

A central biophysical constraint in cellular life becomes striking once physical scales are considered: the genomic DNA of a single eukaryotic cell (*e.g.*, ~3 Gbp in *H. sapiens*) extends over metres when linearized, yet it must reside within a nucleus that is a few microns in diameter. Eukaryotes overcome this genome–volume pressure through the evolution of chromatin - a sophisticated nucleoprotein assembly that achieves extreme compaction while preserving regulated access to the underlying genetic information. At the core of this system lies the nucleosome, the fundamental repeating unit of chromatin. Approximately 146 base pairs of DNA wrap 1.7 turns around an octamer of histones which allows for dense local packaging yet maintains periodic exposure of DNA sequences for transcriptional and replicative machinery[1, 2]. Beyond the nucleosome, chromatin forms hierarchical structures - fibers, loops, and chromosome territories - coordinated by architectural proteins, noncoding RNAs, and nuclear scaffolds. Regulation of chromatin dynamics is achieved through three interdependent mechanisms: (i) post-translational modifications of histone tails, which encode epigenetic information; (ii) ATP-dependent chromatin remodelers that reposition or evict nucleosomes; and (iii) incorporation of histone variants that confer specialized chromatin states [3, 4, 5, 6]. These mechanisms together allow eukaryotes to compact genomes that are orders of magnitude larger than those of prokaryotes while orchestrating precise spatiotemporal control of transcription, replication timing, and DNA repair. High-resolution structural and biochemical studies firmly establish the nucleosome's role as both the physical scaffold and regulatory nexus of eukaryotic genome organization (Figure 1).

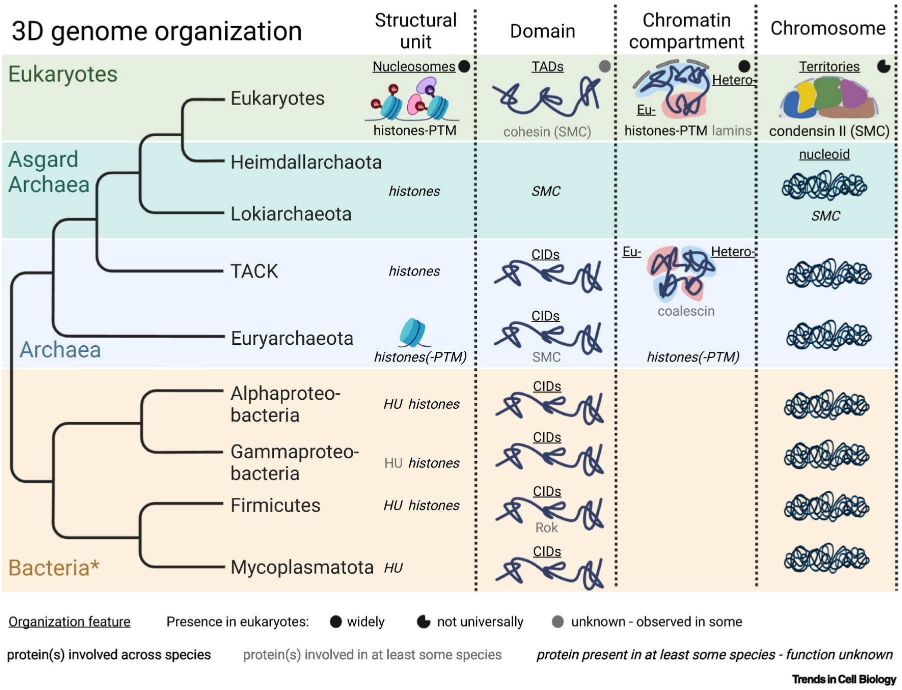


Figure 1: Simplified species tree illustrating clades with available information on 3D genome organization, as well as those potentially informative for the evolution of eukaryotic genome architecture. Features are arranged from smaller to larger scales (left to right). The absence of a feature indicates it has not yet been detected, not that it is necessarily missing. CIDs - chromosomal interaction domains; PTM - post-translational modification; TADs - topologically associating domains. Figure sourced from [7] with permission and used unedited.

Bacterial chromosomes, typically circular and megabase-scale, are compacted in a membrane-less region known as the nucleoid. Here, DNA organization is achieved primarily through negative supercoiling and small, abundant nucleoid-associated proteins (NAPs) such as HU, Fis, H-NS, and IHF[8, 9]. These proteins bend, bridge, or constrain DNA to generate a hierarchical architecture of loops and macrodomains (Figure 1) [10, 11, 12]. This form of organization is inherently dynamic—allowing rapid remodeling to support fast replication and transcriptional responses to environmental change. Electrostatic neutralization and topoisomerase activity provide additional layers of control, enabling an energetically efficient yet flexible solution to the genome-packing problem.

Archaea present a variety of approaches, tailored to their environments and evolutionary groups. Members of the Crenarchaeota employ a suite of small, abundant chromatin proteins that shape DNA without relying on histones. The main among these are Cren7 and Sul7d, both of which bind the minor groove of DNA and induce sharp bends, enabling the formation of compact yet flexible chromatin domains (Figure 2)[13]. AFM and biochemical studies have shown that Cren7 not only bends DNA but can also bridge distant DNA segments, creating higher-order chromatin compaction[14]. Sul7d, a homologous protein, performs a similar function but contributes more to local stiffening and stabilization of DNA structures[14]. Another central player in crenarchaeal chromatin is Alba (Acetylation lowers binding affinity), a small basic protein capable of binding both DNA and RNA[15]. Alba proteins can bridge duplexes, form nucleoprotein filaments, and modulate transcription by competing with transcription factors or RNA polymerase for DNA access[15]. In *Sulfolobus solfataricus*, the interplay between Alba and Cren7 governs chromatin folding and accessibility—Alba promotes filament formation, while Cren7 enhances compaction into 60-nm globules visualized by AFM[16]. Together, these proteins illustrate a modular and cooperative model of chromatin organization, where small non-histone proteins coordinate to produce dynamic and reversible higher-order structures.

Other archaeal groups, particularly halophilic and thermophilic species, employ distinct sets of NAPs adapted to their physicochemical environments. Haloarchaea, which thrive in hypersaline conditions, encode unique proteins such as MC1 (Metallochromatin protein 1) and HU-like proteins, which stabilize DNA supercoiling and maintain chromosomal integrity under osmotic stress. These proteins contribute to DNA bridging and domain formation, allowing the chromosome to remain compact yet functional in extreme ionic environments(Figure 2)[13]. Similarly, thermophilic and acidophilic archaea rely on HU-like and Alba-family proteins to maintain structural integrity at high temperatures or low pH. These proteins often possess enhanced thermostability and DNA-binding flexibility, allowing rapid reorganization of chromatin in response to environmental fluctuations. The diversity of NAPs among extremophilic archaea underscores an evolutionary trend toward environment-specific chromatin remodeling strategies, each fine-tuned to preserve genomic function under stress (Figure 2)[13].

In contrast to these non-histone systems, members of the Euryarchaeota typically encode histones, which assemble into nucleosome like complexes known as hypernucleosomes[17]. These histones share the conserved histone-fold domain but generally lack extended N or C terminal tails. Despite their simplicity, these histones wrap DNA in multiples of ~30 bp per dimer, producing ladder-like MNase digestion patterns characteristic of regularly spaced chromatin units[17, 18] AFM imaging of reconstituted complexes of the archaeal histone (HMfB) revealed beads-on-string structures morphologically indistinguishable from eukaryotic chromatin, and matched with micrococcal-nuclease digestion patterns showing regular DNA protection intervals[18]. Importantly, many euryarchaeal genomes harbor multiple histone variants or paralogs, providing an additional layer of chromatin complexity. For instance, Methanosphaera and related species encode up to 12 histone paralogs within a single genome[19, 20, 21]. These variants differ in amino acid composition and charge distribution, leading to distinct DNA-binding affinities and oligomerization capacities. Combinatorial assembly of these histones may modulate chromatin stability, accessibility, and transcriptional regulation. This diversity of archaeal histone variants reflects an evolutionary experimentation phase preceding the emergence of canonical eukaryotic chromatin, suggesting that the functional plasticity of chromatin predates the appearance of histone post-translational modifications.

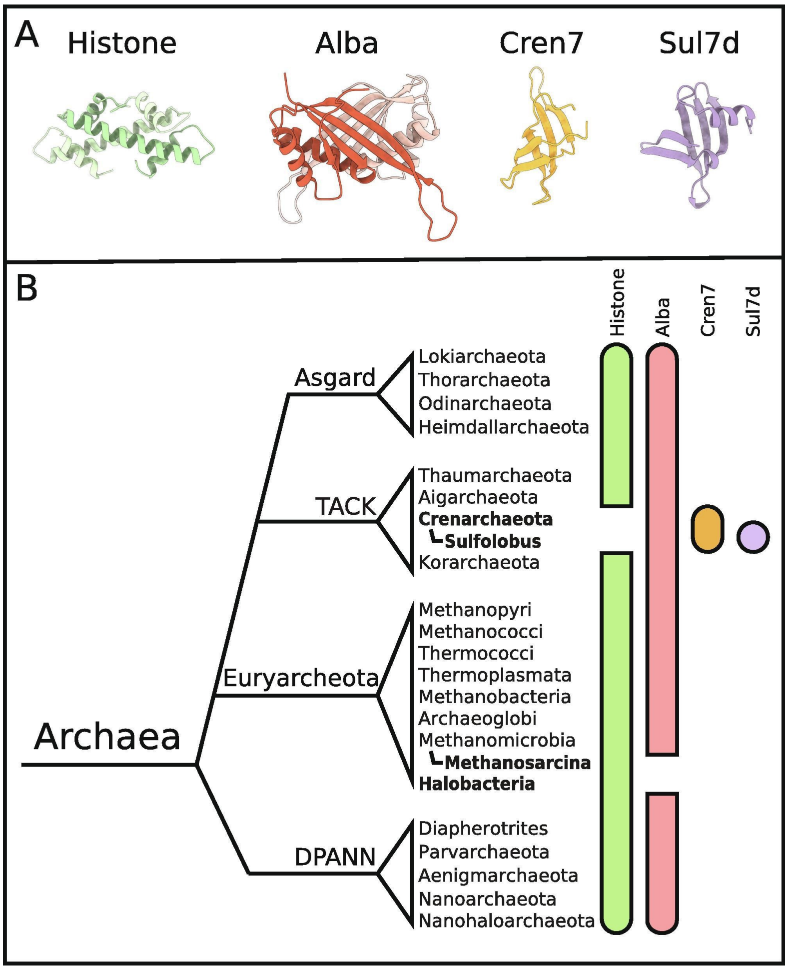


Figure 2: (A) The crystal structures of the four major archaeal DNA-packaging proteins discussed in this review are shown. (B) A simplified phylogenetic tree illustrates how chromatin-structuring proteins are distributed across different archaeal lineages : Asgard, TACK, Euryarchaeota, and DPANN. Colored bars alongside each branch indicate which chromatin-organizing proteins are present within those taxa. Many archaeal species possess both histones and Alba proteins, but taxa shown in bold lack identified genes for either histone or Alba, represented by gaps in the corresponding green and red bars. The proteins Cren7 and Sul7d are found exclusively in Crenarchaeota, which generally do not contain histones. Figure sourced from [13] with permission and used unedited.

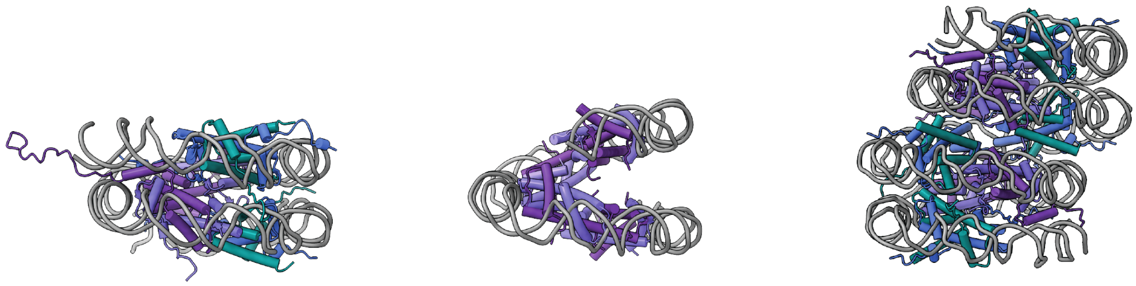
The recent discovery of the Asgard archaea, which encode multiple histone variants—including both tail-less and tail-containing forms—suggests that the emergence of eukaryotic-like chromatin features was already underway in these lineages[19, 22, 23]. Understanding the structural organization of Asgard chromatin is therefore central not only to archaeal biology but also to reconstructing the evolutionary trajectory of genome organization from simple DNA-protein assemblies to the complex nucleosomal architecture of eukaryotes.

# 2 Eukaryotic Nucleosome

Eukaryotic DNA is packaged within the nucleus in the form of chromatin, a dynamic nucleoprotein complex that ensures the compaction, accessibility, and functional regulation of the genome. The fundamental repeating unit of chromatin is the nucleosome, which serves as both a structural scaffold and a regulatory interface for DNA-templated processes such as transcription, replication, and repair. The canonical nucleosome core particle, as elucidated through X-ray crystallography in 1997, consists of ~147 base pairs (bp) of DNA wrapped in 1.65 left-handed superhelical turns around an octamer of core histone proteins—two copies each of H2A, H2B, H3, and H4 [1]. These histones are aranged as obligate heterodimers of H2A-H2B and H3-H4. Each of the histones share a defining structural feature known as the histone fold, composed of three α-helices (α1–α3) connected by two loops (L1, L2). This fold mediates the "handshake" dimerization of histones and serves as the foundation for higher-order chromatin assembly. This highly conserved nucleosomal architecture provides the basic level of chromatin organization across eukaryotes. Histones, through their globular domains and intrinsically disordered and charged N-terminal tails, mediate both DNA contacts and inter-nucleosomal interactions, thereby contributing to higher-order chromatin compaction. However, the nucleosome is not a static entity; it is a dynamic structural platform whose composition, post-translational modification, and stability vary in response to cellular context and environmental cues. This dynamic plasticity underlies the complexity of chromatin regulation and highlights the necessity to explore chromatin states beyond the canonical nucleosome [3, 4, 5, 6].

Histone variants expand the structural and functional repertoire of chromatin beyond the canonical histone complement. Replacement of canonical histones with variants such as H3.3, CENP-A (H3 variant), or H2A.Z can modulate nucleosome stability, DNA accessibility, and chromatin-associated signaling. These substitutions often occur in specific genomic regions, conferring specialized roles in transcriptional regulation, centromeric identity, DNA repair, or epigenetic inheritance[24, 25]. Along with histone variants, there have also been studies elucidating non canonical nucleosomes - these alternative structures include hexasomes [26], tetrasomes[27], and more recently, H3–H4 octasomes[28], each deviating from the canonical histone composition and arrangement. The H3-H4 octasome is composed of four H3–H4 dimers in the absence of H2A–H2B, this structure forms an octameric core that can wrap DNA *in-vitro*. Cryo-electron microscopy (cryo-EM) and biophysical studies have revealed that the H3–H4 octasome displays open, intermediate and closed conformations defined by the pitch of DNA across the turn [28].

Telomeric chromatin represents another distinct form of non-canonical chromatin structure. Eukaryotic telomeres consist of repetitive G-rich DNA sequences that terminate linear chromosomes, serving as protective caps against nucleolytic degradation and end-to-end fusion. High-resolution cryo-EM and X-ray crystallography have shown that telomeric DNA can distort canonical nucleosome geometry, leading to stacked DNA superhelicity and histone-DNA contacts[29, 30].



**Eukaryotic nucleosome**
*X. laevis* (PDB: 1AOI)

**Eukaryotic H3-H4 octasome**
*S. cerevesiae* (PDB: 7X58)

**Eukaryotic telomeric dinucleosome**
*H. sapiens* (PDB: 7V9K)

Figure 3: Eukaryotic nucleosome structures showin in a cylinder (protein) -tube (DNA) representation. From left to right: canonical nucleosome (PDB 1AOI), non-canonical H3-H4 octasome in open conformation (PDB 7X58) and telomeric dinucleosome, part of a trinucleosome model (PDB 7V9K). Histones H2A, H2B, H3 and H4 are colored in teal, blue, light and dark purple, respectively. DNA is denoted in grey.

# 3 Origin and Evolution of Histones

Each of the four eukaryotic histone proteins exhibit a histone fold, with three α-helices enabling a defined structure, obligate dimerisation and acidic patches for regulation. The helices are connected by two loops (L1, L2) that are rich in positively charged residues like arginine and lysine, enabling contact with DNA phosphate groups. The histone fold, despite its structural simplicity, represents one of the most ancient and conserved protein motifs in biology [31]. Understanding the origin and diversification of histone folds and histone sequences is central to uncovering the molecular evolution of chromatin and the emergence of complex genome regulation.

The four eukaryotic core histones display exceptionally high sequence conservation, underscoring their essential structural and regulatory roles (Figure 4). The transition from a single ancestral histone gene to the four distinct eukaryotic core histones involved repeated cycles of gene duplication followed by sequence divergence. Comparative genomic and structural analyses suggest that intermediate forms such as fused histone doublets and tandem histone repeats existed prior to the establishment of the canonical nucleosome [22, 32, 33]. Such intermediates are observed in certain viruses and archaeal lineages, providing valuable clues about the stepwise assembly of the eukaryotic chromatin machinery.

In archaeal systems, histone architecture is retained in a simplified form. Typically, a single histone gene encodes a short protein of ∼70–100 amino acids containing one histone fold domain capable of forming dimers or tetramers that wrap DNA. Unlike eukaryotic nucleosomes, archaeal histone–DNA complexes lack fixed boundaries, resulting in variable-length wrapping and more flexible chromatin organization [21]. Most archaeal histones also lack the extended N- and C-terminal “tails” that, in eukaryotes, are extensively modified by post-translational modifications (PTMs) such as methylation, acetylation, and phosphorylation. The absence of these tails suggests that histone-mediated transcriptional regulation via PTMs evolved later, coinciding with the emergence of more elaborate chromatin dynamics.

Despite their structural simplicity, it is suggested from comparative and structural analyses that even among archaeal species, histone sequences can diverge considerably while preserving the same three-dimensional fold and dimerization interface (Figure 4)[19, 34]. This phenomenon—structural conservation accompanied by sequence drift—is a defining feature of ancient protein families under strong evolutionary constraint. Sequence variability is especially pronounced among species of the Asgard archaeal group, which are considered the closest prokaryotic relatives of eukaryotes[35].
Together, these observations display a discontinuous evolutionary trajectory from simple, flexible and polymeric archaeal histone assemblies to the highly specialized and regulatory chromatin systems characteristic of eukaryotic genomes.

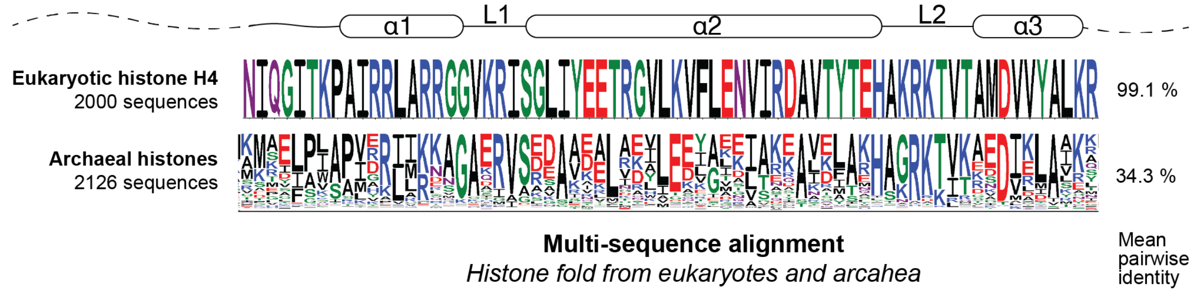


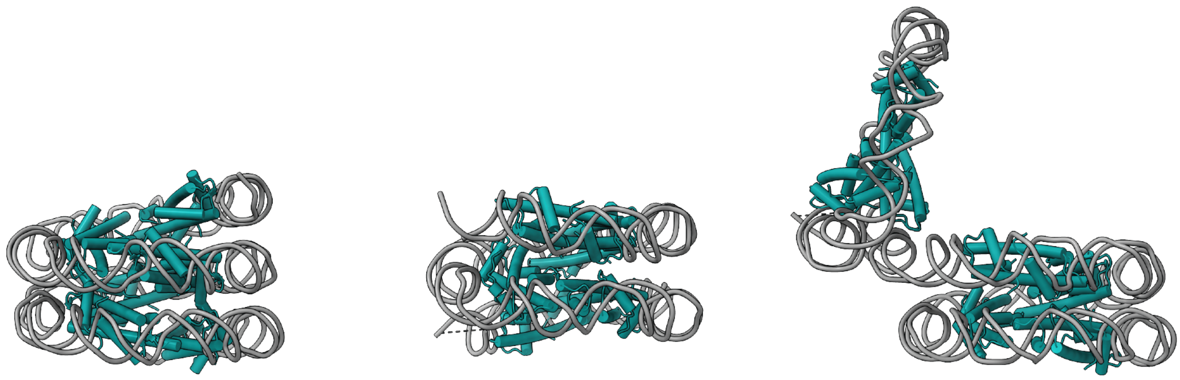
Figure 4: Sequence logos of the histone fold showing conservation of residues from eukaryotic histone H4 (top) and archaeal (bottom)sequences. The height of each stack represents the degree of conservation at that position, while mean pairwise sequence identities for each group are shown to the right. A cartoon representation of the histone fold with three alpha helices and two loops is shown above the sequence alignment.

The eukaryotic linker histone H1 binds to nucleosomes at the entry and exit sites of DNA, stabilizing higher-order chromatin structures and promoting chromatin compaction. H1 itself exists as a diverse family of variants, each with distinct affinities and regulatory roles. Unlike the highly conserved eukaryotic core histones, H1 variants show significant sequence divergence, especially in their C-terminal tails, which modulate chromatin folding and gene expression[36, 37]. Comparative analyses indicate that H1 linker histones did not originate in archaea. Instead, the earliest H1-like proteins appear in eubacteria, where lysine-rich, basic proteins share compositional similarity with the C-terminal domain of eukaryotic H1. The globular domain characteristic of metazoan H1s, known as the ”winged helix” motif, is evolutionarily conserved and appears in several divergent protist lineages, but not in archaea[38].

# 4 Histone based chromatin in archaea

The first structures of of archaeal histone–DNA complexes were revealed with the X-ray crystallographic analysis of the *Methanothermus fervidus* histone HMfB bound to 90 bp of DNA. This structure demonstrated that archaeal histones can form polymers of homodimers around which DNA wraps in a left-handed superhelix strikingly similar in geometry of packing to the eukaryotic nucleosome[17]. Each HMfB dimer contributes ~30 bp of wrapped DNA, and the dimers stack via electrostatic interactions between non-neighbouring histones. Mutation of these residues disrupt filament continuity and compromise transcriptional regulation in vivo [39, 40]. The HMfB crystal structure captured three dimers and a 90 bp DNA fragment, but crystal-contract modelling indicated that additional dimers could polymerize indefinitely to form a 'hypernucleosome', forming a superhelical filament rather than a discrete octameric core (Figure 5). This was further studied using force spectroscopy and the thermodynamics of dimer–dimer stacking were eluidated, revealing weak cooperative interactions that underpin filament flexibility[39]. Despite its significance, the HMfB structure had intrinsic limitations. As a crystallographic model, it represented an *in-vitro* complex under a single ionic and DNA-sequence condition. The crystal captured a static conformation of a short DNA fragment, inherently preventing observation of conformational variability or possibility of filament bending.

Following the HMfB hypernucleosome discovery, the understanding of archaeal nucleosomes was extended using *Thermococcus kodakarensis* histone HTkA and different DNA fragments (up to 207 bp) using sedimentation velocity analytical ultracentrifugation, molecular dynamics simulations and cryo-electron microscopy (cryo-EM)[41]. This work introduced the concept of the “slinky-like” archaeasome - a dynamic, deformable filament in which stacked histone dimers wrap DNA but retain substantial flexibility. *T. kodakarensis* encodes two histone paralogs, HTkA and HTkB, which generate MNase digestion footprints at ~60 bp increments, consistent with tetrameric and higher-order filament assemblies[42, 43]. These observations suggest that archaeal chromatin can be modular and dynamic. However, the moderate resolution (9–10 Å) precludes side-chain visualization and fine mapping of DNA contacts, and there are no high-resolution in situ cryo-electron tomography data of archaeal nucleoids. Thus, although known experimental studies portray archaeal chromatin as inherently flexible, the extent and regulation of this dynamics in vivo are still unresolved.



**Closed hypernucleosome**
*M. fervidus* (PDB: 5T5K)

**Nucleosome**
*T. kodakarensis*

**“Slinky” nucleosome**
*T. kodakarensis*

Figure 5: Archaeal nucleosome structures showin in a cylinder (protein) -tube (DNA) representation. From left to right: hypernucleosome of histone HMfB (PDB 5T5K, 6 histone dimers wrapping 180 bp DNA), closed (4 histone dimers wrapping 120 bp DNA) and slinky-like (7 histone dimers wrapping 207 bp DNA) nucleosomes of histone HTkA (PDB from [41]) . Histones and DNA are colored in teal and grey respectively.

Homology modelling with phylogenetic analysis showed that archaeal histone paralogs have diversified through gene duplication and subfunctionalization, yet the DNA-binding footprint of archaeal histones is deeply conserved, reaffirming the robustness of the ~30 bp-per-dimer wrap architecture[44, 19]. Based on homology modelling and molecular dynamics simulations, some of these histone variants are predicted to act as “capstones” - specialized paralogs that may interrupt or terminate the polymeric stacking of histone dimers into extended archaeal hypernucleosomes. In canonical archaeal histones (group I), as typified by HMfA and HMfB, the proteins contain both the dimer-dimer interface and the histone stacking interface needed for multimerisation. In contrast, those of group II lack stacking interactions but retain dimer-dimer contacts - they can form tetramers and bind DNA but show delayed or reduced hypernucleosome assembly, thereby functioning as modulators of chromatin fibre length. The group III of variants lack even the dimer-dimer connection required for multimer extension. They may act as true capstones by capping a hypernucleosome unit and prevent further addition. Structurally, the capstone concept implies that histone polymers on archaeal DNA are not unconstrained rods, but segmented fibres, with capstone-containing dimers defining boundaries[19]. However, the study was limited by its taxonomic sampling, focusing primarily on euryarchaeal genomes, and by reliance on a single modelling approach. Experimental data on in vivo protein expression, including stoichiometry, potential heteromeric assemblies, and structural validation, are needed to confirm the predicted capstone functions.

Some studies hint toward non-canonical geometries of histone multimerisation and DNA bending. A prominent example is MJ1647 from *Methanocaldococcus jannaschii* [45]. MJ1647 encodes a unique C-terminal tetramerization module, enabling formation of a stable histone tetramer that can simultaneously bind two DNA duplexes. X-ray crystallography revealed that this module forms an α-helical handshake interface connecting two dimers, producing a compact tetrameric core distinct from the linear polymeric filament of HMfB[45, 17]. Predictive structural studies using AlfaFold3 suggest that prokaryotic histone-fold proteins assemble into three general classes: nucleosomal (DNA-wrapping), bridging (DNA–DNA linking), and bending (DNA-curving)[46]. However, without complementary structural and biophysical studies, the mechanistic basis of these distinct architectures and their physiological relevance in archaeal chromatin remains unresolved.

# 5 Asgard archaea

Asgard archaea are a superphylum of archaea first described in 2015, and present a bridge to the evolutionary gap between prokaryotes and eukaryotes[47]. They encode numerous Eukaryotic-Signature Proteins (ESPs) like genes, including those involved in cytoskeleton formation and membrane trafficking. This unique genetic repertoire has positioned Asgard archaea as key organisms for understanding the origins of eukaryotic cellular complexity[48, 47]. Despite their profound evolutionary significance, Asgard archaea remain largely inaccessible to classical cell biology studies. The vast majority are known only from Metagenome-Assembled Genomes (MAGs) retrieved from diverse environments, particularly deep-sea sediments and hydrothermal vents, such as the initial source of the Lokiarchaeota from Loki's Castle[49, 47, 50]. The formidable difficulty in culturing these organisms - attributable to complex, often symbiotic, metabolic requirements and specialized physiochemical conditions - results in a lack of knowledge of cellular information[51, 52]. Direct microscopic observation of cellular structures is scarce, though recent studies have begun to reveal unusually large, complex cell morphologies and the presence of internal actin filaments in lineages like Lokiarchaeota[51]. The evolutionary tree of Asgard archaea is dynamic and refined with every new genome and analysis. Initially defined by the Lokiarchaeota, the superphylum expanded to include Thor-, Odin-, and Heimdallarchaeota, and later, the Hodarchaeales were identified as the lineage most closely related to eukaryotes (Figure 6)[35]. Furthermore, the ecological range of Asgard lineages suggests a complex evolutionary trajectory. Early analyses inferred that the Asgard ancestor of eukaryotes was likely a moderate thermophile (optimal growth temperature $\sim$ 53 °C) [53]. However, the phylum itself encompasses a broad spectrum, including lineages with signatures of thermophilic adaptations (e.g., Njordarchaeales and Panguiarchaeum) and others likely adapted to low temperature conditions, such as the initial Lokiarchaeum isolates, indicating a dynamic adaptation of growth temperatures throughout the superphylum's evolution[53, 54]. This ecological and physiological diversity hints to potential variability in core molecular structures, including the histone-DNA complex.

Asgard genomes frequently encode a greater diversity of histone-fold proteins compared to other archaea. They often possess multiple variants in a single genome, some of which feature eukaryotic-like positively charged N-terminal tails, a critical feature for regulation via PTMs in eukaryotes. Phylogenetic analyses also suggest the presence of ancient fused doublets (histone-fold proteins linked together). Critically, Asgard genomes also encode putative candidate readers and writers of histone PTMs, suggesting the potential for a rudimentary regulatory system predating the eukaryotic nucleus[22].

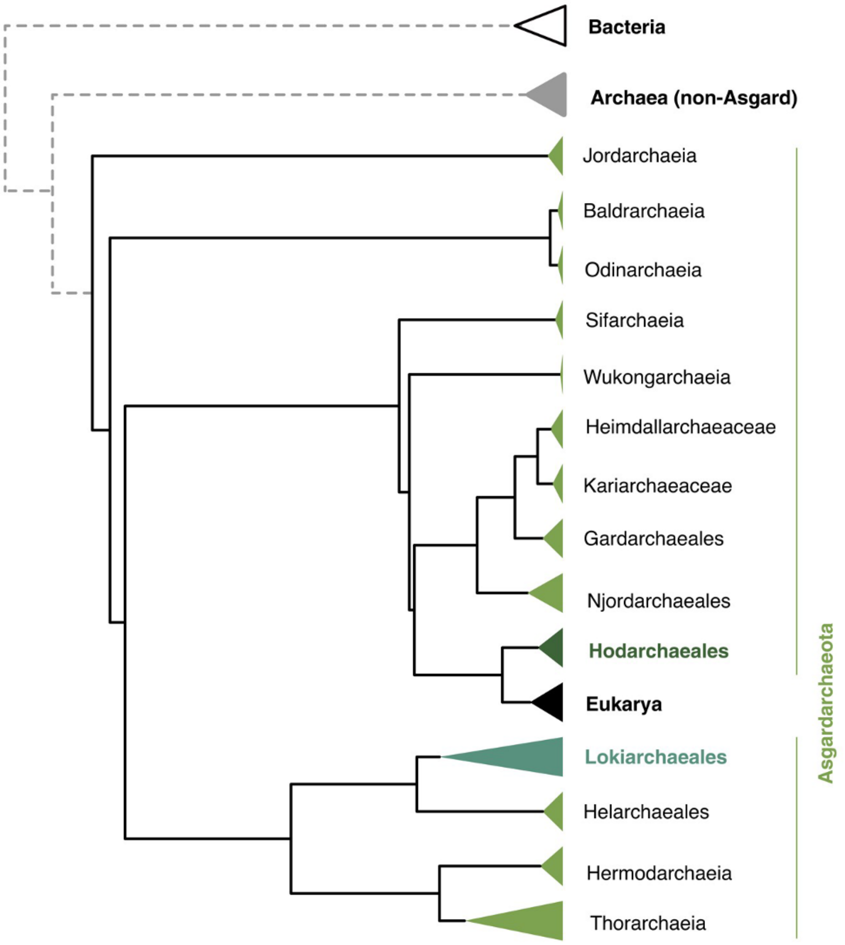


Figure 6: Phylogram showing the placement of eukaryotes within the Asgardarchaeota. Dashed branches representing non-Asgard archaea and Bacteria were manually added to indicate their approximate phylogenetic position; branch lengths are not to scale. Figure sourced from [21] with permissions and unedited.

Given the severe experimental limitations imposed by the difficulties in culturing these organisms, a comprehensive understanding of their cellular function and genome organization remains stalled. This gap in cellular information makes the need for new structural and functional information to be published on Asgards critically important. Specifically, in vitro work -using recombinant expression of Asgard histone proteins with defined DNA templates - provides the most viable path forward. This methodology allows for the deconvolution and precise structural characterization of individual histone variants, potential heterodimer combinations, and their dynamics under controlled conditions. This thesis lays out the first efforts in elucidating *in-vitro* structures and biophysics of asgard histone based chromatin, using Hodarchaeon LC_3 (previously Heimdallarchaeon LC_3) as the model system. The LC_3 metagenome codes for 10 histone variants, putative readers and writers and at the time of study, is a close relative to eukaryotes from phylogenetic analysis (Figure 6) [35, 21]. I present the first cryo-EM structures of asgard archaeal chromatin and the role that divalent ions can play in their conformational space.

# 6 Single particle cryo-EM and heterogeneity analysis

Unlike X-ray crystallography, which requires crystalline samples, cryo-EM allows the study of macromolecular assemblies in their native, hydrated state and makes it possible to capture multiple conformational states from a single dataset. Thousands to millions of noisy 2D projection images of individual particles, randomly oriented within the ice layer, are collected using a transmission electron microscope operated under cryogenic conditions. Each image represents a projection of the molecule's Coulomb potential, modulated by the microscope's contrast transfer function (CTF). Reconstruction proceeds by estimating the relative orientations and in-plane shifts of particles through iterative alignment and classification, typically using maximum-likelihood or Bayesian refinement frameworks. These projections are then combined via Fourier-space methods to reconstruct a 3D electrostatic potential map. Correction for CTF effects, beam-induced motion, and anisotropic magnification are applied to enhance accuracy[55].

Advances in instrumentation and image processing have expanded the applicability of cryo-EM to smaller macromolecules, with reconstructions achieved for proteins as small as ~40 kDa[56] and resolutions approaching 1.25 Å[57, 58]. Complementing single-particle analysis, cryo-electron tomography (cryoET) allows visualization of proteins and complexes within intact cells, offering insights into their structural organization in a native environment[59, 60].

The foundation of cryo-EM lies in the observation that water can be rapidly cooled into a vitreous (non-crystalline) state when frozen at approximately $10^6$ K/s[61, 62]. This vitrification process preserves biological specimens without ice crystal formation, enabling direct imaging by electrons. For pioneering contributions in developing this technique, Jacques Dubochet, Joachim Frank, and Richard Henderson were awarded the 2017 Nobel Prize in Chemistry[63].

Protein samples are prepared on thin metallic grids—typically 3 mm in diameter and made of copper or gold—that feature regularly spaced holes to support ice formation and facilitate automated image acquisition[64]. A small volume (3–4 µL) of the sample is applied to the grid under controlled humidity and temperature. Excess liquid is then blotted away, leaving a thin aqueous film (~100 nm thick) suitable for electron transmission. The grid is immediately plunge-frozen into liquid ethane or an ethane–propane mixture to achieve the high cooling rate required for vitreous ice formation. The sample preparation process is inherently iterative - grids are screened in the microscope to assess ice thickness, particle distribution, and integrity of complexes. Variables such as protein concentration, buffer composition, and support materials are optimized to obtain uniform particle distribution. While carbon, graphene, or

graphene oxide supports can enhance sample stability, they can also introduce preferred particle orientations, reducing reconstruction completeness[65, 66]. To minimize orientation bias and protect proteins from denaturation at the air–water interface, mild detergents or surfactants are often added[67]. Once optimal conditions are identified, a dataset of thousands of micrographs is acquired on a high-end transmission electron microscope operating at 300 keV. Electron exposure is carefully balanced to maximize contrast while minimizing radiation damage. To enhance image contrast, data are typically collected under a controlled defocus (0.5–3 µm), which is later corrected computationally through contrast transfer function (CTF) estimation.
Data processing is performed using specialized software such as CryoSPARC[68], cisTEM[69] or open source software like RELION[70]. Although implementation details differ, the general workflow remains consistent (Figure 7) [71, 72]: Pre-processing – motion correction, frame alignment, and CTF estimation; Particle picking – identification of individual protein particles, often assisted by neural network–based tools[73]; 2D classification – sorting and averaging of particle projections to remove false positives and damaged particles; 3D reconstruction and classification – initial model generation and refinement to identify distinct conformations; High-resolution refinement and post-processing – map sharpening and local resolution estimation.

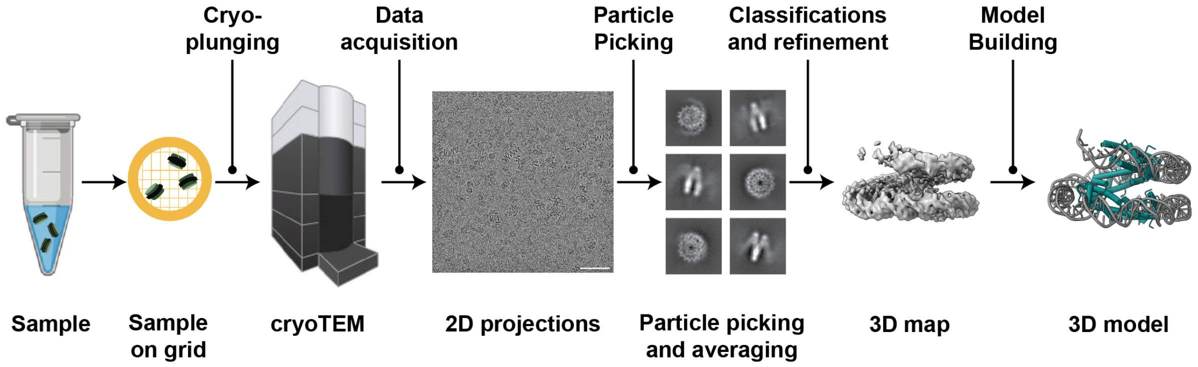


Figure 7: Single particle Cryo-EM workflow showing sample preparation, cryo pluinging, a cryo transmission electron microscope, representative micrograph, 3D map and corresponding model. The processes involved are annotated on the top, and the stages are mentioned on the bottom. Adapted from[74].

Modern cryo-EM not only provides static structures but is also data rich to capture conformational variability. In CryoSPARC, this capability is implemented through 3D Variability Analysis (3DVA), which decomposes particle images into orthogonal components of motion or structural change[68]. 3DVA effectively reveals discrete conformations and transitions between them - for example, domain movements or ligand-induced shifts - allowing for exploring the dynamic landscape of a macromolecule directly from experimental data[75]. A complementary approach is offered by CryoDRGN, a neural-network-based framework that models continuous structural heterogeneity. CryoDRGN uses variable a to map each 2D particle image to a latent representation of its 3D conformation, enabling reconstruction of entire conformational trajectories rather than discrete states (Figure 8)[76]. This approach is particularly valuable for flexible or multi-domain complexes, providing a detailed view of conformational energy landscapes and intermediate states that are often inaccessible to traditional classification methods.

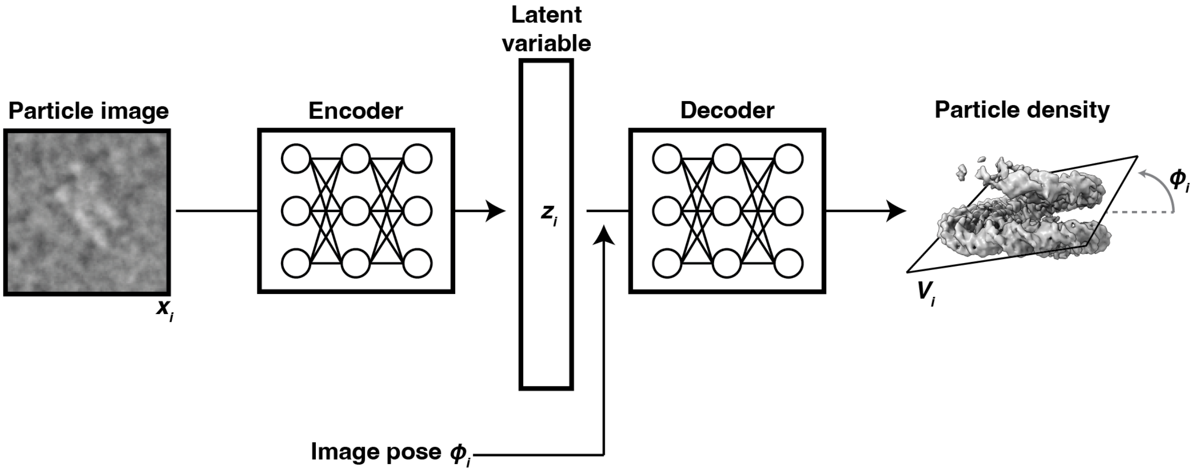


Figure 8: The cryoDRGN framework employs two interconnected neural networks organized in an image encoder–volume decoder architecture. It models structural heterogeneity through a continuous latent variable representation. During training, each particle image is mapped into a low-dimensional latent space by the encoder and subsequently reconstructed by the decoder as a corresponding 3D slice, following the principles of the Fourier slice theorem. For illustration purposes, both image and volume representations are shown in real space. Adapted from [76]

.

Once a high-resolution density map is obtained, an atomic model can be built and refined. In the past, de novo model building or homology modeling were the main approaches to interpret EM maps. Today, AI-based prediction tools such as AlphaFold[77, 78] provide highly accurate templates that can be directly fitted into EM densities. ModelAngelo[79] automates this process by combining AI predictions with experimental maps. Manual refinement can then be performed using COOT[80, 81] or ChimeraX ISOLDE[82, 83] followed by computational refinement in Phenix[84] and validation using MolProbity[85]. While cryo-EM reveals static or quasi-continuous ensembles of structures, molecular dynamics (MD) simulations provide an complementary tool for understanding the atomic-level motions and energetics underlying these conformational changes. MD simulations compute the time-dependent behavior of biomolecules based on physical force fields, allowing exploration of protein flexibility, conformational transitions, and ligand interactions that are often averaged out in cryo-EM classifications. MD simulations have revealed processes like DNA breathing, unwrapping and sliding and effect of post translational modifications in eukaryotic nucleosome models[86, 87, 88].

# 7 Force spectroscopy of chromatin

Force spectroscopy has emerged as a powerful tool to probe the mechanical properties and dynamic behavior of nucleosomes and chromatin fibers at the single-molecule level. Eukaryotic single nucleosomes exhibit stepwise unwrapping under tension: the outer DNA turn typically unwraps at low forces (~3 pN), whereas the inner turn requires higher forces up to ~20 pN, reflecting the hierarchical stability of the nucleosome structure[89, 90]. Embedding nucleosomes within chromatin fibers increases their mechanical stability, often by ~10 $k_BT$, and modulates their unfolding pathways depending on nucleosome repeat length (NRL). Short NRLs (~167 bp) favor zig-zag, two-start folding, whereas longer NRLs (~197 bp) adopt solenoidal, one-start conformations[91].

In the archaeal field, single molecule force spectroscopy has offered a powerful approach to probe the mechanical and structural properties of hypernucleosome complexes formed by archaeal histones from *M.fervidus*[39]. A long DNA molecule was tethered between a fixed surface and a force-measuring magnetic bead, then hypernucleosomes were assembled using histones HMfA or HMfB. By applying controlled stretching force to the DNA–histone tether, force-extension curves were recorded that deviate from bare DNA behaviour, compacting the DNA upto seven fold. The force-extension curves reflect the stacking, intermediate and unwrapping transitions of histone DNA complexes. HMfB showed higher stacking energies compared to HMfA, hinting at differential hypernucleosome stability[39].

Another recent study investigated the role of magnesium ions in the modulation of mechanical properties of archaeal hypernucleosomes by using form by histones HTkA and HTkB from *Thermococcus kodakarensis* and HMfA and HMfB[92]. Using single molecule tethered particle motion assays, it was shown that all four histones form filamentous hypernucleosomes, with HTkB assembling more at lower molar amounts than HTkA in the absence of $Mg^{2+}$. Addition of magnesium significantly ehanced filament compaction for all paralogues, with the strongest effect observed for the HTk histones. Force spectroscopy experiments reveal that $Mg^{2+}$ increased the mechanical stability and stiffness of the HTkA/B hypernucleosomes, indicating tighter stacking and wrapping interactions, while HMfA/B complexes were less affected[92].

While these studies demonstrate the mechanically dynamic nature of archaeal chromatin as influenced by histone variants and divalent ions, their conclusions are largely based on the HMfB hypernucleosome structure. The absence of high-resolution structures for the other histone variants limits structural interpretation across paralogues. Moreover, both studies focus on euryarchaeal histones, leaving open the possibility that histones from other archaeal lineages may adopt distinct assemblies. This highlights the need for structural and functional exploration of diverse archaeal histones to fully understand chromatin dynamics.

# Chapter II

# Aims of this Thesis

While the eukaryotic nucleosome has been extensively studied structurally and functionally, there is limited knowledge of archaeal histone-DNA complexes, which may display a higher range of variability. The knowledge base is even lower for Asgard archaea, widely accepted to represent the family of organisms that contributed to, among others, the informational processing machinery in the origin of LECA and eukaryotes. With Asgard metagenomes encoding multiple histone paralogs, including those with charged N-terminal tails, and putative reader and writer enzymes, the study of asgard chromatin is of relevance in understanding this group of organisms and in the light of evolution. The demonstrated difficulties in culturing these complex organisms emphasise the need for *in-vitro* structural studies to enabling exploring the space of possibilities.

The overall goal of this doctoral work is to structurally and biophysically characterize the histone-DNA complexes from the Asgard superphylum of archaea. I use the Hodarchaeon LC_3, belonging to the group of asgards phylogenetically closes to eukaryotes, as the model system. By resolving the fundamental architecture and mechanical properties of these ancestral complexes, this work aims to gain the first insights into asgard archaeal chromatin and their place in the path to eukaryotic origins.

## 1 Structures of asgard archaeal chromatin

First, recombinant production and biochemical characterisation of a sample of histones from the LC_3 metagenome, including one long tailed histone, two small tailed histone and one tailless minimal histone. This part involves optimising expression and purificaiton of histones and checking for their folded state. Additionally, DNA substrates of 147 bp Wid 601 and 420bp of native genomic DNA from the Hodarchaeon LC_3 were optimised for PCR amplification and purification. They serve as substrates for recombination and structural characterisation. Using purified histones and DNA substrates, I screen them for *in-vitro* reconstitution using electrophoretic mobility shift assays, quantify using microscale thermophoresis and compare to known binding behaviours from archaeal histones. For complexes that reconstitute successfully, I aim to determine high-resolution cryo-EM structures. This includes optimisation of samples by screening, data collection and single particle analysis of histone-DNA complexes. Using modern techniques I aim to resolve heterogeneity in samples to highlight the conformational space that they may occupy. With high resolution maps I aim to build atomic models and elucidate key interactors involved in the structure using mutational analysis. Comparing the determined models to kown archaeal and eukaryotic chromatin structures will reveal the position of asgard chromatin in the evolutionary path.

# 2 Biophysical characterisation

I aim to implement and perform single-molecule force spectroscopy measurements to quantify the mechanical stability of the Asgard histone-DNA complexes. This will involve measuring the force required to physically unwrap the DNA from the histone core, providing key insights into complex stability between conformations. Additionally, I will employ molecular dynamics simulations, informed by the cryo-EM structures, to analyze the dynamic behavior, flexibility, specific residue interactions and ions that govern the assembly and stability of the Asgard complexes.

# Chapter III

# Materials and Methods

## 1 Plasmids and Strains

Full length histone sequences of the Asgard archaeon Hodarchaeon LC_3 (GCA_001940645.1) were codon-optimized for expression in Escherichia coli and cloned into the LIC 1B vector (pET His6 TEV LIC cloning vector 1B, Addgene #29653). The resulting HHoB construct (Uniprot A0A1Q9NRY6, 7.5 kDa) contained an N-terminal 6×His-tag followed by a TEV protease cleavage site. Similarly, HHoF (Uniprot A0A1Q9N8N6, 7.7 kDa), HHoG (Uniprot A0A1Q9NAM9, 7.9 kDa), and HMfB from Methanothermus fervidus (Uniprot P19267, 7.7 kDa) were cloned into LIC 1B plasmids for expression.

Site-specific mutations of HHoB were generated as follows: HHoB mut1 (Y44A I48Y), HHoB mut2 (K29A E32A D36A), and HHoB-Atto647N, which carries a C-terminal GGGC extension for labeling. Constructs for mut1 and mut2 were obtained from IDT as pET-IDT expression vectors. HHoB-Atto647N was generated by PCR amplification, followed by transformation into E. coli XL-blue cells. Positive colonies were identified by colony PCR, plasmids were isolated using a Qiagen Miniprep Kit, and all mutations were verified by Sanger sequencing (Eurofins Genomics).

## 2 Protein Expression and Purification

HHoB, mut1, mut2, and HHoB-Atto647N were expressed in E. coli BL21-CodonPlus (DE3)-RIL. Cultures were induced with 0.5 mM IPTG at an OD600 of 0.5–0.6 and incubated at 37 °C for 3–4 hours. Cells were harvested by centrifugation at 4000 rpm at 4 °C and stored at −20 °C. Pellets were resuspended in lysis buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 5% glycerol, 1% Triton X-100, 1× protease inhibitor tablet, Roche) for 20 min at 4 °C and further lysed by sonication (Branson Sonifier, 30% power, 0.4 s duty cycle, 10 min). Lysates were clarified by centrifugation at 27,000 g for 20 min at 4 °C. His-tagged proteins were enriched using Ni-NTA affinity beads (ROTI®Garose-His/Ni NTA-HP, Carl Roth). The His-tag was cleaved during overnight dialysis against buffer B (20 mM HEPES pH 7.5, 200 mM NaCl, 5% glycerol) using GST-tagged TEV protease. TEV protease was removed with Glutathione Sepharose 4 Fast Flow beads (Cytiva), and proteins were further purified via cation exchange (HiTrap SP 5 ml, Cytiva). Final protein concentrations were measured with a Nanodrop One (Thermo Fisher Scientific), and proper folding was confirmed by circular dichroism. Purified proteins were aliquoted, snap-frozen in liquid nitrogen, and stored at −80 °C. HHoF and HHoG were expressed and purified using the same protocol. HMfB was expressed in E. coli LOBSTR cells and purified identically, except that the clarified lysate underwent an additional heat step at 80 °C for 15 min to enrich for the thermostable protein.

# 3 DNA Sequences and Preparation

For EMSAs and cryo-EM studies, a 147 bp Widom60122-derived sequence was used: CTG GAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTT AAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTA GTCTCCAGGCACGTGTCAGATATATACATCCTGT

A 33 bp random DNA labeled with 6-FAM at the 3' end was used for microscale thermophoresis (MST): AGGGTCACATGGGTGTTTGGCACTACCGACAGT-6-FAM Oligos were obtained from Sigma (HPLC-purified) and annealed by heating to 98 °C followed by gradual cooling. A 420 bp native LC_3 sequence from gene HeimC3_31310 was also used for EMSA and cryo-EM: ACCAGTTTTATTAGATCAACATATCAAGAAGTTACTAAAATC-CGTTAAAAAAATAA... ...CTTAGATCGGCTAGATCTACGTCGACTT

DNA constructs were ordered from GeneArt as pMA or pMK vectors, amplified by PCR, and purified via ion exchange chromatography (Resource Q 5 ml, Cytiva). For force spectroscopy, biotinylated λ DNA (48.5 kb, Lumicks, SKU 00001) was used.

# 4 *In-vitro* Reconstitution and Binding Assays

Histone-DNA complexes were assembled in 20 mM HEPES pH 7.5, 100 mM NaCl, 5% glycerol. DNA was maintained at 18 nM, while protein concentration ranged from 9 - 540 nM. Reactions were incubated at room temperature for 20 min, placed on ice, and loaded onto 5.5% 0.5× TBE PAGE gels. Electrophoresis was performed at 80 V for 90 min at 4 °C. Gels were stained with SYBR Gold (Invitrogen) and imaged using a Typhoon FLA 9500. For magnesium-dependent binding studies, 10 mM $MgCl_2$ was added to reaction and running buffers, and gels were run for 120 min. For DNA substrates of varying lengths, GeneRuler Ultra Low Range Ladder was mixed with HHoB at indicated weight ratios in 20 mM HEPES pH 7.5, 100 mM NaCl, 5% glycerol, with or without 10 mM $MgCl_2$. Reactions were incubated at room temperature for 30 min and analyzed by 10% TBE-PAGE at 120 V for 60–105 min at 4 °C depending on $Mg_{2+}$ presence. HHoB binding affinity was quantified using MST with a 33 bp 6-FAM-labeled oligo at 15.5 nM. Samples were prepared in 20 mM HEPES pH 7.5, 100 mM NaCl, 0.05% Tween-20, with an additional set containing 1 mM $MgCl_2$. Measurements were conducted in triplicate using Monolith NT.115 (NanoTemper Technologies) with blue excitation at 40% LED power in MST Premium Coated Capillaries (K005). Data were fitted using the Kd model in MO.Affinity analysis software.

# 5 Cryo-EM Sample Preparation

Nucleosomes were assembled by combining DNA and histone proteins at a 1:20 molar ratio, with a DNA concentration of 1.9 µM, in buffer A (20 mM HEPES, pH 7.5, 100 mM NaCl). For magnesium concentration screening, nucleosome reconstitution was performed in buffer A supplemented with 1, 20, 40, 60, 80, or 100 mM $MgCl_2$. All samples were incubated at room temperature (RT) for 20 min and subsequently kept on ice until plunge-freezing. Cryo-EM grids were prepared using Quantifoil R 2/1 Cu 200 mesh grids. Grids were glow-discharged for 20 s at 0.26 mbar and 25 mA using a PELCO easiGlow device (Ted Pella). Three microliters of sample were applied to the grids inside a Vitrobot Mark IV (Thermo Fisher Scientific) maintained at 100% humidity and 20 °C. Excess liquid was blotted with a blot force of 7 for 2.5 s, and grids were vitrified by plunging into liquid ethane.

# 6 Cryo-EM Data Collection

HHoB–DNA complex in buffer A: Imaging was performed on a 300 kV Titan Krios microscope (FEI) equipped with a K2 Summit direct electron detector (Gatan) in counting mode, using a Gatan Quantum energy filter with a slit width of 20 eV. A total of 3,660 movies were acquired using SerialEM software, with defocus values ranging from −1.5 to −2.5 µm, at a nominal magnification of 130,000× (pixel size 1.04 Å). To reduce preferred particle orientation, a constant stage tilt of 25° was applied. The total electron dose of 46.63 $e^-/Å^2$ was distributed over 40 frames. Magnesium titration datasets (20–80 mM $MgCl_2$): Micrographs were collected on the same Titan Krios microscope at the same nominal magnification (130 k×) and defocus range (−0.5 to −1.75 µm), with a total dose of 65.04 $e^-/Å^2$.

HHoB mutants: HHoB mut1 (1 mM $MgCl_2$) was collected at 130 k× magnification with a pixel size of 1.04 Å. A total dose of 62.44 $e^-/Å^2$ was fractionated over 40 frames. HHoB mut2 (100 mM $MgCl_2$) was collected at 105 k× magnification with a pixel size of 1.33 Å, total dose 51.87 $e^-/Å^2$ over 35 frames. Defocus range for both datasets was −0.5 to −1.75 µm.

HHoB–DNA complex in buffer A + 1 mM $MgCl_2$: Data were acquired on a G4 Titan Krios (FEI) equipped with a Falcon4i direct electron detector and Selectris X energy filter (slit width 20 eV), at a defocus range of −0.5 to −1.75 µm, nominal magnification 165 k× (pixel size 0.73 Å), and total dose 59.44 $e^-/Å^2$ distributed over 40 EER movie frames.

HHoB–DNA complex in buffer A + 20 mM $MgCl_2$: Data were collected on a 300 kV Titan Krios with a K3 detector and Quantum energy filter, at 105 k× magnification (0.82 Å/pixel), defocus −0.5 to −1.5 µm, energy filter slit width 20 eV, and total dose 63.84 $e^-/Å^2$ distributed over 40 frames.

HHoB–DNA complex in buffer A + 100 mM $MgCl^2$: Imaging was performed on the same microscope at 130 k× magnification (0.645 Å/pixel), defocus −0.5 to −1.5 µm, energy filter slit 20 eV, with a total dose of 39.93 $e^-/Å^2$ distributed across 40 frames.

# 7 Data Processing and Analysis

All datasets were processed in cryoSPARC v.4.4.1. Dataset 1 (buffer A, no $MgCl_2$): Initial particle selection using blob picking yielded 4,300 curated particles for Topaz training. Full-dataset particle picking and a second Topaz model generated 318,063 particles. After two rounds of 2D classification to remove junk, ab-initio reconstruction and 3D classification into three classes were performed. One well-resolved class containing 96,361 particles was selected for non-uniform refinement and per-particle CTF correction, yielding a 4.4 Å map. Isotropy was verified with the 3DFSC server.

Dataset 2 (buffer A + 1 mM $MgCl_2$): Blob and template picking curated 50,000 particles. Due to preferred orientation and conformational heterogeneity, two separate Topaz models were trained (top views vs. side/oblique views), resulting in 3,108,514 particles. After two rounds of 2D classification, particles were separated into open and closed conformations. Ab-initio reconstruction and 3D classification produced 96,738 open and 117,260 closed particles. Non-uniform refinement, CTF, and reference-based motion correction yielded final maps at 3.6 Å (open) and 3.5 Å (closed). Additionally, 546,206 particles were refined to obtain low-resolution maps of mixed conformations by 3D classification in cryoSPARC. The same dataset were downsampled by fourier cropping from a box size of 128px and used for heterogeneity analysis in cryoDRGN[76]. First, a network of 1024 nodes in three layers each of encoder and decoder and a latent space of dimension 8 was trained for 10 epochs to eliminate junk particles. They were clearly separated as cluster 0. Next, all the remaining particles were used for training a model of 512 nodes in three layers each of encoder and decoder and a bigger latent space of 16 dimensions. This network was trained for 30 epochs, and analyzed to determine the learnt heterogeneity of the dataset. Particles from clusters were re-imported into cryoSPARC for 2D classificaion.

Dataset 3 (buffer A + 20 mM $MgCl_2$): Blob picking produced 7,816 curated particles to train a Topaz model. Picking of the full dataset and 2D classification resulted in 214,840 particles. Ab-initio reconstruction and 3D classification produced a single well-resolved class of 40,918 particles. These open hypernucleosomes were highly flexible, with mixed conformations. Final maps after non-uniform refinement and per-particle CTF correction were resolved to 10.5 Å.

Dataset 4 (buffer A + 100 mM $MgCl_2$): Filament tracer picking identified 4,964,843 particles. Three rounds of 2D classification yielded 484,756 particles for ab-initio volume generation. A cylindrical template (outer diameter 13 nm, inner diameter 1 nm) was used for refinement without imposing helical parameters. Helical symmetry search was applied, and 353,468 par-

ticles from the well-resolved class underwent helical refinement with non-uniform refinement. Per-group CTF refinement and reference-based motion correction produced the final closed hypernucleosome map at 2.6 Å. The helical twist, rise, and pitch were 77.9°, −19.4 Å, and 89.6 Å, respectively, with symmetry order 3.
Dataset 5 ($MgCl_2$ concentration screen): Thirty micrographs from the 40 mM $MgCl_2$ dataset were split into two subsets to train separate Topaz models. These models were used to pick particles from 30 micrographs for each $MgCl_2$ condition (0, 1, 20, 40, 60, 80, 100 mM). Pixel sizes were adjusted for each dataset (0.64–1.04 Å/pixel). After 2D classification, particle numbers ranged from 1,349 to 7,575 per dataset. Particles were categorized as open, mixed, or closed based on 2D classes. Representative 3D reconstructions from the 1 mM $MgCl_2$ dataset are shown in Figure S4. FFT analyses of the micrographs were performed using the FFT function in FIJI (v2.16.0), with power spectra presented in Figures S5 and S10.

# 8 Model Fitting and Refinement

The AlphaFold2-predicted structure of the HHoB histone dimer was initially used for rigid-body fitting into cryo-EM density maps in UCSF ChimeraX v1.7.1. Real-space refinement was performed in Phenix v1.21.1, followed by per-residue adjustment in Coot v0.9.8.93 EL, applying all-atom self-restraints. Final relaxation into the density map was carried out using ISOLDE within ChimeraX. For the open conformation, DNA was fit into the density by ISOLDE relaxation using the H3-H4 open eukaryotic octasome structure (PDB 7X58) as a starting model, with hydrogen-bond restraints applied to the DNA. For the closed conformation, the procedure was repeated using DNA from the canonical eukaryotic nucleosome (PDB 1AOI). DNA backbone angles were refined with B-form restraints in Coot. RMSD and solvent-accessible surface area calculations were performed in ChimeraX. Reported RMSD values are precise to 0.1 Å, though the actual Cα positioning error at the operating resolution (2.6–3.6 Å) may be slightly higher.

Table III.1: Cryo-EM data collection, refinement, and validation statistics for HHoB nucleosome and hypernucleosome structures.

| **Histone HHoB + 147 bp Wid601** | **Open nucleosome** | **Open nucleosome** | **Closed nucleosome** | **Open hypernucleosome** | **Closed hypernucleosome** |
|---|---|---|---|---|---|
| Deposited Map ID | EMDB-53390 | EMDB-53388 | EMDB-53386 | EMDB-53389 | EMDB-53387 |
| Deposited PDB ID |  | PDB 9QV7 | PDB 9QV5 |  | PDB 9QV6 |
| Concentration of $MgCl_2$ in buffer | 0 | 1 mM | 1 mM | 20 mM | 100 mM |

Continued on next page

**Table continued from previous page**

| **Histone HHoB + 147 bp Wid601** | **Open nucleosome** | **Open nucleosome** | **Closed nucleosome** | **Open hypernucleosome** | **Closed hypernucleosome** |
|---|---|---|---|---|---|
| **Data collection and processing** | | | | | |
| Magnification | 130k | 165k | 165k | 105k | 130k |
| Voltage (kV) | 300 | 300 | 300 | 300 | 300 |
| Electron exposure ($e^{-}/Å^{2}$) | 46.63 | 59.44 | 59.44 | 63.84 | 39.93 |
| Defocus range (μm) | -1.5 to -2.5 | -0.5 to -1.75 | -0.5 to -1.75 | -0.5 to -1.75 | -0.5 to -1.5 |
| Pixel size (Å) | 1.04 | 0.73 | 0.73 | 0.82 | 0.645 |
| Symmetry imposed | C1 | C1 | C1 | C1 | C1 |
| Initial particle images (no.) | 318,063 | 3,108,514 | | 214,840 | 4,964,843 |
| Final particle images (no.) | 96,361 | 96,738 | 117,260 | 40,918 | 353,468 |
| Map resolution (Å) | 4.4 | 3.6 | 3.5 | 10.5 | 2.6 |
| FSC threshold | 0.143 | 0.143 | 0.143 | 0.143 | 0.143 |
| Map resolution range (Å) | 2-Jun | 2.5-7.5 | 2.5-7.5 | Aug-16 | 2-Apr |
| **Refinement** | | | | | |
| Map : Model resolution (Å) | – | 4 | 3.9 | – | 2.9 |
| FSC threshold | – | 0.5 | 0.5 | – | 0.5 |
| Map sharpening B factor ($Å^2$) | – | -50 | -50 | – | -50 |
| Model composition | | | | | |
| Non-hydrogen atoms | – | 8628 | 8628 | – | 13932 |
| Protein residues | – | 544 | 544 | – | 852 |
| B factors ($Å^2$) | | | | | |
| Protein | – | 128.6 | 82.74 | – | 65.36 |
| R.m.s. deviations | | | | | |
| Bond lengths (Å) | – | 0.013 | 0.012 | – | 0.012 |
| Bond angles (°) | – | 1.872 | 1.672 | – | 1.911 |

Continued on next page

**Table continued from previous page**

| **Histone HHoB + 147 bp Wid601** | **Open nucleosome** | **Open nucleosome** | **Closed nucleosome** | **Open hypernucleosome** | **Closed hypernucleosome** |
|---|---|---|---|---|---|
| Validation | | | | | |
| MolProbity score | – | 0.81 | 1.42 | – | 0.91 |
| Clashscore | – | 1.08 | 7.73 | – | 1.62 |
| Poor rotamers (%) | – | 0.96 | 0.64 | – | 0 |
| Ramachandran plot | | | | | |
| Favored (%) | – | 99.62 | 98.48 | – | 100 |
| Allowed (%) | – | 0.38 | 1.52 | – | 0 |
| Disallowed (%) | – | 0 | 0 | – | 0 |

# 9 Protein Labeling with Fluorescent Dye

HHoB-Atto647N was labeled by incubating the protein with Atto647N-maleimide (FP-202-647N, Jena Biosciences) at a molar ratio of 1:5 overnight at room temperature on a shaker in buffer A. Free dye was removed using a PD-10 desalting column (Cytiva). The degree of labeling, determined using a Nanodrop spectrophotometer (Thermo Fisher Scientific), was 0.99 (99%). Labeled protein aliquots were snap-frozen in liquid nitrogen and stored at −80 °C. EMSA experiments confirmed that the DNA-binding properties of labeled HHoB were equivalent to those of the unlabeled protein (data not shown).

# 10 Force Spectroscopy Measurements and Analysis

Single-molecule force-extension experiments were performed using a high-resolution correlative fluorescence optical tweezers instrument (C-trap, LUMICKS). A microfluidic flow cell with five parallel laminar channels allowed controlled movement of the optical traps between different solutions. Streptavidin-coated polystyrene beads (4.35 µm diameter) were trapped in 1× PBS and used to anchor biotinylated double-stranded λ DNA (48.5 kb) for force measurements. Trap stiffness was calibrated to 0.3–0.4 pN/nm. An initial baseline measurement was taken in the buffer channel (Buffer C: 20 mM HEPES pH 7.5, 100 mM NaCl, 0.1% BSA) prior to DNA anchoring. Force-extension measurements were performed by stretching DNA from a relaxed length of 8 µm to a stretched length of 16.7 µm at 0.5 µm/s. DNA was then incubated in the protein channel containing 200 nM histone for 5 minutes at an inter-bead distance of 8 µm. Force-extension measurements were repeated on the histone-DNA complex by stretching from relaxed to stretched states at 0.05 µm/s. Following measurements, the DNA was ruptured by further bead separation to remove all histones, and a second baseline was recorded in the protein channel. Force-extension curves were corrected using the first baseline for free DNA and the second for histone-DNA complexes. Data were processed and analyzed in Python using the LUMICKS Pylake library v1.6.1. Statistical comparisons between datasets were performed using paired t-tests in JASP v0.19.3.

# 11 Molecular Dynamics Simulations

The HHoB hypernucleosome in the open conformation (EMDB: 53389, PDB: 9QV7) was used as the starting model for simulations. The structure was prepared for simulation using ChimeraX, in which the terminal phosphate groups of each DNA strand were removed to prevent errors during later simulations, and the model was protonated. All-atom molecular dynamics simulations with explicit solvent were performed using AMBER and the ff14SB, bsc1, TIP3P, and ionsjc_tip3p force fields for protein, DNA, water, and magnesium ions, respectively. Structures were protonated again through TLEAP, and hydrogen mass repartitioning was performed using PARMED. The structures were placed in cubic boxes extending 10 Å beyond the solute, solvated, and charge-neutralized with sodium ions. Additional ions were added to achieve concentrations equivalent to 100 mM NaCl and, when required, 100 mM $MgCl_2$. Energy minimization was carried out in two steps: the first step restrained protein and DNA molecules to allow solvent relaxation (1,000 cycles), and the second step allowed full system relaxation (2,500 cycles). The minimized structures were heated to 300 K over 100 ps (2,500 steps) and equilibrated at atmospheric pressure (1.01325 atm). Production simulations were then performed for 1,000 ns using 4 fs time steps. Simulations were carried out in triplicate, starting from the step of solvation and ion assignment with random initial velocities assigned at minimization. To assess nucleosome compaction during the simulations, distances between ten pairs of phosphate residues located at the center of DNA strands in neighboring superhelical turns were calculated as a proxy for structural compaction.

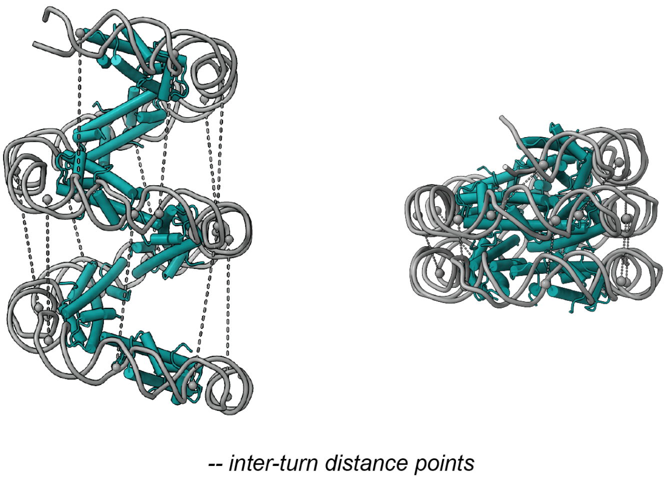


Figure 1: (A) Representative raw micrograph of HHoB nucleosomes in 1 mM $MgCl_2$. Scale bar: 50 nm. (B) Representative 2D class averages showing top, side and oblique views. Side views corresponding to open (purple), mixed (blue) and closed (orange) conformations are highlighted in the micrograph and class averages.

# 12 Multi-Sequence Alignment

Histone sequences were obtained from recent studies and the UniProt database (IDs provided in Figure S12). A total of 2,126 archaeal histones were aligned using MAFFT, and cross-referenced with the GTDB taxonomy database to classify sequences into Euryarchaeota (673 sequences), Asgard archaea (684 sequences), and a subset of Asgard lineages closest to eukaryotes (138 sequences). Additionally, 2,000 canonical eukaryotic histone H4 sequences were retrieved from HistoneDB v2.0. Sequence visualization was performed in JalView v2.11.4.1, and sequence logos were generated using WebLogo v2.8.2.

Table III.2: Reagents, resources, and software used in this study.

| **REAGENT or RESOURCE** | **SOURCE** | **IDENTIFIER** |
|---|---|---|
| **Bacterial and virus strains** | | |
| E.coli BL21 CodonPlus (DE3) RIL | Agilent Technologies | 230245 |
| E.coli BL21 LOBSTR (DE3) | Kerafast | EC1002 |
| **Chemicals, peptides, and recombinant proteins** | | |
| Protease inhibitor cocktail | Roche | 4693116001 |
| Kanamycin | Carl Roth | T832.2 |
| IPTG | Carl Roth | 367-93-1 |

Continued on next page

**Table continued from previous page**

| **REAGENT or RESOURCE** | **SOURCE** | **IDENTIFIER** |
|---|---|---|
| Roti Garose-his/Ni NTA-HP beads | Carl Roth | 805,1 |
| Glutathione sepharose 4 fast flow beads | Sigma Aldrich | GE17-5132-01 |
| HiTrap SP column | Sigma Aldrich | GE17-1152-01 |
| Phusion Polymerase | NEB | M0530L |
| Phusion HF Buffer Pack | NEB | B0518S |
| dNTP Mix (10mM each) | Thermo Fisher Scientific | Cat# R0192 |
| SYBR GOLD nucleic acid stain | Invitrogen | Cat# S11494 |
| Atto647N-maledimide | Jena Biosciences | FP-202-647N |
| Triton X-100 | Sigma Aldrich | CAT# X100-1L |
| Glycerol | Carl Roth | 7530.4 |
| GeneRuler Ultra Low Range DNA Ladder | Thermo Fisher Scientific | SM1213 |
| PD-10 desalting column | Cytiva | Cat# 17085101 |
| Tween-20 | Sigma Aldrich | Cat# P7949 |
| Streptavidin coated beads | Lumicks | SKU 11288 |
| Miniprep kit | Qiagen | Cat# 27104 |
| Quantifoil R 2/1 Cu 200 mesh grids | Quantifoil | N/A |
| **Deposited data** | | |
| HHoB nucleosome in open state (MAP) | This study | EMD-53390 |
| HHoB open hypernucleosome (MAP) | This study | EMD-53389 |
| HHoB closed hypernucleosome (MAP) | This study | EMD-53387 |
| HHoB nucleosome in closed state in 1 mM $MgCl_2$ (MAP) | This study | EMD-53386 |
| HHoB nucleosome in open state in 1 mM $MgCl_2$ (MAP) | This study | EMD-53388 |
| HHoB open state model | This study | PDB 9QV7 |
| HHoB closed state model | This study | PDB 9QV5 |

Continued on next page

**Table continued from previous page**

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| HHoB closed hypernucleosome model | This study | PDB 9QV6 |
| **Oligonucleotides** | | |
| 6-FAM-Labeled 33 nucleotide DNA oligo | IDT DNA | N/A |
| Unlabeled 33 nucleotide DNA oligo | IDT DNA | N/A |
| **Recombinant DNA** | | |
| Plasmid: pETHis6TEVLic1B | Addgene | Cat# 29653 |
| Plasmid: pETHis6TEVLic1B-HHoB | This study | N/A |
| Plasmid: pETHis6TEVLic1B-HHoF | This study | N/A |
| Plasmid: pETHis6TEVLic1B-HHoG | This study | N/A |
| Plasmid: pETIDT-HHoBMut1 | IDT | N/A |
| Plasmid: pETIDT-HHoBMut2 | IDT | N/A |
| Plasmid: pETHis6TEVLic1B-HHoB-GGGC | This study | N/A |
| Plasmid: pMA/pMK-420bp | GeneArt | N/A |
| Plasmid: pMA/pMK-147 bp_Wid601 | GeneArt | N/A |
| Biotinylated Lambda DNA | Lumicks | SKU 00001 |
| **Software and algorithms** | | |
| SerialEM | University of Colorado Boulder [93] | `https://bio3d.colorado.edu/SerialEM/` |
| cryoSPARC v4.4.1 | Punjani et al [68] | `https://cryosparc.com/` |
| UCSF ChimeraX v.1.7.1 | Goddard et al [94] | `https://www.rbvi.ucsf.edu/chimerax/` |
| Coot v0.9.8.93 EL | Casanal et al [81] | `https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/` |
| ISOLDE v1.3 | Croll [82] | `https://isolde.cimr.cam.ac.uk/` |

Continued on next page

**Table continued from previous page**

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Phenix v1.21.1 | Afonine et al [84] | `https://phenix-online.org/` |
| FIJI v2.16.0 | Schindelin et al [95] | `https://imagej.net/software/fiji/` |
| JASP v0.19.3 | JASP team | `https://jasp-stats.org` |
| Python + Pylake v1.6.1 | LUMICKS | DOI 10.5281/zenodo.4280788 |
| Jalview v2.11.4.1 | Waterhouse et al [96] | `https://www.jalview.org` |
| Alphafold2 | Jumper et al [77] | `https://github.com/google-deepmind/alphafold` |
| MO.AffinityAnalysis | NanoTemper Technologies GmbH | `http://www.nanotempertech.com` |
| MAFFT v7 | Katoh K & Stankley [97] | `https://mafft.cbrc.jp` |
| Weblogo v2.8.2 | Crooks et al [98] | `https://weblogo.berkeley.edu` |
| **Other** | | |
| Monolith NT.115 MST Premium Coated Capillaries | NanoTemper Technologies | K005 |

# 13 Use of large-language models

Large language model based tools were utilized to support aspects of this research. The application Consensus[99] was employed to aid in identifying relevant scholarly literature. ChatGPT[100] was used to assist with grammatical refinement during the preparation of this thesis. All text has been carefully reviewed and verified by the author, who assumes full responsibility for the accuracy and integrity of the content.

# Chapter IV

# Results

## 1 Protein, DNA preparation and *in-vitro* reconstitutions

In this section, I will present how I implemented the recombinant protein production of asagrd histones and DNA substrates involved in this project. They serve as the basis for structural determination and analysis of the complexes in the next chapters. The initial cloning was carried out by Svetlana and Tina Bohstedt.

### 1.1 Expression and Purification of Hodarchaeon LC_3 histones

The metagenome assembly of Hodarchaeaon LC_3 codes for 10 histone proteins (Figure 1). Based on the sequence length and composition, they can be classified into tail-less minimal histones (HHoB, HHoI, HHoJ), histones with short tails (HHoE, HHoF, HHoG, HHoH) and histones with tails (HHoA, HHoD). In this project, I worked with representatives from said classes - HHoA, HHoB, HHoF and HHoG.

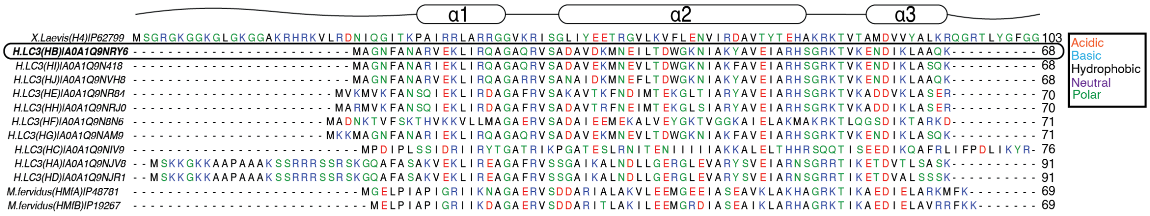


Figure 1: Multiple sequence alignment (MSA) of histones from Hodarchaeon LC_3 Asgard metagenome, the histones from euryarchaeon *M. fervidus*, and histone H4 from eukaryon *X. laevis*. Histone fold arrangement is shown schematically above.

My supervisor Svetlana Dodonova had previously cloned HHoA and HHoB pETLIC1B expression vectors and the cloning for HHoF and HHoG into same backbone vectors were carried out by me. I optimised the expression of these proteins in *E.coli* cells using LOBSTR and CODON-PLUS strains, grown in LB or TB media and induced for expression by 0.5 mM IPTG for 3h at 37°C or 16h at 18°C. The best expression conditions were optimised for soluble protein expression - chosen and large scale growth cultures were done for the HHoB, HHoF and HHoG histones. Since the proteins were His-tagged, an affinity purification step (His-trap HP, Cytiva) was used to enrich the protein. The His-Tag was then cleaved using TEV-Protease during overnight dialysis into base buffer containing 20 mM HEPES pH7.5, 100 mM NaCl, 5% glycerol. The protein was further purified via cation-exchange (HiTrap SP column, Cytiva). Soluble proteins of high purity were obtained.

## 1.2 DNA substrate preparation and nucleosome reconstitution

The DNA template used for *in-vitro* reconstitutions was 147 bp Widom 601 - an artificial DNA sequence that was engineered for its ability to position eukaryotic histone octamers, and form stable nucleosomes [101], and has been used in the majority of eukaryotic nucleosome structural studies [102, 103] as well as in the latest study of HTkA archaeal nucleosomes [41]. Therefore, I used the same Widom601 DNA to allow direct comparison with other structures. Additionally, one native DNA sequence from the LC_3 metagenome (GeneBank: HeimC3_31310) of length 420bp was also purified as template for reconstruction. Gel extracted PCR products were used as templates for successive PCR reactions, and based on these optimizations, large scale PCR reactions were performed and products were purified using anion exchange (Resource Q column). The resulting DNA products were pure.

Minimal nucleosome reconstitutions were carried out using purified proteins and 147 bp Widom 601 DNA. The basic protocol involved mixing DNA and different amounts of protein (1-30 times the DNA molar amount) for 20 mins at room temperature in buffer containing 20mM HEPES, 100 mM NaCl, and 5% glycerol. Samples were analyzed for complex formation by running EMSAs on 5% acrylamide gels.

When using proteins HHoB, HHoF and HHoG distinct ladder like dimer binding events could be observed in gels till saturation of the highest band indicating a full complex (Figure 2). This is in contrast to the euryarchaeal histone HMfB and the canonical eukaryotic nucleosome, which show a smear like shift indicating a more cooperative binding behaviour of histones to DNA [104, 1]. As can be seen in Figure 2 for 18 nM of DNA, amounts of protein required for complex formation were comparable between HHoB (x15), HF (x10) and HG (x10).

Since the pattern of binding for HHoB, HHoF and HHoG was similar, HHoB was chosen for the first electron microscopy screening to determine structures. The binding affinity of HHoB to DNA was determined by micro-scale thermophoresis (MST) using a FAM labelled 33bp fragment derived from the Widom601 sequence. The $K_d$ was determined to be 30.8 ± 3.7 nM (mean ± standard deviation, N=3).

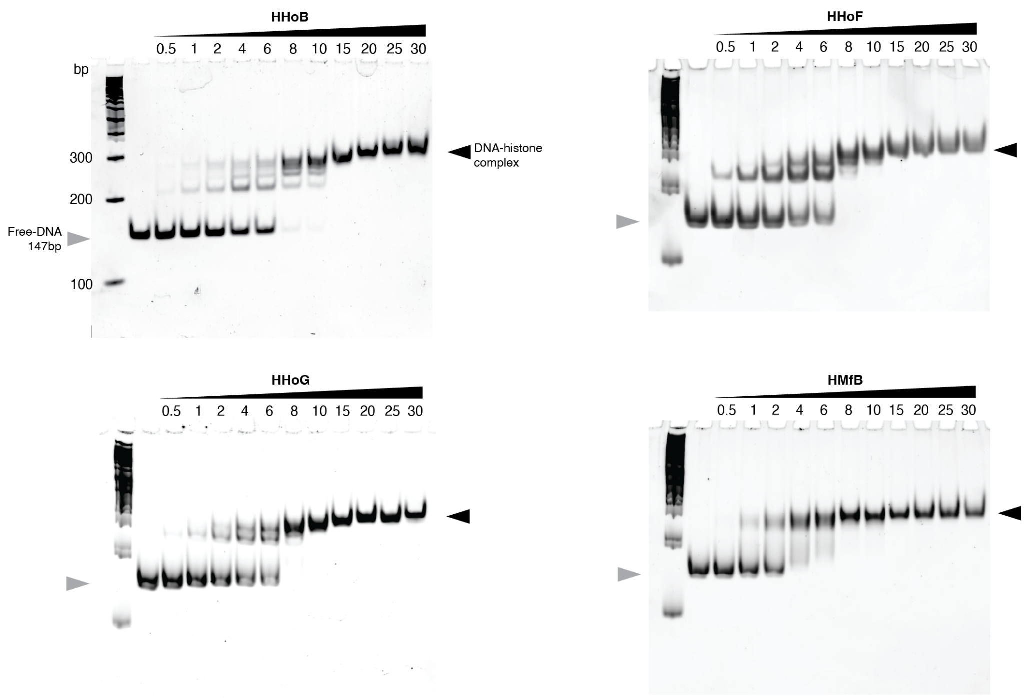


Figure 2: Electrophoretic mobility shift assay (EMSA) gels showing *in-vitro* reconstitutions of 147 bp Widom 601 DNA with different histones (denoted on the top of gels) in buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl and 5 % glycerol. Increasing molar ratios of histone to DNA are shown on top of the gels. DNA concentration was constant (20 nM). First lane shows 100 bp DNA ladder (peqGOLD 100 bp Plus).

## 2 Histone HHoB forms open and closed nucleosomes

Following the screening of three Hodarchaeon LC_3 histones for *in-vitro* DNA binding in the previous section, I chose the minimal histone HHoB for subsequent experiments. In this chapter, I present the structural characterization and analysis of HHoB–DNA complexes, highlighting the different conformational states observed. While the text has been rephrased for this dissertation, certain figures correspond in part to those included in the associated publication [105].

Minimal nucleosome complexes of HHoB and 147 bp Widom601 were imaged by cryo-EM. Unwrapping of DNA from nucleosomes and preferred orientations have been observed in earlier studies. Therefore, different conditions were tested for optimizing the particle densities and orientations on glow-discharged Quantfoil R2/1 200-mesh Cu grids. Concentration of complexes, blotting time during plunging and presence of detergents in the mixture (0.05% Tween-20, 0.4% CHAPS, 0.005% β-OG). Grid preparation using a vitrojet was also tested. The micrograph in Figure 3A shows samples prepared using buffer without any additives during screening at a 200kV Talos Arctica cryo transmission electron microscope (TEM) – the orientation bias of nucleosomes is present.

High resolution micrographs were collected at a 300kV Krios cryo TEM, at 130000x with a pixel size of 1.0415Å and dose rate of 46.5/A2 per movie across 40 frames. The grid was tilted at 25° during acquisition to compensate for the orientation bias of the complexes. Analysis of the acquired movies was carried out in cryoSPARC v3[68]. From 3660 movies (Figure 3), 120 movies were used for blob picking of 28,068 particles which were then used to train a topaz model for picking particles from the whole dataset. 310,826 particles were selected after 2D classification (Figure 3) and used for ab-initio reconstruction of two classes.

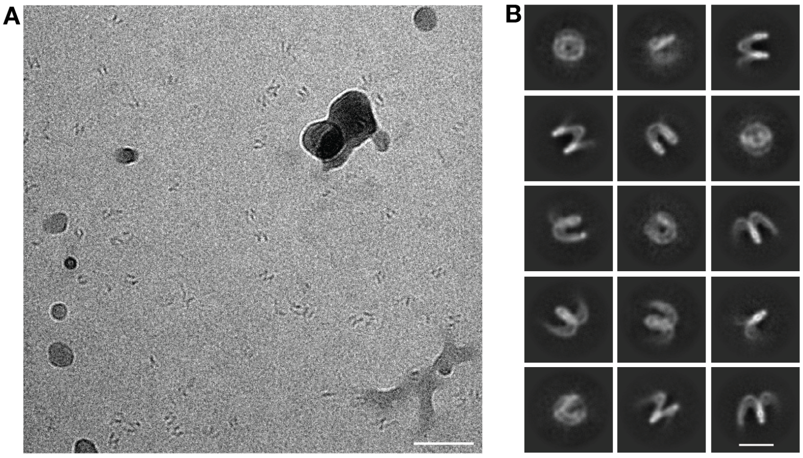


Figure 3: (A) Representative raw micrograph of HHoB nucleosomes. Scale bar: 50 nm. (B) Representative 2D class averages showing open conformations of nucleosomes. Scale bar: 10 nm.

214,088 particles were classified into one class with a full complex, and these were processed for non-uniform refinement, per-particle CTF refinement and local refinement using a particle wide mask. The resultant structure was validated using 3DFSC to be isometric in resolution at 4.44Å. Many features can already be inferred from the EM volume density (Figure 4). The complex consists of a left-handed DNA superhelix wound around four HHoB homodimers occupying 120bp of DNA, with flanks on both ends spanning ~13-14 bp. The density is well resolved that Alpha-fold prediction of the histone homodimer can be fit (Figure 4 right). Stacking interactions between histone dimers are not present, unlike the X-ray crystallography structure of the archaeal nucleosome from M. fervidus (PDB 5T5K) and canonical eukaryotic nucleosomes (PDB 1AOI). This is also contrary to an earlier prediction of stacking interactions between HHoB homodimers [17]. However, similarities can be found when comparing this structure to the recently published eukaryotic H3-H4 octasome structure (Figure 3), namely the clamshell like shape [28]. Owing to the insufficient resolution for determining positions of side chains, this map was not used for atomic level model building and inference.

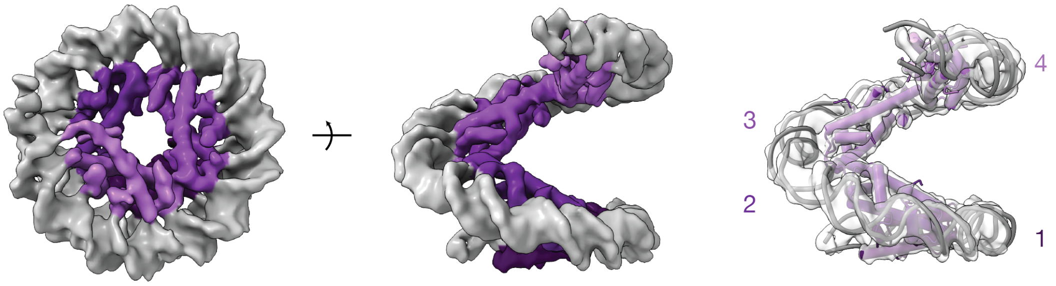


Figure 4: HHoB-DNA complex in open conformation. Top and side views of the EM map (left) are presented with DNA in grey and histone dimers in shades of purple. On the right is a transparent map with a rigid body fit of 120 bp of DNA and four histone dimer copies from AlphaFold2 predictions.

## 2.1 Nucleosome preparation in magnesium

$Mg^{2+}$ ions are both essential and abundant in eukaryotic cells, typically present at concentrations of 1–10 mM [**27**, 106], and in archaeal cells, where the euryarchaeno *Thermococcus kodakarensis* have been reported to contain up to 120 mM [107]. These ions are also known to influence chromatin organization in bacteria [108, 109]. However, due to the extreme difficulty of isolating and cultivating Asgard archaea [49, 51], there is currently no direct data on their intracellular $Mg^{2+}$ levels. Consequently, their physiological $Mg^{2+}$ concentration can only be inferred to fall somewhere within the known ranges for eukaryotes, bacteria, and other archaea (1–120 mM). Importantly, previous structural studies in this field have consistently included $Mg^{2+}$ in their samples [17, 41]. In this section, I describe the biochemical characterization and structural analysis of histone – DNA complexes from Hodarchaeon LC_3 in the presence of magnesium ions.

I previously described the reconstitution and non-cooperative binding patterns of HHoB, HHoF and HHoG to DNA in Figure 2. I performed reconstitutions by supplementing the reconstitution and running buffer with 10mM $MgCl_2$. Samples were run on 5% native-PAGE gels at 80V for 120min. The gel images are shown in Figure 5.

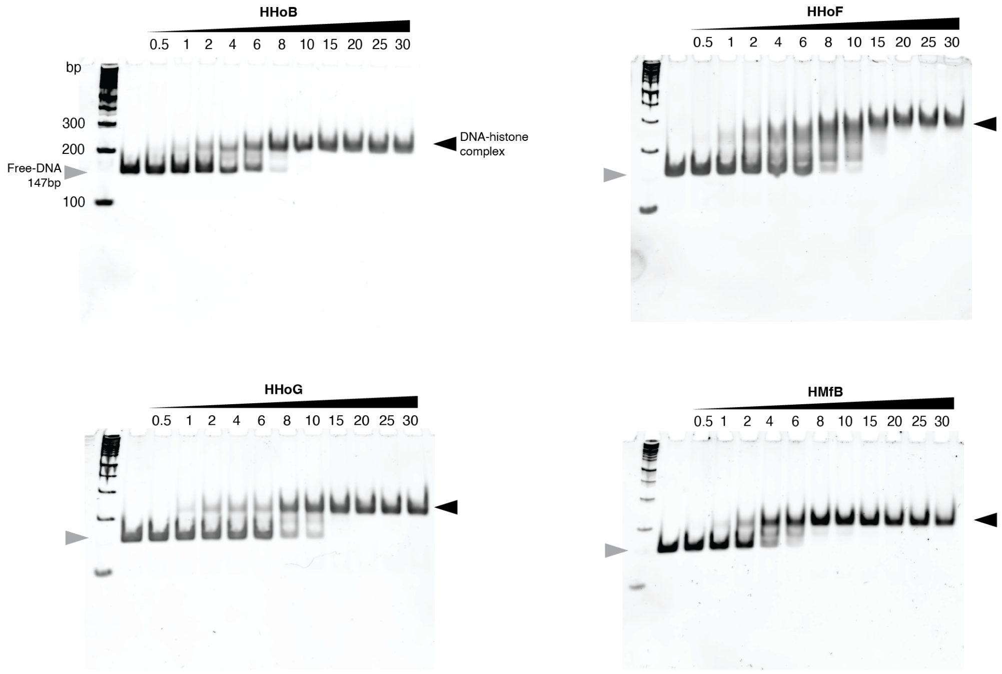


Figure 5: Electrophoretic mobility shift assay (EMSA) gels showing *in-vitro* reconstitutions of 147 bp Widom 601 DNA with different histones (denoted on the top of gels) in buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$ and 5 % glycerol. Increasing molar ratios of histone to DNA are shown on top of the gels. DNA concentration was constant (20 nM). First lane shows 100 bp DNA ladder (peqGOLD 100 bp Plus).

When compared to EMSA patterns without magnesium in Figure 2, a smear-like shift for all three proteins instead of ladder like behaviour were observed. This hinted toward cooperative binding of histones to the DNA in the presence of magnesium. Since all three proteins showed similar types of behaviour in the presence of magnesium, I decided to pursue HHoB-DNA complexes for characterisation of binding behaviour and cryo-EM in the presence of magnesium. MST was performed using the same concentrations of samples as before in a buffer supplemented with 1 mM $MgCl_2$, and resultant $K_d$ was measured to be 32.6 ± 3.1 nM (mean ± standard deviation, N=3).

## 2.2 Cryo-EM analysis of HHoB nucleosomes in 1 mM $MgCl_2$

HHoB nucleosomes were reconstituted in buffer containing 1 mM $MgCl_2$ and incubated at room temperature for 20min. I prepared cryo-EM grids and collected data on a Titan Krios, and analysed it in cryoSPARC. A representative micrograph is shown in Figure 6A. After pre-processing and filtering out junk particles, a Topaz model was trained to pick particles from a subset of 200 movies. Then this model was used to pick particles from the entire dataset. Already at the 2D classes level (Figure 6B), a greater range of conformations could be seen when compared to the dataset without magnesium chloride.

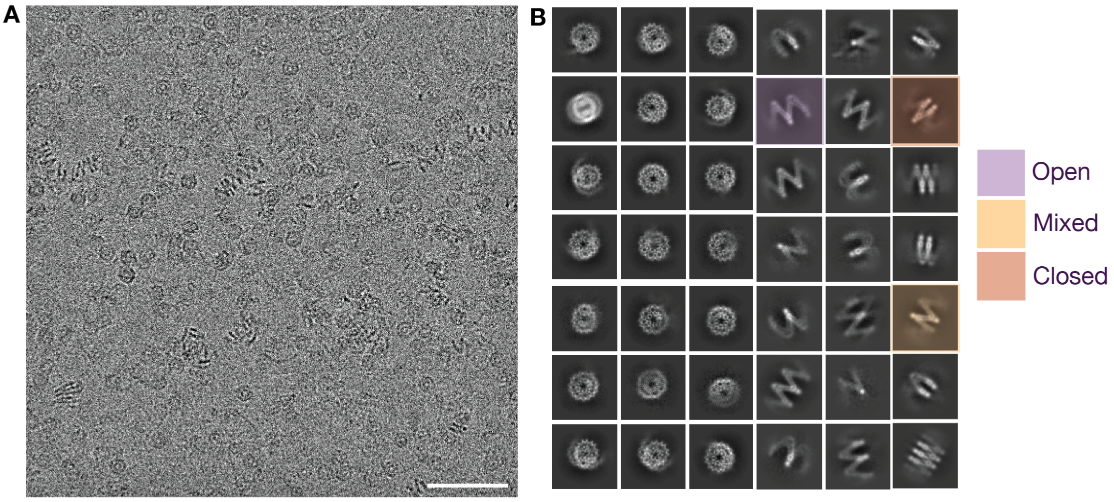


Figure 6: (A) Representative raw micrograph of HHoB nucleosomes in 1 mM $MgCl_2$. Scale bar: 50 nm. (B) Representative 2D class averages showing top, side and oblique views. Side views corresponding to open (purple), mixed (blue) and closed (orange) conformations are highlighted in the micrograph and class averages.

After further 2D and 3D classifications, two main classes of EM maps were obtained - largely differing in the pitch of opening. I named this 'open' and 'closed' with 96,738 and 117,260 particles respectively. After refinements, CTF correction and polishing I obtained the corresponding refined maps at 3.6Å and 3.4Å (Figure 7 A and B for open and closed respectively).

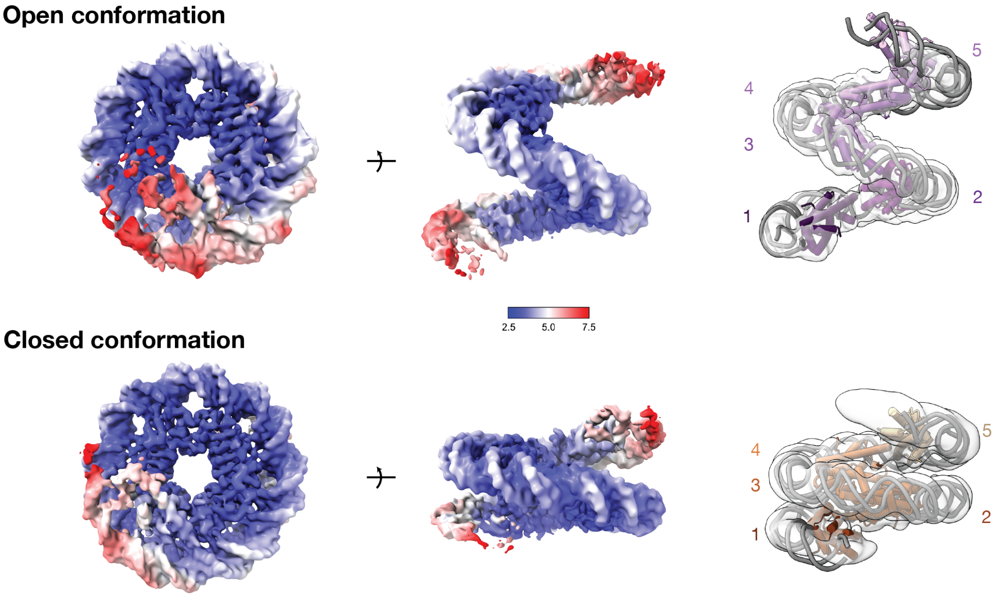


Figure 7: Local resolution representations of the refined maps of the HHoB nucleosome in open and closed conformations. Resolution ranges from 2.5 Å (blue) to 7.5 Å (red). The right panels show transparent maps with rigid-body fits of 147 bp of DNA and five histone dimer copies, colored in shades of purple for the open conformation and orange for the closed conformation.

## 2.3 Open and closed conformations differ in key interfaces and residues

I built atomic models for both volumes using Phenix [110], ChimeraX-ISOLDE [82] and Coot [80], using DNA from previous nucleosome models of H3-H4 octasome (PDB 7X58) and eukaryotic nucleosome (PDB 1AOI) for the open and closed conformation respectively. The resulting models include 4 histone dimers binding 120bp of DNA in a left handed superhelix (Figure 8 ), indicating the 30bp footprint of every histone dimer. As can be noted in the local resolution map in Figure 7 C, the resolution at the edges of the maps are lower. Indeed when set to a lower threshold, density corresponding to a fifth histone dimer can be seen (Figure 7). The "closed" nucleosome structure (PDB 9QV5) closely resembles the previously characterized HMfB euryarchaeal nucleosome (PDB 5T5K), with a Cα RMSD of 1.09 Å. In contrast, the "open" nucleosome conformation (PDB 9QV7) represents a new structural state that is notably more extended than the closed form (Figure 8). In the closed state, the helical pitch is 29.5 Å, calculated between DNA bases A30 (N1) and T105 (N3). By comparison, the open state shows a nearly twofold increase in pitch, reaching 63.0 Å, as measured between A30 (N1) and A103 (N3).

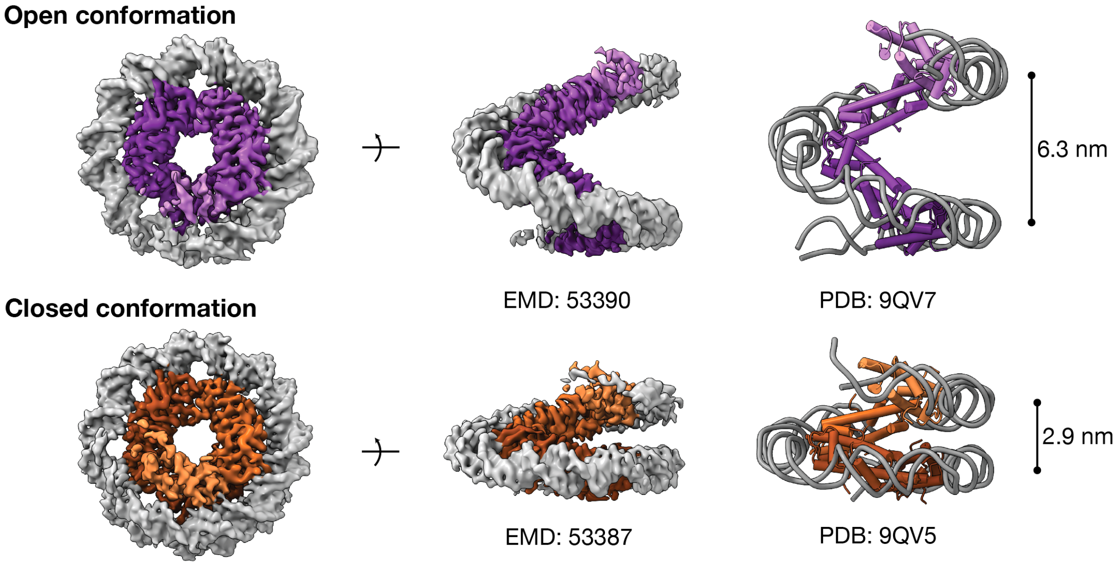


Figure 8: HHoB-DNA complex in 1 mM $MgCl_2$, shown in open (top) and closed (bottom) conformations. Top and side views of the EM map (left) and molecular model (right) are presented, with DNA in grey and histone dimers in shades of purple (PDB 9QV7) and orange (PDB 9QV5). Superhelical pitch (distance between DNA gyres) is indicated with a bar.

In both the models, every histone dimer is involved in three main interfaces - the histone-DNA interface (between DNA phosphate groups and positively charged residues in the histone), the histone dimer-dimer interface (at the intersection of two histone dimers) and the stacking interface (between non-neighbouring histones). The hisone-DNA interface, visualised with respect to one histone dimer in Figure 9, is similar in the closed(1959 $Å^2$ area) and open (1966 $Å^2$ area). Indeed this is reflected in the low RMSD of 0.4 Å between open and closed conformations in the residues at the histone-DNA interface - Arg9, Arg15, Arg21, Lys55, Lys58. The histone-DNA binding motif exists on the loop domains of the histone fold, and is fairly conserved across the tree of life.

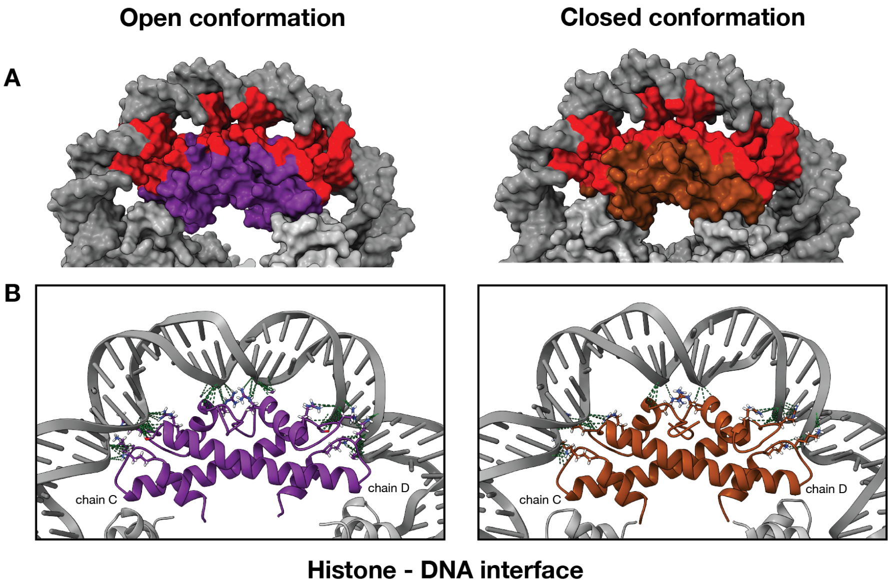


Figure 9: Histone–DNA interface in the HHoB nucleosome open (purple) and closed (closed) conformations. (A) Surface representations of the nucleosome showing the histone–DNA interface in red. (B) Ribbon representations of chains C and D in complex with DNA, highlighting the key residues involved at the interface: R9, R15, R21, K55, K58. Dashed lines indicate electrostatic and hydrogen bond interactions.

The histone dimer-dimer interface undergoes significant changes between the two models. In the open conformation, residues Tyr44, Glu47, Ile48, His51, Arg54, Asp61, and Lys68 mediate contacts between helices $\alpha 2$ and $\alpha 3$ and the L2 loop of neighboring histones (Figure 10). In the closed conformation, additional contributions come from Leu64 and Gln67 at the dimer–dimer interface, and several of the previously listed residues reorganize into a different interaction network. Notably, the interactions Tyr44–Glu47, Tyr44–His51, and Glu47–Lys68 are specific to the open state. Additionally, the orientations of the His51 and Tyr44 side chains differ markedly between the open and closed conformations (Figure 10). The overall rearragnement of residues at the dimer-dimer interface is also reflected in the larger interface area in closed state (727 Å$^2$) as compared to the open state (558 Å$^2$). The equivalent position of His51 is largely conserved across species, whereas several of the other residues show lower conservation. The stacking interface will be explored in subsection 4.1.

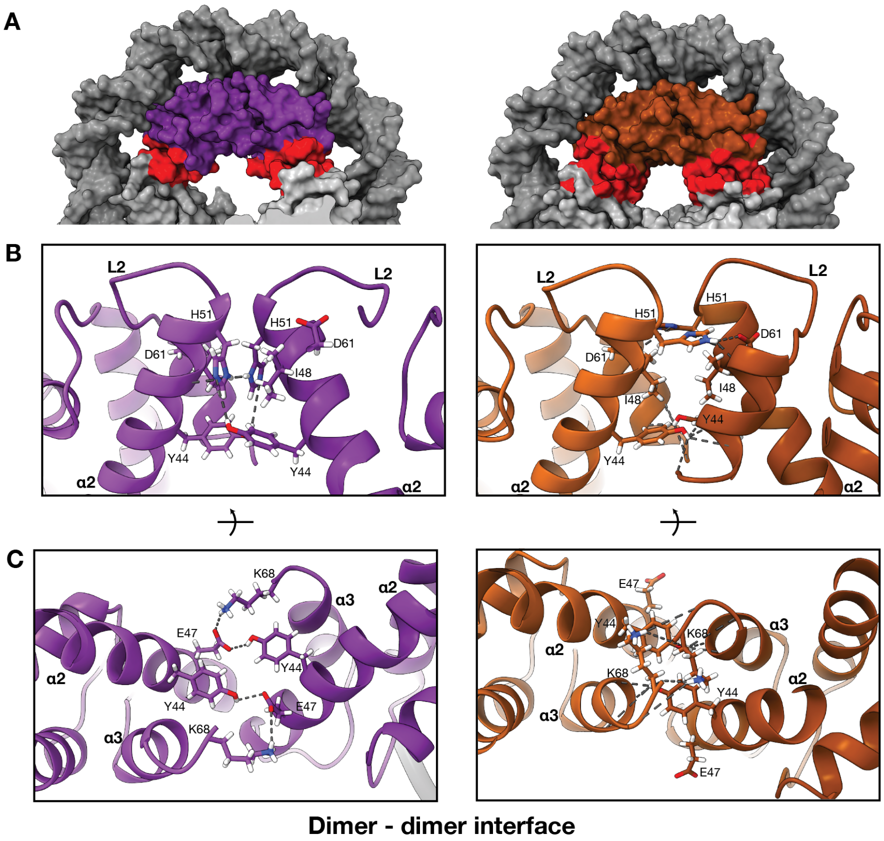


Figure 10: Histone dimer-dimer interface in the HHoB nucleosome open (purple) and closed (orange) conformations. (A) Surface representations of the nucleosomes showing the histone dimer-dimer interfaces in red. (B,C) Ribbon representations of histones showing key interactions involved at the dimer-dimer interface: Y44, E47, I48, H51, R54, D61, K68. Dashed lines indicate electrostatic and hydrogen bond interactions.

## 2.4 Conformational heterogeneity of HHoB-DNA complexes

As illustrated by the diversity of 2D classes (Figure 6B), the HHoB–DNA dataset in 1 mM $MgCl_2$ exhibited substantial conformational heterogeneity. In addition to closed hypernucleosomes, I also identified mixed classes - composites of open and closed conformations based on classes from 2D averages (Figure 6). To further characterize this heterogeneity, I carried out complementary analyses in discrete space using 3D classification in cryoSPARC, and in continuous space using CryoDRGN [76].

Following 2D classification, 546,206 particles were aligned to a low-resolution mixed conformation map and subjected to 3D classification in cryoSPARC. This was performed through random initialization of volumes followed by principal component–based dimensionality reduction. The resulting reconstructions revealed significant classes corresponding to open, closed, and mixed conformations, including complexes with more than one turn (Figure 11). Owing to the complexity of conformations and inadequate distributions of particle views, higher resolution could not be achieved upon further refinements of these classes. The mixed conformation class IV could be fit with open and closed models determined from single nucleosomes (Figure 11 right).

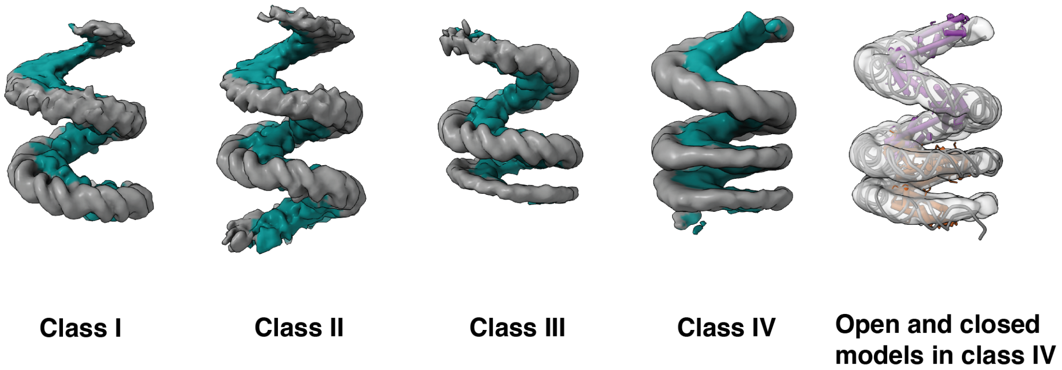


Figure 11: EM density maps showing four classes with variable number of turns, and open and closed conformations. The classes were obtained from 3D classification in PCA model using cryoSPARC. The maps are false colored with grey (DNA) and cyan (HHoB). On the right a transparent map of class IV is shown with rigid-body fits of the closed (orange) and open (purple) nucleosomes.

CryoDRGN utilizes a neural network–based variational autoencoder to capture conformational variability in a continuous, non-linear latent space, avoiding the linear constraints imposed by PCA. For this analysis, the dataset was trained for 30 epochs using a network architecture with three layers in both the encoder and decoder, each comprising 512 nodes, and a 16-dimensional latent space. The resulting latent space (Figure 12, left) provided a compact yet informative representation of conformational heterogeneity across the dataset. Sampling volumes from clusters within the latent space revealed conformations corresponding to open (Figure 12, cluster 5) and mixed states, as well as a class of closed hypernucleosomes (Figure 12, cluster 1) that were not clearly resolved by cryoSPARC-based 3D classification. These findings demonstrate CryoDRGN's ability to capture subtle or continuous structural variations that may be overlooked by conventional methods. Representative 2D classes from these clusters (Figure 12) display conformational diversity; however, this variability is less apparent in the reconstructed volumes. This discrepancy can be attributed to CryoDRGN's architecture and the incorporation of pose information after the latent space embedding (Figure 8) [76]. Because the particle set was imported from refinement of the mixed conformation in cryoSPARC, an inherent bias toward that reference state was present. Neither ab initio pose estimation nor pose refinement by gradient descent within CryoDRGN produced meaningful reconstructions, limiting the independence of the analysis. To reduce computational cost and memory usage, particles were downsampled by a factor of five, which likely reduced the effective resolution of the reconstructions. Taken together, these constraints suggest that, although the latent space effectively captures non-linear conformational variability prior to pose incorporation, it was not suitable for quantitative or high-resolution structural interpretation. Nevertheless, both CryoDRGN and conventional analyses underscore the conformational heterogeneity of the sample.

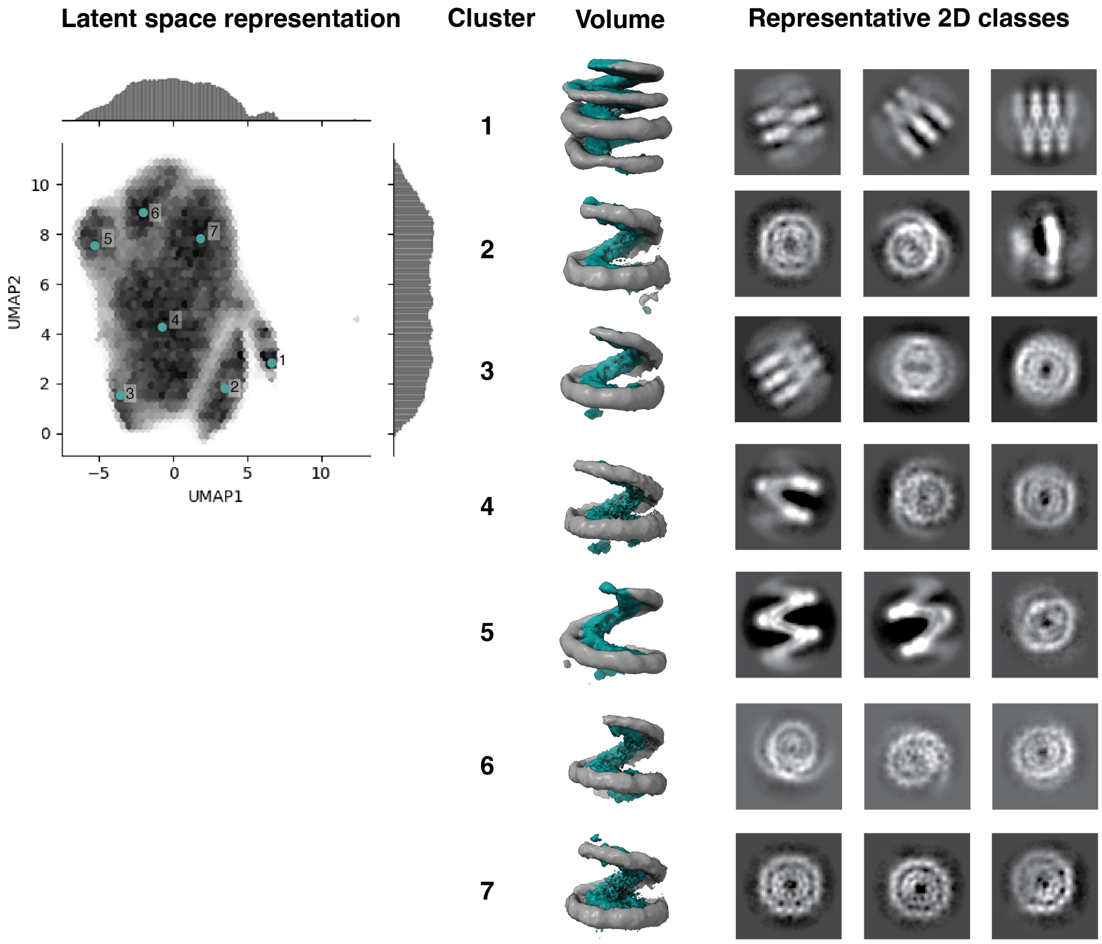


Figure 12: CryoDRGN analysis of the 1 mM $MgCl_2$ dataset of HHoB-DNA complex. On the left is the latent space representation and four clusters are highlighted in cyan. The corresponding density maps are shown on the right, showing the heterogeneity of conformations. The maps are false colored with grey (DNA) and cyan (HHoB).

## 2.5 Role of Tyr44 in the open conformation

An open conformation of histone-DNA complexes had been observed in the H3-H4 octasome [28] in addition to an intermediate and closed form. Notably, the degree of opening in the HHoB octasome is larger (63 Å) when compared to the open state of H3-H4 octaosme (53 Å). Interestingly, the dimer-dimer interface at the dyad of this structure involves tyrosine residues at the equivalent position of 48 when compared to the Tyr44 in HHoB. This corresponds to one alpha helical turn toward the His51. I hypothesised that the position of the tyrosine along the dimer-dimer interface could play a role in the degree of opening of the nucleosome. Hence, I developed a double mutant of the HHoB (Y44A-I48Y) - henceforth called HHoB mut1. After recombinant protein expression and purification using the same protocol as wild type HHoB, I performed *in-vitro* reconstitution with 147 bp Wid601 DNA and visualised it on EMSA, both in the presence and absence of magnesium chloride in the buffer (Figure 13). The EMSAs show similar patterns to wild type HHoB - with a ladder like binding behaviour in the absence of magnesium ions and a smear shift to saturation in the presence of magnesium ions. This is expected since the histone-DNA interface was unchanged in the mutation.

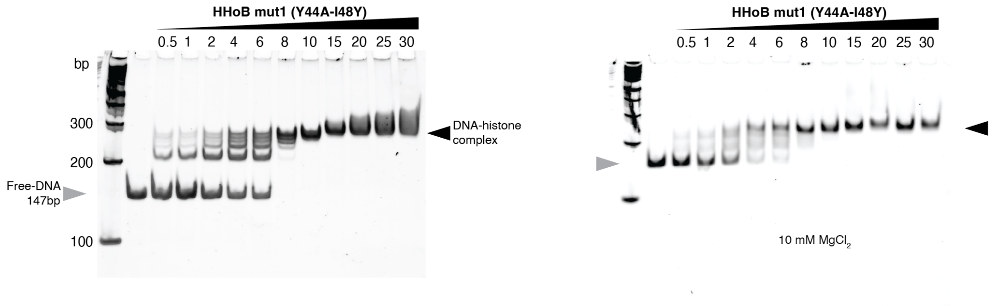
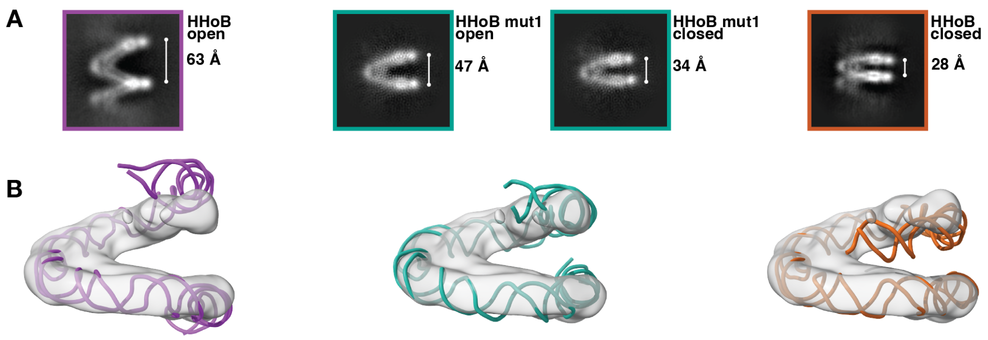


Figure 13: EMSA gels showing *in-vitro* reconstitutions of 147 bp Widom 601 DNA with HHoB mut1 in buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl and 5 % glycerol (left) and with an additive of 10 mM $MgCl_2$ (right). Increasing molar ratios of histone to DNA are shown on top of the gels. DNA concentration was constant (20 nM). First lane shows 100 bp DNA ladder (peqGOLD 100 bp Plus).

Next I prepared samples and collected cryo-EM data on the HHoB mut1 - DNA complexes in a buffer containing 1 mM $MgCl_2$. After pre-processing and filtering out of junk particles, I obtained 2D classes of nucleosomes in both open and closed conformations. However, when comparing the 2D classes, I determined that the open conformation had a smaller ’pitch’ than that of wild type HHoB nucleosome (Figure 14 A). The closed conformation from HHoB mut1 and HHoB nucleosome were not significantly different in the degree of opening (Figure 14 A). After further processing I obtained a low resolution EM map of the HhoB mut1 nucleosome, highlighting the instability of the nucleosome upon mutation of the dimer-dimer interface or it leading to a larger heterogeneity in the dataset. Regardless, the EM density seemed to have a lower pitch of opening than the native HHoB open structure. To determine this, I fit the density with 90 bp of DNA from the HHoB open and closed models, as well as from the H3-H4 octasome open model using rigid body fit in ChimeraX. As can be seen in the Figure 14B, only H3-H4 octasome showed a convincing overlap of 81.6% between the model and the density. This demonstrates that the open conformation is stabilised by the Tyr44, and that its position along the dimer-dimer interface can have an effect on the ’pitch’ of opening.



DNA from HHoB
open conformation
**(PDB 9QV7)**
overlap: 60.87

DNA from H3-H4 octasome
open conformation
(PDB 7X58)
overlap: 81.61

DNA from HHoB
closed conformation
**(PDB 9QV5)**
overlap: 52.43

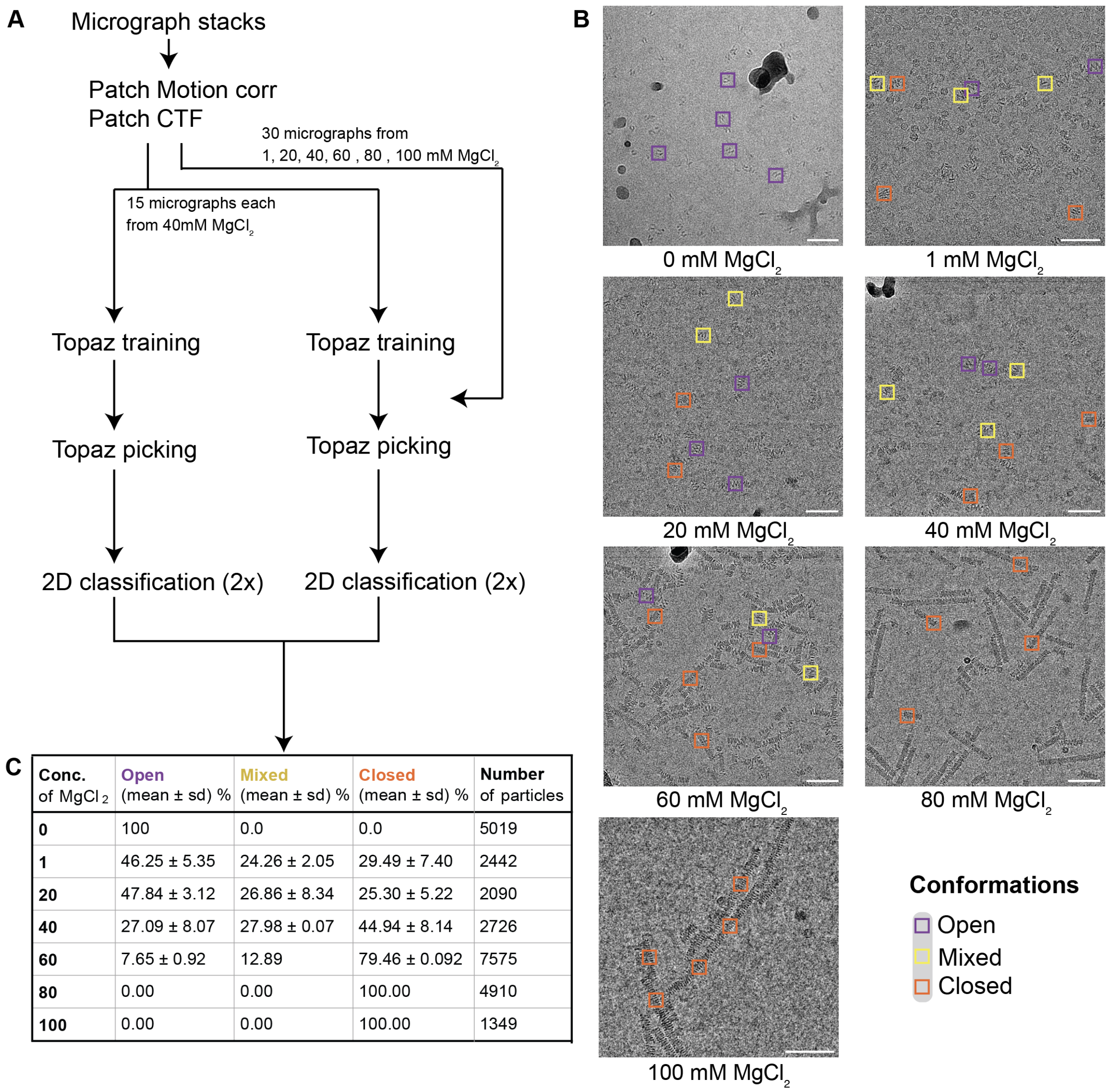
Figure 14: (A) 2D classes of HHoB mut1 – DNA in 1 mM $MgCl_2$ (cyan) are shown alongside corresponding view from open (purple) and closed (orange) conformations of the wild type complex. The superhelical pitch was intermediate. (B) A low-resolution density map of the HHoB mut1–DNA complex (transparent grey) was fitted with DNA models from the open (purple) and closed (orange) wild type molecular models, as well as the open conformation of the H3-H4 octasome (cyan)

# 3 HHoB nucleosome conformational screens

The concentration-dependent influence of magnesium ions on chromatin compaction has been well established in both eukaryotic and archaeal systems. After examining the HHoB–DNA complex under conditions containing 1 mM $MgCl_2$, I extended the analysis to a broader range of concentrations to assess how divalent ions shape its conformational landscape. This section presents the results of conformational screening across varying magnesium chloride concentrations, as well as comparative tests with other divalent ions. While the description here is rephrased for the purposes of this thesis, certain figures correspond closely to those reported in our publication [105].

## 3.1 Magnesium screen

Based on the limited known knowledge of intracellular ionic concentrations in archaea [107], I reconstituted HHoB nucleosomes in buffer A supplemented with magnesium chloride in steps of 20 mM i.e. 0, 1, 20, 40, 60, 80, 100 mM $MgCl_2$. After incubating for 20 min at room temperature, they were cryo-plunged and subjected for cryo-EM SPA. At least 30 micrographs were collected for every condition. Then two independent Topaz models were trained from blob picked and 2D classified particles from the 40mM $MgCl_2$ dataset, since this dataset showed the largest variety of conf ormations (Figure 15). Then, the two Topaz models were used to pick particles from all datasets in the conformation screen and 2D classified to eliminate junk particles (ice contamination, false positives, etc). The 2D classes were then labelled open, closed and mixed based on the pitch of opening, they are visualised in Figure 16A.



| **Conc.** of $MgCl_2$ | **Open** (mean ± sd) % | **Mixed** (mean ± sd) % | **Closed** (mean ± sd) % | **Number** of particles |
|---|---|---|---|---|
| **0** | 100 | 0.0 | 0.0 | 5019 |
| **1** | 46.25 ± 5.35 | 24.26 ± 2.05 | 29.49 ± 7.40 | 2442 |
| **20** | 47.84 ± 3.12 | 26.86 ± 8.34 | 25.30 ± 5.22 | 2090 |
| **40** | 27.09 ± 8.07 | 27.98 ± 0.07 | 44.94 ± 8.14 | 2726 |
| **60** | 7.65 ± 0.92 | 12.89 | 79.46 ± 0.092 | 7575 |
| **80** | 0.00 | 0.00 | 100.00 | 4910 |
| **100** | 0.00 | 0.00 | 100.00 | 1349 |

Figure 15: Single-particle analysis of the HHoB–DNA complex was carried out under buffer conditions containing different $MgCl_2$ concentrations. (A) illustrates the image-processing workflow. Panel (B) shows representative micrographs, where selected particle picks are labeled according to their conformational class (purple = open, yellow = mixed, orange = closed). Scale bar = 50 nm. Panel (C) provides a summary table of particle distribution among the conformations, with values expressed as mean ± standard deviation, calculated from technical replicates obtained using two independently trained Topaz particle-picking models.

Overall, with an increase in magnesium ion concentration, the population shifted toward closed conformations. The open conformation was present in samples upto a magnesium concentration of 60mM, demonstrating a range of physiological conditions over which this novel conformation might be stable. At higher concentrations of magnesium chloride, long fibres of closed conformation were prominent (Figure 16). Even in the highest magnesium chloride concentrations, I did not observe precipitation and aggregation, which would be expected for eukaryotic systems even beyond 5mM [41] - this can hint toward the ability of stable complexes being formed in archaeal systems with measured intracellular concentrations in the same range.

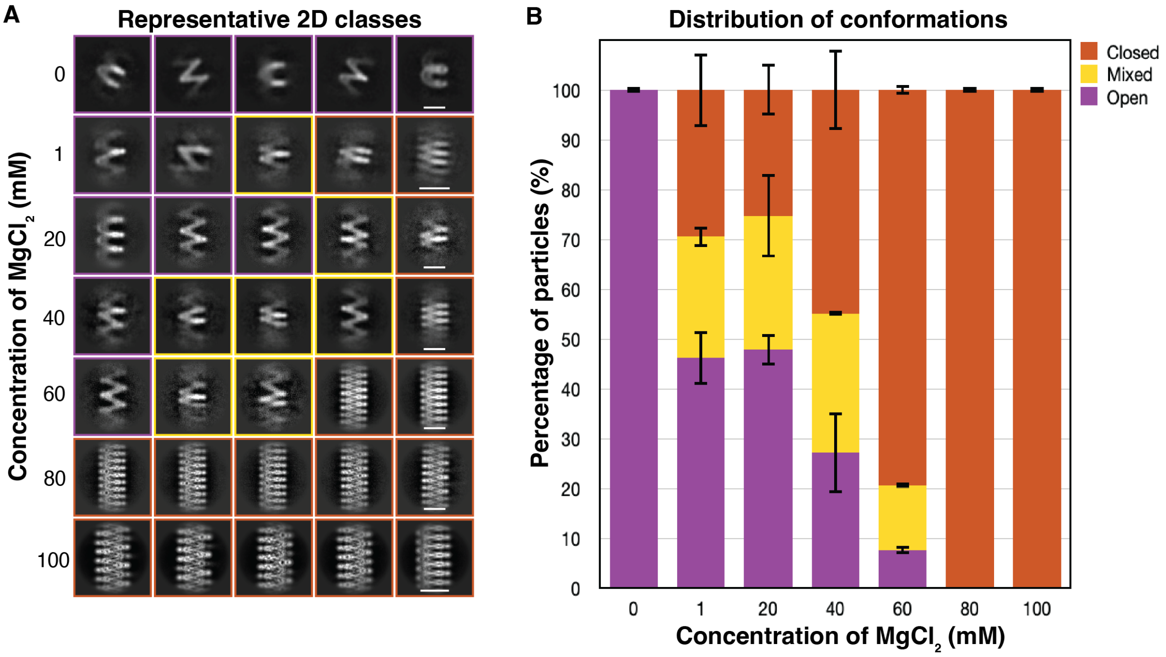


Figure 16: The influence of $MgCl_2$ concentration on HHoB-DNA complex conformation. (A) Representative 2D) class averages from cryo-EM data, sorted into open (purple), mixed (yellow), and closed (orange) conformations. A 10 nm scale bar is provided. (B) The bar chart illustrates the percentage distribution of these conformations across different datasets. Error bars show the standard deviation from two technical replicates, each analyzed using an independent particle-picking model.

Notably, the length of DNA used as template for reconstitution is 147 bp - which has enough footprint only for 4 histone dimers, corresponding to 1 superhelical turn. Yet, in samples containing magnesium chloride, I repeatedly observed complexes with multiple continuous turns - alluding to the end to end interactions of DNA in the presence of magnesium, and lack of sequence bias in the histone binding to the DNA. These hypernucleosomes were present in both the open and the closed state - with their populations most enriched in the 20mM $MgCl_2$(~40%) and the 100 mM $MgCl_2$ (100%) datasets respectively. They will be explored further in the next section.

## 4 HHoB Hypernucleosomes

After identifying that HHoB hypernucleosomes adopt both open and closed conformational states, with their relative abundance varying according to the magnesium chloride concentration in solution, I proceeded to investigate these assemblies in greater detail. My investigations sought to determine the structure and the role of key residues involved in these higher order assemblies, and to finally characterise them biophysically. Although the results are described independently in this thesis, certain figures partially overlap with those previously published in our article [105].

## 4.1 Cryo-EM of open and closed hypernucleosomes

I collected a dataset on HHoB - DNA complexes reconstituted in 20mM $MgCl_2$ since it had the largest enrichment of open hypernucleosome classes in the magnesium screen (subsection 3.1). As can be seen in the micrograph and 2D classes in Figure 17A, open hypernucleosomes corresponding to more than 1 superhelical turns of DNA were present, and often flexible in the number of turns and bendability of the fibre. Consequently, the EM density I obtained was resolved at a resolution of ~10 Å (Figure 17A). Still, I could fit two copies of the HHoB open nucleosome (built from the 3.6Å map) with high confidence (90% overlap of map to model). The resultant model contains two DNA molecules connecting end to end, and 6 histone dimers coating them with only histone dimer-dimer contacts, which are enough for driving the hypernucleosome formation.

Unlike the open hypernucleosome dataset at 20mM $MgCl_2$, the dataset at 100 mM $MgCl_2$ contained exclusively closed hypernucleosomes, with individual fibre lengths up to $1\mu$ m (Figure 17B). The closed hypernucleosomes formed rigid helical assemblies, and were processed accordingly in cryoSPARC [68]. After 3D classification and polishing, I obtained the closed hypernucleosome map at 2.6Å, and built the model with high confidence of side-chain placements. The model comprises of 180 bp of DNA bound by 6 histone dimers in a continuous left handed superhelix (Figure 17B). Again, while the length of DNA used as template was only 147 bp, the end-to-end joining of DNA could not be resolved due to averaging during helical reconstruction.

The HHoB closed hypernucleosome structure is very similar to the X-ray crystallography structure of the HMfB hypernucleosome (RMSD 0.82 Å across three histone dimers at C$\alpha$)[17]. The HHoB closed conformations from the single nucleosome and hypernucleosome dataset are similar too (RMSD 0.87 Å across three histone dimers at C$\alpha$). The pitch of the rigid closed hypernucleosome (24.6 Å) is smaller than the closed nucleosome (29.5 Å), where the DNA ends have more flexibility. The position of side-chains at the dimer-dimer interface of the closed hypernucleosomes resemble that of the closed nucleosome state.

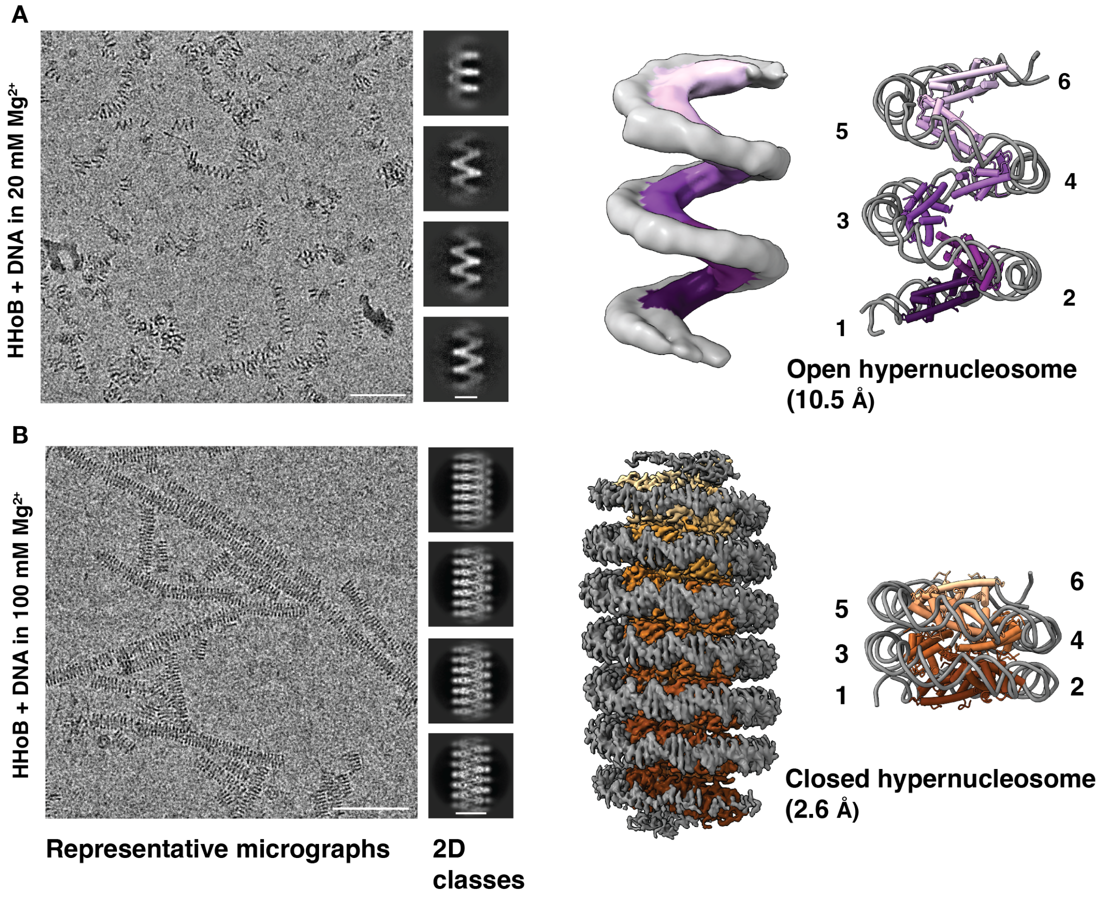


Figure 17: Structural characterization of HHoB hypernucleosomes in open (A) and closed (B) states. (A) Cryo-EM micrograph and 2D class averages of the HHoB–Widom601 147 bp complex in 20 mM MgCl_2 (scale bars: 50 nm and 10 nm), and the open-state cryo-EM map with a rigid-body fit of two models showing 180 bp of DNA (gray) and six histone dimers (shades of purple). (B) Cryo-EM micrograph and 2D class averages of the complex in 100 mM MgCl_2 (scale bars: 50 nm and 10 nm), and the refined cryo-EM map with molecular model of the closed hypernucleosome (PDB 9QV6) containing 180 bp of DNA (gray) and six histone dimers (shades of orange).

Additionally, the stacking interface (1674 + 1676 Å² per histone dimer) could be determined between non-neighbouring histones (Figure 18), which is absent in the open conformation. These occur between the α3 and α2 helices of histone dimer N1 and the α2 helix of dimer N3, as well as between the $\alpha 2$ and $\alpha 1$ helices and the L1 loop of dimer N3. The structural analysis identified Gln16, Lys29, Glu32, Asp36, Arg50, and Lys63 as the main contributors to these stacking interactions (Figure 18), forming a ’velcro-like’ network of complementary charged residues.

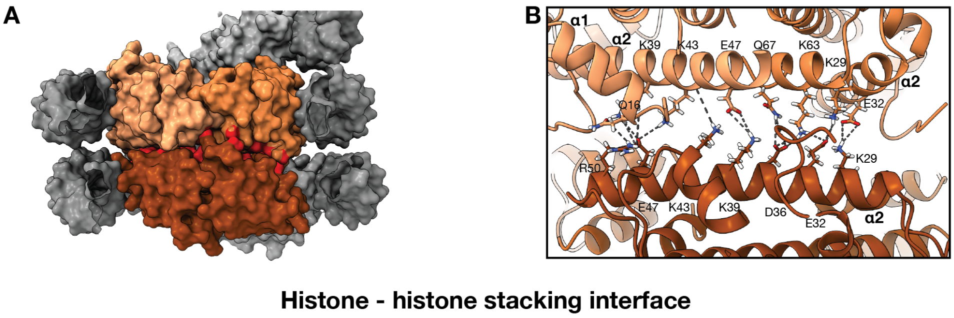


Figure 18: Histone stacking interface in the HHoB hypernucleosome closed (orange) state between the N, N+2 and N+3 histone dimers. (A) Surface representation of the hypernucleosome showing the histone–stacking interface in red. (B) Ribbon representations of the histones highlighting the key residues involved at the interface: Q16, K29, E32, K39, K43, E47, R50, K63, Q67. Dashed lines indicate electrostatic and hydrogen bond interactions.

Although the map resolution was sufficient to identify potential ion-binding sites, no corresponding electron densities for magnesium ions were detected. This suggests that magnesium ions do not directly associate with the histones, and must act by non-specifically binding to the phosphate backbone of DNA leading to reduced replusion in the closed state. Additional support to this was provided when I observed that hypernucleosome formation under the “high Mg” condition (100 mM) could also be achieved by replacing $Mg^{2+}$ with other divalent cations, such as $Zn^{2+}$ and $Ca^{2+}$ (Figure 20).

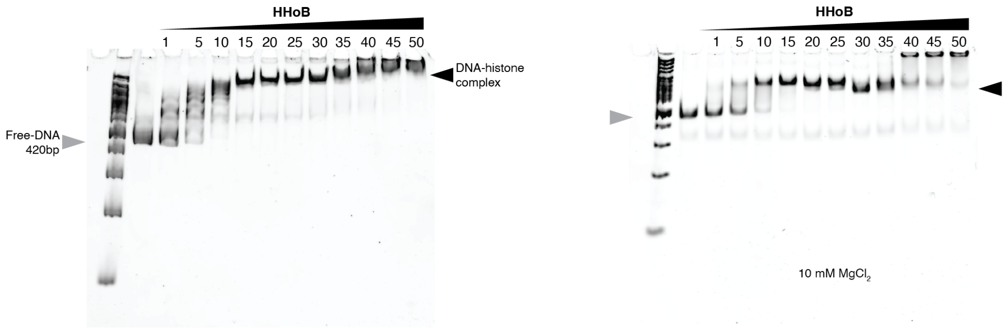


Figure 19: EMSA gels showing *in-vitro* reconstitutions of 420 bp HeimC3_31310 native DNA with HHoB in buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl and 5 % glycerol (left) and with an additive of 10 mM $MgCl_2$ (right). Increasing molar ratios of histone to DNA are shown on top of the gels. DNA concentration was constant (20 nM). First lane shows 100 bp DNA ladder (peqGOLD 100 bp Plus).

To rule out the possibility that hypernucleosome formation was specifically influenced by the SELEX-derived Widom601 DNA sequence, I performed reconstitutions using a native 420 bp genomic fragment from the HeimC3_31310 gene of the LC_3 metagenome. Electrophoretic mobility shift assays (EMSA) revealed binding characteristics similar to those observed with Widom601 DNA (Figure 19). As expected, the longer native DNA displayed more shifted bands compared to the 147 bp Widom601 fragment, reflecting the presence of additional binding sites. Notably, with 100 mM $MgCl_2$, cryo-EM analysis showed hypernucleosome assembly comparable to that seen with Widom601 DNA (Figure 20). The pitch of the native-sequence hypernucleosome was determined to be 25.2 Å, closely matching the 25.4 Å pitch observed for the closed hypernucleosome assembled using Widom601 DNA.

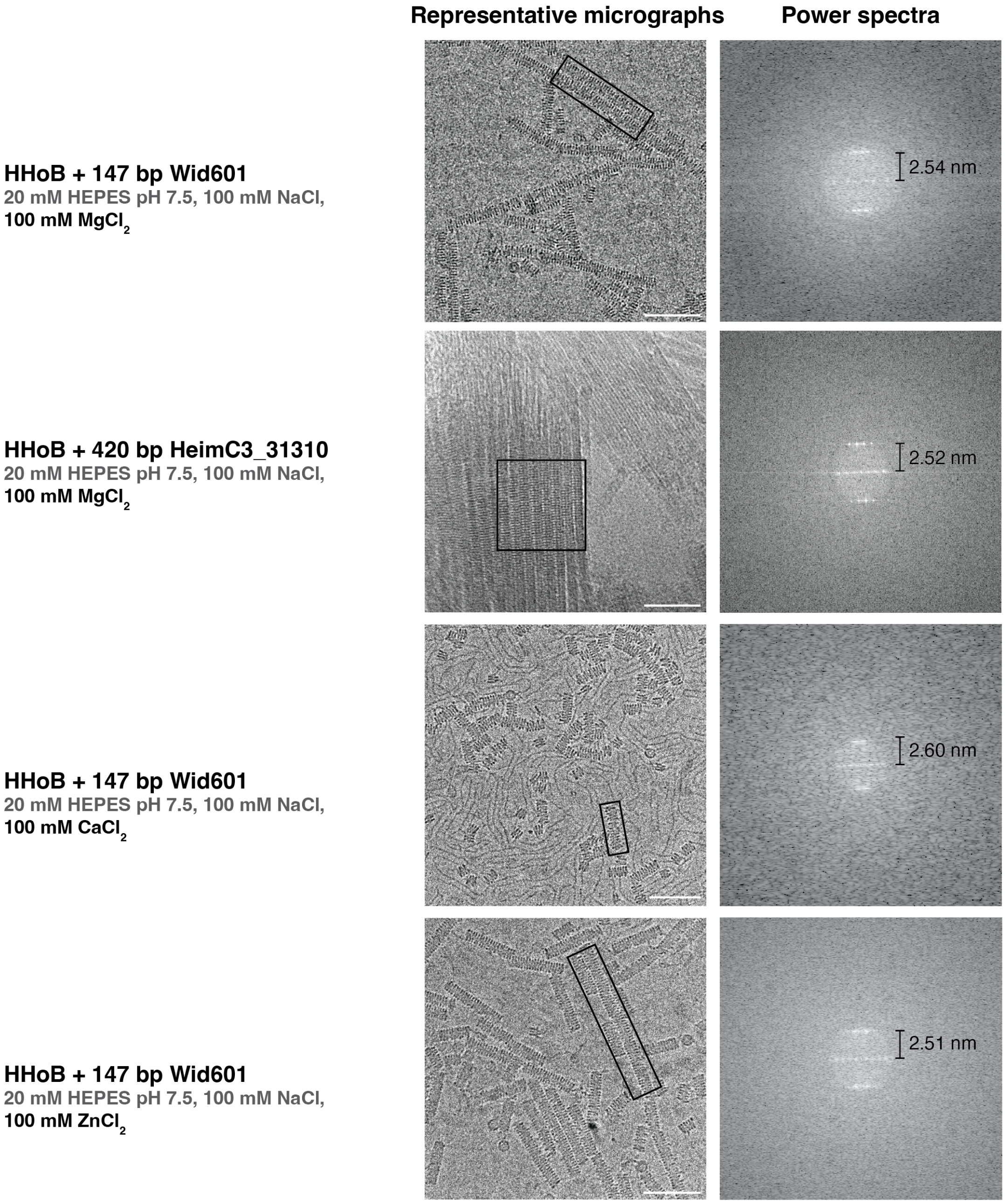


Figure 20: Representative HHoB closed hypernucleosomes formed under varying buffer conditions (indicated on the left). Shown are micrographs with highlighted regions of interest, illustrating the presence of closed hypernucleosomes across conditions. The corresponding power spectra from fast Fourier transform (FFT) analysis are displayed on the right, with annotated inter-turn distances ("pitch"). Scale bar: 50 nm.

## 4.2 Role of the stacking interface in closed hypernucleosome

To furthe relucidate the role of the residues involved in the stacking interface, I prepared a triple mutant (K29A E32A D36A) called HHoB mut2. After recombinant protein expression and purification using the same protocol as wild type HHoB, I performed *in-vitro* reconstitution with 147 bp Wid601 DNA and visualised it on EMSA, both in the presence and abscence of magnesium chloride in the buffer (Figure 21). The EMSAs show similar patterns to wild type HHoB - with a ladder like binding behaviour in the absence of magnesium ions and a smear shift to saturation in the presence of magnesium ions. This is expected since the histone-DNA interface was unchanged in the mutations.

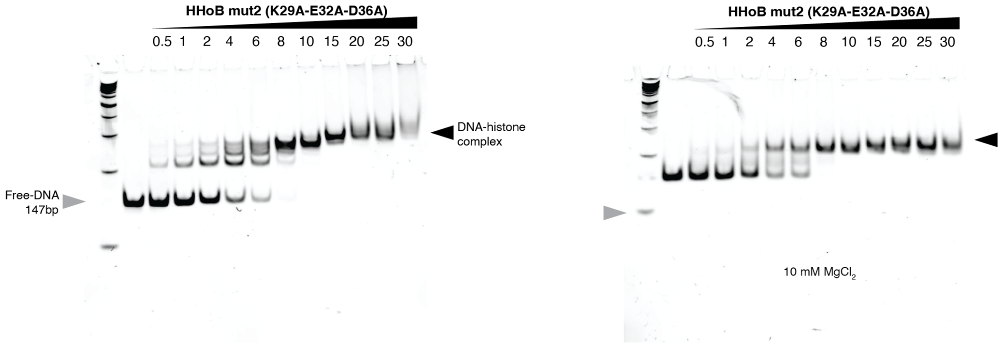


Figure 21: EMSA gels showing *in-vitro* reconstitutions of 147 bp Widom 601 DNA with HHoB mut2 in buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl and 5 % glycerol (left) and with an additive of 10 mM $MgCl_2$ (right). Increasing molar ratios of histone to DNA are shown on top of the gels. DNA concentration was constant (20 nM). First lane shows 100 bp DNA ladder (peqGOLD 100 bp Plus).

Next, I reconstituted HHoB mut2 nucleosomes in the presence of 100 mM $MgCl_2$ to examine using cryo-EM the effect of disrupting stacking residues on hypernucleosome stability. From Figure 22, it is evident that while the HHoB mut2 forms hypernucleosomes, the fibres are flexible with a variable pitch that is larger than the native closed hypernucleosome. Hence, while three out of five residues involved in the stacking interface are not enough to disrupt hypernucleosome formation completley, they do destabilise the stacking interface and hence the rigidity of the fibres.

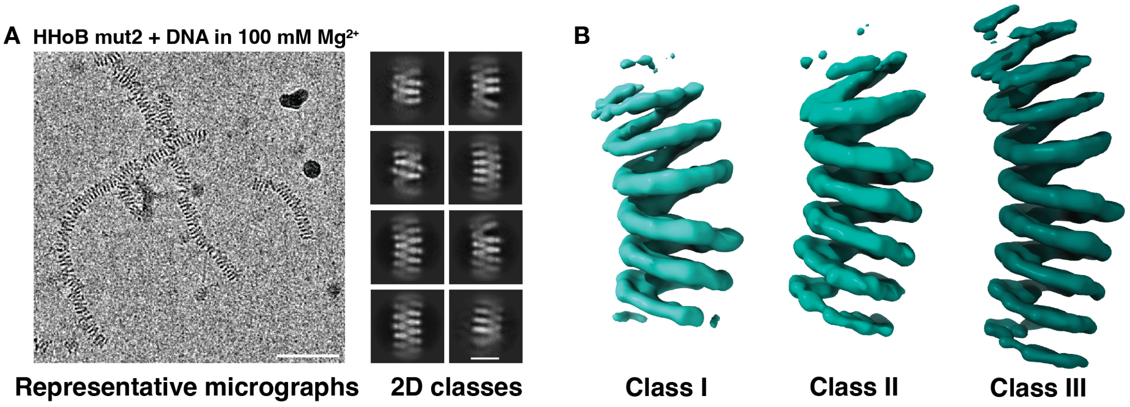


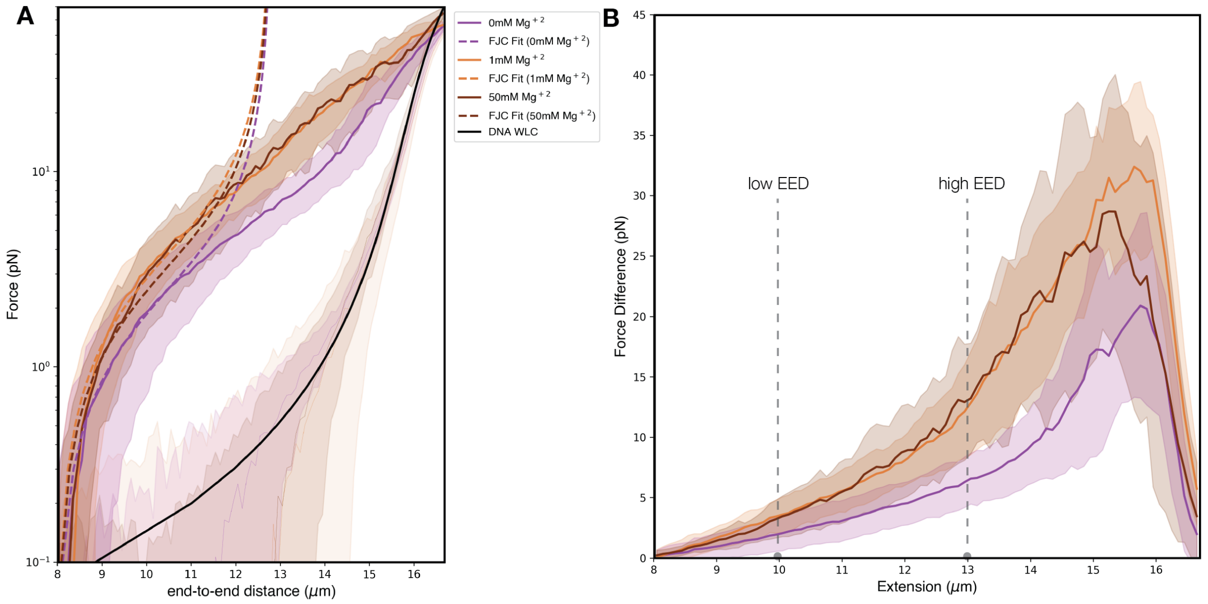
Figure 22: (A) Representative raw micrograph of HHoB mut2 - DNA complex in 100 mM $MgCl_2$. Scale bar: 50 nm. Representative 2D class averages showing side views. (B) Shades of cyan densities are volumes from 3D classification showing flexible fibres.

### 4.3 Effect of magnesium ions on HHoB - DNA complexes by force spectroscopy

To examine the mechanical response of the HHoB hypernucleosome to $Mg^{2+}$, I employed optical-tweezer-based force spectroscopy using long biotinylated λ-phage DNA (48,502 bp) as the substrate. In these experiments, I used $Mg^{2+}$ concentration as a proxy for chromatin compaction, consistent with cryo-EM observations showing that the histone–DNA complex adopts exclusively an open state in the absence of $Mg^{2+}$, whereas mixtures of open and closed conformations are found at intermediate $Mg^{2+}$ levels. At concentrations above ~80 mM $Mg^{2+}$, only the closed conformation is observed. I therefore carried out force-extension measurements of the HHoB–DNA complex in buffers containing 0 mM, 1 mM, and 50 mM $MgCl_2$, with each curve representing the mean ± standard deviation from 10 independent force-extension traces (Figure 23). The protein concentration in the channel (HHoB-Atto647N, 200 nM) was approximately 6.5-fold higher than the measured dissociation constant (Kd = 30.8 ± 3.7 nM), ensuring stable nucleoprotein complex formation. Theoretical worm-like chain (WLC) fits for free DNA are also plotted and are in excellent agreement with the experimental free-DNA curves, confirming that magnesium does not alter the intrinsic mechanical properties of DNA alone.

Interestingly, I observed that the overall force–extension curves are highly similar to those reported previously for HMfB hypernucleosomes [39]. Since the closed HHoB hypernucleosome is structurally very close to HMfB (RMSD = 0.8 Å over three histone dimers at Cα atoms), I applied the free-jointed chain (FJC) model originally developed for HMfB to describe HHoB as well. In the low-force regime (« $k_BT$), this model describes the fibre as a Hookean spring, where stiffness is inversely proportional to the Kuhn length. Fitting the FJC model to the HHoB force–extension curves (average FJC curves shown as dotted lines in Figure 23, N=10) yielded stiffness values of 1.77 ± 0.64 pN/nm (1 mM $MgCl_2$) and 1.45 ± 0.63 pN/nm (50 mM $MgCl_2$), both substantially higher than the value obtained in the absence of magnesium (0.76 ± 0.51 pN/nm). These findings demonstrate that $Mg^{2+}$ markedly stiffens HHoB–DNA fibres. However, given the relatively small dataset, I consider these results to be preliminary and not yet statistically definitive.

The effect of magnesium is further evident when examining force differences between HHoB–DNA complexes and free DNA across the full extension range (Figure 23B). At short extensions (10 µm), forces were 2.01 ± 0.36 pN (0 mM), 3.52 ± 0.59 pN (1 mM), and 3.27 ± 0.68 pN (50 mM). At longer extensions (13 µm), forces increased to 6.63 ± 0.56 pN, 12.73 ± 2.03 pN, and 13.41 ± 1.94 pN, respectively (N = 10). Paired t-tests confirmed that forces at 0 mM were significantly lower than those at both 1 mM and 50 mM $MgCl_2$ at both extension regimes ($p < 0.01$). While clogging in the microfluidic channels at higher protein concentrations prevented exploration of fully saturated binding conditions, the consistent, statistically significant $Mg^{2+}$-dependent increase in force strongly supports the conclusion that HHoB promotes DNA compaction. This effect is likely mediated by cooperative histone binding and/or hypernucleosome stacking interactions.



**Force-extension curves** **Force difference curves**

Figure 23: Biophysical characterization of the HHoB–DNA complex in the presence of magnesium ions. (A) Force-extension curves of HHoB–Lambda phage DNA (48.5 kb) in 0 mM, 1 mM and 50 mM MgCl_2 (purple, orange, and brown, respectively). Solid lines and shading indicate the mean and standard deviation (n=10). The black solid line represents the theoretical Worm-Like Chain (WLC) for DNA at 25°C. Dotted lines represent Free-Jointed Chain (FJC) models fitted for the three mean curves in the low force regime (force « $k_BT$). (B) Force-difference curves between histone-DNA complexes and free DNA. Vertical dashed lines indicate the low and high end-to-end distances at which the datasets were compared.

## 4.4 Molecular Dynamics simulations of HHoB hypernucleosome

To investigate the effect of magnesium ions on the conformational dynamics of HHoB, I performed all-atom molecular dynamics (MD) simulations using the AMBER package. As the initial model, the open-state HHoB hypernucleosome structure (PDB ID: 9QV7) containing histone dimers wrapping 180bp of DNA was employed. Each simulation was carried out for 1000 ns in triplicate, providing sufficient sampling for structural relaxation and potential large-scale conformational transitions. To quantify DNA compaction during the trajectories, I measured the distances between ten pairs of corresponding phosphate groups along the helical pitch of the hypernucleosome. Two sets of conditions were compared: (i) in the absence of magnesium ions, and (ii) in the presence of $MgCl_2$ at a concentration corresponding to 100 mM in solution. The resulting DNA compaction profiles are presented as mean values with standard deviations in Figure 24. In the absence of magnesium ions (purple trace), the trajectories display local breathing fluctuations; however, the overall degree of DNA compaction remains comparable to the open state resolved by cryo-EM. These findings support the notion that the open conformation of the HHoB hypernucleosome is intrinsically stable in the absence of divalent cations. By contrast, in simulations including 100 mM $MgCl_2$, a pronounced deviation from the no-magnesium trajectories is observed beyond ~200 ns. Here, the complex undergoes progressive compaction, reaching values of ~27 nm within the final 400 ns of simulation. Strikingly, this is in close agreement with the compaction observed in the cryo-EM structure of the closed hypernucleosome fibre (24.6 nm).

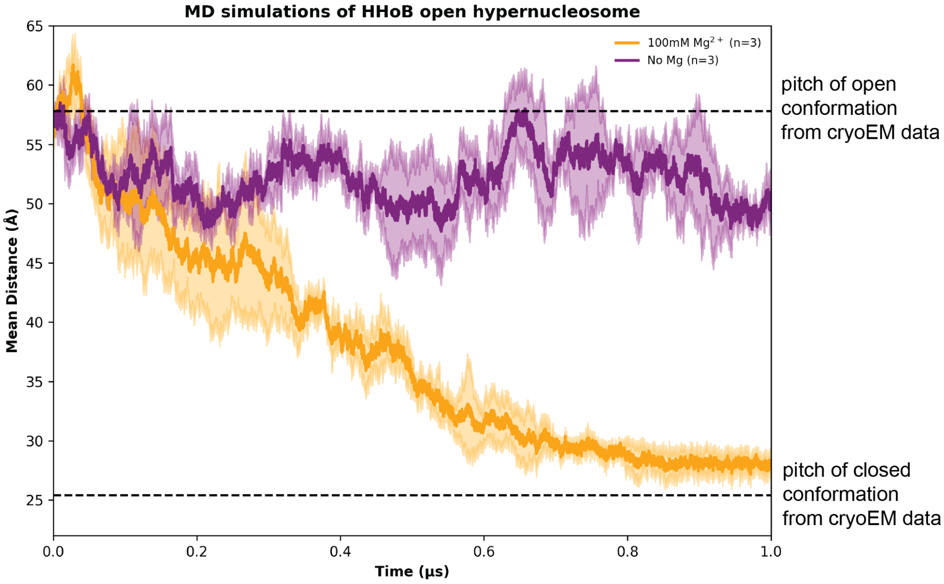
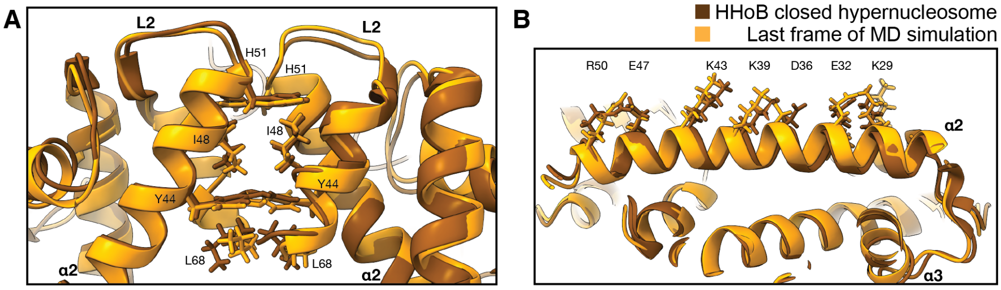


Figure 24: Molecular dynamics traces without (purple) and with 100 mM $Mg^{2+}$ (orange), using HHoB hypernucleosome in open conformation as the starting model. The mean distance from 10 pairs of phosphtate groups are plotted as mean and standard deviation (n=3). The horizontal lines denote the corresponding mean distances from the closed hypernucleosome and open nucleosome models determined using cryo-EM.

Structural comparisons further support this conformational transition. When the final frame of the $Mg^{2+}$ - containing simulation is overlaid with the cryo-EM closed state, the agreement extends beyond global DNA compaction and is evident at the level of side-chain conformations. At the dimer–dimer interface (Figure 25A), the key residues His51, Ile48, Tyr44, and Leu68 adopt conformations matching those in the experimental closed structure. A similar correspondence is observed at the histone-stacking interface (Figure 25B), where the defining side-chain orientations are recapitulated. Together, these results demonstrate that $Mg^{2+}$ - driven MD simulations not only capture the correct degree of DNA compaction but also reproduce the atomic-scale rearrangements that underpin the high-resolution architecture of the HHoB closed hypernucleosome.



**Dimer - dimer interface**

**Histone stacking interface**

Figure 25: Comparing key side-chain residues from the last frame of the MD simulation of HHoB open hypernucleosome in the presence of 100 mM $Mg^{2+}$ (orange) and that from the closed hypernucleosome model determined using cryo-EM (brown). Both the dimer-dimer interface (A) and the stacking interface (B) residues display considerable overlap in conformations.

To further investigate the mechanism of $Mg^{2+}$ - driven DNA compaction, I calculated the nearest distances between ten pairs of $Mg^{2+}$ ions to $PO_4$ of the DNA at the first and last frames of the MD simulations. The resulting distributions, aggregated as mean counts with standard deviations across the three runs (n=3), are presented in the attached bar plot Figure 26. The most striking finding is the presence of a significantly large population of $Mg^{2+}$ ions in close proximity to the $PO_4$ groups of the DNA backbone in the final state. Specifically, the bin corresponding to the $0-5$ Å distance shows a mean count of $\sim$ 72 ions (teal bar), which is higher than any other bin. This short distance, less than 5 Å, makes these ions valid candidates for direct or outer-shell interactions with the phosphate groups. I did not observe any direct specific interactions between $Mg^{2+}$ ions and the HHoB protein during the MD runs. The high occupancy at the DNA surface strongly supports the hypothesis I presented earlier (subsection 4.1) that $Mg^{2+}$ ions interact non-specifically with the DNA backbone, providing electrostatic screening and facilitating the compaction of the hypernucleosome.

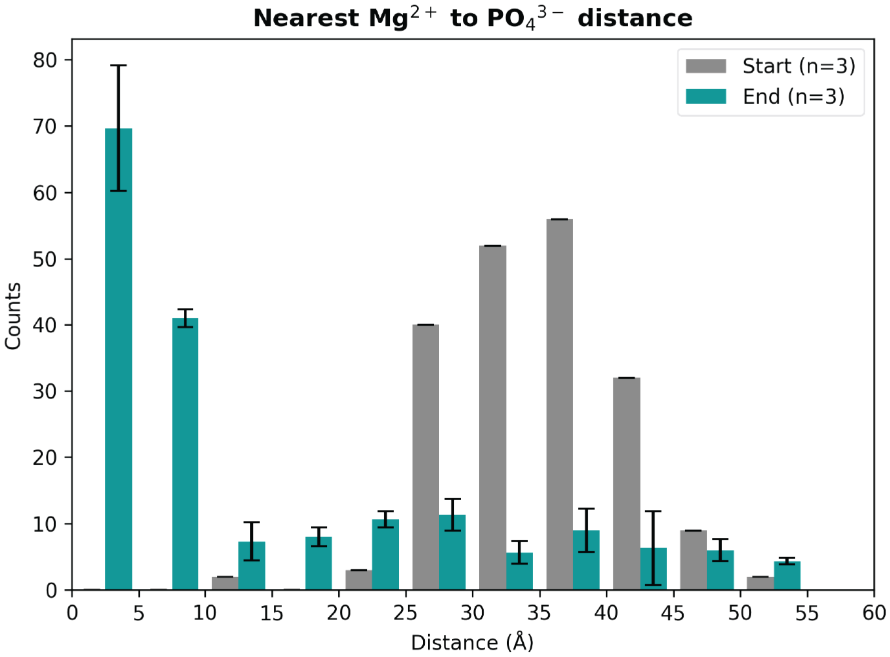


Figure 26: Relocation of magnesium ions in the MD run of HHoB open hypernucleosome in the presence of 100 mM $Mg^{2+}$ between the first (gray) and the last frame (cyan). Distances were calculated as the nearest neighbours of magnesium ions and phophate groups from DNA. The columns are plotted as mean and standard deviation (n=3).

# 5 Scouting Experiments

In this section, I describe some of the extra experiments done in parallel with the HHoB studies - these include attempts with the tailed histone HHoA, and histones HHoF and HHoG from Hodarcheon LC_3. Additionally, I present a conformation of nucleosomes that is fascinating but probably an artefact of the DNA substrate.

## 5.1 Histone HHoA from Hodarcaheon LC_3

HHoA is the 91 amino acid histone from the Hodarchaeal LC_3 metagenome with an N-terminal tail with sequence generally comparable with H4 from eukaryotes. I carried out expression tests and it was determined that inclusion body purification would yield the highest protein amount. Hence His-tag based affinity purification was carried out in 6M urea, and pure protein was obtained. Multiple strategies were tested for refolding the protein – quick dilution, overnight dialysis, dialysis into ammonium sulfate followed by dialysis into base buffer. The presence of alpha helices as an indication of the folded state of the protein was then measured using circular dichroism (Figure 27). As a positive control, the unfolded protein was treated with 40% tetrafluoroethylene to induce alpha helix formation – yielding 30 % alpha helical content in comparison to the predicted 50.1%. Unfortunately, none of the refolded proteins from inclusion body purifications contained alpha helices. The next strategy was to express the protein in the soluble fraction to avoid unfolding and refolding. Hence, constructs with MBP, GST, NUS tags on the N-termini were tested and an increase in solubility was achieved. Leaky expression of the N-terminally His-tagged construct also yielded small amounts of soluble protein. Mutants were made and tested as well – HA without a tail, HHoB with the HA tail on the N-terminus. While the core HA protein could be expressed and purified, it was also revealed to be unfolded (Figure 27). Mutants of HHoB with the HA tail on the N-terminal did not express in any of the tested conditions.

To test folding upon binding to DNA, extensive screening was carried out for reconstitution conditions by testing high protein concentrations (1 – 100 times molar excess), salt concentrations (30-150 mM NaCl, 30 mM KCl, 1 mM $MgCl_2$), presence of 1 mM $MgCl_2$, presence of crowding agents (5% 1,6 hexane-diol, 0.1 mg/ml BSA, 1.5 mM CHAPS, 250 mM glycine) and reconstitutions at 4°C instead of 25°C in the protocol for HHoB, HHoF and HHoG. An optimal reconstitution mixture contained 20 mM HEPES, 1 mM EDTA, 5% glycerol, 30 mM NaCl and 250 mM glycine incubated at 4°C for 30 min. The corresponding EMSA for this reaction can be seen in Figure 28A, showing smear like shifts to full complexes at a high molar ratio (x40). Smear-like shifts indicating co-operative binding have been shown for other archaeal histones [5], however the requirement for crowding agents, reconstitution at 4°C and high molar ratio (x40) for shift in the band highlights the low DNA binding affinity of the protein. I prepared cryo-EM grids with this sample and upon screening, only free DNA was seen on the grid holes (Figure 28). Ultimately, the protein was not pursued further.

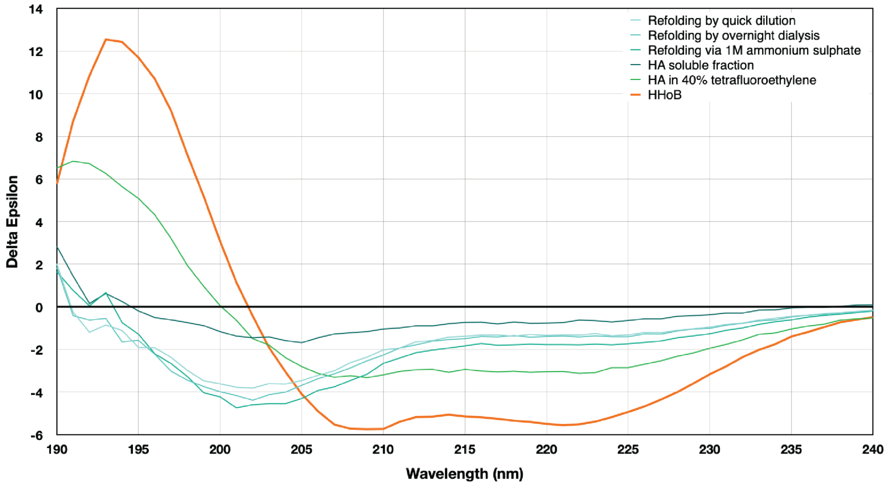


Figure 27: Circular dichroism spectra of histones HHoB (red) (typical of a protein with alpha helices) and HHoA prepared in various methods to test for folded state of the protein. A positive control of HHoA in 40% tetrafluoroethylene (green) indicates an ideal folded state signature for that histone.

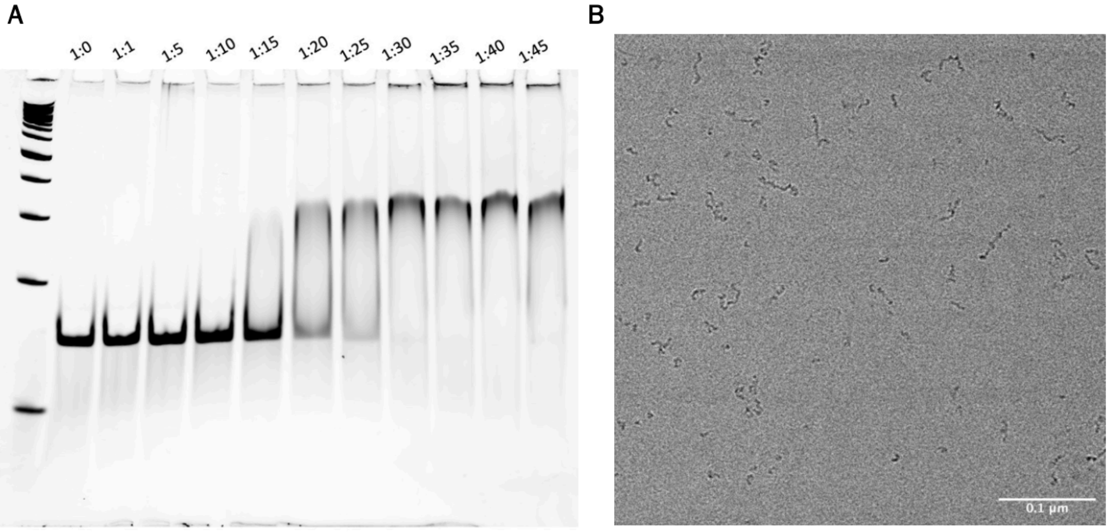


Figure 28: (A)EMSA gel showing *in-vitro* reconstitutions of 147 bp Widom 601 DNA with histone HHoA in buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl, 50mM glycine and 5 % glycerol. Increasing molar ratios of histone to DNA are shown on top of the gels. DNA concentration was constant (20 nM). First lane shows 100 bp DNA ladder (peqGOLD 100 bp Plus). (B) Screening micrograph HHoA - DNA sample with 40 times molar excess of the protein.

## 5.2 Histones HHoF and HHoG

Upon large scale protein purificaiton and EMSAs, I performed cryo-EM screening of HHoF and HHoG proteins with 147 bp Widom 601 DNA in equivalent amounts as for optimised HHoB grids. Tested a range of magnesium chloride concentrations (20, 60, 100 mM $MgCl_2$). While HHoF proteins did not show nucleosome formations in any of the tested regimes, HHoG presented closed hypernucleosomes (Figure 29) in the presence of 60mM $MgCl_2$. From the power spectrum it was clear that the pitch of the HHoG of 2.55 nm closely resembles that of the HHoB hypernucleosome at 2.56 nm (Figure 20). Indeed, histone HHoG shows identical residues at key positions as in HHoB - at the DNA binding interface, stacking interface and the dimer-dimer interface, albeit with a phenylalanine instead of tyrosine at the position 44 (HHoB numbering)(Figure 1).

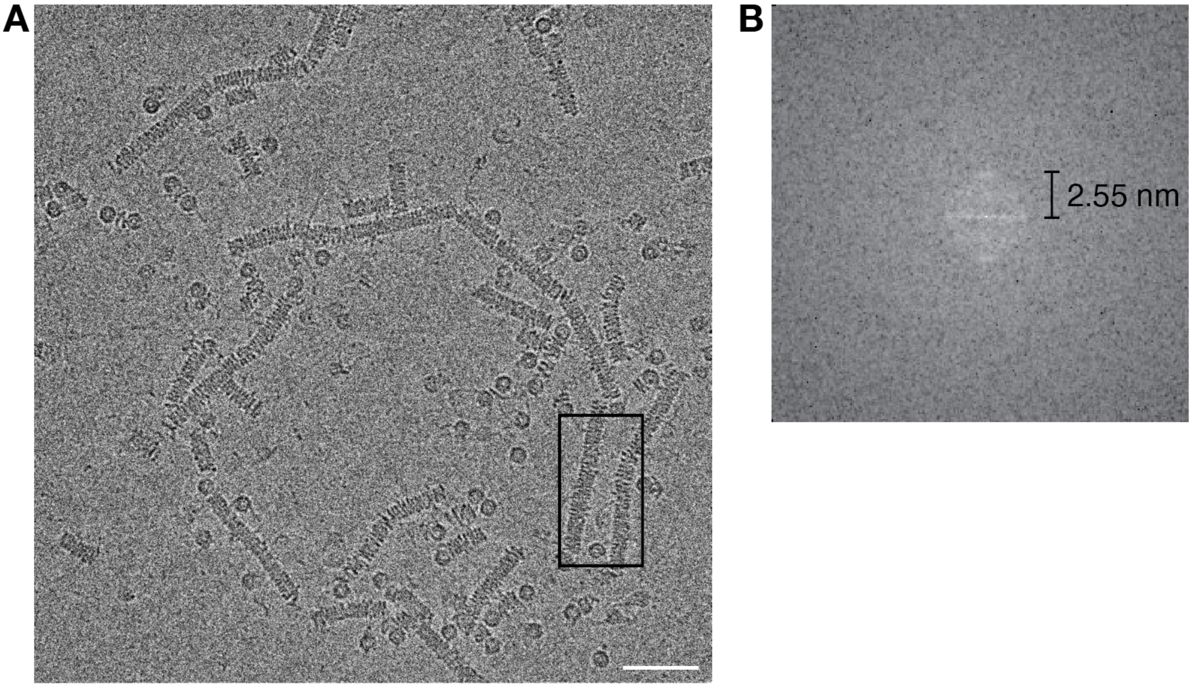


Figure 29: Representative HHoG closed hypernucleosomes formed in a buffer containing 60 mM $MgCl_2$. (A) Shows a micrograph with highlighted regions of interest, illustrating the presence of closed hypernucleosomes. The corresponding power spectrum from fast Fourier transform (FFT) analysis is displayed in (B), with annotated inter-turn distances ("pitch"). Scale bar: 50 nm.

### 5.3 Infinity shaped nucleosome conformation

In addition to the Hodarchaeal LC_3 histones, I also tested histones from some other Asgard metagenomes, selecting for thermostability to explore for potentially stable histones with tails. After small scale protien purifications and EMSA screenings, histone HB from Heim_SZ_4 (length 72 aa) was selected for cryo-EM screenings. Upon reconstituting with 147 bp Wid601 DNA in a minimal buffer (20mM HEPES pH 7.5, 100 mM NaCl) and testing a range of concentrations ( 1.9 µM – 2.5 µM), I observed the formation of continuous fibre-like structures with an approximate width of ~20 nm. This dimension is nearly twice that for hypernucleosome fibres formed by histone HHoB from Hodarcheon LC_3. Using the most concentrated sample grids, I collected cryo-EM data on the Krios cryo-TEM, revealing micrographs with a high density of fibres (Figure 30). Following pre-processing of 15,044 movies, I initially traced filamentous side views and subsequently trained a Topaz model to identify particles in the top-view orientation. Two rounds of 2D classification were performed to eliminate artefacts and low-quality particles. The resulting 2D classes revealed distinct assemblies (Figure 30B), which contrasted markedly with the assemblies of HHoB hypernucleosomes (Figure 17). Notably, the fibres did not display continuous helical symmetry. Instead, they appeared to consist of two parallel protein fibres shifted by one turn, with DNA densities observed only on every alternate turn. Consequently, standard helical reconstruction approaches were not suitable for refinement.

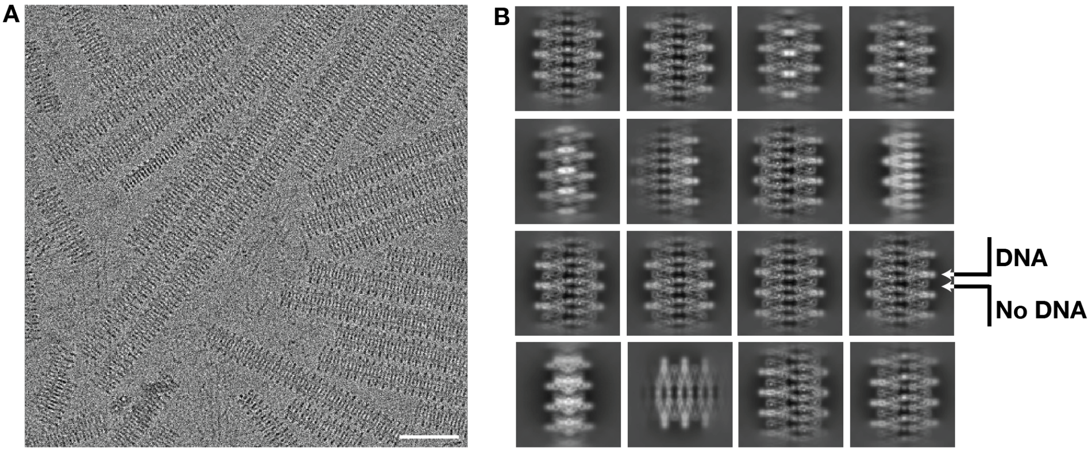


Figure 30: (A) Representative micrograph showing Heim_SZ_4 histone HA closed hypernucleosomes formed in minimal buffer without divalent ions. Scale bar: 50 nm. (B) Shows 2D classes from particle picking showing different views of the fibre. Annotated are turn with and without signal corresponding to DNA wrapping.

Subsequent non-uniform refinement produced a 3D reconstruction at sufficient resolution to allow fitting of AlphaFold-predicted histone dimers into the density (Figure 31). The DNA densities within the map, however, were poorly resolved, likely due to the limited angular distribution of particle views. Nevertheless, the DNA backbone could be modelled, and the arrangement was consistent with the assemblies observed in 2D classes. Specifically, the complex comprised two protein fibres, with DNA wrapping around each fibre in a manner resembling an infinity (‘∞’) symbol in the top view. The fitted DNA length corresponded to ~147 bp per infinity turn, which matched the DNA fragment size used in reconstitution experiments. Interestingly, within every three helical turns of the protein fibres, two histone dimers appeared unoccupied by DNA. These positions were presumably stabilised through stacking and dimer–dimer interactions, in the absence of divalent ions, hinting at the propensity for this histone to make closed conformation assemblies.

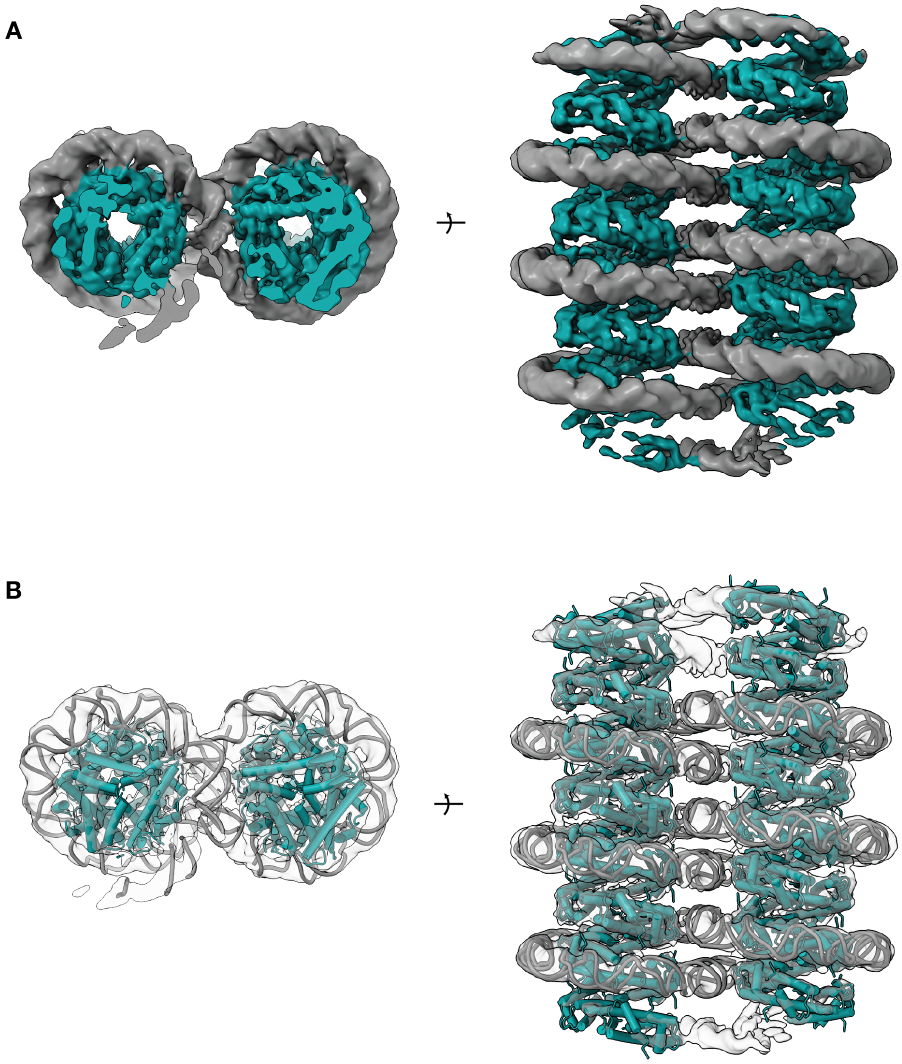


Figure 31: Heim_SZ_4 histone HA - DNA complex. Top and side views of the EM map are presented with DNA in grey and histone dimers in cyan (A). (B) shows transparent maps in the same views with a rigid body fit of 3 copies of 147 bp of DNA and histone dimer copies from AlphaFold2 predictions.

To test whether the observed infinity-shaped fibres were influenced by DNA fragment length, I reconstituted the histone with 420 bp DNA and analysed the assemblies by cryo-EM. Due to limited availability of the DNA fragment, screening could not be performed at identical concentrations but was carried out under conditions comparable to those used for HHoB. Under these conditions, complexes were less abundant in the vitrified ice. Still, I observed hypernucleosome fibres with widths comparable to those of HHoB hypernucleosomes, and no infinity-shaped assemblies were detected Figure 32. These results suggest that the infinity-shaped architecture is favoured only with DNA fragments of 147 bp, implicating DNA length as a key determinant of supramolecular assembly.

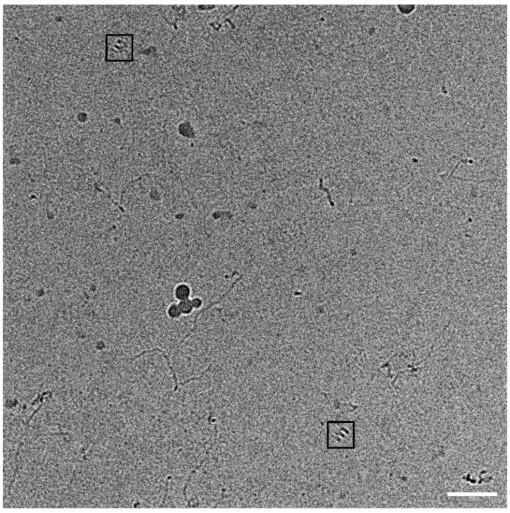


Figure 32: Heim_SZ_4 histone HA - DNA complex. Top and side views of the EM map are presented with DNA in grey and histone dimers in cyan (A). (B) shows transparent maps in the same views with a rigid body fit of 3 copies of 147 bp of DNA and histone dimer copies from AlphaFold2 predictions.

# Chapter V

# Discussion

## 1 Closed conformation is conserved

This investigation into the chromatin structure of Asgard archaea, specifically focusing on the tail-less histone HHoB from LC_3 Hodarchaea, provides the first structural insights into this lineage, significantly expanding our comprehension beyond the more extensively studied Euryarchaeota. Given that Asgard archaea, particularly Hodarchaea, represent the closest known archaeal relatives to Eukaryotes, the organization of their genome is critically important for understanding the evolutionary transition towards eukaryotes[35]. The HHoB histone was demonstrated to form nucleosomes in two structurally distinct conformations: closed and open. All previously reported archaeal nucleosome structures, including those formed by the euryarchaeal histones HMfB and HTkA—have been observed solely in the closed conformation. This conservation is striking: the key histone dimer-dimer and stacking interfaces observed in the L3 HHoB closed state are consistent with those found in the classical and ”slinky” arrangements of Euryarchaeal nucleosomes[41, 17]. The identical pitch of DNA wrapping between HHoB, euryarchaeal and eukaryotic nucleosomes in the closed conformation strongly suggests that this structural configuration is conserved[1, 41, 17]. Furthermore, these experimental findings provide validation for earlier computational predictions that HHoB is capable of forming closed hypernucleosomes, which necessarily engage these stacking interactions. Crucially, the discovery of the open Asgard HHoB nucleosome conformation is entirely novel and was not anticipated by previous structural or predictive studies[21].

## 2 Open hypernucleosomes are asgard specific

The newly characterized HHoB open state represents a major advancement in our understanding of archaeal chromatin structural diversity. This configuration is structurally distinct, featuring a unique dimer-dimer interface that differs from the closed state and all other previously reported archaeal structures. Most notably, the open state is characterized by a complete absence of histone stacking. This open structure is stabilized by key internal interactions involving Tyr44-Glu47 and Tyr44-His51. These three residues mediating these interactions are conserved across a subset of LC_3 histones, including HHoB, HHoJ, HHoE, HHoH. From combining recent datasets and analysing the prevalence of tyrosine at the residue corresponding to 44 in HHoB, it becomes apparent that the archaeal histones have a low frequency (3.5%), is enriched in asgard histones (11.5%) and substantially enriched in the Hodarchaeal lineage closest to eukaryotes (44.7%). This group includes Heimdall, Hod, Njord, Gerd, Kari, Wukong and Hermodarchaeal histones [35].

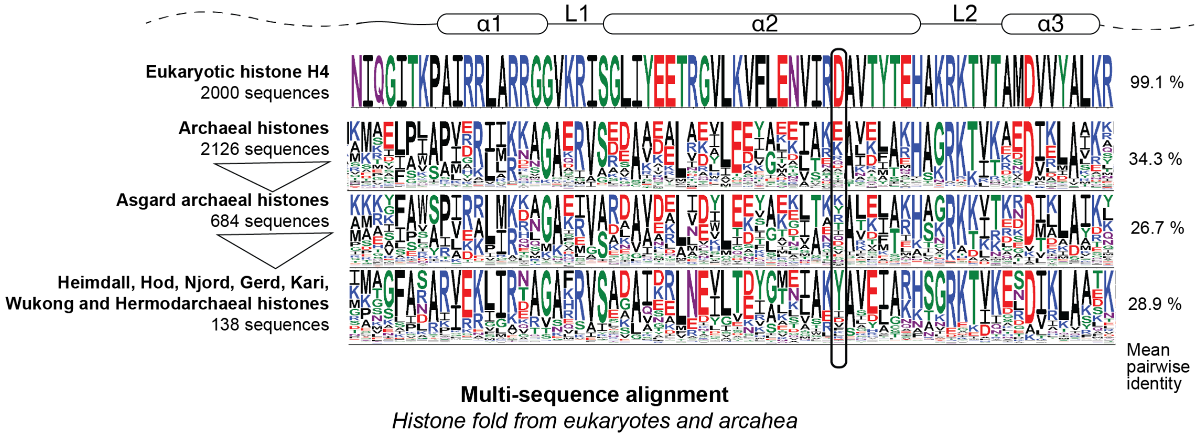


Figure 1: Histone dimer-dimer interface in the HHoB nucleosome open (purple) and closed (orange) conformations. (A) Surface representations of the nucleosomes showing the histone dimer-dimer interfaces in red. (B,C) Ribbon representations of histones showing key interactions involved at the dimer-dimer interface: Y44, E47, I48, H51, R54, D61, K68. Dashed lines indicate electrostatic and hydrogen bond interactions.

The near-total absence of Tyr44 in Euryarchaeal histones (0.1%) strongly suggests that the capacity to form this open nucleosome state could be a unique and potentially widespread innovation within the Asgard archaea, while it should be largely absent in Euryarchaeota. The open Asgard nucleosome exhibits a striking structural resemblance to the putative ancestral eukaryotic H3–H4 octasome assembly, which is likewise capable of adopting both open and closed conformations[28]. Using the HHoB mut1 variant, bearing a tyrosine residue at the position corresponding to histone H4, I observed an identical degree of DNA wrapping to that of the H3–H4 octasome in its open conformation (Figure 14). This structural congruence strongly suggests that the open state may represent an evolutionary intermediate in the transition toward the canonical eukaryotic nucleosome. Notably, current predictive algorithms such as AlphaFold3 display a pronounced bias toward the closed, one-turn nucleosome configuration, even for HHoB, and fail to reproduce the experimentally observed open or closed hypernucleosome conformations under equivalent conditions.

# 3 Implications of open hypernucleosomes

The open hypernucleosome present stable yet dynamic assemblies, showing breathing in MD simulations (subsection 4.4), and are stabilised by only the dimer-dimer interface, with the absence of stacking interactions like in the closed conformation(subsection 2.3). The absence of the stacking interface in the open assembly would make the energy barrier for disruption lower than closed hypernucleosome states. Indeed, in my analysis of the force spectroscopy data, the stiffness of HHoB-DNA complexes in the presence of $Mg^{2+}$ was substantially higher than in the open nucleosome condition without $Mg^{2+}$ (Figure 23). This implies the possibility of histone exchanges with higher effency in open conformations, which could be particularly important for Asgards since they typically encode a larger repertoire of histone variants compared to the one or two variants in many Euryarchaea [19]. The open chromatin state may also minimize steric clashes associated with extended structural elements, such as the N- or C-terminal α-helices or tails that are more common in Asgard histones. This structural tolerance could have facilitated the ”evolutionary exploration” of histone extensions in this lineage [22]. The large solvent accessible area in open conformation might be access to crucial for chromatin modification factors, such as the predicted ”readers” and ”writers” in Asgard archaea, and could also facilitate the passage of DNA-machinery by reducing the number of necessary contacts to be broken compared to the closed state.

The divalent ion driven dual-state system (open and closed) suggests a complex regulatory interplay, where closed hypernucleosomes provide maximal stability and compaction, potentially limiting access, while the open state offers enhanced accessibility and dynamics. This balance of stability and accessibility is particularly vital for thermophilic and hyperthermophilic Asgard species the suggested phenotype of the last common ancestor of this group where DNA must be thermally stabilized by histone association while maintaining functional accessibility.

# 4 Role of $Mg^{2+}$ ions in regulation of conformations

My investigation into the effect of $Mg^{2+}$ concentrations reveal a probable mechanism for regulating the open-closed structural duality. While I determined structures of the HHoB nucleosomes at 1 mM, 20 mM and 100 mM $Mg^{2+}$ for enrichment of respective conformations, both closed and open hypernucleosomes were present across a wide range of $Mg^{2+}$ concentrations (1–60 mM). This suggests an $Mg^{2+}$ dependent regulation analogous to a ”salt-in” mechanism often utilized by archaea to adjust intracellular ion levels [111]. Archaeal cell have shown high cellular amounts of $Mg^{2+}$ (120mM in *T. kodakarensis*) while eukaryotes have a lower abundance (1-5 mM)[107, 106, 112]. It must be noted here that the lack of knowledge in cellular biochemsitry of Asgard cells limits our interpretive liberty.

I propose that $Mg^{2+}$ regulates the chromatin state by binding to the negatively charged DNA phosphate backbone, effectively shielding these charges. At elevated $Mg^{2+}$ levels, this charge shielding allows the DNA gyres within the hypernucleosome to draw closer together, which in turn promotes the closed state by stabilizing the protein-protein interactions at both the histone dimer-dimer and stacking interfaces. This is further highlighted by MD simulations with a high proportion of $Mg^{2+}$ ions relocating to close proximity ($< 5$ Å) of phosphate groups of the DNA as the open hypernucleosome transitioned to the closed state(subsection 4.4). Furthermore, Electrophoretic Mobility Shift Assays (EMSA) demonstrated that the cooperativity of HHoB-DNA binding is generally lower than that of Euryarchaeal histone HMfB in the absence of $Mg^{2+}$. Critically, the cooperativity of HHoB binding increases with higher $Mg^{2+}$ concentrations, showing a direct correlation with the increased formation of closed hypernucleosomes(subsection 2.1). This implies that $Mg^{2+}$ plays a key role in regulating chromatin states through its influence on histone-DNA binding cooperativity, which is defined by the synergy of three factors: $Mg^{2+}$ shielding of DNA charges, and the formation of the dimer-dimer and stacking interfaces.

# 5 Emerging model of Asgard chromatin

The findings presented in this study provide an expanded understanding of the “variable beads on a string” model previously proposed for archaeal chromatin. According to this model, closed hypernucleosomes of variable length are formed by the association of a variable number (N) of histone dimers wrapping the DNA[17]. In the LC_3 metagenome, which encodes ten distinct histone variants, my work predicts that hypernucleosomes are not only variable in length but may also adopt both open and closed conformations locally, depending on the specific combination of histone types incorporated and the microenvironment of the complexes.

Histone variants with a lower intrinsic propensity for hypernucleosome formation, as well as other nucleoid-associated proteins, are likely to function as capstones and roadblocks, respectively, to regulate the size, stiffness, and overall DNA accessibility of these hypernucleosome domains[19]. Importantly, my structural analyses suggest that the presence of a tyrosine residue at the dimer-dimer interface introduces a layer of conformational plasticity: histones previously designated as “partial capstones” (lacking a stacking interface but retaining a dimer-dimer interface) may form closed fibres in the presence of appropriate microenvironments (like high concentration of divalent ions), while in their absence, these histones could assemble into open hypernucleosomes capable of polymerizing along the DNA. The prevalence of Tyr44 (position in HHoB) in Asgard histones relative to other archaea further hints that open hypernucleosomes may be a widespread and physiologically relevant structure in these organisms.

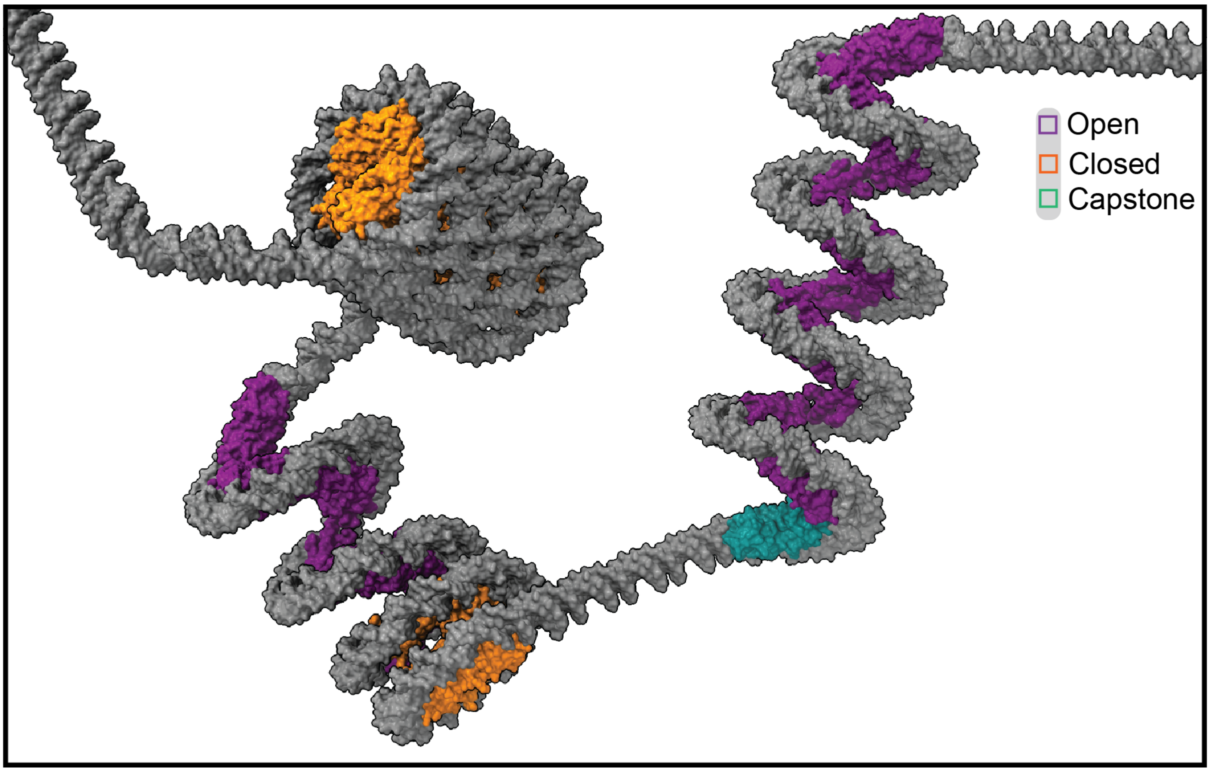


Figure 2: Proposed asgard chromatin model based on open (purple) and closed (orange) hypernucleosomes and capstone (teal) hsitones that may act as roadblocks.

Comparative sequence analysis underscores the evolutionary significance of Asgard histones. Unlike eukaryotic histones, which display extreme conservation (99.1% mean identity), Asgard histone-fold domain proteins exhibit remarkable sequence variability (26.7% mean identity) (Figure 1). The presence of multiple histone paralogs -including histone-fold fusions, histones with variable tail lengths, and a lack of obligate heterodimer formation - suggests that the Asgard lineage may be in an evolutionary transitional state, preceding the definitive histone complement established in the last eukaryotic common ancestor[22]. This diversity provides both challenges and opportunities for understanding chromatin organization in these organisms.
While this study focuses on in vitro analyses of HHoB homodimers, several open questions remain. The roles of heterodimers, both structurally and as potential capstones or fibre interruptors, are yet to be elucidated. Similarly, the contribution of histone tails to hypernucleosome stability, fiber compaction, and regulation by lysine and arginine charges remains unexplored. Cellular and environmental contexts will be critical to addressing these questions. For instance, biochemical studies in Asgard cells could reveal physiologically relevant ionic concentrations, differential expression patterns of histone variants, and potential obligatory heterodimer formations. Environmental factors, including temperature, likely influence hypernucleosome stability; one can hypothesize that open hypernucleosomes are less thermostable than their closed counterparts, highlighting the importance of studying chromatin in thermophilic Asgards. Furthermore, the interplay between histones and other nucleoid-associated proteins is likely crucial in shaping local chromatin architecture and DNA accessibility.

Taken together, despite the limited knowledge of archaeal and Asgard histone-DNA complexes, the model proposed here provides a framework for understanding Asgard chromatin. It emphasizes variable hypernucleosome length, conformational heterogeneity, and the regulatory potential of histone variants and accessory proteins. Moving forward, a combination of structural, biophysical, biochemical, and cellular studies will be essential to refine this model, elucidate the full conformational landscape of Asgard chromatin and its adaptations to the environment their organisms survive in.

# Chapter VI

# Conclusion

This thesis provides the first structural and biophysical insights into chromatin organization in Asgard archaea - organisms that represent the closest known relatives of eukaryotes. Through an integrative approach combining cryo-electron microscopy, single-molecule force spectroscopy and molecular dynamics simulations, I characterized how minimal histone HHoB from Hodarchaeon LC_3 assembles with DNA to form nucleosome and hypernucleosome complexes that exhibit a remarkable degree of conformational flexibility.

At the core of this work lies the discovery that the LC_3 histone HHoB forms both open and closed nucleosome conformations in vitro. The open state, seemingly unique to Asgard archaea, contrasts sharply with the tightly packed, closed hypernucleosomes observed with Euryarchaeal histones such as HMfB and HTkA. High resolution cryo-EM reconstructions revealed that these two conformations are differ in the opening pitch of the DNA, and I identifiied residues relvant in the dimer-dimer and stacking interfaces differing in the two states. This conformational duality suggests that Asgard histones posses intrinsic plasticity, enabling a tunable balance between genome compaction and accessibility.

The functional relevance of this conformational equilibrium was supported by structural and force spectroscopy measurements, which demonstrated that magnesium ions act as a biophysical regulator of chromatin compaction. Increasing $Mg^{2+}$ concentration stabilizes the closed conformation and stiffens the histone - DNA fiber. Molecular dynamics simulations further substantiated these observations, revealing ion-mediated charge neutralisation of DNA leading to stabilisation of the closerd hypernucleosome. Together, these findings establish divalent cations as potential modulators of Asgard chromatin architecture, hinting at an early mechanism for chromatin regulation predating eukaryotic post-translational modification systems.

While this thesis explores homo-dimers of four histones (out of ten) in the LC_3 metagenome, the possibility of heterodimerisation adds a layer of unexplored combinatorial complexity. The lack of cellular biochemical information on asgard archaeal cells limits the relevance of minimal buffer components and the variety and concentrations of divalent ions tested. Lastly, in-situ data is needed to confirm the presence of long hypernucleosome fibres akin to those observed in the single particle samples presented in the thesis (upto 0.5 $\mu$m , corresponding to ~15kbp of DNA).

The Asgard lineage also exhibits extraordinary histone diversity, with multiple paralogs and variable tail lengths encoded within single genomes. This diversity, together with the coexistence of open and closed assemblies, suggests that Asgard chromatin remains in an evolutionary transitional phase - one that may have laid the foundation for the emergence of the canonical eukaryotic nucleosome. The structural and functional variability observed here supports a “variable beads-on-a-string” model, wherein hypernucleosomes of different lengths and conformations coexist, regulated by ion concentration, histone paralogs, and potentially nucleoid-associated proteins acting as regulators of fibre length. Such modular organization would allow local tuning of DNA accessibility, offering a plausible evolutionary precursor to the elaborate chromatin-based regulation seen in modern eukaryotes.

# Bibliography

[1] Karolin Luger et al. “Crystal structure of the nucleosome core particle at 2.8 Å resolution”. In: *Nature* 389.6648 (Sept. 1997), pp. 251–260. ISSN: 1476-4687. DOI: `10.1038/38444`. URL: `http://dx.doi.org/10.1038/38444`.

[2] Amber R. Cutter and Jeffrey J. Hayes. “A brief review of nucleosome structure”. In: *FEBS Letters* 589.20PartA (May 2015), pp. 2914–2922. ISSN: 1873-3468. DOI: `10.1016/j.febslet.2015.05.016`. URL: `http://dx.doi.org/10.1016/j.febslet.2015.05.016`.

[3] Tadasu Nozaki et al. “Dynamic Organization of Chromatin Domains Revealed by Super-Resolution Live-Cell Imaging”. In: *Molecular Cell* 67.2 (July 2017), 282–293.e7. ISSN: 1097-2765. DOI: `10.1016/j.molcel.2017.06.018`. URL: `http://dx.doi.org/10.1016/j.molcel.2017.06.018`.

[4] Cedric R. Clapier et al. “Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes”. In: *Nature Reviews Molecular Cell Biology* 18.7 (May 2017), pp. 407–422. ISSN: 1471-0080. DOI: `10.1038/nrm.2017.26`. URL: `http://dx.doi.org/10.1038/nrm.2017.26`.

[5] Tejas Yadav and Iestyn Whitehouse. “Replication-Coupled Nucleosome Assembly and Positioning by ATP-Dependent Chromatin-Remodeling Enzymes”. In: *Cell Reports* 15.4 (Apr. 2016), pp. 715–723. ISSN: 2211-1247. DOI: `10.1016/j.celrep.2016.03.059`. URL: `http://dx.doi.org/10.1016/j.celrep.2016.03.059`.

[6] Thomas O. Tolsma and Jeffrey C. Hansen. “Post-translational modifications and chromatin dynamics”. In: *Essays in Biochemistry* 63.1 (Mar. 2019). Ed. by Nick Gilbert and James Allan, pp. 89–96. ISSN: 1744-1358. DOI: `10.1042/ebc20180067`. URL: `http://dx.doi.org/10.1042/ebc20180067`.

[7] Jolien J.E. van Hooff. “Towards unraveling the origins of eukaryotic nuclear genome organization”. In: *Trends in Cell Biology* 33.10 (Oct. 2023), pp. 820–823. ISSN: 0962-8924. DOI: `10.1016/j.tcb.2023.07.008`. URL: `http://dx.doi.org/10.1016/j.tcb.2023.07.008`.

[8] Remus Thei Dame and Nora Goosen. “HU: promoting or counteracting DNA compaction?” In: *FEBS Letters* 529.2–3 (Sept. 2002), pp. 151–156. ISSN: 1873-3468. DOI: `10.1016/s0014-5793(02)03363-x`. URL: `http://dx.doi.org/10.1016/s0014-5793(02)03363-x`.

[9] Kazutoshi Kasho, Shogo Ozaki, and Tsutomu Katayama. “IHF and Fis as Escherichia coli Cell Cycle Regulators: Activation of the Replication Origin oriC and the Regulatory Cycle of the DnaA Initiator”. In: *International Journal of Molecular Sciences* 24.14 (July 2023), p. 11572. ISSN: 1422-0067. DOI: `10.3390/ijms241411572`. URL: `http://dx.doi.org/10.3390/ijms241411572`.

[10] Archit Gupta et al. “The bacterial nucleoid-associated proteins, HU and Dps, condense DNA into context-dependent biphasic or multiphasic complex coacervates”. In: *Journal of Biological Chemistry* 299.5 (May 2023), p. 104637. ISSN: 0021-9258. DOI: `10.1016/j.jbc.2023.104637`. URL: `http://dx.doi.org/10.1016/j.jbc.2023.104637`.

[11] Sara Moutacharrif et al. “The Power Duo: How the Interplay Between Nucleoid□Associated Proteins and Small Noncoding RNAs Orchestrates the Cellular Regulatory Symphony”. In: *Molecular Microbiology* 123.6 (Apr. 2025), pp. 531–546. ISSN: 1365-2958. DOI: `10.1111/mmi.15359`. URL: `http://dx.doi.org/10.1111/mmi.15359`.

[12] Haley M. Amemiya, Jeremy Schroeder, and Lydia Freddolino. “Nucleoid-associated proteins shape chromatin structure and transcriptional regulation across the bacterial kingdom”. In: *Transcription* 12.4 (Aug. 2021), pp. 182–218. ISSN: 2154-1272. DOI: `10.1080/21541264.2021.1973865`. URL: `http://dx.doi.org/10.1080/21541264.2021.1973865`.

[13] Shawn P. Laursen, Samuel Bowerman, and Karolin Luger. “Archaea: The Final Frontier of Chromatin”. In: *Journal of Molecular Biology* 433.6 (Mar. 2021), p. 166791. ISSN: 0022-2836. DOI: `10.1016/j.jmb.2020.166791`. URL: `http://dx.doi.org/10.1016/j.jmb.2020.166791`.

[14] Zhenfeng Zhang et al. “Archaeal Chromatin Proteins Cren7 and Sul7d Compact DNA by Bending and Bridging”. In: *mBio* 11.3 (June 2020). Ed. by Christa M. Schleper. ISSN: 2150-7511. DOI: `10.1128/mbio.00804-20`. URL: `http://dx.doi.org/10.1128/mbio.00804-20`.

[15] Manish Goyal et al. “The Alba protein family: Structure and function”. In: *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1864.5 (May 2016), pp. 570–583. ISSN: 1570-9639. DOI: `10.1016/j.bbapap.2016.02.015`. URL: `http://dx.doi.org/10.1016/j.bbapap.2016.02.015`.

[16] Marc Kenneth M. Cajili and Eloise I. Prieto. “Interplay between Alba and Cren7 Regulates Chromatin Compaction in Sulfolobus solfataricus”. In: *Biomolecules* 12.4 (Mar. 2022), p. 481. ISSN: 2218-273X. DOI: `10.3390/biom12040481`. URL: `http://dx.doi.org/10.3390/biom12040481`.

[17] Francesca Mattiroli et al. “Structure of histone-based chromatin in Archaea”. In: *Science* 357.6351 (Aug. 2017), pp. 609–612. ISSN: 1095-9203. DOI: `10.1126/science.aaj1849`. URL: `http://dx.doi.org/10.1126/science.aaj1849`.

[18] Miroslav Tomschik et al. “The Archaeal Histone-Fold Protein HMf Organizes DNA into Bona Fide Chromatin Fibers”. In: *Structure* 9.12 (Dec. 2001), pp. 1201–1211. ISSN: 0969-2126. DOI: `10.1016/s0969-2126(01)00682-7`. URL: `http://dx.doi.org/10.1016/s0969-2126(01)00682-7`.

[19] Kathryn M. Stevens et al. “Histone variants in archaea and the evolution of combinatorial chromatin complexity”. In: *Proceedings of the National Academy of Sciences* 117.52 (Dec. 2020), pp. 33384–33395. ISSN: 1091-6490. DOI: `10.1073/pnas.2007056117`. URL: `http://dx.doi.org/10.1073/pnas.2007056117`.

[20] Liesbeth Lemmens et al. “DNA-Binding Properties of a Novel Crenarchaeal Chromatin-Organizing Protein in Sulfolobus acidocaldarius”. In: *Biomolecules* 12.4 (Mar. 2022), p. 524. ISSN: 2218-273X. DOI: `10.3390/biom12040524`. URL: `http://dx.doi.org/10.3390/biom12040524`.

[21] Bram Henneman et al. “Structure and function of archaeal histones”. In: *PLOS Genetics* 14.9 (Sept. 2018). Ed. by Petra Anne Levin, e1007582. ISSN: 1553-7404. DOI: `10.1371/journal.pgen.1007582`. URL: `http://dx.doi.org/10.1371/journal.pgen.1007582`.

[22] A. Hocher and T. Warnecke. “Nucleosomes at the Dawn of Eukaryotes”. In: *Genome Biol Evol* 16.3 (2024). DOI: `10.1093/gbe/evae029`. URL: `https://www.ncbi.nlm.nih.gov/pubmed/38366053`.

[23] Xavier Grau-Bové et al. “A phylogenetic and proteomic reconstruction of eukaryotic chromatin evolution”. In: *Nature Ecology amp; Evolution* 6.7 (June 2022), pp. 1007–1023. ISSN: 2397-334X. DOI: `10.1038/s41559-022-01771-6`. URL: `http://dx.doi.org/10.1038/s41559-022-01771-6`.

[24] Shuxiang Li, Tiejun Wei, and Anna R. Panchenko. “Histone variant H2A.Z modulates nucleosome dynamics to promote DNA accessibility”. In: *Nature Communications* 14.1 (Feb. 2023). ISSN: 2041-1723. DOI: `10.1038/s41467-023-36465-5`. URL: `http://dx.doi.org/10.1038/s41467-023-36465-5`.

[25] Sara Martire and Laura A. Banaszynski. “The roles of histone variants in fine-tuning chromatin organization and function”. In: *Nature Reviews Molecular Cell Biology* 21.9 (July 2020), pp. 522–541. ISSN: 1471-0080. DOI: `10.1038/s41580-020-0262-8`. URL: `http://dx.doi.org/10.1038/s41580-020-0262-8`.

[26] Yasuhiro Arimura et al. “Structural Analysis of the Hexasome, Lacking One Histone H2A/H2B Dimer from the Conventional Nucleosome”. In: *Biochemistry* 51.15 (Apr. 2012), pp. 3302–3309. ISSN: 1520-4995. DOI: `10.1021/bi300129b`. URL: `http://dx.doi.org/10.1021/bi300129b`.

[27] Georgy N. Rychkov et al. “Partially Assembled Nucleosome Structures at Atomic Detail”. In: *Biophysical Journal* 112.3 (Feb. 2017), pp. 460–472. ISSN: 0006-3495. DOI: `10.1016/j.bpj.2016.10.041`. URL: `http://dx.doi.org/10.1016/j.bpj.2016.10.041`.

[28] Kayo Nozawa et al. “Cryo–electron microscopy structure of the H3-H4 octasome: A nucleosome-like particle without histones H2A and H2B”. In: *Proceedings of the National Academy of Sciences* 119.45 (Nov. 2022). ISSN: 1091-6490. DOI: `10.1073/pnas.2206542119`. URL: `http://dx.doi.org/10.1073/pnas.2206542119`.

[29] Aghil Soman, Nikolay Korolev, and Lars Nordenskiöld. “Telomeric chromatin structure”. In: *Current Opinion in Structural Biology* 77 (Dec. 2022), p. 102492. ISSN: 0959-440X. DOI: `10.1016/j.sbi.2022.102492`. URL: `http://dx.doi.org/10.1016/j.sbi.2022.102492`.

[30] Aghil Soman et al. “Columnar structure of human telomeric chromatin”. In: *Nature* 609.7929 (Sept. 2022), pp. 1048–1055. ISSN: 1476-4687. DOI: `10.1038/s41586-022-05236-5`. URL: `http://dx.doi.org/10.1038/s41586-022-05236-5`.

[31] Kathleen Sandman and John N Reeve. “Archaeal histones and the origin of the histone fold”. In: *Current Opinion in Microbiology* 9.5 (Oct. 2006), pp. 520–525. ISSN: 1369-5274. DOI: `10.1016/j.mib.2006.08.003`. URL: `http://dx.doi.org/10.1016/j.mib.2006.08.003`.

[32] Nicholas A. T. Irwin and Thomas A. Richards. “Self-assembling viral histones are evolutionary intermediates between archaeal and eukaryotic nucleosomes”. In: *Nature Microbiology* 9.7 (May 2024), pp. 1713–1724. ISSN: 2058-5276. DOI: `10.1038/s41564-024-01707-9`. URL: `http://dx.doi.org/10.1038/s41564-024-01707-9`.

[33] Albert J. Erives. “Phylogenetic analysis of the core histone doublet and DNA topo II genes of Marseilleviridae: evidence of proto-eukaryotic provenance”. In: *Epigenetics amp; Chromatin* 10.1 (Nov. 2017). ISSN: 1756-8935. DOI: `10.1186/s13072-017-0162-0`. URL: `http://dx.doi.org/10.1186/s13072-017-0162-0`.

[34] Haiqing Zhao et al. “The role of cryptic ancestral symmetry in histone folding mechanisms across Eukarya and Archaea”. In: *PLOS Computational Biology* 20.1 (Jan. 2024). Ed. by Changbong Hyeon, e1011721. ISSN: 1553-7358. DOI: `10.1371/journal.pcbi.1011721`. URL: `http://dx.doi.org/10.1371/journal.pcbi.1011721`.

[35] Laura Eme et al. “Inference and reconstruction of the heimdallarchaeial ancestry of eukaryotes”. In: *Nature* 618.7967 (June 2023), pp. 992–999. ISSN: 1476-4687. DOI: `10.1038/s41586-023-06186-2`. URL: `http://dx.doi.org/10.1038/s41586-023-06186-2`.

[36] Dmitry V. Fyodorov et al. “Emerging roles of linker histones in regulating chromatin structure and function”. In: *Nature Reviews Molecular Cell Biology* 19.3 (Oct. 2017), pp. 192–206. ISSN: 1471-0080. DOI: `10.1038/nrm.2017.94`. URL: `http://dx.doi.org/10.1038/nrm.2017.94`.

[37] Kinga Rutowicz et al. “Linker histones are fine-scale chromatin architects modulating developmental decisions in Arabidopsis”. In: *Genome Biology* 20.1 (Aug. 2019). ISSN: 1474-760X. DOI: `10.1186/s13059-019-1767-3`. URL: `http://dx.doi.org/10.1186/s13059-019-1767-3`.

[38] HAROLD E. KASINSKY et al. “Origin of H1 linker histones”. In: *The FASEB Journal* 15.1 (Jan. 2001), pp. 34–42. ISSN: 1530-6860. DOI: `10.1096/fj.00-0237rev`. URL: `http://dx.doi.org/10.1096/fj.00-0237rev`.

[39] Bram Henneman et al. “Mechanical and structural properties of archaeal hypernucleosomes”. In: *Nucleic Acids Research* 49.8 (Dec. 2020), pp. 4338–4349. ISSN: 1362-4962. DOI: `10.1093/nar/gkaa1196`. URL: `http://dx.doi.org/10.1093/nar/gkaa1196`.

[40] Amanda M. Erkelens et al. “Specific DNA binding of archaeal histones HMfA and HMfB”. In: *Frontiers in Microbiology* 14 (Apr. 2023). ISSN: 1664-302X. DOI: `10.3389/fmicb.2023.1166608`. URL: `http://dx.doi.org/10.3389/fmicb.2023.1166608`.

[41] Samuel Bowerman, Jeff Wereszczynski, and Karolin Luger. “Archaeal chromatin ‘slinkies’ are inherently dynamic complexes with deflected DNA wrapping pathways”. In: *eLife* 10 (Mar. 2021). ISSN: 2050-084X. DOI: `10.7554/elife.65587`. URL: `http://dx.doi.org/10.7554/eLife.65587`.

[42] Narasimharao Nalabothula et al. “Archaeal nucleosome positioning in vivo and in vitro is directed by primary sequence motifs”. In: *BMC Genomics* 14.1 (June 2013). ISSN: 1471-2164. DOI: `10.1186/1471-2164-14-391`. URL: `http://dx.doi.org/10.1186/1471-2164-14-391`.

[43] K.M. Stevens, A. Hocher, and T. Warnecke. “Deep conservation of histone variants in Thermococcales archaea”. In: *bioRxiv* (2021). DOI: `doi.org/10.1101/2021.09.07.455978`.

[44] Shawn P. Laursen and Karolin Luger. “Histone diversity in the archaeal domain of life”. In: (June 2025). DOI: `10.1101/2025.06.26.661778`. URL: `http://dx.doi.org/10.1101/2025.06.26.661778`.

[45] Sapir Ofer et al. “DNA-bridging by an archaeal histone variant via a unique tetramerisation interface”. In: *Communications Biology* 6.1 (Sept. 2023). ISSN: 2399-3642. DOI: `10.1038/s42003-023-05348-2`. URL: `http://dx.doi.org/10.1038/s42003-023-05348-2`.

[46] Samuel Schwab et al. “Histones and histone variant families in prokaryotes”. In: *Nature Communications* 15.1 (Sept. 2024). ISSN: 2041-1723. DOI: `10.1038/s41467-024-52337-y`. URL: `http://dx.doi.org/10.1038/s41467-024-52337-y`.

[47] Anja Spang et al. “Complex archaea that bridge the gap between prokaryotes and eukaryotes”. In: *Nature* 521.7551 (May 2015), pp. 173–179. ISSN: 1476-4687. DOI: `10.1038/nature14447`. URL: `http://dx.doi.org/10.1038/nature14447`.

[48] Katarzyna Zaremba-Niedzwiedzka et al. “Asgard archaea illuminate the origin of eukaryotic cellular complexity”. In: *Nature* 541.7637 (Jan. 2017), pp. 353–358. ISSN: 1476-4687. DOI: `10.1038/nature21031`. URL: `http://dx.doi.org/10.1038/nature21031`.

[49] Hiroyuki Imachi et al. “Isolation of an archaeon at the prokaryote–eukaryote interface”. In: *Nature* 577.7791 (Jan. 2020), pp. 519–525. ISSN: 1476-4687. DOI: `10.1038/s41586-019-1916-6`. URL: `http://dx.doi.org/10.1038/s41586-019-1916-6`.

[50] Victor Tobiasson et al. “Dominant contribution of Asgard archaea to eukaryogenesis”. In: (Oct. 2024). DOI: `10.1101/2024.10.14.618318`. URL: `http://dx.doi.org/10.1101/2024.10.14.618318`.

[51] Thiago Rodrigues-Oliveira et al. “Actin cytoskeleton and complex cell architecture in an Asgard archaeon”. In: *Nature* 613.7943 (Dec. 2022), pp. 332–339. ISSN: 1476-4687. DOI: `10.1038/s41586-022-05550-y`. URL: `http://dx.doi.org/10.1038/s41586-022-05550-y`.

[52] Muhammad Rafiq et al. “Challenges and Approaches of Culturing the Unculturable Archaea”. In: *Biology* 12.12 (Dec. 2023), p. 1499. ISSN: 2079-7737. DOI: `10.3390/biology12121499`. URL: `http://dx.doi.org/10.3390/biology12121499`.

[53] Zhongyi Lu et al. “Evolution of optimal growth temperature in Asgard archaea inferred from the temperature dependence of GDP binding to EF-1A”. In: *Nature Communications* 15.1 (Jan. 2024). ISSN: 2041-1723. DOI: `10.1038/s41467-024-44806-1`. URL: `http://dx.doi.org/10.1038/s41467-024-44806-1`.

[54] Wen-Cong Huang et al. “Phylogenomic Analyses Reveal that Panguiarchaeum Is a Clade of Genome-Reduced Asgard Archaea Within the Njordarchaeia”. In: *Molecular Biology and Evolution* 42.9 (Aug. 2025). Ed. by Purificacion Lopez-Garcia. ISSN: 1537-1719. DOI: `10.1093/molbev/msaf201`. URL: `http://dx.doi.org/10.1093/molbev/msaf201`.

[55] Yifan Cheng et al. “A Primer to Single-Particle Cryo-Electron Microscopy”. In: *Cell* 161.3 (Apr. 2015), pp. 438–449. ISSN: 0092-8674. DOI: `10.1016/j.cell.2015.03.050`. URL: `http://dx.doi.org/10.1016/j.cell.2015.03.050`.

[56] Dari Kimanius et al. “Data-driven regularization lowers the size barrier of cryo-EM structure determination”. In: *Nature Methods* 21.7 (June 2024), pp. 1216–1221. ISSN: 1548-7105. DOI: `10.1038/s41592-024-02304-8`. URL: `http://dx.doi.org/10.1038/s41592-024-02304-8`.

[57] Ka Man Yip et al. “Atomic-resolution protein structure determination by cryo-EM”. In: *Nature* 587.7832 (Oct. 2020), pp. 157–161. ISSN: 1476-4687. DOI: `10.1038/s41586-020-2833-4`. URL: `http://dx.doi.org/10.1038/s41586-020-2833-4`.

[58] Takanori Nakane et al. “Single-particle cryo-EM at atomic resolution”. In: *Nature* 587.7832 (Oct. 2020), pp. 152–156. ISSN: 1476-4687. DOI: `10.1038/s41586-020-2829-0`. URL: `http://dx.doi.org/10.1038/s41586-020-2829-0`.

[59] Lindsey N. Young and Elizabeth Villa. “Bringing Structure to Cell Biology with Cryo-Electron Tomography”. In: *Annual Review of Biophysics* 52.1 (May 2023), pp. 573–595. ISSN: 1936-1238. DOI: `10.1146/annurev-biophys-111622-091327`. URL: `http://dx.doi.org/10.1146/annurev-biophys-111622-091327`.

[60] Eva Nogales and Julia Mahamid. “Bridging structural and cell biology with cryo-electron microscopy”. In: *Nature* 628.8006 (Apr. 2024), pp. 47–56. ISSN: 1476-4687. DOI: `10.1038/s41586-024-07198-2`. URL: `http://dx.doi.org/10.1038/s41586-024-07198-2`.

[61] J. Dubochet and A.W. McDowall. “VITRIFICATION OF PURE WATER FOR ELECTRON MICROSCOPY”. In: *Journal of Microscopy* 124.3 (Dec. 1981), pp. 3–4. ISSN: 1365-2818. DOI: `10.1111/j.1365-2818.1981.tb02483.x`. URL: `http://dx.doi.org/10.1111/j.1365-2818.1981.tb02483.x`.

[62] Robert M. Glaeser and Bong-Gyoon Han. “Opinion: hazards faced by macromolecules when confined to thin aqueous films”. In: *Biophysics Reports* 3.1–3 (July 2016), pp. 1–7. ISSN: 2364-3420. DOI: `10.1007/s41048-016-0026-3`. URL: `http://dx.doi.org/10.1007/s41048-016-0026-3`.

[63] Daniel Cressey and Ewen Callaway. “Cryo-electron microscopy wins chemistry Nobel”. In: *Nature* 550.7675 (2017), pp. 167–167. ISSN: 1476-4687. DOI: `10.1038/nature.2017.22738`.

[64] Yixin Xu and Shangyu Dang. “Recent Technical Advances in Sample Preparation for Single-Particle Cryo-EM”. In: *Frontiers in Molecular Biosciences* 9 (June 2022). ISSN: 2296-889X. DOI: `10.3389/fmolb.2022.892459`. URL: `http://dx.doi.org/10.3389/fmolb.2022.892459`.

[65] Radosav S. Pantelic et al. “Oxidative Doping Renders Graphene Hydrophilic, Facilitating Its Use As a Support in Biological TEM”. In: *Nano Letters* 11.10 (Sept. 2011), pp. 4319–4323. ISSN: 1530-6992. DOI: `10.1021/nl202386p`. URL: `http://dx.doi.org/10.1021/nl202386p`.

[66] Ying Liu, Xing Meng, and Zheng Liu. “Deformed grids for single-particle cryo-electron microscopy of specimens exhibiting a preferred orientation”. In: *Journal of Structural Biology* 182.3 (June 2013), pp. 255–258. ISSN: 1047-8477. DOI: `10.1016/j.jsb.2013.03.005`. URL: `http://dx.doi.org/10.1016/j.jsb.2013.03.005`.

[67] Edoardo D’Imprima et al. “Protein denaturation at the air-water interface and how to prevent it”. In: *eLife* 8 (Apr. 2019). ISSN: 2050-084X. DOI: `10.7554/elife.42747`. URL: `http://dx.doi.org/10.7554/eLife.42747`.

[68] Ali Punjani et al. “cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination”. In: *Nature Methods* 14.3 (Feb. 2017), pp. 290–296. ISSN: 1548-7105. DOI: `10.1038/nmeth.4169`. URL: `http://dx.doi.org/10.1038/nmeth.4169`.

[69] Timothy Grant, Alexis Rohou, and Nikolaus Grigorieff. “cisTEM, user-friendly software for single-particle image processing”. In: *eLife* 7 (Mar. 2018). ISSN: 2050-084X. DOI: `10.7554/elife.35383`. URL: `http://dx.doi.org/10.7554/eLife.35383`.

[70] Dari Kimanius et al. “New tools for automated cryo-EM single-particle analysis in RELION-4.0”. In: *Biochemical Journal* 478.24 (Dec. 2021), pp. 4169–4185. ISSN: 1470-8728. DOI: `10.1042/bcj20210708`. URL: `http://dx.doi.org/10.1042/BCJ20210708`.

[71] Helen R. Saibil. “Cryo-EM in molecular and cellular biology”. In: *Molecular Cell* 82.2 (Jan. 2022), pp. 274–284. ISSN: 1097-2765. DOI: `10.1016/j.molcel.2021.12.016`. URL: `http://dx.doi.org/10.1016/j.molcel.2021.12.016`.

[72] Jose Luis Vilas, Jose Maria Carazo, and Carlos Oscar S. Sorzano. “Emerging Themes in CryoEM—Single Particle Analysis Image Processing”. In: *Chemical Reviews* 122.17 (July 2022), pp. 13915–13951. ISSN: 1520-6890. DOI: `10.1021/acs.chemrev.1c00850`. URL: `http://dx.doi.org/10.1021/acs.chemrev.1c00850`.

[73] Tristan Bepler et al. “Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs”. In: *Nature Methods* 16.11 (Oct. 2019), pp. 1153–1160. ISSN: 1548-7105. DOI: `10.1038/s41592-019-0575-8`. URL: `http://dx.doi.org/10.1038/s41592-019-0575-8`.

[74] Microscopy Australia. *Introducing Single Particle Analysis*. `https://myscope.training/CRYO_Introducing_Single_Particle_Analysis`. MyScope training module. Licensed under CC BY-SA 4.0. 2025.

[75] Ali Punjani and David J. Fleet. “3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM”. In: *Journal of Structural Biology* 213.2 (June 2021), p. 107702. ISSN: 1047-8477. DOI: `10.1016/j.jsb.2021.107702`. URL: `http://dx.doi.org/10.1016/j.jsb.2021.107702`.

[76] Ellen D. Zhong et al. “CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks”. In: *Nature Methods* 18.2 (2021), pp. 176–185. DOI: `10.1038/s41592-020-01049-4`. URL: `https://doi.org/10.1038/s41592-020-01049-4`.

[77] John Jumper and Demis Hassabis. “Protein structure predictions to atomic accuracy with AlphaFold”. In: *Nature Methods* 19.1 (Jan. 2022), pp. 11–12. ISSN: 1548-7105. DOI: `10.1038/s41592-021-01362-6`. URL: `http://dx.doi.org/10.1038/s41592-021-01362-6`.

[78] Josh Abramson et al. “Accurate structure prediction of biomolecular interactions with AlphaFold 3”. In: *Nature* 630.8016 (May 2024), pp. 493–500. ISSN: 1476-4687. DOI: `10.1038/s41586-024-07487-w`. URL: `http://dx.doi.org/10.1038/s41586-024-07487-w`.

[79] Kiarash Jamali et al. “Automated model building and protein identification in cryo-EM maps”. en. In: *Nature* 628.8007 (Apr. 2024), pp. 450–457.

[80] Paul Emsley and Kevin Cowtan. “Coot: model-building tools for molecular graphics”. In: *Acta Crystallographica Section D Biological Crystallography* 60.12 (Nov. 2004), pp. 2126–2132. ISSN: 0907-4449. DOI: `10.1107/s0907444904019158`. URL: `http://dx.doi.org/10.1107/S0907444904019158`.

[81] Ana Casañal, Bernhard Lohkamp, and Paul Emsley. “Current developments in Coot for macromolecular model building of Electron Cryo□microscopy and Crystallographic Data”. In: *Protein Science* 29.4 (Mar. 2020), pp. 1055–1064. ISSN: 1469-896X. DOI: `10.1002/pro.3791`. URL: `http://dx.doi.org/10.1002/pro.3791`.

[82] Tristan Ian Croll. “ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps”. In: *Acta Crystallographica Section D Structural Biology* 74.6 (Apr. 2018), pp. 519–530. ISSN: 2059-7983. DOI: `10.1107/s2059798318002425`. URL: `http://dx.doi.org/10.1107/S2059798318002425`.

[83] Elaine C. Meng et al. “<scp>UCSF ChimeraX</scp>: Tools for structure building and analysis”. In: *Protein Science* 32.11 (Oct. 2023). ISSN: 1469-896X. DOI: `10.1002/pro.4792`. URL: `http://dx.doi.org/10.1002/pro.4792`.

[84] Pavel V. Afonine et al. “Real-space refinement in PHENIX for cryo-EM and crystallography”. In: *Acta Crystallographica Section D Structural Biology* 74.6 (May 2018), pp. 531–544. ISSN: 2059-7983. DOI: `10.1107/s2059798318006551`. URL: `http://dx.doi.org/10.1107/S2059798318006551`.

[85] Vincent B Chen et al. "MolProbity: all-atom structure validation for macromolecular crystallography". en. In: *Acta Crystallogr. D Biol. Crystallogr.* 66.Pt 1 (Jan. 2010), pp. 12–21.

[86] Jan Huertas and Vlad Cojocaru. "Breaths, twists, and turns of atomistic nucleosomes". en. In: *J. Mol. Biol.* 433.6 (Mar. 2021), p. 166744.

[87] Rutika Patel et al. "Conformational dynamics of the nucleosomal histone H2B tails revealed by molecular dynamics simulations". en. In: *J. Chem. Inf. Model.* 64.12 (June 2024), pp. 4709–4726.

[88] Giovanni B Brandani et al. "Molecular dynamics simulations for the study of chromatin biology". en. In: *Curr. Opin. Struct. Biol.* 77.102485 (Dec. 2022), p. 102485.

[89] He Meng, Kurt Andresen, and John van Noort. "Quantitative analysis of single-molecule force spectroscopy on folded chromatin fibers". In: *Nucleic Acids Research* 43.7 (Mar. 2015), pp. 3578–3590. ISSN: 0305-1048. DOI: `10.1093/nar/gkv215`. URL: `http://dx.doi.org/10.1093/nar/gkv215`.

[90] César Díaz-Celis et al. "Assignment of structural transitions during mechanical unwrapping of nucleosomes and their disassembly products". In: *Proceedings of the National Academy of Sciences* 119.33 (Aug. 2022). ISSN: 1091-6490. DOI: `10.1073/pnas.2206513119`. URL: `http://dx.doi.org/10.1073/pnas.2206513119`.

[91] Artur Kaczmarczyk et al. "Single-molecule force spectroscopy on histone H4 tail-cross-linked chromatin reveals fiber folding". In: *Journal of Biological Chemistry* 292.42 (Oct. 2017), pp. 17506–17513. ISSN: 0021-9258. DOI: `10.1074/jbc.m117.791830`. URL: `http://dx.doi.org/10.1074/jbc.M117.791830`.

[92] Ilias Zarguit et al. "MODULATION OF ARCHAEAL HYPERNUCLEOSOME STRUCTURE AND STABILITY BY Mg2+". In: (May 2025). DOI: `10.1101/2025.05.26.656202`. URL: `http://dx.doi.org/10.1101/2025.05.26.656202`.

[93] David N. Mastronarde. "Automated electron microscope tomography using robust prediction of specimen movements". In: *Journal of Structural Biology* 152.1 (Oct. 2005), pp. 36–51. ISSN: 1047-8477. DOI: `10.1016/j.jsb.2005.07.007`. URL: `http://dx.doi.org/10.1016/j.jsb.2005.07.007`.

[94] Thomas D. Goddard et al. "UCSF ChimeraX: Meeting modern challenges in visualization and analysis". In: *Protein Science* 27.1 (Sept. 2017), pp. 14–25. ISSN: 1469-896X. DOI: `10.1002/pro.3235`. URL: `http://dx.doi.org/10.1002/pro.3235`.

[95] Johannes Schindelin et al. "Fiji: an open-source platform for biological-image analysis". In: *Nature Methods* 9.7 (June 2012), pp. 676–682. ISSN: 1548-7105. DOI: `10.1038/nmeth.2019`. URL: `http://dx.doi.org/10.1038/nmeth.2019`.

[96] Andrew M. Waterhouse et al. “Jalview Version 2—a multiple sequence alignment editor and analysis workbench”. In: *Bioinformatics* 25.9 (Jan. 2009), pp. 1189–1191. ISSN: 1367-4803. DOI: `10.1093/bioinformatics/btp033`. URL: `http://dx.doi.org/10.1093/bioinformatics/btp033`.

[97] K. Katoh and D. M. Standley. “MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability”. In: *Molecular Biology and Evolution* 30.4 (Jan. 2013), pp. 772–780. ISSN: 1537-1719. DOI: `10.1093/molbev/mst010`. URL: `http://dx.doi.org/10.1093/molbev/mst010`.

[98] Gavin E. Crooks et al. “WebLogo: A Sequence Logo Generator: Figure 1”. In: *Genome Research* 14.6 (June 2004), pp. 1188–1190. ISSN: 1088-9051. DOI: `10.1101/gr.849004`. URL: `http://dx.doi.org/10.1101/gr.849004`.

[99] Consensus Inc. *Consensus: AI-Powered Academic Search Engine*. [software]Version used: 2.0. 2025. URL: `https://consensus.app/`.

[100] OpenAI. *ChatGPT (Large language model)*. Version used: 5. 2025. URL: `https://chat.openai.com/`.

[101] P.T Lowary and J Widom. “New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning”. In: *Journal of Molecular Biology* 276.1 (Feb. 1998), pp. 19–42. ISSN: 0022-2836. DOI: `10.1006/jmbi.1997.1494`. URL: `http://dx.doi.org/10.1006/jmbi.1997.1494`.

[102] Svetlana O. Dodonova et al. “Nucleosome-bound SOX2 and SOX11 structures elucidate pioneer factor function”. In: *Nature* 580.7805 (Apr. 2020), pp. 669–672. ISSN: 1476-4687. DOI: `10.1038/s41586-020-2195-y`. URL: `http://dx.doi.org/10.1038/s41586-020-2195-y`.

[103] Silvija Bilokapic, Mike Strauss, and Mario Halic. “Histone octamer rearranges to adapt to DNA unwrapping”. In: *Nature Structural amp; Molecular Biology* 25.1 (Dec. 2017), pp. 101–108. ISSN: 1545-9985. DOI: `10.1038/s41594-017-0005-5`. URL: `http://dx.doi.org/10.1038/s41594-017-0005-5`.

[104] K. A. Bailey, C. S. Chow, and J. N. Reeve. “Histone stoichiometry and DNA circularization in archaeal nucleosomes”. In: *Nucleic Acids Res* 27.2 (1999), pp. 532–6. DOI: `10.1093/nar/27.2.532`. URL: `https://www.ncbi.nlm.nih.gov/pubmed/9862976`.

[105] Harsh M. Ranawat et al. “Cryo-EM reveals open and closed Asgard chromatin assemblies”. In: *Molecular Cell* (Oct. 2025). ISSN: 1097-2765. DOI: `10.1016/j.molcel.2025.10.001`. URL: `http://dx.doi.org/10.1016/j.molcel.2025.10.001`.

[106] Andrea Romani. “Regulation of magnesium homeostasis and transport in mammalian cells”. In: *Archives of Biochemistry and Biophysics* 458.1 (Feb. 2007), pp. 90–102. ISSN: 0003-9861. DOI: `10.1016/j.abb.2006.07.012`. URL: `http://dx.doi.org/10.1016/j.abb.2006.07.012`.

[107] Mariko Nagata et al. “The Cdc45/RecJ-like protein forms a complex with GINS and MCM, and is important for DNA replication in Thermococcus kodakarensis”. In: *Nucleic Acids Research* 45.18 (Aug. 2017), pp. 10693–10705. ISSN: 1362-4962. DOI: `10.1093/nar/gkx740`. URL: `http://dx.doi.org/10.1093/nar/gkx740`.

[108] Ramon A van der Valk et al. “Mechanism of environmentally driven conformational changes that modulate H-NS DNA-bridging activity”. In: *eLife* 6 (Sept. 2017). ISSN: 2050-084X. DOI: `10.7554/elife.27369`. URL: `http://dx.doi.org/10.7554/eLife.27369`.

[109] Fatema-Zahra M. Rashid et al. “The environmentally-regulated interplay between local three-dimensional chromatin organisation and transcription of proVWX in E. coli”. In: *Nature Communications* 14.1 (Nov. 2023). ISSN: 2041-1723. DOI: `10.1038/s41467-023-43322-y`. URL: `http://dx.doi.org/10.1038/s41467-023-43322-y`.

[110] Paul D. Adams et al. “PHENIX: a comprehensive Python-based system for macromolecular structure solution”. In: *Acta Crystallographica Section D Biological Crystallography* 66.2 (Jan. 2010), pp. 213–221. ISSN: 0907-4449. DOI: `10.1107/s0907444909052925`. URL: `http://dx.doi.org/10.1107/S0907444909052925`.

[111] A. Oren. “Microbial life at high salt concentrations: phylogenetic and metabolic diversity”. In: *Saline Syst* 4 (2008), p. 2. ISSN: 1746-1448 (Electronic) 1746-1448 (Linking). DOI: `10.1186/1746-1448-4-2`. URL: `https://www.ncbi.nlm.nih.gov/pubmed/18412960`.

[112] A. Romani and A. Scarpa. “Regulation of cell magnesium”. In: *Archives of Biochemistry and Biophysics* 298.1 (Oct. 1992), pp. 1–12. ISSN: 0003-9861. DOI: `10.1016/0003-9861(92)90086-c`. URL: `http://dx.doi.org/10.1016/0003-9861(92)90086-c`.