Structural and Functional Characterization of Bacterial Histones

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Yimin Hu (胡伊旻)

aus Jiangsu, China

Tübingen

2025

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	16.07.2025
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter/-in:	Prof. Dr. Andrei N. Lupas
2. Berichterstatter/-in:	Prof. Dr. Doron Rapaport

"A turn of a page, may our quiet sorrows be greeted with laughter,
Transforming shadows into dawn's gentle light.

A turn of a page, may our tears turn to sweet,
Each drop a testament to resilience and newfound joy.

A turn of a page, frowns will smile against the hopeful light of the new year,
Embracing the promise of beginnings yet to unfold.

A turn of a page, may our hearts find peace in the symphony of change,
Harmonizing past melodies with future aspirations.

A turn of a page, let love's enduring flame guide our way,
Illuminating paths of compassion and understanding.

A turn of a page, as we write our stories anew,
May each chapter be filled with grace, courage, and dreams come true."

- Wayne Lim Junjie, 2024

Table of contents

List of publications

Publications included in this dissertation:

Hu, Y.[†], Schwab, S.[†], Deiss, S., Escudeiro, P., van Heesch, T., Joiner, J.D., Vreede, J., Hartmann, M.D., Lupas, A.N., Hernandez Alvarez, B.[*], Alva, V.[*] and Dame, R.T.[*] (2024) Bacterial histone HBb from *Bdellovibrio bacteriovorus* compacts DNA by bending. *Nucleic Acids Res*, **52**, 8193-8204.

Hu, Y., Schwab, S.[§], Qiu, K.[§], Zhang, Y., Bär, K., Reichle, H., Panzera, A., Lupas, A.N., Hartmann, M.D., Dame, R.T., Alva, V.[*] and Hernandez Alvarez, B.[*] (2025). DNA wrapping by a tetrameric bacterial histone. bioRxiv, 2025-05.

Publications not included in this dissertation:

Schwab, S., **Hu, Y.**, van Erp, B., Cajili, M.K.M., Hartmann, M.D., Hernandez Alvarez, B., Alva, V., Boyle, A.L. and Dame, R.T.[*] (2024) Histones and histone variant families in prokaryotes. *Nature Communications*, **15**, 7950.

Schwab, S., **Hu, Y.**, Hernandez Alvarez, B., Alva, V. and Dame, R. T.[*] (2025). Histone-mediated chromatin organization in prokaryotes and viruses. *Trends in Biochemical Sciences*.

Kretsch, R. C., Albrecht, R., Andersen, E. S., Chen, H. A., Chiu, W., Das, R., Gezelle, J. G., Hartmann, M. D., Höbartner, C., **Hu, Y.**, Jadhav, S., Johnson, P. E., Jones, C. P., Koirala, D., Kristoffersen, E. L., Largy, E., Lewicka, A., Mackereth, C. D., Marcia, M., Nigro, M., Ojha, M., Piccirilli, J. A., Rice, P. A., Shin, H., Steckelberg, A. L., Su, Z., Srivastava, Y., Wang, L., Wu, Y., Xie, J., Zwergius, N. H., Moult, J. and Kryshtafovych, A. (2025). Functional relevance of CASP16 nucleic acid predictions as evaluated by structure providers. bioRxiv, 2025-04.

Note: [†] The first two authors should be regarded as joint first authors.

[§] The two authors should be regarded as joint second authors.

[*] The corresponding author(s).

List of figures

List of abbreviations

Abbreviation	Meaning
3D domain swapping	Three-dimensional domain swapping
AAA+ protein	ATPases associated with diverse cellular activities protein
ADS	Antecedent domain segment
ASU	Asymmetric unit
ATP	Adenosine triphosphate
BN-PAGE	Blue native polyacrylamide gel electrophoresis
bp	Base pair
CD spectroscopy	Circular dichroism spectroscopy
Cryo-EM	Cryogenic electron microscopy
DAPI	4′,6-diamidino-2-phenylindole
DAP-seq	DNA affinity purification sequencing
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
EMSA	Electrophoretic mobility shift assay
Fis	Factor for inversion stimulation
FtF group	Face-to-face group
GFP	Green fluorescent protein
H1	Histone H1
H2A	Histone H2A
H2B	Histone H2B
H3	Histone H3
H4	Histone H4
H5	Histone H5
HBb	Histone of *Bdellovibrios bacteriovorus*
HLp	Histone of *Leptospira perolatii*

HMfA	Histone A of *Methanothermus fervidus*
HMfB	Histone B of *M. fervidus*
HMk	Histone of *Methanopyrus kandleri*
H-NS	Histone-like nucleoid structuring protein
HSH motif	Helix-strand-helix motif
HTkA	Histone A of *Thermococcus kodakarensis*
HTkB	Histone B of *T. kodakarensis*
HTkC	Histone C of *T. kodakarensis*
HU	Heat-stable protein found in *Escherichia coli* strain U93
IHF	Integration host factor
IPTG	Isopropyl-β-d-thiogalactoside
LACA	Last archaeal common ancestor
LECA	Last eukaryotic common ancestor
LUCA	Last universal common ancestor
MD simulation	Molecular dynamics simulation
MM	*Medusavirus medusae*
MNase	Micrococcal nuclease
MR	Molecular replacement
MSA	Multiple sequence alignment
MST	Microscale thermophoresis
MV	Melbournevirus
MW	Molecular weight
NAP	Nucleoid-associated protein
NF-Y	Nuclear factor Y
nt	nucleotide
PDB	Protein Data Bank
PTM	Post-translational modification
RMSD	Root-mean-square deviation

RNA	Ribonucleic acid
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SEC-MALS	Size exclusion chromatography coupled with multi-angle light scattering
SOS protein	Son of sevenless protein
ssDNA	Single-stranded DNA
TAF	TATA-binding protein-associated factor
TPM experiment	Tethered particle motion experiment

Abstract

Histones are a class of highly conserved DNA-binding proteins found in eukaryotes and archaea that play crucial roles in chromatin organization and gene regulation. Their evolutionary conservation and structural features make them ideal models for studying protein evolution. Histones are characterized by a core structural motif – the "histone fold" – comprising two short flanking α-helices and a long central α-helix linked by two strap loops.

In eukaryotes, histones form heterodimers that assemble into octameric complexes, wrapped by approximately 150 bp of double-stranded DNA, constituting nucleosomes – the fundamental units of chromatin. This nucleosomal structure compacts the eukaryotic genome to fit within the cell nucleus and regulates genomic DNA accessibility, contributing to transcriptional control and genome stability. In contrast, archaeal histones assemble into homo- or heterodimers that bind shorter DNA segments, forming compact and stable protein-DNA complexes. In some extremophilic archaea, histones further oligomerize into extended superhelical structures, termed hypernucleosomes, enhancing DNA protection under harsh conditions.

Although histones were previously thought to be hallmarks of eukaryotes and archaea, a recent comprehensive bioinformatic analysis identified approximately 600 histone homologs across diverse bacterial lineages. This unexpected finding suggests that histone-based DNA organization may be an evolutionary conserved feature across all domains of life. However, the structure, function, and DNA-binding mechanisms of bacterial histones remain largely unexplored. To address this knowledge gap, this study presents the structural and functional characterization of two bacterial histone candidates: HBb (Histone of *Bdellovobrio bacteriovorus*) and HLp (Histone of *Leptospira perolatii*). Using an integrative approach combining biophysical, biochemical, and computational methods, HBb has been identified as a homodimeric histone adopting the canonical histone fold. Despite its low sequence similarity to eukaryotic and archaeal histones, HBb binds DNA through a conserved interface

characteristic of classical histones. Unlike nucleosomal histones, HBb induces DNA bending reminiscent of bacterial nucleoid-associated proteins such as HU, implying a role in genome compaction and dynamics in *B. bacteriovorus*. In contrast, HLp assembles into unconventional homotetramers – an oligomeric state not previously observed among classical histones. Nevertheless, HLp retains the fundamental DNA wrapping and compaction properties similar to its eukaryotic and archaeal counterparts, indicating a potential structural role in *L. perolatii* chromatin.

These findings provide compelling evidence that histone-based DNA organization extends to bacteria, challenging the long-held view that bacteria lack histones. This work offers new insights into the evolutionary history of histones and expands our understanding of chromatin organization mechanisms across all domains of life.

Zusammenfassung

Histone sind eine Klasse hochkonservierter DNA-bindender Proteine, die in Eukaryoten und Archaeen vorkommen und eine entscheidende Rolle in der Chromatinorganisation und Genregulation spielen. Aufgrund ihrer evolutionären Konservierung und ihrer strukturellen Merkmale sind sie ideale Modelle für evolutionäre Studien zu Proteinen. Histone zeichnen sich durch ein zentrales strukturelles Motiv aus, die sogenannte "Histon-Faltung (*histone fold*)", die aus zwei kurzen flankierenden α-Helices und einer langen zentralen α-Helix besteht, welche durch zwei Schlaufen miteinander verbunden sind.

In Eukaryoten bilden Histone Heterodimere, die sich zu oktameren Komplexen zusammenschließen. Diese werden von ungefähr 150 Basenpaaren doppelsträngiger DNA umwickelt und bilden die Nukleosomen, die Grundeinheiten des Chromatins. Diese nukleosomale Struktur komprimiert das Genom, damit es in den Zellkern passt, und reguliert die Zugänglichkeit der DNA, was zur Transkriptionskontrolle und genomischen Stabilität beiträgt. Im Gegensatz dazu lagern sich archaische Histone zu Homo- oder Heterodimeren zusammen, die kürzere DNA-Segmente binden und kompakte sowie stabile Protein-DNA-Komplexe bilden. In einigen extremophilen Archaeen assemblieren Histone in superhelikale Strukturen, die als Hypermukleosomen bezeichnet werden und den Schutz der DNA unter extremen verbessern.

Obwohl man früher davon ausging, dass Histone ein Merkmal von Eukaryoten und Archaeen sind, wurden bei einer kürzlich durchgeführten umfassenden bioinformatischen Analyse etwa 600 Histon-Homologe in verschiedenen Bakterienlinien nachgewiesen. Dieser unerwartete Befund deutet darauf hin, dass die histonbasierte DNA-Organisation ein evolutionär konserviertes Merkmal sein könnte, das sich über alle Domänen des Lebens erstreckt. Dennoch bleiben die Struktur, Funktion und DNA-Bindungsmechanismen bakterieller Histone weitgehend unerforscht. Um diese Wissenslücke zu schließen, präsentiert diese Studie die

strukturelle und funktionelle Charakterisierung zweier bakterieller Histon-Kandidaten: HBb (Histon von *Bdellovibrio bacteriovorus*) und HLp (Histon von *Leptospira perolatii*). Mit einem integrativen Ansatz, der biophysikalische, biochemische und computergestützte Methoden kombiniert, wurde HBb als homodimeres Histon identifiziert, das die kanonische Histon-Faltung annimmt. Trotz geringer Sequenzähnlichkeit zu eukaryotischen und archaischen Histonen ähnelt die Bindungsschnittstelle zwischen DNA und HBb der klassischer Histone. Anders als nukleosomale Histone biegt HBb die gebundene DNA auf eine Weise, die an bakterielle Nukleoid-assoziierte Proteine wie HU erinnert, was auf eine Rolle bei der Genomkompaktierung und -dynamik in *B. bacteriovorus* hindeutet. Im Gegensatz dazu assembliert HLp zu ungewöhnlichen Homotetrameren, einem oligomeren Zustand, der bei klassischen Histonen bislang nicht beobachtet wurde. Dennoch zeigt HLp die grundlegenden DNA-Umhüllungs- und Verdichtungseigenschaften, die denen seiner eukaryotischen und archaischen Gegenstücke ähneln, was auf eine potentielle regulatorische Rolle im Chromatin von *L. perolatii* hindeutet.

Die vorliegenden Ergebnisse liefern überzeugende Beweise dafür, dass die histonbasierte DNA-Organisation auch in Bakterien existiert, und stellen damit die langgehegte Annahme in Frage, dass Bakterien keine Histone besitzen. Diese Erkenntnisse bieten neue Perspektiven der Histonevolution und erweitert unser Verständnis der Chromatin-Organisationsmechanismen in allen Lebensbereichen.

1. Introduction

“生命中的全部偶然，其实都是命中注定。是为宿命。”

—— 东野圭吾《宿命》

1.1. The evolution of proteins

Proteins are large, complex biomolecules composed of one or more chains of amino acid residues linked together by peptide bonds. They are encoded by genes and undertake a spectrum of significant responsibilities in living organisms, acting as cell structural components, enzymes to catalyze biochemical reactions, signaling molecules, and transporters, among many others. A plausible hypothesis of protein evolution is that it is shaped by a variety of genetic events that introduce diversity and enable adaptation. Large-scale genetic processes like lateral (horizontal) gene transfer gene duplication, followed by gene fusion play significant roles in expanding protein families and introducing novel functions (1,2). Small-scale genetic changes, such as mutations, insertions, and deletions, fine-tune protein structures and functions over time (1,2). Additional mechanisms, including domain shuffling, gene fission, and chromosomal rearrangements, also lead to the origin of new protein structures (1,2). These processes act as a driving force, adding to the complexity and versatility of proteins (1,2). Despite the enormous sequence diversity of modern proteins, their structural repertoire appears to be constrained, with most proteins exhibiting a restricted range of folds, leading to the conceivable hypothesis that protein structures may be constructed based on a fundamental “vocabulary” of repetitive supersecondary structural elements (1-3). A growing number of evidence, both at the sequence and structural levels, has identified an increasing number of these shared basic elements (2). This observation supports the notion that modern proteins may have evolved from a relatively small set of ancestral domain segments, referred to as antecedent domain segments (ADSs) (1). During the

era of the “RNA world”, short peptides presumably existed as cofactors, supporting RNA replication and catalysis rather than forming stable folds (2). Peptides naturally exhibit a tendency to form secondary structures, a property that may have been further stabilized and enhanced through interactions with nucleic acids (2). Eventually, the association and fusion of these peptides into larger assemblies enabled the emergence of folded proteins – macromolecules that could perform diverse and specialized functions beyond their original role as RNA cofactors (2).

In spite of their apparent dissimilarity, the recurring motifs observed across proteins are plausibly relics of ADSs, which might have been established at the time of the last universal common ancestor (LUCA) (1-3). These sequences and structures hold information about the evolutionary events that shaped them. This continuity allows researchers to study and attempt the reconstruction of the evolutionary pathways of proteins. In this sense, understanding protein evolution relies heavily on the comparative analysis of sequences and structures (1). Much like paleontology tries to uncover the history of life through the study of fossils, studying protein evolution involves identifying “living fossils” within the modern protein world – highly conserved proteins that play essential roles in cellular processes, such as ribosomal proteins and ubiquitin, acting as a window into their evolutionary history (2).

Protein structures diverge much slower than sequences, making structural similarity a valuable marker for identifying distant evolutionary relationships (1). However, identical structures can arise not only from shared ancestry but also from the limited number of stable folded conformations available to polypeptide chains. This is particularly the case at the supersecondary structure level or as a result of convergent evolution driven by functional constraints (1,2). Consequently, sequence similarity often serves as a more reliable marker for identifying common ancestors (3). Moreover, it is important to note that protein sequences with very high similarity can sometimes fold into distinct structures (4,5). With these aspects in mind, a comprehensive approach that incorporates both sequence and structural analyses is essential in protein evolution

studies. Thus, bioinformatic analyses can help identify promising research targets, yet experimental validation remains critical to attain robust evidence. This integrative strategy ensures a more accurate reconstruction of protein evolutionary pathways.

1.2. The origin of the histone fold

Histones are a group of highly conserved, basic proteins that are essential for chromatin organization and gene regulation (6,7). They share a characteristic structural motif, known as the "histone fold", which consists of approximately 65 amino acids arranged into three α-helices – one long helix in the middle of the chain, flanked by two short ones – connected by two strap loops (8,9). This conserved structure facilitates dimerization through interactions along the second α-helices, forming a distinctive "handshake" motif, while the loop regions align to generate short parallel β-bridges (10) (Fig. 1.1).

Comparative sequence and structural analyses have been employed to investigate the origins of the histone fold, in the light of the ADS-based protein evolution hypothesis. The analysis results suggested that, through divergent evolution, an ADS akin to a helix-strand-helix (HSH) motif, might have given rise to both the helical region of the extended ATPase domain (C-domain) found in AAA+ proteins and the N-terminal substrate recognition domain of Clp/Hsp100 proteins (11). The histone fold is thought to have originated from the duplication of a primordial HSH motif, followed by three-dimensional (3D) domain swapping (11,12) (Fig. 1.1). Though the N-terminal and C-terminal HSH motifs of the C-domain of AAA+ proteins display high structural similarity, it remains ambiguous whether the two HSH motifs within the histone fold arose from the duplication of a single ancestral HSH motif or from the fusion of two distinct, heterologous HSH motifs (10,13,14).

While studying the origin of the histone fold, this structural motif was found to be present in archaeal histones, as well as in many other eukaryotic proteins, mainly involved in DNA binding (14), such as son of sevenless (SOS) proteins (15) and TATA-

binding protein-associated factors (TAFs) (16). Interestingly, histone-like proteins have also been identified in approximately ten bacterial species (14). These bacterial proteins were predicted to display a pseudodimeric structure, a characteristic previously observed only in archaeal histones (17). Thus, it was hypothesized that histones originated before the split between archaea and eukaryotes, and that some bacteria acquired these proteins from archaea later on, through lateral (horizontal) gene transfer (11,18).

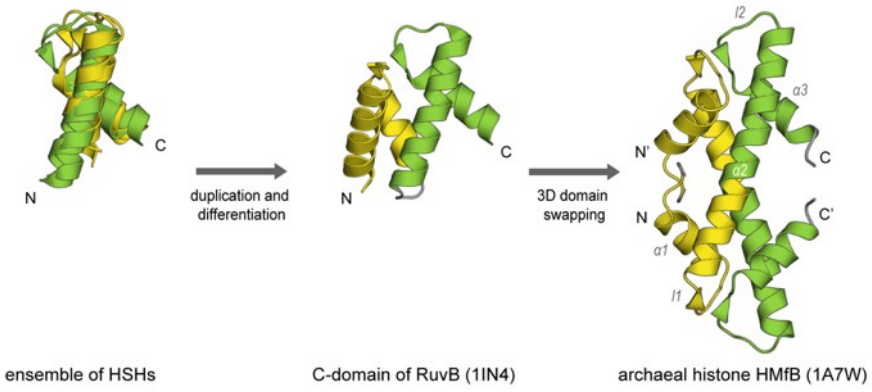


Figure 1.1 Evolutionary scenario for the origin of the histone fold.

The N-terminal and C-terminal HSH motifs in each subunit are colored in yellow and green, respectively. The superimposed ensemble of HSH motifs contains the N-terminal HSH motif in the C-domain of RuvB (PDB: 1IN4, residues 181-212), the N-terminal HSH motif in HMfB (PDB: 1A7W, residues 3-32), the C-terminal HSH motif in the C-domain of RuvB (PDB: 1IN4, residues 216-251), and the C-terminal HSH motif in HMfB (PDB: 1A7W, residues 33-65). *This figure was adapted from the Figure 4 in Alva et al., 2007 (11).*

1.3. The discovery of bacterial histones

With the rapid expansion of protein sequence databases, and the advent of highly sensitive sequence search tools, like HHpred (19), an increasing number of histone homologs have been identified. This is achieved using representative histones and their distant homologs as queries for comprehensive sequence searches (20).

A recent search seeded with the distant histone homolog Aq_328 uncovered a novel group of prokaryotic pseudodimeric histone fold proteins, including about 360 bacterial homologs (20). These proteins feature two consecutive histone folds within a single polypeptide chain, yet no significant sequence homology between the two folds

has been detected. Most of these homologs were identified in the ancient and deep-branching bacterial class Actinobacteria, with the additional homologs sparsely distributed across other classes, including Aquificia (previous name: Aquificae), Bacteroidia (Bacteroidetes), Chloroflexia (Chloroflexi), and the genus Deinococcus, and the phyla Cyanobacteriota (Cyanobacteria), Bacillota (Firmicutes) and Pseudomonadota (Proteobacteria) (20). Another novel group of prokaryotic single histone fold-like proteins, including approximately 230 bacterial homologs of 60-70 residues in length, was identified using the bacterial histone homolog found in *Leptospira interrogans* as the seed (20). These homologs were found across diverse bacterial classes such as Bdellovibrionia, Spirochaetia, Deltaproteobacteria, Planctomycetia (Planctomycetes), Elusimicrobia and Chlamydiia (Chlamydiae), as well as in the phylum Candidatus Parcubacteria (Parcubacteria) (20).

The discovery of a substantial number of histone homologs across a wide range of deep-branching bacterial lineages indicates that histone-like proteins are likely present in bacteria. Furthermore, these results suggest that the histone fold may have originated before the divergence of bacteria and archaea, possibly at the time of LUCA (20).

In addition to this comprehensive bioinformatic study, only one experimental study has previously focused on bacterial histones. This study demonstrated by crystallography that the histone homolog Aq_328, found in the hyperthermophilic bacterium *Aquifex aeolicus*, possesses an ancestral histone fold (21). However, its proposed DNA-binding function is solely supported by sequence similarity to other histones, and it has not been experimentally validated (21).

1.4. Research aims

The need to research bacterial histones stems from their critical role in promoting our understanding of protein evolution, along with a notable gap in our knowledge regarding these proteins within the domain of bacteria. Compared to the extensively characterized roles of histones in DNA organization and gene regulation in eukaryotes

and archaea, the existence, structures, and functions of histone homologs in bacteria remain poorly understood and largely unexplored (Fig. 1.2).

This research aims to fill the knowledge gap by characterizing bacterial histone homologs both structurally and functionally, using a combination of computational, structural, and functional analyses. Specifically, it focuses on: (1) determining whether bacteria possess *bona fide* histones; (2) analyzing their DNA binding properties; and (3) comparing them to their eukaryotic and archaeal counterparts. By investigating the functional roles and evolutionary relationships of bacterial histones, this research seeks to uncover their contributions to genome organization and regulation, as well as to highlight both conserved and divergent features. By exploring these homologs, this study aims to advance our understanding of the evolutionary trajectory of histones across all domains of life.

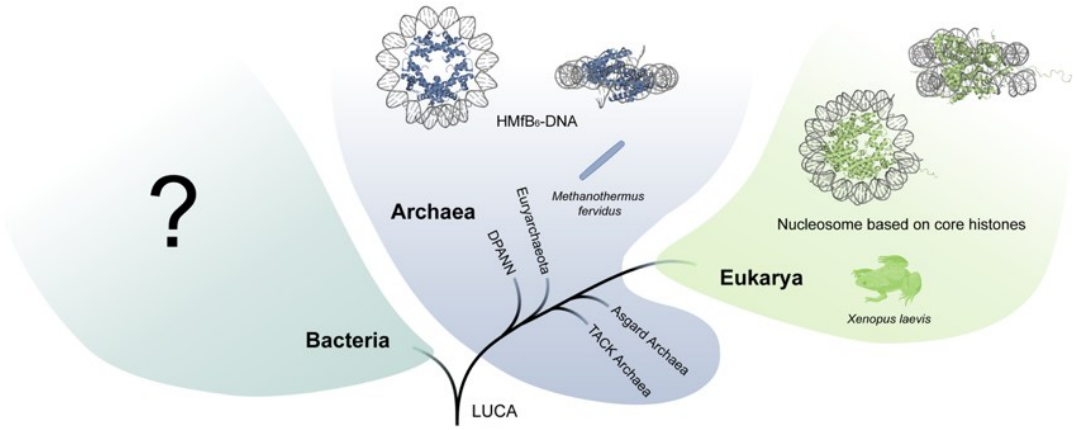


Figure 1.2 The simplified tree of life, including the representative histone-based nucleosomes in archaea and eukaryotes.

The representative archaeal nucleosome is formed based on the archaeal histone HMfB (in blue) found in *Methanothermus fervidus* (PDB: 5T5K). The representative eukaryotic nucleosome is assembled based on the core octamer (in light green), containing H2A, H2B, H3 and H4, found in *Xenopus laevis* (PDB: 1AOI). The icon of *X. laevis* is from BioRender.com. *This figure was adapted from the graphical abstract in Hu et al., 2024 (22).*

2. Background

“‘这个世界里，’汤川说道，‘有些谜是无法用现代科学来解释的。但是，随着科技的进步，迟早一天，那些谜也会被人们解开的。那么，科学是否有极限呢？如果有的话，那么这极限到底是从哪里来的呢？’

……

汤川用手指着恭平的额头，说道：‘那就是人。’”

—— 东野圭吾《盛夏的方程式》

2.1. Eukaryotic histones

The study of histones has spanned nearly one and a half centuries, reflecting the incremental progress of scientific exploration. In 1884, Albrecht Kossel isolated a basic protein from red blood cells of goose blood, noted their association with nucleic acids, and named it “Histone”, marking the starting point of histone research (23,24). This groundbreaking observation established histones as integral components of chromatin, setting the stage for subsequent research into their structural and functional roles in eukaryotic cells.

A pivotal breakthrough in chromatin biology occurred in 1974 with the proposal of the “beads-on-a-string” model, identifying the nucleosome as the fundamental repeating unit (25). Nucleosomes consist of a central H3-H4 tetramer flanked by two H2A-H2B dimers, upon which approximately 147 base pair (bp) of double-stranded DNA (dsDNA) wrap around it in a left-handed superhelical manner (8,26) (Fig. 2.1A and C). The core histones, H2A, H2B, H3, and H4, are generally essential for the structural and functional integrity of nucleosomes. The lack or defect of individual core histones causes dysregulated transcription, thus impairing cell cycle progression and subsequently leading to cell death (27-30). In addition to the core histones, linker histones, such as H1 and H5, play an auxiliary role in chromatin organization (31,32).

Although not part of the nucleosome core and dispensable under certain circumstances, linker histones bind to the entry and exit points of dsDNA on the nucleosome, stabilizing the nucleosomes and promoting the compaction of higher-order chromatin (31,32).

The most distinctive feature of eukaryotic histones is their unstructured terminal tails, which serve as sites for post-translational modifications (PTMs) that are associated with gene regulation. Since the 1960s, PTMs such as acetylation, methylation, phosphorylation, and ubiquitination have been discovered. They occur primarily in the lysine-rich N-terminal tails of the histones (33,34). These modifications influence chromatin structure by altering histone-DNA interactions or recruiting other chromatin-associated proteins, thereby affecting DNA-dependent processes (35). For instance, acetylation of lysine residues typically correlates with transcriptional activation (36), whereas methylation can signal either activation or repression (37,38). The “histone code” hypothesis proposes that distinct PTM combinations establish a regulatory framework directing specific genomic processes (39,40).

Histone modifiers, including acetyltransferases, kinases, and methyltransferases, catalyze covalent modifications on the histone tails, and therefore play an essential role in gene transcription (39,40). In addition to histone modifiers, chromatin remodelers, histone variants, and chaperones also contribute to nucleosome organization (39,40). Histone remodelers reposition or evict nucleosomes to modulate transcription, and histone chaperones regulate histone deposition and exchange, ensuring nucleosome integrity during DNA replication, damage, repair, recombination and transcription (41-43). Collectively, these factors coordinate chromatin dynamics, and maintain genome stability cooperatively.

2.2. Archaeal histones

Histones were long thought to be exclusive to eukaryotes until this notion was overturned in 1990 with the discovery of the archaeal histones HMfA and HMfB in the hyperthermophilic archaeon *Methanothermus fervidus* (44) (Fig. 2.1B). These proteins retain the conserved histone fold, dimerize, and bind DNA. Some of them form

continuous chromatin-like structures, termed “hypernucleosomes” or “archaeasomes”, upon DNA binding (45-49) (Fig. 2.1D). This arrangement offers enhanced chromatin stability and protection, which is important and favorable for archaea to survive in extreme environments. These features exhibit significant structural and functional similarities to eukaryotic histones. Another intriguing archaeal histone, HMk, was later identified in *Methanopyrus kandleri* (17). In contrast to conventional histones, HMk is a pseudodimeric histone, harboring two histone folds separated by a short linker region within a single polypeptide chain (17).

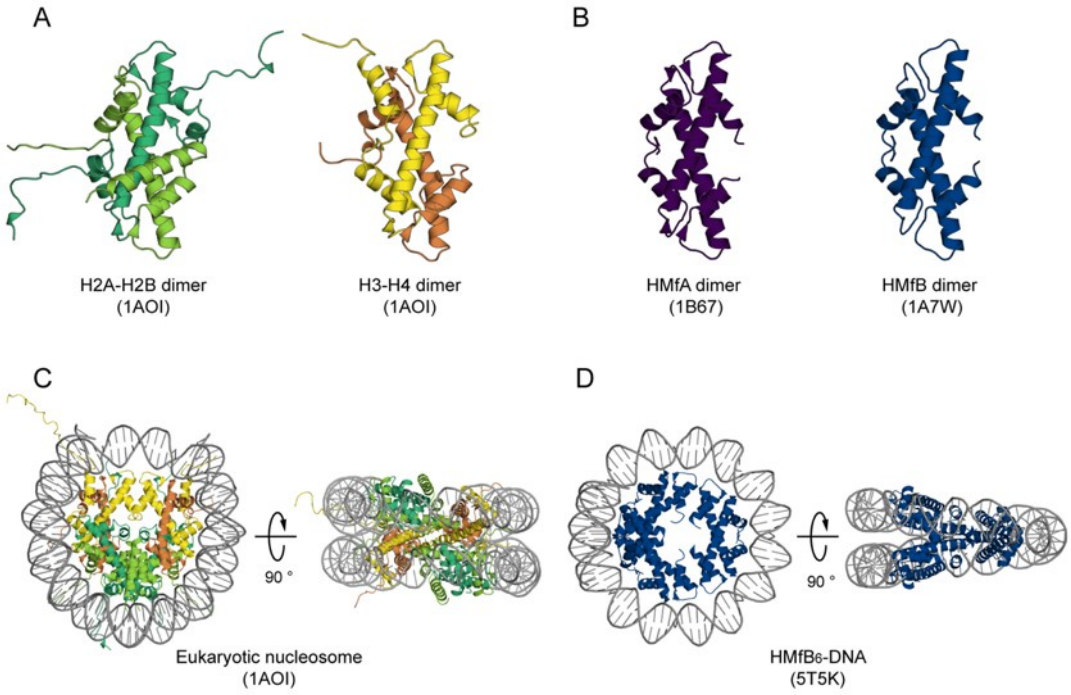


Figure 2.1 The representatives of eukaryotic and archaeal histones.

A. Structures of eukaryotic histones H2A (in dark green), H2B (in light green), H3 (in yellow) and H4 (in orange) from *X. laevis*. **B.** Structures of archaeal HMfA (in dark purple) and HMfB (in dark blue) from *M. fervidus*. **C.** The structure of eukaryotic chromosome consists of a 147 bp DNA wrapping around a core histone octamer (PDB: 1AOI). **D.** The structure of HMfB-based hypernucleosome comprises three HMfB dimers bound to a 90 bp DNA fragment (PDB: 5T5K).

The characterization of diverse archaeal species and the increasing availability of metagenomic data have led to the discovery of more histone homologs across most archaeal lineages. Archaeal genomes typically encode two to three histones, with some species possessing an expanded repertoire (20,50-52). These homologs and variants

display considerable sequence variability and structural diversity, with some incorporating additional domains. While most archaeal histones dimerize into homo- or heterodimers, certain homologs are predicted to assemble into DNA-independent tetramers (53). Previously, archaeal histones were believed to lack the extended tails characteristic of their eukaryotic counterparts. However, recent bioinformatic analyses have identified several histone homologs in the Asgard superphylum, a clade of archaea considered the closest prokaryotic relatives of eukaryotes, that feature long, disordered N-terminal tails akin to those found in eukaryotic histones (47,53-55). Whether these positively charged tails in archaea and eukaryotes share a direct evolutionary link remains an open question (52,53).

The abundance and functional significance of histones vary across archaeal taxa, where they operate alongside nucleoid-associated proteins (NAPs), and probably other histone variants (53,56-62). These observations indicate that archaeal histones may fulfill roles beyond chromatin organization. For instance, HpyA, a histone-like protein found in the halophilic archaeon *Halobacterium salinarum*, is proposed to function as a transcription factor, binding discrete genomic loci to regulate gene expression in response to environmental salt concentrations (60,63). Furthermore, the mechanisms by which archaeal histones compact genomic DNA and modulate transcription appear to differ fundamentally from those in eukaryotes. Structural and biochemical studies suggest that archaeasomes exist in highly dynamic equilibrium between open and closed states while maintaining histone-DNA interactions (49). This “slinky”-like conformational flexibility enables a balance between genome compaction and accessibility in the absence of ATP-dependent chromatin remodelers, providing an alternative strategy for histone-based chromatin regulation in archaea (49).

The widespread presence of histone homologs and variants in archaea, along with their high structural and functional conservation relative to eukaryotic histones, suggests that histones were established at the time of the last archaeal common ancestor (LACA) (53). The bioinformatic analysis revealed a close evolutionary relationship

between HMfB-like histones and eukaryotic histones, further supporting this notion (20). Additionally, phylogenomic studies positioning Eukarya as a sister group to Archaea reinforce the hypothesis that eukaryotic histones likely evolved from archaeal ancestors (9).

2.3. Viral histones

Histone-like proteins have also been discovered in several representatives of nucleocytoplasmic large DNA viruses. Among Nucleocytoviricota virus metagenomes, these proteins are found in diverse configurations, ranging from singlets to quadruplets, incorporating various combinations of remote homologs of the four eukaryotic core histones in linear arrays (64,65).

Some viral histone homologs assemble into nucleosome-like particles, resembling the chromatin architecture of eukaryotes. For instance, Melbournevirus (MV) encodes two histone doublets, MV-H2A-H2B and MV-H4-H3. Despite displaying low sequence similarity (< 30%) to their eukaryotic counterparts, these doublets are necessary for viral propagation and form a pseudooctameric complex that wraps approximately 180 bp of dsDNA in a manner similar to eukaryotic nucleosomes (66). Likewise, *Medusavirus medusae* (MM) encodes four histone singlets (MM-H2A, MM-H2B, MM-H3, and MM-H4), which assemble into octameric core histones (Fig. 2.2) and further organize into tri-nucleosome-like arrays (67). The putative linker protein MM-H1, containing two linker histone repeats, exhibits low sequence conservation compared to eukaryotic H1 and appears dispensable for chromatin compaction (67).

In other cases, viral histone quadruplets, such as HQA (H2B-H2A-H3-H4) and HQC (H4-H3-H2B-H2A), have been biochemically characterized and shown to protect dsDNA fragments of about 150 bp from MNase digestion (64). They are predicted to assemble into compact higher-order structures upon DNA binding, reminiscent of hypernucleosomes/archaeasomes, possibly through the sequential addition of monomers or dimers (64).

The previous bioinformatic study has shown that viral histones lie within the eukaryotic clusters, with characterized viral histones displaying high structural and functional similarity to their eukaryotic counterparts (20,66,68). This suggests viruses likely acquired histone genes from eukaryotes through horizontal gene transfer (68,69). In contrast, viral histone doublets, which exhibit low sequence identity to eukaryotic histones, might result from coevolution with their hosts (68-70). Nonetheless, phylogenetic analyses support the possibility that viral histones originated from the last eukaryotic common ancestor (LECA) (64,68). The emergence of viral histone repeats, including doublets, triplets, and quadruplets, might have been driven by multiple genetic events during eukaryogenesis, such as (ir)reversible gene fusion, fission, and domain shuffling (64).

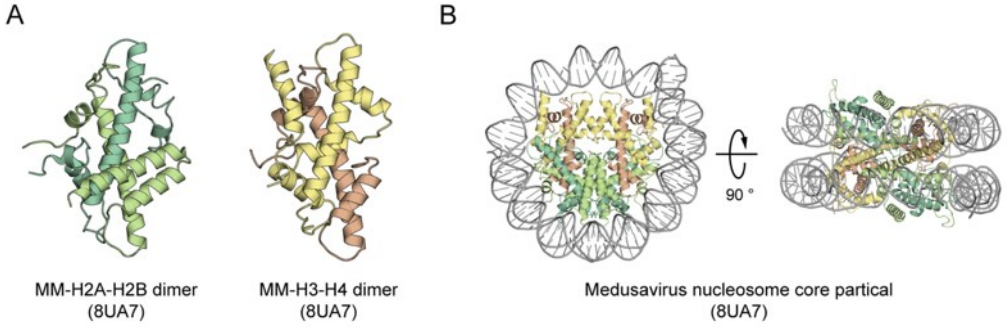


Figure 2.2 The histones found in *M. medusae*.

A. The structure of viral histone heterodimers MM-H2A-H2B and MM-H3-H4 (PDB: 8UA7). MM-H2A is colored in dark green, MM-H2B is colored in light green, MM-H3 is colored in yellow, and MM-H4 is colored in orange. **B.** The structure of the reconstituted Medusavirus nucleosome core particle (PDB: 8UA7).

2.4. Bacterial nucleoid-associated proteins

In bacteria, genetic material and associated proteins are colocalized within an irregularly shaped, membrane-free region, which is known as the nucleoid (71). This highly dynamic structure is constrained by the interplay of multiple factors, including DNA topology, macromolecular crowding, and NAPs (72,73). Initially referred to as "histone-like proteins" because of their functional resemblance to eukaryotic histones, NAPs are now categorized according to their role in chromatin architecture (73-76).

Functionally, NAPs bind DNA and modulate its architecture by bending, bridging, and/or wrapping the nucleic acid, thereby influencing genome compaction and gene regulation. The most abundant and well-characterized NAPs are histone-like nucleoid structuring protein (H-NS), heat-stable protein found in *Escherichia coli* strain U93 (HU), factor for inversion stimulation (Fis), and integration host factor (IHF) (72,76). H-NS, a highly conserved protein predominantly found in Gram-negative bacteria, consists of an N-terminal dimerization domain and a C-terminal DNA-binding domain, joined by a disordered linker (77,78). While existing as dimer in solution, H-NS forms higher-order oligomers upon DNA binding, bridging and stiffening non-specific DNA (72,79). This process alters DNA topology and represses transcription by restricting promoter accessibility or trapping RNA polymerase (76,80). HU, another abundant NAP, is composed of HUα and HUβ subunits and forms homo- or heterodimers depending on the bacterial growth phase (Fig. 2.3A) (81). It exhibits a preference for A/T-rich sequences, supercoiled DNA, and certain structures, such as curved DNA, replication forks, or four-way junctions (82-86). At low concentrations, HU reduces DNA stiffness by inducing short-range bending (Fig. 2.3A); whereas at high concentrations, it increases local DNA stiffness through filament formation (72,87). IHF, a sequence homolog of HU, possesses both sequence-specific and non-specific DNA binding ability (Fig. 2.3B) (88,89). In *Escherichia coli*, approximately 1000 IHF-binding sites have been mapped, most of which are located near the promoter regions, implying a role in transcription regulation (90,91). Primarily, by introducing U-turn into the DNA, IHF can occlude promoter regions or create DNA loops, preventing RNA polymerase access and leading to repression (Fig. 2.3B) (92). In some cases, its DNA bending ability assists the correct positioning of σ^{54}-containing RNA polymerase on promoters, thus resulting in activation (93). Fis, a dimeric NAP composed of helix-turn-helix DNA-binding motifs, also induces DNA bending (72). Besides its non-specific DNA-binding affinity, it has been reported to recognize a poorly conserved binding site of 15 bp in length (72). In addition to these four major NAP families, there are other

NAPs that actively participate in genomic DNA-related cellular activities, such as StpA, CbpA, MukB, and so forth (76,94-96).

Although NAPs act as key players in bacterial chromatin organization and gene regulation, studies have shown that not all NAPs are strictly essential for bacterial viability. For instance, *E. coli* can maintain the nucleoid structure after deleting HU, H-NS, Fis or IHF individually, probably due to the cooperative effect of remaining NAPs (97,98). The antagonistic functions of NAPs, ranging from DNA stiffening and bridging to bending and wrapping, reflect the intrinsic dynamic nature of bacterial chromatin, enabling a delicate balance between genome integrity and rapid gene regulation in response to environmental changes (72).

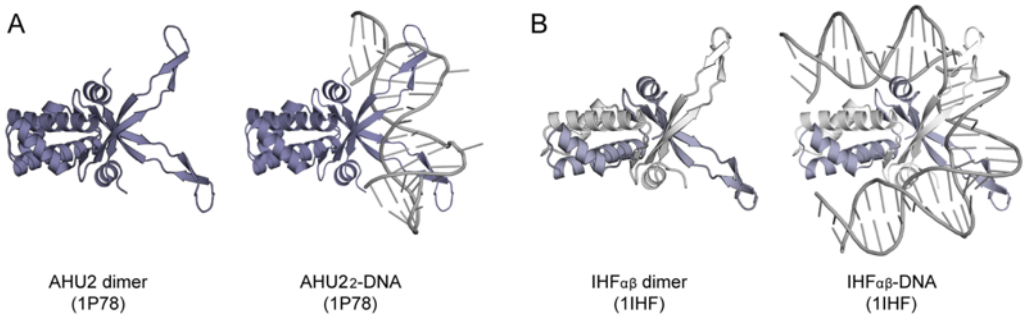


Figure 2.3 Two examples of bacterial NAPs.

A. The structures of homodimeric AHU, found in cyanobacterium *Anabaena*, with and without DNA (PDB: 1P78) (99). **B.** The structures of heterodimeric $IHF_{\alpha\beta}$, from *E. coli*, with and without DNA (PDB: 1IHF) (100). The α subunit is colored in light lavender, and the β subunit is colored in white.

2.5. Potential bacterial histone candidates for characterization

As discussed in section 1.3, a comprehensive bioinformatic analysis identified about 600 histone homologs across various bacterial lineages, subverting the long-held concept that bacteria lack histones (20). A subsequent search for histone homologs, conducted by Vikram Alva (Max Planck Institute for Biology Tübingen), expanded the known repertoire of bacterial histone candidates, which can be categorized into the previously described groups: the single histone fold-like proteins and the pseudodimeric histone fold proteins (20). In the pseudodimeric group, two histones

folds that are encoded within a single polypeptide chain often exhibit high sequence diversity, and also lack key conserved residues that are critical for DNA binding (20). In contrast, proteins within the single histone fold-like group share a comparatively higher sequence similarity, and have a higher conservation of residues that are essential for protein-DNA interactions in eukaryotes and archaea (20). This group includes not only small proteins containing a solitary histone fold (about 60 residues in length) but also larger variants incorporating N- and/or C-terminal extensions or additional structured domains. Given these variations, the present study primarily focuses on the simplest bacterial histone homologs, comprising only the core histone fold.

Considering the availability of genetic tools for gene manipulation and the feasibility of cultivating the bacterial species, HBb, a hypothetical protein identified from *Bdellovibrio bacteriovorus*, was selected for initial characterization. The detailed investigation of HBb is presented in Chapter 3.

Intriguingly, another candidate from the single histone fold-like protein group, HLp, found in *Leptospira perolatii*, was predicted to form homotetramers in the absence of DNA, which has never been observed for histones before. Due to this interesting potential structure, HLp was chosen for experimental characterization to confirm its structure and investigate the functional properties resulting from this unique fold. Genome annotation identified two potential transcription start sites for the HLp-encoding gene in *L. perolatii.* Based on the position of the TATA box, the shorter open reading frame was selected for protein expression. Since *L. perolatii* is a pathogenic microorganism, all analyses were conducted based on heterologous expression of the protein in *E. coli*. The detailed characterization of HLp is described in Chapter 4.

3. The dimeric histone HBb from *Bdellovibrio bacteriovorus*

“没有亲自动手做过实验的人，是不会明白的。”

—— 东野圭吾《盛夏的方程式》

Abstract

Histones are essential for genome compaction and transcription regulation in eukaryotes, where they assemble into octamers to form the nucleosome core. In contrast, archaeal histones assemble into dimers that form hypernucleosomes upon DNA binding. Although histone homologs have been identified in bacteria recently, their DNA-binding characteristics remain largely unexplored. Our study reveals that the bacterial histone HBb (Bd0055) is indispensable for the survival of *Bdellovibrio bacteriovorus*, suggesting critical roles in DNA organization and gene regulation. By determining the crystal structures of free and DNA-bound HBb, we unveil its distinctive dimeric assembly, which differs from those of eukaryotic and archaeal histones. We also elucidate how HBb binds and bends DNA through interaction interfaces reminiscent of its eukaryotic and archaeal counterparts. Building on this, by employing various biophysical and biochemical approaches, we further substantiated the ability of HBb to bind and compact DNA by bending in a sequence-independent manner. Finally, using DNA affinity purification and sequencing, we reveal that HBb binds along the entire genomic DNA of *B. bacteriovorus* without sequence specificity. These distinct DNA-binding properties of bacterial histones, showcasing remarkable similarities yet significant differences from their archaeal and eukaryotic counterparts, highlight the diverse roles histones play in DNA organization across all domains of life.

This publication and the supplementary information are available in Appendix I.

Author contributions

Author	Author position	Scientific ideas (%)	Data generation (%)	Analysis and interpretation (%)	Paper writing and figure making (%)
Yimin Hu[†]	1	25	48	30	25
Samuel Schwab[†]	1	10	45	30	15
Silvia Deiss	3	N/A	5	N/A	N/A
Pedro Escudeiro	4	N/A	N/A	10	3
Thor van Heesch	5	N/A	N/A	3	N/A
Joe D. Joiner	6	N/A	N/A	3	2
Jocelyne Vreede	7	N/A	N/A	N/A	N/A
Marcus D. Hartmann	8	N/A	N/A	4	2
Andrei N. Lupas	9	5	N/A	N/A	N/A
Birte Hernandez Alvarez[*]	10	27	N/A	10	30
Vikram Alva[*]	11	23	2	5	20
Remus T. Dame[*]	12	10	N/A	5	3

Title of the paper	Bacterial histone HBb from *Bdellovibrio bacteriovorus* compacts DNA by bending
Status in publication process	Published

Note: [†]The first two authors should be regarded as joint first authors.
[*]The corresponding authors.

The original idea of this project was offered by Andrei N. Lupas and Vikram Alva (Dept. of Protein Evolution, MPI for Biology Tübingen), and the experiments were designed by Birte Hernandez Alvarez (Dept. of Protein Evolution, MPI for Biology

Tübingen) and Remus T. Dame (Leiden Institute of Chemistry; Centre for Microbial Cell Biology; Centre for Interdisciplinary Genome Research, Leiden University).

Yimin Hu performed cloning, protein expression and purification, CD spectroscopy, SEC-MALS, crystallization, data collection, structure determination, MST, EMSA (with 80 bp DNA fragment), as well as the experimental part of DAP-seq. All figures were made by Y. Hu, and the initial paper draft was written by Y. Hu with the help of Joe D. Joiner. The paper revision was done by Y. Hu with contributions from other authors.

Samuel Schwab (Leiden Institute of Chemistry; Centre for Microbial Cell Biology; Centre for Interdisciplinary Genome Research, Leiden University) conducted DNA binding assays, including EMSA (with DNA marker), MNase digestion assay, DNA topology assay, ligase-mediated circularization assay, and TPM experiment. S. Schwab also performed the MD simulation of HBb-DNA complex with the help of Thor van Heesch and Jocelyne Vreede (Van't Hoff Institute for Molecular Sciences, University of Amsterdam). S. Schwab wrote the corresponding method and result sections.

Silvia Deiss (Dept. of Protein Evolution, MPI for Biology Tübingen) performed *hbb* gene deletion in *B. bacteriovorus* HD100.

Pedro Escudeiro (Dept. of Protein Evolution, MPI for Biology Tübingen) analyzed DAP-seq data, plotted corresponding figures, and wrote the according part of the result section.

Marcus D. Hartmann (Dept. of Protein Evolution, MPI for Biology Tübingen; Interfaculty Institute of Biochemistry, University of Tübingen) helped with the structure determination and deposition. He also proofread the structure part of the result section.

B. Hernandez Alvarez and V. Alva revised the paper draft, and R. T. Dame proofread the manuscript.

4. The tetrameric histone HLp from *Leptospira perolatii*

“无路可走的时候，就不断回到原点。”

—— 东野圭吾《麒麟之翼》

Abstract

Histones are conserved DNA-packaging proteins found across all domains of life. In eukaryotes, canonical histones form octamers that wrap ~147 base pairs of DNA into nucleosomes — the fundamental units of chromatin. In archaea, histones form dimers that further multimerize into extended hypernucleosomes along DNA. Although bacteria were long thought to lack histones, recent studies have uncovered histone homologs in diverse bacterial lineages, many of which possess key DNA-binding features. We previously characterized HBb, a bacterial histone from *Bdellovibrio bacteriovorus*, which binds DNA as a dimer and induces bending. Here, we describe HLp from *Leptospira perolatii*, a representative of a distinct bacterial histone group. Crystallographic and biophysical analyses reveal that HLp forms stable tetramers. Like HBb, HLp binds DNA non-specifically; however, it adopts a different mode of interaction — wrapping ~60 base pairs of DNA around its tetrameric core. This wrapping mode, supported by molecular dynamics simulations and DNA-binding assays, promotes DNA compaction and alters its topology. When expressed heterologously in *Escherichia coli*, HLp reorganizes nucleoid morphology, consistent with a role in chromatin organization. These findings expand the known repertoire of histone–DNA interaction in bacteria and underscore the structural and functional diversity of histone-based genome organization across the tree of life.

This manuscript and the supplementary information are available in Appendix II.

Author contributions

Author	Author position	Scientific ideas (%)	Data generation (%)	Analysis and interpretation (%)	Paper writing and figure making (%)
Yimin Hu	1	50	60	58	28
Samuel Schwab[§]	2	N/A	3	3	2
Kaiyu Qiu[§]	2	N/A	15	15	17
Yunsen Zhang	4	N/A	6	5	1
Kerstin Bär	5	N/A	2	N/A	N/A
Heidi Reichle	6	N/A	2	N/A	N/A
Aurora Panzera	7	N/A	6	5	2
Andrei N. Lupas	8	5	N/A	N/A	N/A
Marcus D. Hartmann	9	N/A	N/A	4	2
Remus T. Dame	10	N/A	N/A	N/A	N/A
Vikram Alva[*]	11	20	6	5	20
Birte Hernandez Alvarez[*]	12	25	N/A	5	28
Title of the paper		DNA wrapping by a tetrameric bacterial histone			
Status in publication process		Manuscript in preparation			

Note: [§]The two authors should be regarded as joint second authors.
[*]The corresponding authors.

The original idea of this project was offered by Andrei N. Lupas, Vikram Alva, and Birte Hernandez Alvarez (Dept. of Protein Evolution, MPI for Biology Tübingen). The experiments were designed by Yimin Hu, and the collaborations were coordinated by Y. Hu.

V. Alva performed bioinformatic analyses and wrote the corresponding method and result sections.

Y. Hu conducted protein expression and purification, SEC-MALS, crystallization, data collection, and structure determination, as well as DNA binding assays (including EMSA, MNase digestion assay, DNA topology assay, ligase-mediated circularization assay, and MST). Y. Hu also prepared the sample for light microscopy. Most of the figures were made by Y. Hu, and the paper draft was written by Y. Hu with contributions from other authors.

Samuel Schwab (Leiden Institute of Chemistry; Centre for Microbial Cell Biology; Centre for Interdisciplinary Genome Research, Leiden University) performed the TPM experiment and wrote the corresponding method and result sections.

Kaiyu Qiu (Dept. of Protein Evolution, MPI for Biology Tübingen) and Yunsen Zhang (Theoretical and Computational Biophysics Group, Beckman Institute for Advanced Science and Technology, Center of Biophysics and Quantitative Biology, University of Illinois Urbana-Champaign) conducted the MD simulations. K. Qiu plotted all related figures and made Figure 5. He also wrote the corresponding method and result sections. Y. Zhang proofread the sections.

Kerstin Bär and Heidi Reichle (Dept. of Protein Evolution, MPI for Biology Tübingen) performed protein expression and purification. Kerstin conducted the CD spectroscopy.

Aurora Panzera (BioOptics Facility, MPI for Biology Tübingen) imaged all the samples and analyzed images together with Y. Hu. The corresponding method and result sections were written by A. Panzera.

Marcus D. Hartmann (Dept. of Protein Evolution, MPI for Biology Tübingen; Interfaculty Institute of Biochemistry, University of Tübingen) helped with the structure determination and deposition. He also proofread the structure part of the result section.

B. Hernandez Alvarez and V. Alva revised the paper draft, and Remus T. Dame proofread the manuscript.

5. Results

“不管什么问题，都必然存在着答案。但是，答案未必是能立刻引导出的。这一点同样适用于人生。”

—— 东野圭吾《盛夏的方程式》

5.1. The dimeric histone HBb from *Bdellovibrio bacteriovorus*

Bacterial histone homologs have mainly been identified from environmental metagenomes or in marine bacterial lineages, with limited structural and functional characterization (20) *B. bacteriovorus*, a genetically tractable microorganism suitable for laboratory cultivation, offers an opportunity to study its function *in vivo*. Therefore, HBb (UniProtKB accession number: Q6MRM1), the histone homolog from *B. bacteriovorus*, was selected for detailed experimental characterization.

5.1.1. The dimeric histone HBb binds DNA *in vitro*

Recombinant HBb was heterogeneously expressed in *E. coli* Mutant56(DE3) and purified to homogeneity. SDS-PAGE analysis showed a single band at about 10 kDa, in line with the theoretical molecular weight (MW) of 7 kDa of HBb (Fig. 5.1B). BN-PAGE revealed two bands, suggesting that HBb exists in two different oligomeric states (Fig. 5.1B), however, subsequent size exclusion chromatography coupled with multi-angle light scattering (SEC-MALS) analysis indicated a single dominant species with a calculated MW to be 14 ± 0.4 kDa, consistent with a dimeric state of HBb in solution (Fig. 5.1C). Circular dichroism (CD) spectroscopy confirmed a predominantly α-helical secondary structure with a melting temperature of 51.6 °C (Fig. 5.1C). The comparison of the HBb sequence with selected eukaryotic, archaeal and bacterial histones in a multiple sequence alignment (MSA) shows that HBb retains four of six DNA-binding

residues conserved in archaeal histones, implying a potential for DNA binding (Fig. 5.1A).

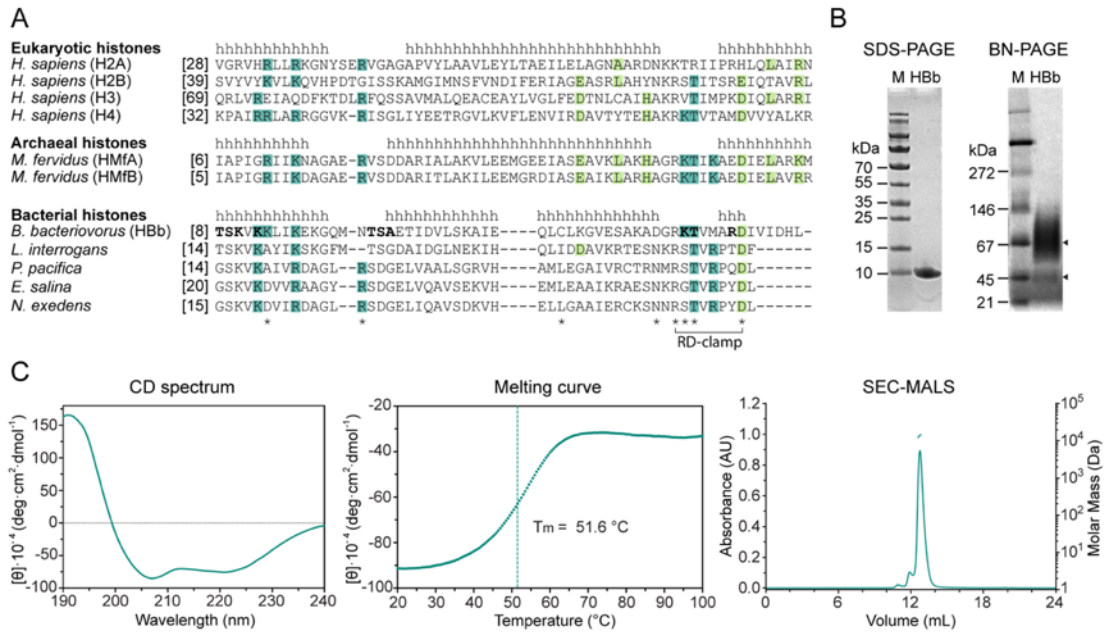


Figure 5.1 Biophysical characterization of HBb.

A. Sequence alignment of representative eukaryotic, archaeal, and bacterial histones from *Homo sapiens* (UniProtKB/Swiss-Prot entry for H2A: P04908; H2B: P62807; H3: P68431; H4: P62805), *Methanothermus fervidus* (HMfA: P48781; HMfB: P19267), *Bdellovibrio bacteriovorus* (Q6MRM1), *Leptospira interrogans* (Q8F3E8), *Plesiocystis pacifica* (A6GF99), *Enhygromyxa salina* (NCBI entry WP_052547863.1), and *Nannocystis exedens* (WP_096326703.1). α-Helices are labelled with "h" according to the crystal structures of H2A, HMfB and HBb, respectively. Conserved residues involved in DNA binding and tetramerization are highlighted in dark and light green, respectively. HBb residues involved in DNA backbone interaction according to the solved structures HBb-DNA_1 (PDB: 9EZZ) and HBb-DNA_2 (PDB: 9F0E) are in bold. Residues represented as sticks in Fig. 5.3 A and B are labelled with *. Residues R51 and D58, forming the RD clamp in HBb, are connected by a bracket. **B.** SDS-PAGE and BN-PAGE showing purified HBb. Triangles mark HBb on BN-PAGE. **C.** Biophysical characterization of HBb by CD spectroscopy (left: single CD spectrum, middle: thermal melting curve) and SEC MALS (right) for determination of secondary structure, thermal stability and oligomeric state. *This figure and the figure description correspond to Fig. 1 in Hu et al., 2024 (22), as shown in Appendix I.*

To assess the DNA-binding ability of HBb, electrophoretic mobility shift assay (EMSA) was performed using the 80 bp DNA substrate previously shown to bind archaeal histone HMfB with high affinity (101). Incubation with HBb resulted in a mobility shift, indicating the formation of HBb-DNA complexes, though the smeared bands may suggest a dynamic equilibrium between different binding stoichiometries or

conformational states (Fig. 5.2A). The DNA-binding affinity of HBb was further quantified by microscale thermophoresis (MST), using the 80 bp DNA in comparison to a nonspecific DNA fragment with 50% GC content (80 bp-GC50) (Fig. 5.2B and Table S2 in Appendix I). HBb exhibited micromolar affinity for both tested DNA substrates, comparable to HMfB (101,102).

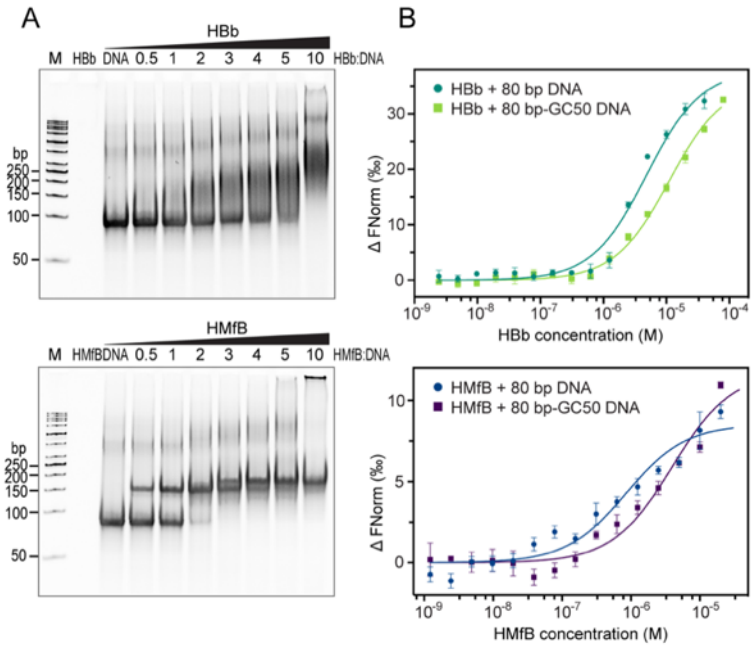


Figure 5.2 HBb binds DNA.

A. EMSA with the 80 bp DNA fragment in the presence of HBb and HMfB as a control. The molar ratios of protein:DNA in lanes 4-10 are indicated. **B.** MST curves showing binding of HBb to the 80 bp DNA fragment and the 80 bp-GC50 DNA fragment (10 nM), as compared to HMfB. *This figure and the figure description correspond to Fig. 2A and B in Hu et al., 2024 (22), as shown in Appendix I.*

5.1.2. Crystal structures of HBb in free and DNA-bound forms

To elucidate the structural basis for DNA interaction, HBb was set up for crystallization. Crystals were obtained within a few days, and the best dataset was processed to a resolution of 1.06 Å (Table S3 in Appendix I). The structure of HBb was solved via molecular replacement (MR) using the archaeal histone HMk (PDB: 1F1E) as the search model. The asymmetric unit (ASU) contains an HBb monomer that assembles into a dimer through crystallographic symmetry. The HBb monomer displays the histone fold, comprising three α-helices (α1, α2, and α3) linked by two short strap loops

(l1 and l2) (Fig. 5.3A and B). Compared with HMfB, HBb features a truncated α2-helix and a very short α3-helix, yet retains the conserved RD clamp formed between R51 and D58, stabilizing the histone fold. An additional salt bridge between D49 and R51 further stabilizes the fold. Two HBb monomers dimerize in a histone-typical "hand-shaking" manner, with the α2-helices antiparallel aligned at an angle of about 25° (Fig. 5.3A). The dimerization is mediated primarily by hydrophobic interactions, with additional contributions from residues located in α1 and α3.

To capture the HBb in complex with DNA, co-crystallization of HBb with a 20 bp nonspecific DNA fragment with 50% GC content (20 bp-GC50) yielded two distinct crystal forms within a few days. The aquired datasets of HBb-DNA_1 and HBb-DNA_2 were processed to a resolution of 1.95 Å and 1.85 Å, respectively (Table S3 in Appendix I). The two structures were solved by MR using the structure of DNA-free HBb as a search model. Both structures contain one HBb dimer and a short DNA segment in the ASU.

In the HBb-DNA_1 structure, in addition to the HBb dimer, a 5-bp double-stranded DNA (dsDNA) segment spans the ASU and forms a continuous dsDNA helix via crystallographic symmetry (Fig. 5.3C). The DNA interacts with residues K12 (α1), T23 and S24 (l1), A25 (α2) from one monomer, and K52, T53 (l2) and R57 (α3) from the opposing monomer. This binding mode closely resembles the structure of Bd0055-DNA reported by Hocher *et al.* (PDB: 8FW7), although the crystal packing differs in the arrangement of HBb dimers (103).

The HBb-DNA_2 structure uncovers a complementary DNA-binding interface, which was not observed previously. Apart from the HBb dimer, the ASU contains a 9-nt single-stranded DNA (ssDNA) fragment forming a dsDNA through crystallographic symmetry (Fig. 5.3D). Here, DNA contacts are mediated by residues T8, S9, and K10 located at the tip of α1. The crystal packing of HBb-DNA_2, including short DNA segments, is different from that of HBb-DNA_1 (Fig. 5.3E and F).

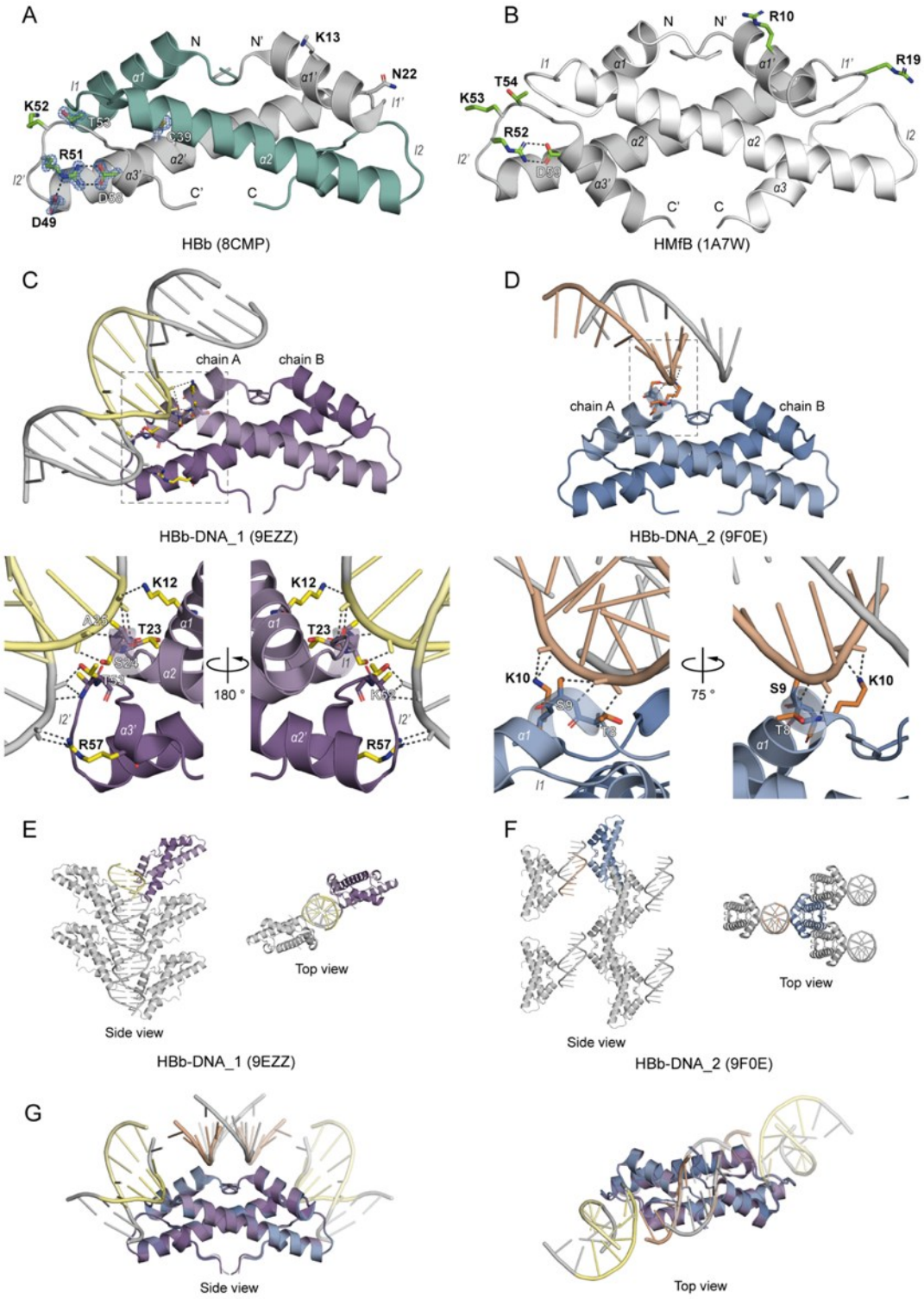


Figure 5.3 Crystal structures of HBb in its free and DNA-bound forms.

A. Crystal structure of the HBb dimer (PDB: 8CMP) in cartoon representation. Selected residues are shown as sticks, and conserved residues predicted to be involved in DNA binding or stabilization of the histone fold are in green. The salt bridges formed between D49, R51, and D58 are indicated as dashed lines. A representative 2Fo – Fc electron density map is shown for selected residues at a contour level of 2.0 σ. **B.** Crystal structure of the HMfB dimer (PDB: 1A7W) in cartoon representation for comparison. Conserved residues playing a role in DNA

binding or stabilization of the histone fold are depicted as sticks. **C.** Crystal structure of HBb-DNA_1 (PDB: 9EZZ) in cartoon representation with residues involved in DNA binding shown as sticks. The dashed frame marks the enlarged image section. **D.** Crystal structure of HBb-DNA_2 (PDB: 9F0E) in cartoon representation with DNA binding residues shown as sticks. The enlarged image section is marked by a dashed square. **E.** Crystal packing of HBb-DNA_1 with selected symmetry mates generated within 4 Å. **F.** Crystal packing of HBb-DNA_2 with selected symmetry mates generated within 20 Å. **G.** Superposition of structures HBb-DNA_1, HBb-DNA_2, and their symmetrical mates, with DNA fragments trimmed for visualization. In all panels, the contents of a single asymmetric unit are shown in color and selected symmetry mates in gray. Hydrogen bonds and salt bridges are indicated as dashed lines. *This figure and the figure description correspond to Fig. 3 in Hu et al., 2024 (22), as shown in Appendix I.*

Given the structures of HBb-DNA_1 and HBb-DNA_2 show complementary interfaces, their superposition elucidates a contiguous DNA-binding interface across the surface of the HBb dimer (Fig. 5.3G). The overlap of these two structures is similar to histone-DNA complexes based on eukaryotic or archaeal histones, implying that HBb dimers can bend DNA upon binding.

5.1.3. *In vitro* and *in vivo* functional analysis of HBb

For comprehensive characterization of HBb function *in vitro* and *in vivo* further experiments and analyses were performed in collaboration with Samuel Schwab (Leiden Institute of Chemistry; Centre for Microbial Cell Biology; Centre for Interdisciplinary Genome Research, Leiden University), Silvia Deiss (Dept. of Protein Evolution, MPI for Biology Tübingen) and Pedro Escudeiro (Dept. of Protein Evolution, MPI for Biology Tübingen).

S. Schwab conducted molecular dynamics (MD) simulations showing that an HBb dimer bends the 36-bp DNA fragment to an angle of about 84°, close to that of HMfB (48). By performing tethered particle motion (TPM) experiments, he confirmed that HBb compacts DNA and achieves the saturation level with an RMS value of about 110 nm, comparable to DNA benders (104-106). Furthermore, additional biochemical assays, including micrococcal nuclease (MNase) digestion assay, DNA topology assay, and ligase-mediated circularization assay, revealed that HBb does not exert DNA-

protective effects, but bends DNA, induces DNA supercoiling, and alters DNA topology *in vitro*.

To examine the physiological role of HBb, S. Deiss tried to replace the *hbb* gene in *B. bacteriovorus* by a kanamycin resistance cassette using the suicide plasmid pT18mobSacB. In several attempts, the plasmid was successfully integrated into the genome by a cross-over event, but selection for a second recombination event resulting in the loss of the *hbb* gene failed. This suggests that HBb is indispensable for the viability of *B. bacteriovorus*.

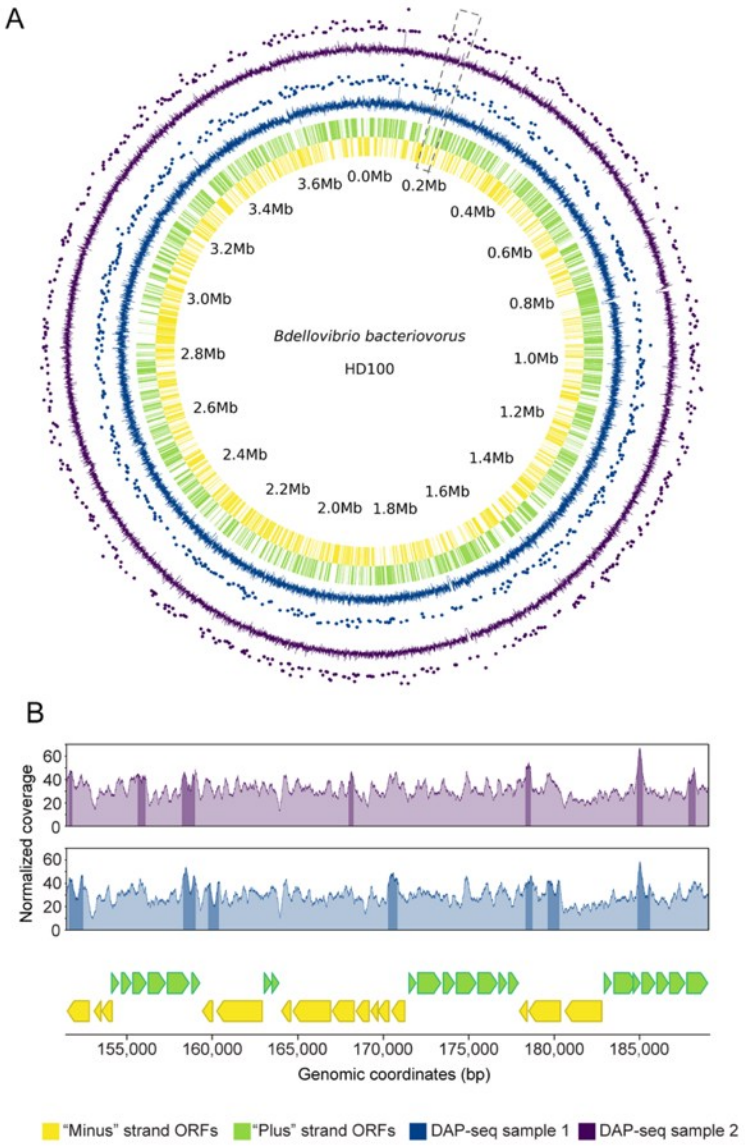


Figure 5.4 The HBb dimer binds DNA nonspecifically.

A. DAP-seq data visualized using a circos plot of the *B. bacteriovorus* HD100 genome. The inner rug plots represent the "minus" (yellow) and "plus" (light green) strand ORFs. Each wedge represents the coordinates of a gene. Outer line plots represent the normalized read coverage values, and scatter plots above show the major peaks for the two DAP-seq samples (blue and purple). Peaks are shown as dots, where the strength of the peak signal correlates positively with the distance from the axis. The enlarged section shown in panel E is indicated by gray dashed lines. **B.** Line plots showing the normalized read coverage values for a randomly selected section of the *B. bacteriovorus* HD100 genome encompassing the genomic region from 151,464 bp to 189,033 bp. The bottom plot shows the genomic context, and the top line plots depict the normalized read coverage values for each DAP-seq sample with the major peaks shaded darker. The color code corresponds to panel A. *This figure and the figure description correspond to Fig. 6 in Hu et al., 2024 (22), as shown in Appendix I.*

Next, DNA affinity purification sequencing (DAP-seq) was performed to evaluate the DNA-binding specificity of HBb. The purified FLAG-tagged HBb was immobilized on anti-FLAG magnetic beads and incubated with a genomic *B. bacteriovorus* DNA library. The protein-bound DNA fragments were recovered, size-selected, amplified, and sequenced. Samples performed in the absence of FLAG-HBb were used as controls. The plot of the sequencing data against the chromosome map showed that HBb binds nonspecifically across the *B. bacteriovorus* genome (Fig. 5.4). The peak-calling analysis, performed by P. Escudeiro, revealed about 465 significant HBb-binding sites of various lengths with a GC content of about 50%, which is close to the genomic GC content of *B. bacteriovorus*, indicating that HBb has no sequence specificity or preference (Fig. S4 and S5 in Appendix I).

5.2. The tetrameric histone HLp from *Leptospira perolatii*

Bioinformatic studies conducted in parallel with this doctoral thesis increasingly revealed that bacterial histones can be assigned to specific groups based on their sequence conservation. In order to further deepen our understanding of the structural and functional diversity of bacterial histones, a second histone homolog from *L. perolatii* (UniProtKB accession number: A0A2M9ZN55), a representative of the face-to-face (FtF) subfamily of α3 prokaryotic histones, was chosen for structural and functional characterization (53).

5.2.1. HLp forms tetramers

According to the bioinformatic analyses performed by Vikram Alva (Dept. of Protein Evolution, MPI for Biology Tübingen), the bacterial FtF cluster is adjacent to the cluster of its archaeal counterpart, implying potential shared structural features (Fig. 5.5A). The MSA shows that HLp retains three out of six conserved residues that are involved in DNA binding in HMfB, indicating potential DNA-binding capability (Fig. 5.5B).

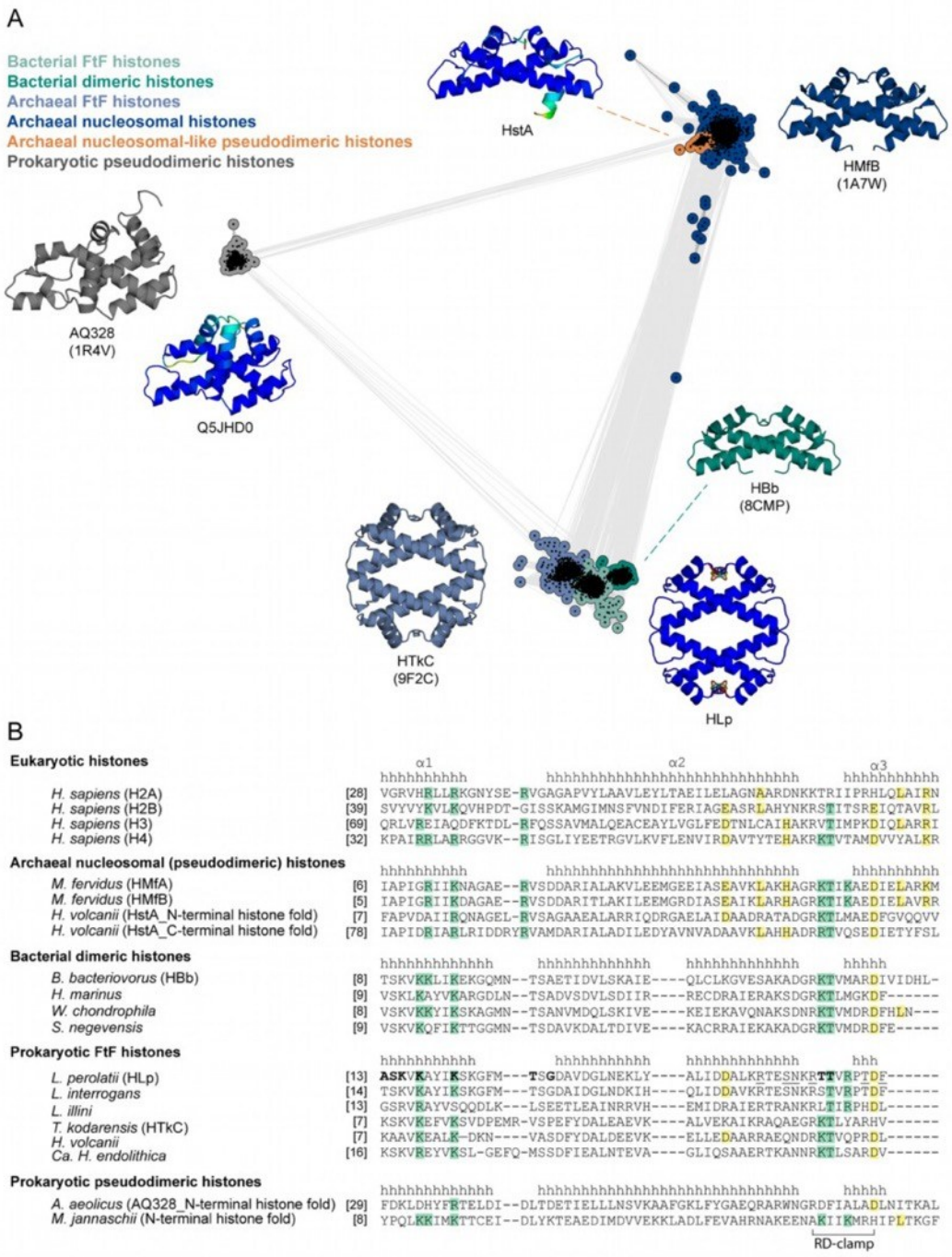


Figure 5.5 Comparative analysis of prokaryotic histones composed exclusively of histone fold domains.

A. Cluster map of ~3,300 prokaryotic histones containing only the histone fold domain. Each dot represents a single protein sequence, edges reflect pairwise sequence similarity. Structures of representative proteins — either crystal structures (PDB IDs in parentheses) or AlphaFold models — are shown in cartoon representation. AlphaFold models are colored by pLDDT confidence scores. Representative proteins include: HstA from *Haloferax volcanii* (D4GS56), an archaeal nucleosomal-like pseudodimeric histone; HMfB from *Methanothermus fervidus* (P19267), an archaeal nucleosomal histone; HBb from *Bdellovibrio bacteriovorus* (Q6MRM1), a bacterial dimeric histone; HLp from *Leptospira perolatii* (A0A2M9ZN55), a bacterial FtF histone; HTkC from *Thermococcus kodakarensis* (Q5JDW7), an archaeal FtF histone; and Q5JHD0 from *T. kodakarensis* (Q5JHD0) and AQ328 from *Aquifex aeolicus* (O66665), both prokaryotic pseudodimeric histones. **B.** Multiple sequence alignment of representative histones from both eukaryotes and prokaryotes, including *Homo sapiens* (H2A: P04908; H2B: P62807; H3: P68431; H4: P62805), *M. fervidus* (HMfA: P48781; HMfB: P19267), *H. volcanii* (HstA: D4GS56), *B. bacteriovorus* (HBb: Q6MRM1), *Hymenobacter marinus* (NCBI accession No.: ATH09486), *Waddlia chondrophila* (D6YWW1), *Simkania negevensis* (F8L7X8), *L. perolatii* (HLp: A0A2M9ZN55), *L. interrogans* (Q8F3E8), *Leptonema illini* (H2CFS2), *T. kodakarensis* (HTkC: Q5JDW7), *H. volcanii* (D4GZE0), Candidatus *Heimdallarchaeum endolithica* (A0A9Y1FPJ9), *A. aeolicus* (O66665), and *Methanocaldococcus jannaschii* (A0A832T4V6). α-helices are annotated with "h" based on crystal structures of H2A, HMfA, HBb, HLp, and AQ328. Conserved residues associated with DNA binding and oligomerization are highlighted in green and yellow, respectively. HLp residues contributing to tetramerization in the HLp crystal structure (PDB: 9QT0) are underlined, while those interacting with the DNA backbone in HLp-DNA complexes (PDB: 9QT1 and 9QT2) are shown in bold. Unless otherwise stated, accession numbers in parentheses refer to UniProtKB. *This figure and the figure description correspond to Fig. 1 in Hu et al., 2025 (107), as shown in Appendix II.*

Since the *hlp* gene is annotated with two translation start sites in available databases, the Shine-Dalgarno sequence was identified and the start methionine was determined accordingly (Fig. S1A in Appendix II). HLp was heterologously expressed in *E. coli* BL21(DE3), and purified to homogeneity. The purity of the protein was confirmed by SDS-PAGE showing a single band of about 10 kDa, matching the HLp theoretical MW of 7.1 kDa (Fig. S2A in Appendix II). Interestingly, the SEC-MALS analysis revealed a predominant species with a MW of 26 ± 0.3 kDa, consistent with a tetrameric assembly of HLp (Fig. S2B in Appendix II). CD spectroscopy confirmed that HLp adopts an α-helical structure and undergoes thermal unfolding with a melting temperature of approximately 57 °C (Fig. S2C and D in Appendix II).

5.2.2. Crystal structure of HLp

Crystallization of purified HLp yielded crystals under several conditions within a few days. The best dataset could be processed to a resolution of 1.30 Å (Table S3 in Appendix II), and the structure was solved by MR using an AlphaFold predicted HLp structure as the search model (Fig. 5.6A) (108). The solved crystal structure of HLp closely matches its AlphaFold prediction, with an RMSD of 0.907 Å.

The ASU of HLp contains a dimer that forms a tetramer through crystallographic symmetry. Each HLp monomer displays the typical histone fold, consisting of three α-helices (α1, α2, and α3) joined by two short strap loops (l1 and l2) (Fig. 5.6A). Like the dimeric histone HBb, the second α-helix of HLp is one turn shorter than that of the archaeal histone HMfB, and the third α-helix comprises only one helical turn. HLp also retains the conserved RD clamp formed by salt bridges between R55 and D62 (Fig. 5.5B and 5.6A).

HLp monomers dimerize in the histone-typical “hand-shaking” configuration, with the second α-helices antiparallelly crossed at an angle of about 40°. The dimerization interface is primarily mediated by hydrophobic interactions, contributed by residues V16 and I20 (α1), L35, L39, L42, and L47 (α2), and F63 (α3), with additional hydrogen bonds formed between residues located in l1 of one monomer and l2 of the opposing monomer (Fig. 5.6B).

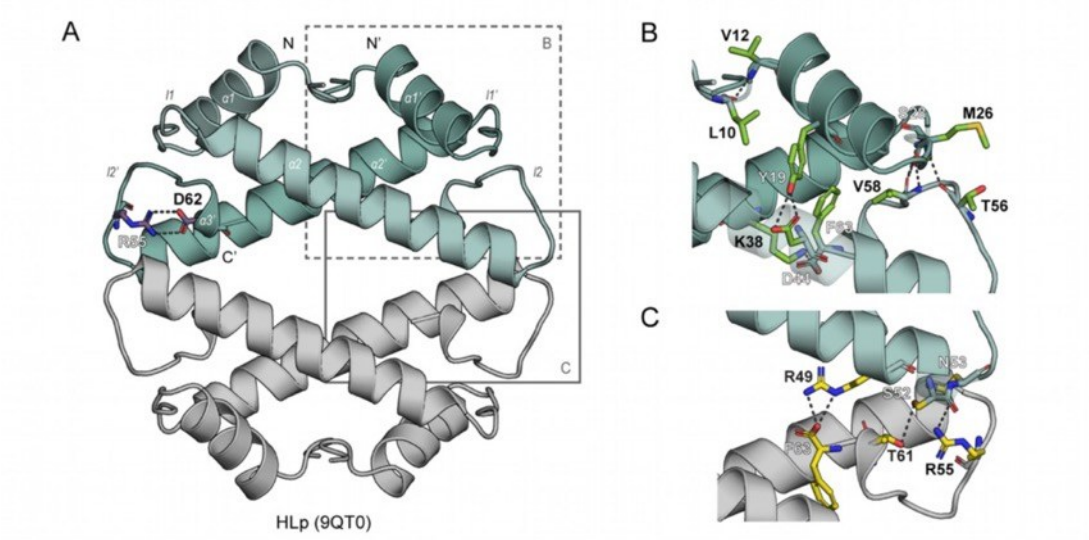


Figure 5.6 Crystal structure of the HLp tetramer.

A. Crystal structure of the HLp tetramer (PDB: 9QT0) shown in cartoon representation. Residues R55 and D62, which form the RD clamp, are shown as sticks. Salt bridges are indicated as dashed lines. **B.** Close-up view of the HLp dimerization interface. Residues involved in dimerization are shown as green sticks, with hydrogen bonds depicted as dashed lines. **C.** Close-up view of the HLp tetramerization interface. Tetramerization-associated residues are shown as yellow sticks. Salt bridges and hydrogen bonds are indicated by dashed lines. In all panels, the contents of the asymmetric unit are shown in color, with selected symmetry mates in gray. *This figure and the figure description correspond to Fig. 2 in Hu et al., 2025 (107), as shown in Appendix II.*

Tetramerization of HLp is facilitated by polar contacts between R49, S52, and N53 (α2) of one monomer, and F63 (C-terminus), T61 (α3), and R55 (l2) of the opposing monomer (Fig. 5.6C). Particularly, the amino acid residues R49, N53, and R55 are highly conserved among the bacterial FtF histone homologs, indicating a shared oligomerization mechanism (Fig. 5.5B). This tetramerization mode is essentially different from the spring-like assembly observed in archaeal nucleosomal histones such as HMfB upon DNA binding (48).

5.2.3. HLp binds non-specifically to DNA *in vitro*

Given the presence of conserved DNA-binding residues and surface-distributed positive charges, HLp was hypothesized to bind DNA. EMSA of HLp with the 80 bp DNA fragment and 30 bp DNA fragments of varying GC content confirmed that HLp binds nonspecifically to dsDNA, with a preference for fragments having a GC content of 40% (Fig. 5.7A, 5.7B, and Fig. S5 in Appendix II). The DNA-binding affinity was further quantified by MST using the 80 bp DNA and a nonspecific DNA fragment of 40% GC content (80 bp-GC40), revealing that HLp has a DNA binding capability comparable to that of HMfB (Fig. 5.7C and Table S2 in Appendix II). SEC-MALS analysis of HLp in the presence of the 30 bp-GC40 DNA fragment revealed an additional peak at 69.2 ± 5.7 kDa, indicating the formation of HLp-DNA complexes composed of an HLp tetramer and two DNA fragments (Fig. 5.7D).

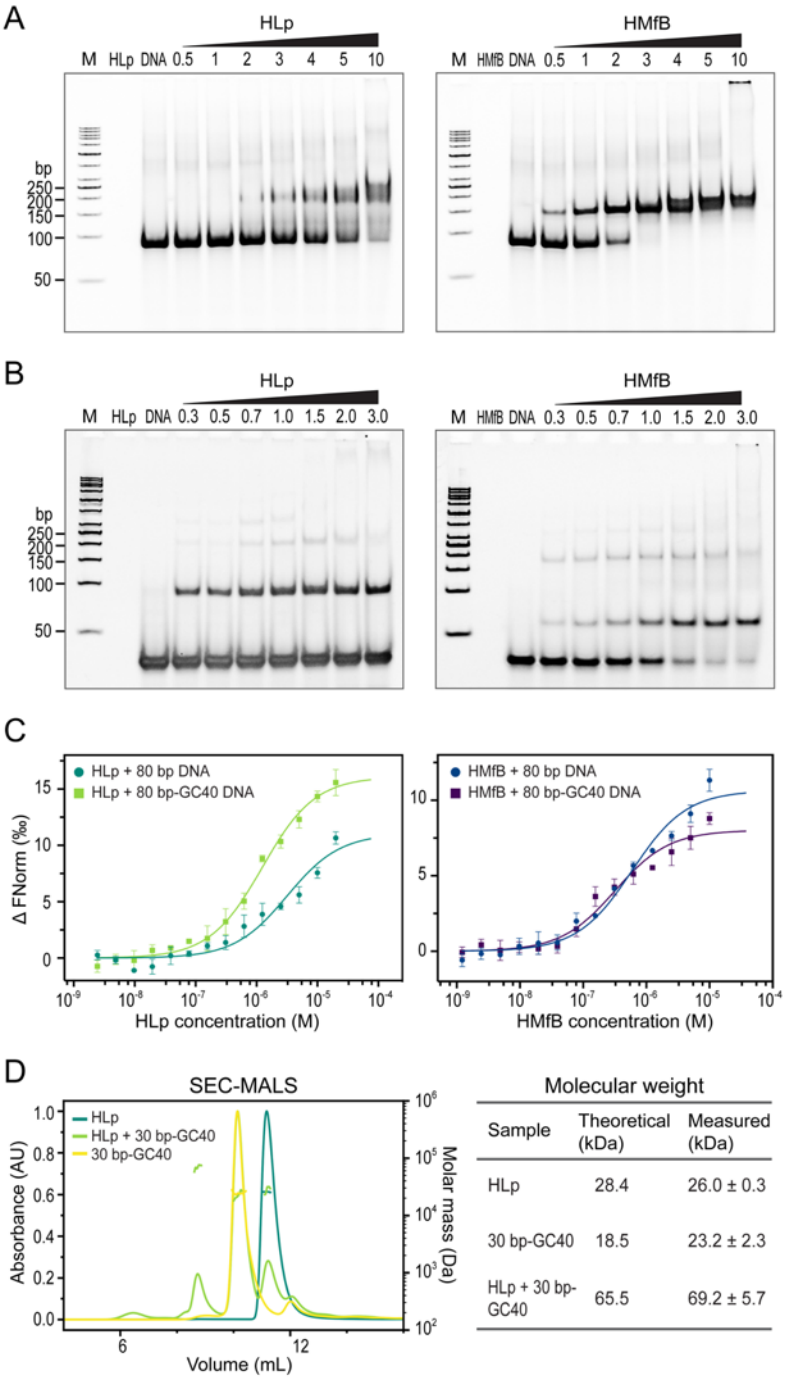


Figure 5.7 HLp binds DNA *in vitro*.

EMSAs showing the binding of HLp and HMfB (control) to the 80 bp DNA fragment (**A**) and the 30 bp-GC40 DNA fragment (**B**). Increasing protein concentrations (lanes 4-10), indicated as molar protein-to-DNA ratios, were incubated with the corresponding DNA and analyzed on a 6% polyacrylamide gel. **C.** MST measurements comparing the binding of HLp to the 80 bp and the 80 bp-GC40 DNA fragments, alongside HMfB. **D**. Profiles of SEC-MALS runs of HLp, the 30 bp-GC40 DNA, and a mixture of both; the table summarizes both theoretical and experimentally determined molecular weights. *This figure and the figure description correspond to Fig. 3 in Hu et al., 2025 (107), as shown in Appendix II.*

5.2.4. Crystal structures of HLp-DNA complexes

To elucidate HLp-DNA interactions, purified HLp was crystallized together with the 30 bp-GC40 DNA fragment. Two diffraction datasets HLp-DNA_1 and HLp-DNA_2 were collected and processed to resolutions of 2.10 Å and 1.90 Å, respectively (Table S3 in Appendix II). Two HLp-DNA structures were solved by MR using the free HLp structure as the search model. In both of the two crystal forms, the DNA fragments appear to form endless double helices extending through the crystal lattice, illustrating two distinct DNA-binding interfaces (Fig. 5.8A and B).

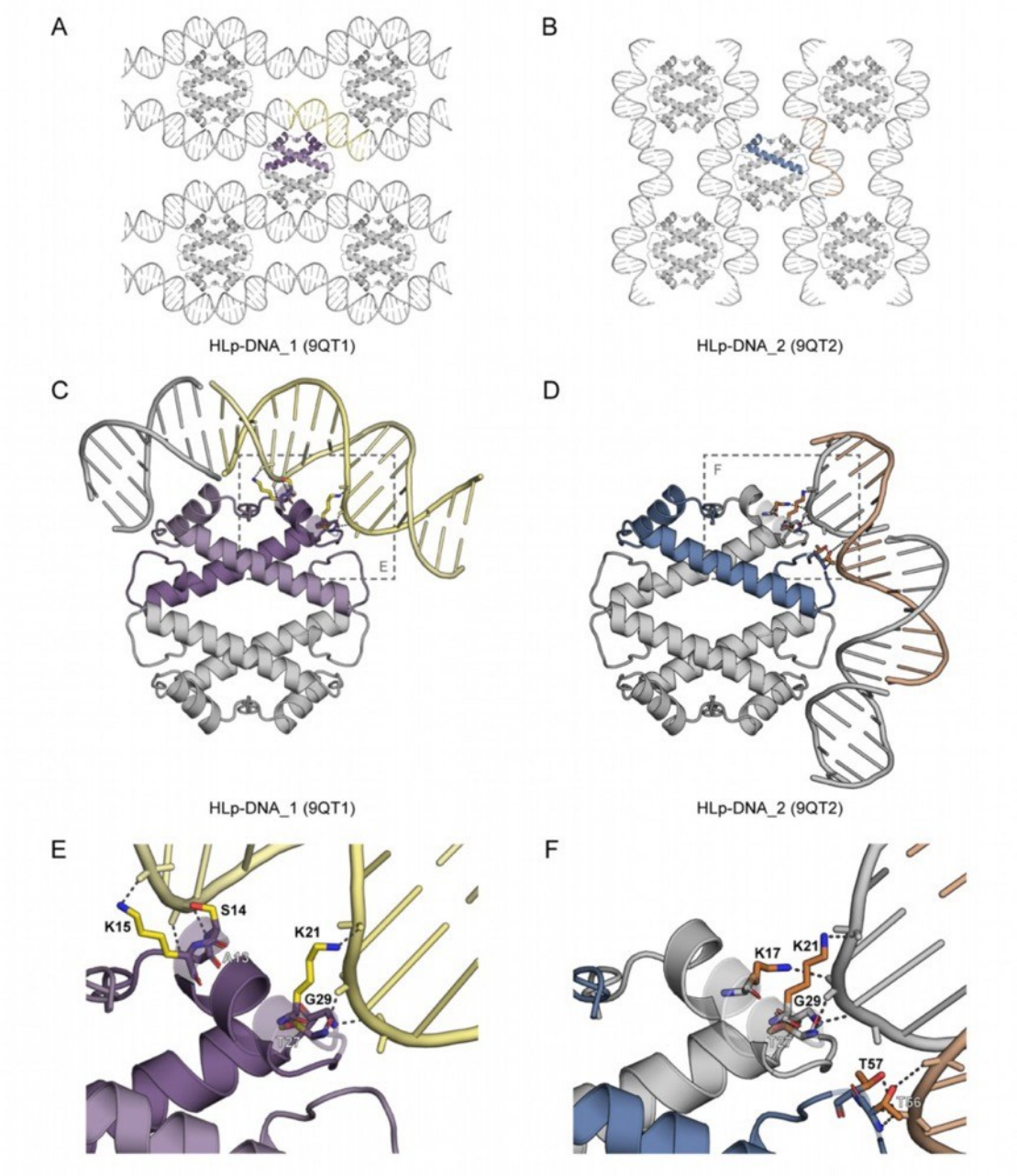


Figure 5.8 Crystal structures of DNA-bound HLp.

A, B. Crystal packing of HLp-DNA_1 (**A**) and HLp-DNA_2 (**B**) showing selected symmetry mates within 20 Å. **C, D.** Crystal structures of HLp-DNA_1 (PDB: 9QT1) (**C**) and HLp-DNA_2 (PDB: 9QT2) (**D**) shown in cartoon representation. The framed regions in panels **C** and **D** correspond to the magnified views shown in **E** and **F**, respectively. **E, F.** Residues involved in DNA binding are shown as sticks, and protein-DNA interactions are depicted as dashed lines. In all panels, the contents of the asymmetric unit are shown in color, and selected symmetry mates in gray. *This figure and the figure description correspond to Fig. 4 in Hu et al., 2025 (107), as shown in Appendix II.*

In HLp-DNA_1, the ASU consists of one HLp dimer, which forms tetramers via crystallographic symmetry, and a 16-bp dsDNA segment of the apparently continuous DNA helices (Fig. 5.8A). Similar to the HBb-DNA complexes, the interaction is facilitated predominantly by contacts between basic or polar side chains of HLp residues and the DNA backbone. Here, residues A13, S14, K15, and K21 (α1), as well as T27 and G29 (l1) of each HLp monomer are involved in DNA binding (Fig. 5.8C and E). In HLp-DNA_2, the ASU comprises an HLp monomer, which assembles into tetramers by crystallographic symmetry, and a 15-nt ssDNA segment of the seemingly infinite DNA (Fig. 5.8B). The residues T27 and G29 (l1), T56 and T57 (l1), as well as K17 and K21 (α1), are involved in DNA binding (Fig. 5.8D and F).

Collectively, these structures of the HLp-DNA complexes uncovered a continuous DNA-binding interface covering the entire ring surface of the HLp tetramer, suggesting a potential DNA-wrapping capacity.

5.2.5. *In vitro* functional characterization of HLp

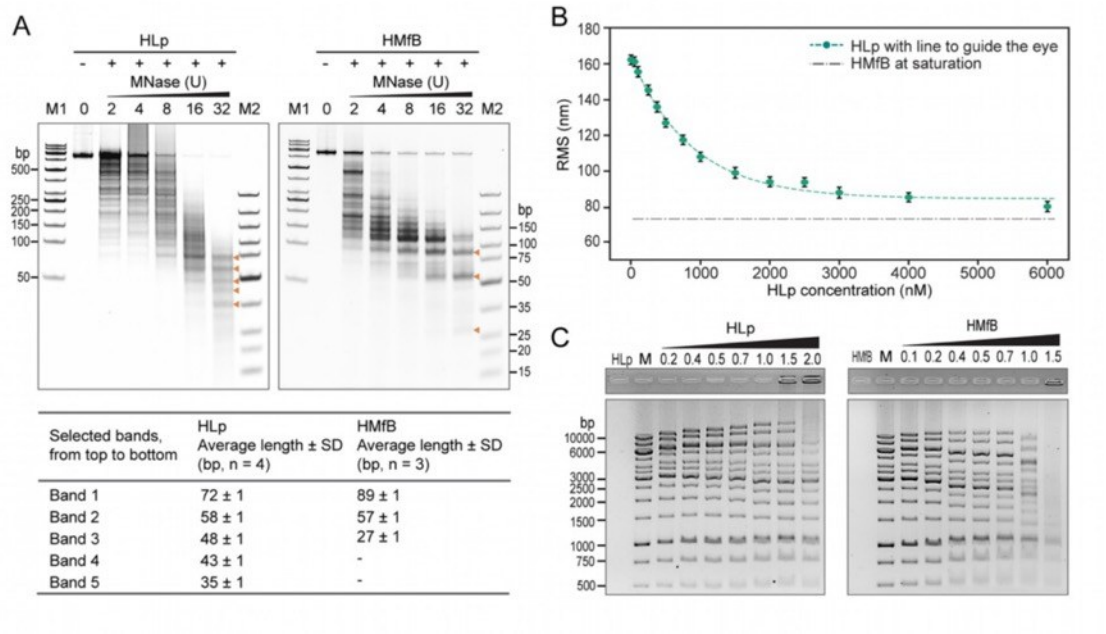
A comprehensive biophysical and biochemical characterization of HLp was performed in collaboration with Kaiyu Qiu (Dept. of Protein Evolution, MPI for Biology Tübingen) and Yunsen Zhang (Theoretical and Computational Biophysics Group, Beckman Institute for Advanced Science and Technology, Center of Biophysics and Quantitative Biology, University of Illinois Urbana-Champaign) performing MD simulations and S. Schwab (Leiden Institute of Chemistry; Centre for Microbial Cell Biology; Centre for Interdisciplinary Genome Research, Leiden University) conducting TPM assays.

While the crystal packing of the HLp-DNA structures suggests that HLp might bridge DNA, the superposition of the two structures strongly supports the possibility of DNA wrapping. To examine the relative stability of both models and potential interconversion, K. Qiu and Y. Zhang conducted all-atom MD simulations using starting models constructed based on the HLp-DNA crystal structures. The results show that both bridging and wrapping modes are theoretically possible, with the DNA-wrapping mode being energetically preferable.

A variety of well-established biophysical and biochemical assays was performed to further explore the function of HLp. The MNase digestion test of HLp using the 600 bp DNA fragment showed a ladder-like digestion pattern, indicative of periodic DNA protection akin to HMfB. However, the HLp-protected DNA fragment varied in length (35-72 bp), contrasting with the 30-bp interval of HMfB (Fig. 5.9A). The TPM assay using the 685 bp DNA fragment, conducted by S. Schwab, showed that HLp can compact DNA to saturation level, with an RMS value of about 80 nm, similar to HMfA and HMfB (Fig. 5.9B) (106,109). Notably, full DNA compaction required high HLp concentrations (6000 nM).

To investigate whether HLp can introduce DNA supercoiling and condensation, ladder EMSA assay, DNA topology assay, and ligase-mediated circularization assay were performed. The results of the ladder EMSA and DNA topology assay showed that HLp binding affects DNA conformation and supercoiling, but its effect is not as obvious as that of HMfB (Fig. 5.9C and Fig. S14A in Appendix II). In the ligase-mediated circularization assay, HLp favored the formation of linear DNA multimers over the formation of circular DNA monomers, suggesting the assembly of open-ended protein-DNA complexes (Fig. S14B in Appendix II).

Collectively, these observations suggest that HLp wraps and compacts DNA via a mechanism that fundamentally differs from the wrapping mode employed by archaeal nucleosomal histones and the bending mode applied by the bacterial histone HBb.



Selected bands, from top to bottom	HLp Average length ± SD (bp, n = 4)	HMfB Average length ± SD (bp, n = 3)
Band 1	72 ± 1	89 ± 1
Band 2	58 ± 1	57 ± 1
Band 3	48 ± 1	27 ± 1
Band 4	43 ± 1	-
Band 5	35 ± 1	-

Figure 5.9 DNA wraps around HLp.

A. MNase digestion assay analyzing the protection of the 600-bp-GC40 DNA fragment by HLp in comparison to HMfB. Increasing MNase concentrations used in the assay are indicated. The lengths of DNA fragments obtained at the highest MNase concentration were determined by densitometric analysis, using the molecular weight marker M2 (GeneRuler Ultra Low Range DNA Ladder, ThermoFisher Scientific) as the standard. **B.** TPM experiment with HLp and the 685-bp DNA. Each measurement point represents the average of triplicate measurements, with error bars indicating standard deviations. The line connecting the points was generated by fitting the means to a logistic function. **C.** EMSA analyzing the binding of HLp and HMfB (control) to the DNA fragments of the GeneRuler 1 kb Ladder (ThermoFisher Scientific). Increasing protein concentrations (lanes 3-9), indicated as protein-to-DNA mass ratios, were incubated with the corresponding DNA and analyzed on a 1% agarose gel. *This figure and the figure description correspond to Fig. 6 in Hu et al., 2025 (107), as shown in Appendix II.*

5.2.6. DNA binding of HLp *in vivo*

Because of the pathogenicity of *Leptospira* strains, HLp was recombinantly expressed in *E. coli* to characterize its function *in vivo* using light microscopy. HLp was overexpressed in *E. coli* BL21(DE3), and the archaeal histone HMfB was overexpressed in *E. coli* Mutant56(DE3) as the positive control. Ubiquitin from *Caldiarchaeum subterraneum* (*Cs*Ub), a non-DNA-binding protein overexpressed in *E. coli* BL21(DE3), and uninduced HLp-expressing *E. coli* BL21(DE3) served as negative controls (110).

To evaluate possible effects of HLp overexpression on cell growth and determine the onset of the stationary phase, the optical density at the wavelength of 600 nm of each cell culture was monitored with and without isopropyl-β-d-thiogalactoside (IPTG) as expression inducing agent. The HLp- and *Cs*Ub-expressing cells exhibited a moderate growth delay and reduced final optical density compared with uninduced controls (Fig. S15 in the Appendix II). These phenomena were also observed in the sample of *E. coli* harboring the empty vector, implying that these changes were not caused by protein overexpression.

For imaging, the cells were harvested in early stationary phase to ensure the consistency of the samples and reduce nucleoid variability. Brightfield imaging and Airyscan revealed cell elongation and increased nucleoid volume in HLp-expressing cells (Fig. 5.10), which may be caused by disturbed chromosome organization and/or cell division triggered by binding of HLp to chromosomal DNA. An increase of the nucleoid volume was also observed in uninduced HLp expressing cells, which may be due to low levels of protein expression driven by the leaky T7 promoter (Fig. 5.10). Generally, the genomic DNA occupied most of the cytoplasm of the HLp-expressing cells, similar to the phenotype of HMfB-expressing cells, indicating that HLp binds to genomic DNA *in vivo* and remodels the nucleoid morphology. In contrast to the sharply defined nucleoid region of HMfB-expressing cells, DAPI staining revealed more diffuse, regionally dispersed signals in HLp-expressing cells (Fig. 5.10). Combined with the TPM result indicating that high HLp concentrations are required for full compaction, these findings suggest that HLp induces localized DNA condensation *in vivo*.

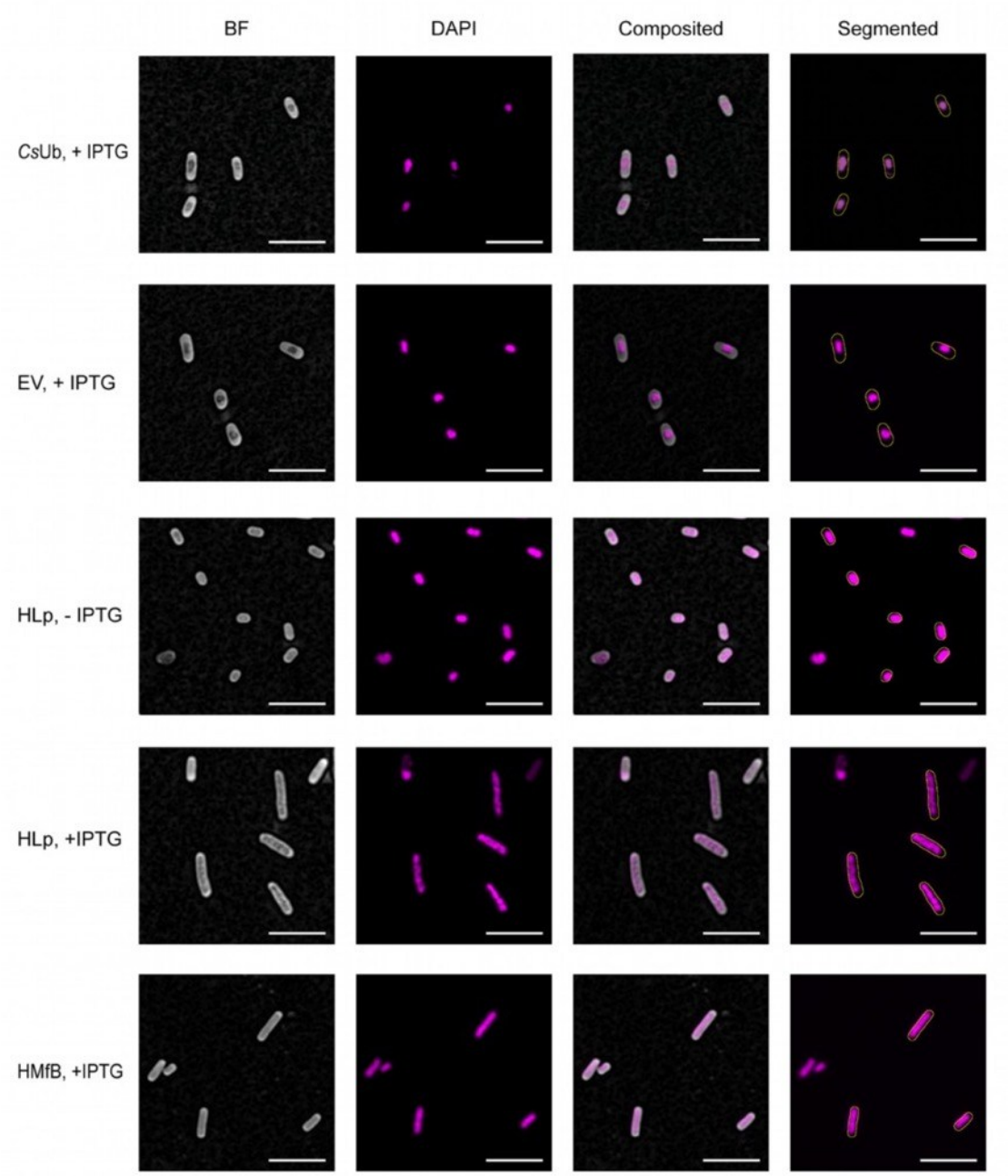


Figure 5.10 HLp binds to genomic DNA *in vivo*.

Airyscan imaging of *E. coli* cells expressing HLp and HMfB (positive control) with DAPI dye, compared to negative controls transformed with the empty vector (EV), the vector expressing *Cs*Ub, and the uninduced HLp sample (-IPTG). The left column presents the brightfield image with subtracted background. The second column displays the maximum intensity projection of fluorescence. The third one is the overlap, and the last column shows the segmented bacteria outline on top of the fluorescence channel. Scale bar: 5 µm. *This figure and the figure description correspond to Fig. 7 in Hu et al., 2025 (107), as shown in Appendix II.*

6. Discussion

“在风暴过后，过去绝望地看不见的东西，如拨云见日般清晰可见。”

—— 东野圭吾《白夜行》

As mentioned in the previous chapters, two bacterial histone candidates have been intensively characterized in this study: the histone homologs HBb found in *Bdellovibrio bacteriovorus*, and HLp identified from *Leptospira perolatii*. Although proteins with identical names already exist, such as the hemoglobin protein HBB (111), the DNA-binding protein Hbb from *Borrelia burgdorferi* belonging to the IHF/HU family (112), and the histone-like protein Hlp found in *Mycobacterium* spp. (113), the two bacterial histones characterized in this study are named HBb (Histone of *B. bacteriovorus*) and HLp (Histone of *L. perolatii*), respectively, to keep consistent with the established nomenclature for prokaryotic histones in the literature (17,44,114).

6.1. Bioinformatic classification of HBb and HLp

The recent bioinformatic investigation has discovered a broader spectrum of prokaryotic histones and histone variants than previously identified, which have been classified into 17 distinct groups based on sequence similarity and structural predictions using AlphaFold2 (53). HBb and HLp belong to the major histone family, the α3 histone group, which is characterized by a short α3 helix of only 3-4 amino acid residues in length, compared to the 10-12 amino acid residue length of the third α helix in nucleosomal histones (53). According to the predicted quaternary structure and the presence of additional domains, HBb is clustered into the dimeric histone group, mainly found across bacterial phyla, such as Elusimicrobiota, Planctomycetota, Chlamydiota, Bdellovibrionota, and Spirochaetota (53). Whereas HLp belongs to the face-to-face (FtF) histone group, the largest subgroup of prokaryotic α3 histones, which are predicted to assemble into tetramers without DNA and are distributed across most of the archaeal

phyla, and several bacterial phyla, including Myxococcota, Planctomycetota, Bdellovibrionota, and Spirochaetota (53).

Beyond the bacterial dimeric histone group and the FtF group, several other histone subgroups have been identified in prokaryotes by the bioinformatic analysis (53). One such group is the ZZ histones, which contain an N-terminal ZZ-type zinc finger domain and a C-terminal histone fold domain, predominantly found in Proteobacteria (53). The ZZ domain harbors two conserved fold domains, one composed of 4 cysteines and the other of 2 cysteines and 2 histidines, suggesting the ability to coordinate 2 zinc ions. Structurally, it resembles ZZ domains found in eukaryotic proteins, such as the histone acetyltransferase p300, the ubiquitin protein ligase HERC2, and the histone H3 reader ZZZ3, all of which play roles in histone PTMs (53,115-117). Another subgroup, the IHF histones, is featured by a histone fold followed by an IHF-like domain, predicted to assemble into dimers (53). Although relatively rare, these histone homologs have been identified in metagenomes from bacterial phyla, such as Elusimicrobiota, Wallbacteria, and Candidatus Omnitrophica (53). Despite not yet experimentally validated, the bioinformatic discovery of these diverse histone subclasses highlights the structural and sequence variability of prokaryotic histones, hinting at their potential to have a broad spectrum of functions *in vivo*.

6.2. Structures of HBb and HLp in the absence and presence of DNA

Structural determination of DNA-free HBb and HLp by crystallography has revealed that, despite their low sequence similarity to classical histones, both proteins adopt the conserved histone fold containing the characteristic RD clamp, a salt bridge, that stabilizes the second loop of the fold (Fig. 6.1). These results underscore the principle, mentioned in the introduction, that protein structures are often more conserved than sequences (2). The truncations observed in the α2 and α3 helices in HBb and HLp, compared to nucleosomal histones, are anticipated to be a hallmark of the prokaryotic α3 histone supergroup (53) (Fig. 6.1).

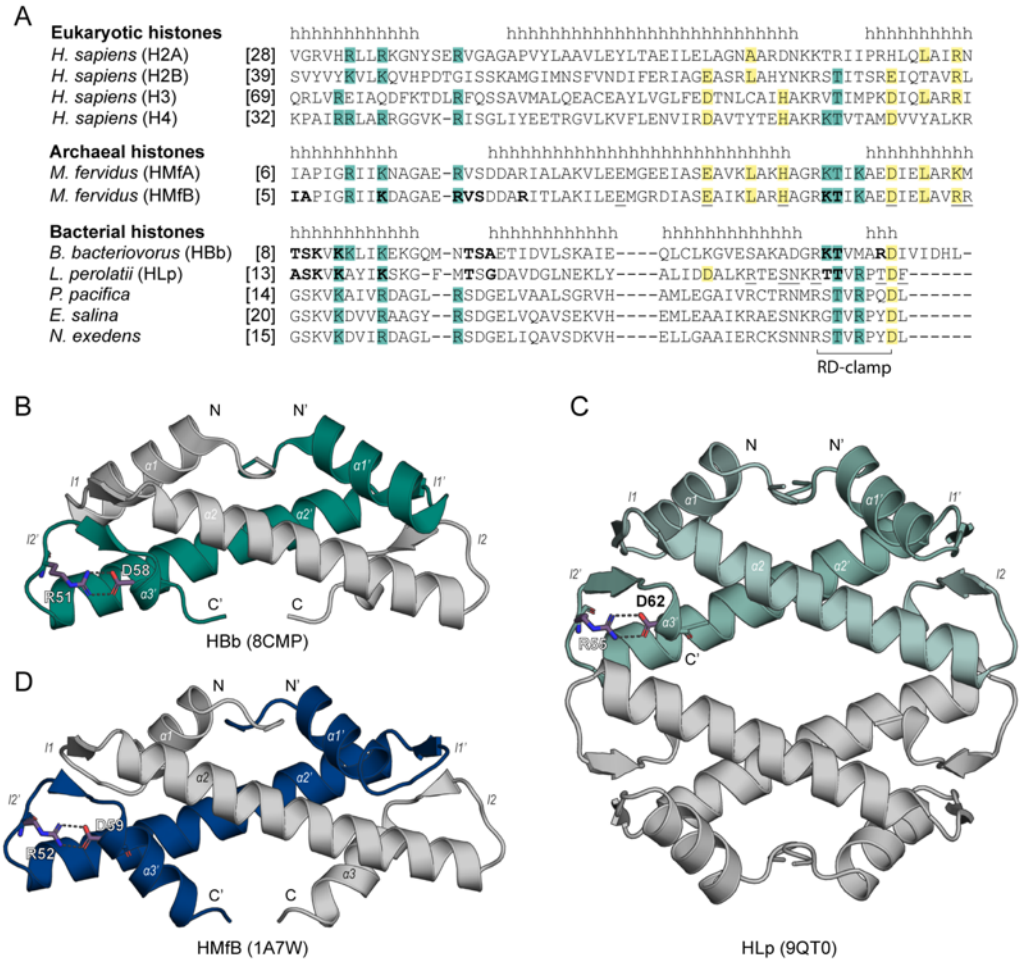


Figure 6.1 Crystal structures of DNA-free HBb and HLp in comparison with the archaeal histone HMfB.

A. Sequence alignment of representative histones and bacterial histones, including eukaryotic histones from *Homo sapiens* (UniProtKB accession number for H2A: P04908; H2B: P62807; H3: P68431; H4: P62805), archaeal histones from *M. fervidus* (HMfA: P48781; HMfB: P19267), and bacterial histone homologs from *B. bacteriovorus* (Q6MRM1), *L. perolatii* (A0A2M9ZN55), *Plesiocystis pacifica* (A6GF99), *Enhygromyxa salina* (NCBI entry WP_052547863.1), and *Nannocystis exedens* (WP_096326703.1), respectively. α-helices are marked by the "h" above the sequences based on crystal structures of H2A, HMfB and HBb, accordingly. Conserved amino acid residues predicted to participate in DNA binding and tetramerization are highlighted in green and yellow, respectively. Residues interact with DNA backbone based on the crystal structures of HMfB (PDB: 5T5K), HBb (PDB: 8CMP), and HLp (PDB: 9QT0), are in bold. Residues involved in tetramerization established on the crystal structures of HMfB (PDB: 5T5K) and HLp (PDB: 9QT0) are underlined. *This panel was adapted from the Fig. 1A in Hu et al., 2024 (22).* **B.** Crystal structure of HBb dimer (PDB: 8CMP) in cartoon representation. **C.** Crystal structure of HLp tetramer (PDB: 9QT0) in cartoon representation. **D.** Crystal structure of HMfB dimer (PDB: 1A7W) in cartoon representation. In panel B, C and D, the contents of a single asymmetric unit are indicated in colors and the selected symmetry mates are in gray. The conserved RD clamps are displayed as sticks. Salt bridges and hydrogen bonds are shown as dashed lines.

The dimerization interfaces of HBb and HLp closely resemble those of eukaryotic

and archaeal histones, although the residues facilitating dimerization are less conserved (Fig. 6.2). Crystallographic and biophysical data suggest that HBb functions as a dimer under physiological conditions, whereas HLp forms unconventional tetramers through C-terminal and α2 helix interactions, generating a torus-like structure akin to the recently characterized archaeal FtF histone HTkC from *Thermococcus kodakarensis* (Fig. 6.3) (53). Despite similar tetramerization interaction surfaces, FtF histones associate into compact tetramers, whereas archaeal histone dimers, such as those of HMfB, which have longer C-termini, assemble into open-ended spiral oligomers upon DNA binding (48,53). The truncated C-termini of prokaryotic FtF histones, as well as the conservation of specific amino acid residues involved in tetramerization, suggest that the tetrameric oligomeric state is a hallmark of this subgroup (53).

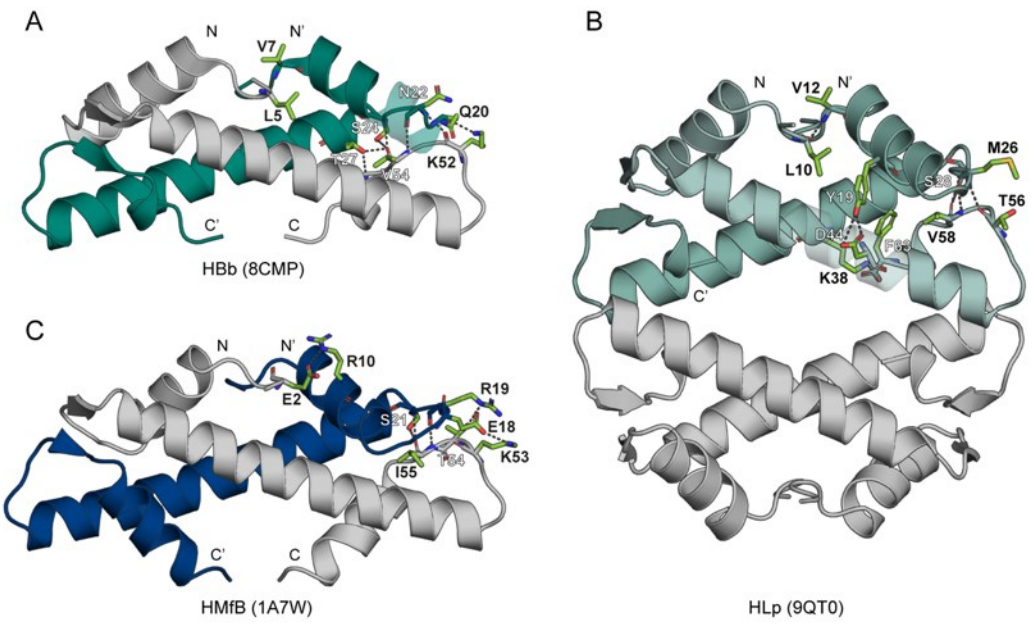


Figure 6.2 Dimerization-associated residues in HBb and HLp, compared with those in the archaeal histone HMfB.

A. Crystal structure of HBb dimer (PDB: 8CMP) in cartoon representation. **B.** Crystal structure of HLp tetramer (PDB: 9QT0) in cartoon representation. **C.** Crystal structure of HMfB dimer (PDB: 1A7W) in cartoon representation. In all panels, the amino acid residues involved in dimerization are depicted as sticks, and their side chains are colored in light green. The contents of a single asymmetric unit are indicated in colors and the selected symmetry mates are in gray. Salt bridges and hydrogen bonds are shown as dashed lines.

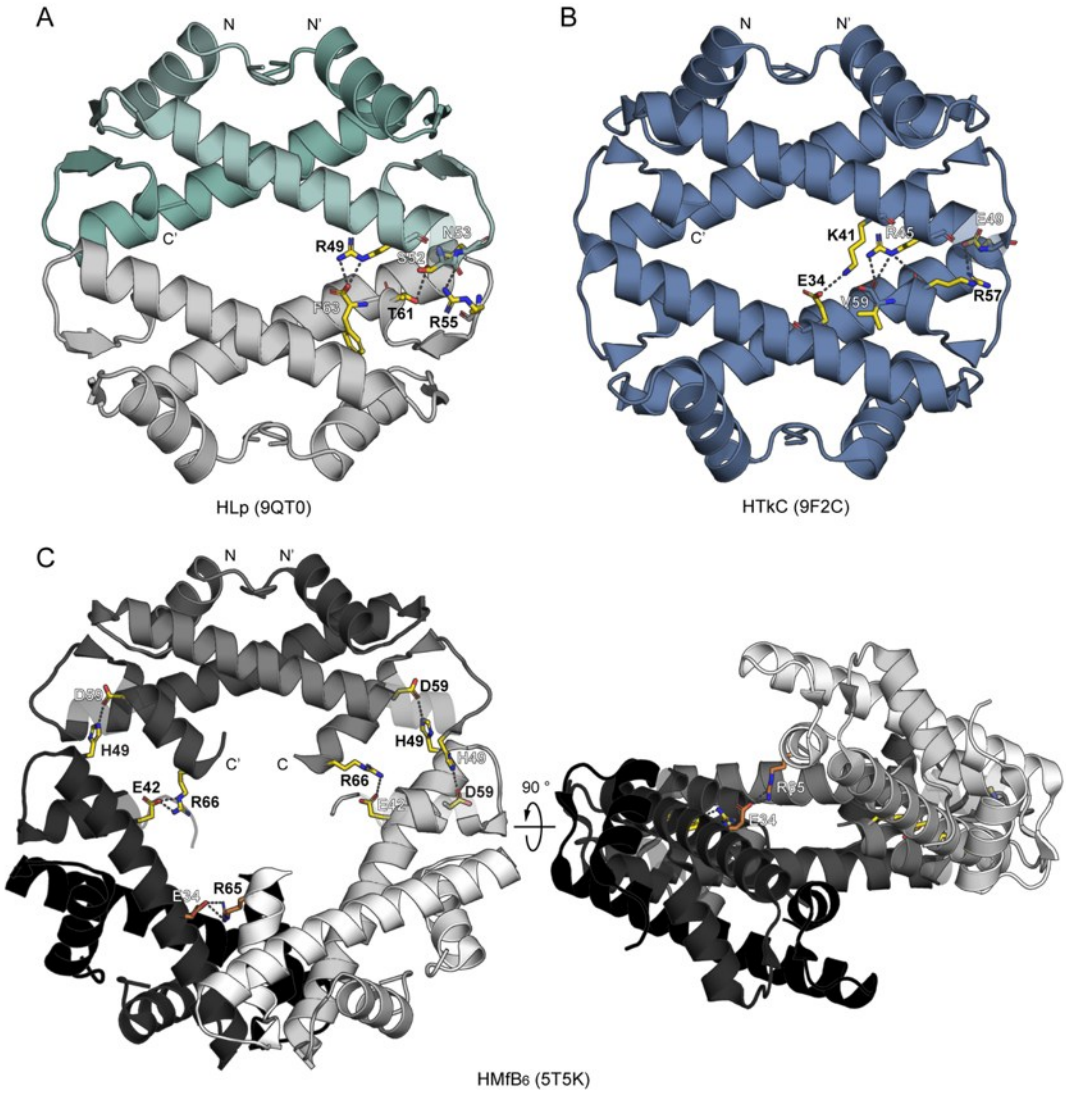


Figure 6.3 Tetramerization-associated residues in HLp, compared to those in the archaeal histones HTkC and HMfB.

A. Crystal structure of HLp tetramer (PDB:9QT0) in cartoon representation. **B**. Crystal structure of HTkC tetramer (PDB: 9F2C) in cartoon representation. **C.** Crystal structure of HMfB hexamer (PDB: 5T5K) in cartoon representation. In panel A and B, the contents of a single asymmetric unit are indicated in colors and the selected symmetry mates are in gray. In all panels, the amino acid residues involved in tetramerization and further oligomerization are exhibited as sticks, and their side chains are colored in yellow and orange, respectively. Salt bridges and hydrogen bonds are shown as dashed lines.

Co-crystallization with 30 bp DNA fragments unveiled two distinct protein-DNA complexes for both HBb and HLp, featuring three DNA binding sites per histone dimer that interact with DNA minor grooves (Fig. 6.4). These sites are akin to the “paired end of helix I” motif and the “parallel beta bridges” motifs described for canonical histones

H2A, H2B, H3 and H4 (13). Although less conserved, the amino acid residues involved in DNA-binding fall into three categories, as is the case in eukaryotic histones (Fig. 6.1A) (13): (1) positively charged residues, typically lysine and arginine residues, found along the entire DNA-binding surface, which interact with the DNA backbone due to the length, flexibility, and positive charged of their side chains; (2) hydroxyl group-containing residues, such as serine and threonine residues, forming contacts with the phosphate groups of the DNA; and (3) hydrophobic residues, for instance the alanine residues located at the tip of the α1 in both HBb and HLp, which also participate in DNA binding (13). All these interactions formed between amino acid residues and the DNA backbone, including phosphate groups and deoxyriboses, strongly support the nonspecific DNA binding mechanism of both HBb and HLp.

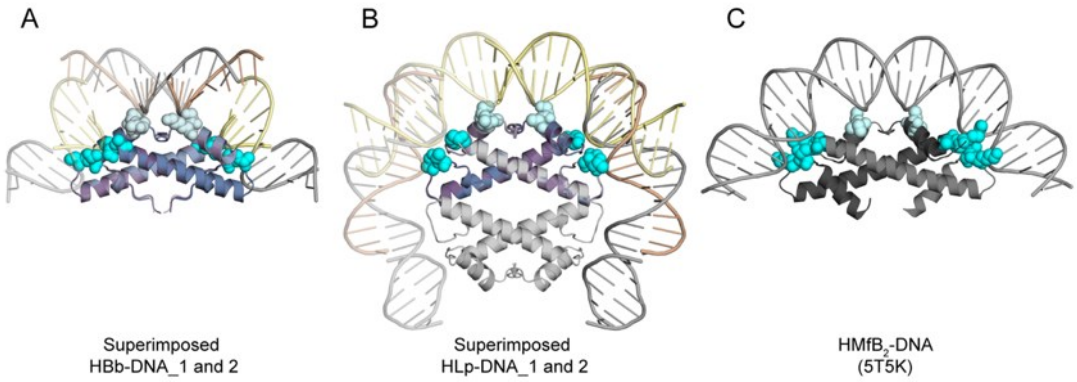


Figure 6.4 Superimposed structures of HBb-DNA and HLp-DNA complexes, respectively, in comparison with the structure of HMfB-DNA.

A. Superposition of HBb-DNA_1 (PDB: 9EZZ) and HBb-DNA_2 (PDB: 9F0E). **B.** Superposition of HLp-DNA_1 (PDB: 9QT1) and HLp-DNA_2 (PDB: 9QT2). **C.** Crystal structure of HMfB dimer bound to DNA (PDB: 5T5K). In panel A and B, the contents of an asymmetric unit are shown in colors and the selected symmetry mates are in gray. The amino acid residues corresponding to "paired end of helix I" motifs and "parallel beta bridges" motifs are indicated as spheres, and colored in pale cyan and cyan, respectively.

The superposition of the obtained HBb-DNA crystal structures suggests a DNA-bending mode by histone binding, analogous to HMfB of archaea (22). This is in contrast to the unconventional DNA binding mode suggested in a recent study (103). However, the results of molecular dynamics (MD) simulation performed on the basis of the HBb-DNA crystal structures support the assumption that HBb is able to bend

DNA in a way similar to HMfB. In the case of HLp, no clear conclusions can be drawn from the crystal structures. Here, the crystal packing of HLp-DNA complexes indicates a DNA-bridging mode, while the superposition of these two structures reveals a potential circular DNA interface across its surface, implying a second DNA-binding mode – wrapping. Performed MD simulations confirm that both DNA-binding modes are theoretically possible. Nevertheless, the HLp wrapping mode is the energetically preferred configuration. The architecture of individual bacterial histone dimers bound to DNA is akin to that of canonical histones, further supporting their role in chromatin organization (Fig. 6.5).

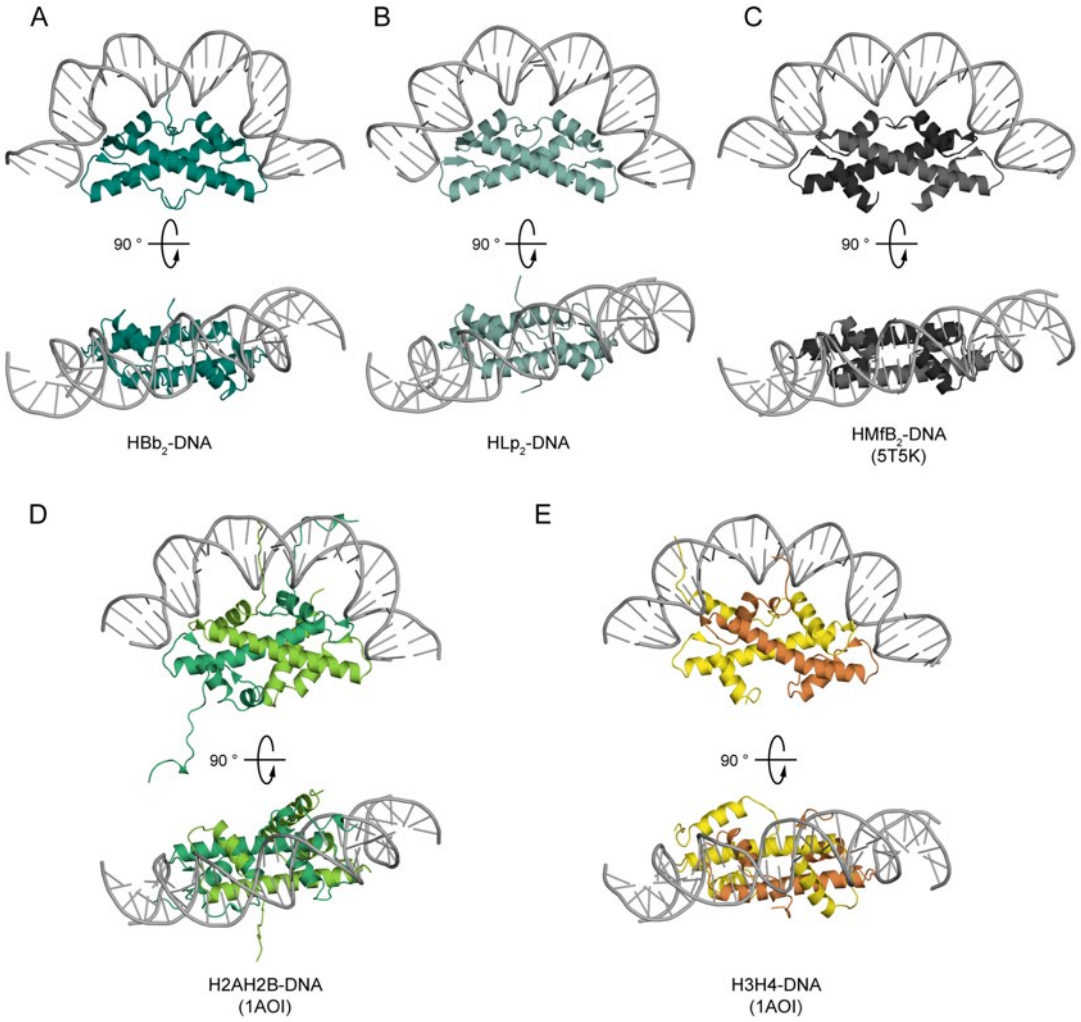


Figure 6.5 Structures of histone dimers bound to 30 bp DNA.

A. The last frame of the HBb-DNA model in MD simulation (22). **B.** The last frame of the HLp-DNA model in MD simulation. **C.** The structure of HMfB dimer interacting with DNA based on the crystal structure of HMfB-based chromatin (PDB: 5T5K). **D** and **E.** The structures of H2AH2B

and H3H4 heterodimers bound to DNA according to the crystal structure of the eukaryotic nucleosome (PDB: 1AOI), respectively.

6.3. Possible functions of HBb and HLp

A series of comprehensive biochemical analyses have elucidated the functional properties of HBb and HLp step-by-step. Microscale thermophoresis (MST) and electrophoretic mobility shift assays (EMSA) demonstrated that both proteins bind DNA with micromolar affinity, comparable to the DNA-binding affinity of archaeal histone HMfB (22). The results obtained from the micrococcal nuclease (MNase) digestion assay, tethered particle motion (TPM) experiment, DNA topology assay, as well as ligase-mediated circularization assay revealed that HBb does not form nucleosomes or even hypernucleosomes, but compacts and supercoils DNA as individual dimers through bending (22). It binds non-specifically to DNA across the entire genome, a function that is essential for the survival of *B. bacteriovorus* (22). Such non-specific binding was also demonstrated after heterologous expression in *E. coli* by showing the co-localization of GFP-fused HBb with the nucleoid (103). Transcriptomic data revealed a high expression level of the HBb gene during the growth phase of *B. bacteriovorus* (103). Quantitative label-free proteomic analysis of cells in the attack phase demonstrated that HBb represents the most abundant protein within the *B. bacteriovorus* nucleoid, which surpasses other putative NAPs, including HU (Bd3382), IHF (Bd1639), HU/IHF (Bd0711), Dps (Bd2620), and HU (Bd2104) (103,118). These findings suggest that HBb is not only involved in genome segregation during the growth phase but also in the extreme chromatin compaction in the attack phase, similar to the bacterial DNA organizer HU (22,72,103). In addition to HBb, *B. bacteriovorus* encodes a second histone fold-like protein, Bd3044, which belongs to the ZZ histone group (53,103). However, Bd3044, which is comparatively less expressed during bacterial growth and attack phases, cannot compensate for the deletion of the HBb-encoding gene in *B. bacteriovorus* (22,103,118). This suggests that HBb and Bd3044 may have distinct functions in genome organization.

In comparison, HLp compacts DNA via wrapping and alters DNA topology through a mechanism that differs from those of both bacterial dimeric histones and canonical histones. The DNA binding model proposed in the bioinformatic study suggests that FtF proteins, when acting as main chromatin organizers, bind DNA of 50-60 bp in length and form complexes with DNA resembling “beads-on-a-string” structures (53). However, MNase digestion of HLp yielded a ladder pattern of protected DNA fragments at low enzyme concentrations, with fragments exceeding 60 bp indicating the formation of hypernucleosome-like structures (119). This hypothesis is supported by TPM experiments showing that HLp induces substantial DNA condensation only at high protein concentrations, implying a local condensing effect under physiological conditions. Light microscopy analysis of HLp-overexpressing *E. coli* further confirmed nucleoid compaction caused by HLp. Although no experiments were conducted in pathogenic *L. perolatii*, a close HLp homolog sharing 79% sequence identity has been studied in *L. interrogans*. This histone is highly abundant in the cell, and attempts to delete its encoding gene were unsuccessful, suggesting it may be essential for *L. interrogans* (103,120). Although the cellular abundance of HLp remains unknown, it is likely to be essential for the viability of *L. perolatii*, especially given that this organism encodes neither other histone homologs nor canonical NAPs. These observations support the hypothesis that HLp functions as a global chromatin organizer with the potential to form hypernucleosomes. Additional analyses, such as nucleosome reconstitution, cryo-EM, and optical tweezers experiments will be useful to fully elucidate the architecture of possible HLp-based hypernucleosomes and their interactions with long DNA fragments.

Despite current findings, the possibility remains that bacterial histones HBb and HLp cooperate with other proteins *in vivo* and may serve alternative roles, such as some histone homologs that have been shown to recognize and bind specific DNA sequences and modulate gene expression by influencing transcriptional activity rather than organizing chromatin. For instance, the eukaryotic transcription factor NF-Y is a

heterotrimer composed of NF-YA, NF-YB, and NF-YC. NF-YB and NF-YC contain H2A/H2B histone-like domains and dimerize to bind DNA non-specifically, while NF-YA facilitates sequence-specific recognition of the CCAAT motif, a common promoter element in eukaryotes (121). This histone-like trimer associates with two types of NF-Y loci, one marked by active histone modifications linked to transcriptional activation and the other enriched in repressive marks, highlighting its dual role as a transcriptional regulator (122). Additionally, the pseudodimeric archaeal histone HpyA, found in *Halobacterium salinarum*, is highly conserved among haloarchaea, and retains key residues critical for eukaryotic histone function. However, instead of facilitating DNA compaction, HpyA regulates cell morphology and growth phase-dependent gene expression (58,63). Beyond transcriptional regulation, certain histones exhibit enzymatic activity. The eukaryotic H3-H4 tetramer has been reported to function as an oxidoreductase *in vitro*, binding divalent copper at the H3-H3' dimerization interface and reducing it to a monovalent state, potentially contributing to cellular copper homeostasis (123). These diverse roles suggest that bacterial histone homologs and variants may similarly participate in functions beyond chromatin organization, warranting further investigation.

In eukaryotes, the PTMs have been extensively studied, as briefly mentioned in section 2.1, and play a crucial role not only in chromatin organization but also in epigenetic regulation. These modifications predominantly occur in lysine-rich N-terminal tails, which are generally absent in archaeal histones. Surprisingly, the acetylation of lysine residues within the histone fold has been identified in the essential archaeal histones HTkA and HTkB from *T. kodakarensis* (124). These modified lysine residues are located at the stacking interface between histone dimers, suggesting a potential mechanism of chromatin regulation mediated by PTMs. Furthermore, bioinformatic analyses have discovered archaeal histones within the Asgard archaea clade, that possess positively charged N-terminal tails with high sequence conservation to those of eukaryotic histones, implying the possibility of PTM-mediated chromatin

regulation in archaea (47,53,125). Whether the bacterial histones are post-translationally modified by “histone writers” or are remodeled by chromatin-remodeling complexes remains an open question.

6.4. The origin of bacterial histones

The presence of histones in bacteria raises a fundamental question regarding their evolutionary origin. One hypothesis proposes that primordial histones, characterized by the histone fold, emerged as DNA-binding proteins during the time of LUCA (11,22). These ancestral proteins may have initially functioned as simple NAPs and gradually acquired adaptive modifications that enabled the formation of nucleosomes and higher-order chromatin architectures. Such evolutionary refinements likely enhanced DNA-binding affinity and structural organization, thereby facilitating more efficient DNA packaging and compaction. This hypothesis is supported by the identification of histone homologs in the deeply branching bacterial phylum Actinobacteria (20). Additionally, the conservation of crucial DNA-binding residues across eukaryotic, archaeal, and bacterial histones further reinforces a shared evolutionary origin (20). The detailed characterization of two bacterial histones in this study provides further support for this scenario.

An alternative hypothesis, however, considers the rarity of histones in bacteria – found in less than 2% of bacterial genomes – as significant (103,126). In contrast, the nucleoid-associated protein HU is nearly ubiquitous, present in approximately 93% of bacterial genomes (103,126). Meanwhile, nucleosomal and FtF histones are universally distributed among archaea, indicating that histones were already present in the LACA (53). This pattern, along with phylogenetic evidence, supports the possibility that bacterial histones originated in archaea and were later acquired by bacteria via horizontal gene transfer (53). Following the acquisition, these proteins may have undergone functional diversification and specialization, adapting to distinct roles in different bacterial lineages. This specialization is reflected in the structural and

architectural diversity of bacterial histone variants – evident in their distinct oligomerization states, DNA-binding modes, and the presence of additional structural domains.

Despite these competing hypotheses, reconstructing the evolutionary history of histones remains challenging due to the low sequence conservation among histone families across all domains of life. Comprehensive comparative genomic and structural analyses will be crucial for elucidating the origins, and structural and functional diversification of bacterial histones.

7. Conclusion and outlook

“没有脚踏实地建立起来的东西，就无法形成精神和物质上的支撑。”

—— 东野圭吾《时生》

The comprehensive structural and functional characterization of HBb (Histone of *Bdellovibrio bacteriovorus*) and HLp (Histone of *Leptospira perolatii*) presented in this study provides compelling evidence for the existence of histones in bacteria, challenging the long-standing belief that histones are restricted to eukaryotes and archaea. This work represents an in-depth investigation of bacterial histones, offering novel insights into their structural and functional properties, and uncovers a previously unrecognized dimension of bacterial chromatin organization.

HBb has been established as a dimeric bacterial histone with distinctive DNA-structuring properties. In spite of its low sequence similarity to canonical histones, HBb adopts the characteristic histone fold and binds non-specific DNA through a conserved interface reminiscent of those present in its archaeal and eukaryotic counterparts. However, its DNA-binding mode diverges significantly, with HBb bending DNA rather than forming nucleosome-like structures. This DNA-bending feature aligns HBb more closely with bacterial NAPs, especially the HU/IHF family proteins. Furthermore, HBb is abundant in the nucleoid and is essential for the viability of *B. bacteriovorus*. These results imply that HBb may function as a chromatin organizer in *B. bacteriovorus*, contributing to genome compaction and chromatin dynamics.

HLp, on the other hand, represents a unique bacterial histone that assembles into homotetramers, an oligomeric state that has never before been observed for classical histones. Despite this unconventional assembly, HLp binds and compacts non-specific DNA in a manner analogous to typical histones. The results of *in vitro* biochemical

assays and the observation of its nucleoid condensation effect *in vivo* further support the hypothesis that HLp plays as a global role in chromatin organization in *L. perolatii*.

Taken together, the findings of this study establish bacterial histones as a distinct and previously underexplored family of chromatin-associated proteins, helping to bridge the longstanding knowledge gap caused by the limited experimental investigation of bacterial histone homologs. Although they share the characteristic histone fold, bacterial histones exhibit notable diversity in their structural architecture and modes of DNA binding. The experimental evidence presented here provides compelling support for the existence and functionality of bacterial histones, underscoring the evolutionary conservation of histones across all domains of life. These results expand our understanding of histone evolution and reveal previously unrecognized strategies of chromatin organization strategies in bacteria (Fig. 7.1).

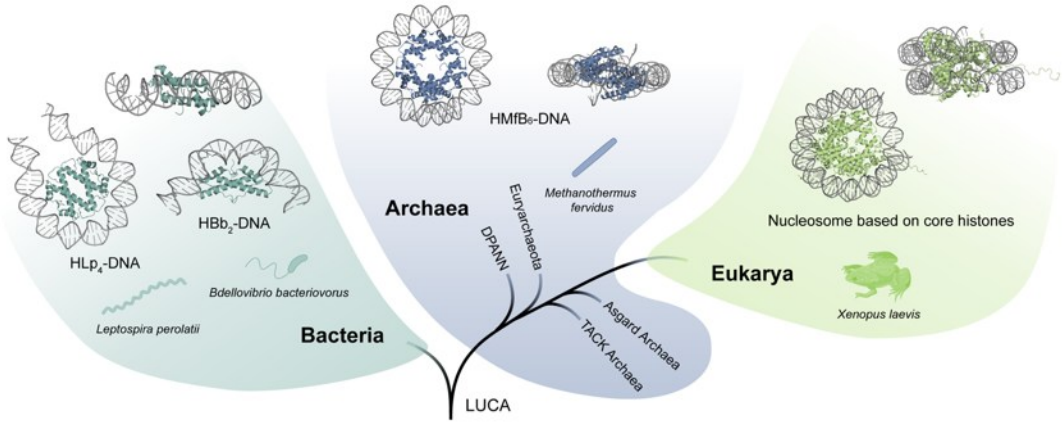


Figure 7.1 The simplified tree of life, including the representative histone-based nucleosomes in archaea and eukaryotes, and the models of histone-DNA complexes in bacteria.

The structures of bacterial histones (in dark green) bound to DNA are the last frames of the bacterial histone-DNA models in MD simulation performed in this study. Detailed information about the representative archaeal and eukaryotic nucleosomes is the same as listed in the legend to Fig.1.1. *This figure was adapted from the graphical abstract in Hu et al., 2024 (22).*

As one of the pioneering studies in the emerging field of bacterial histones, this work paves the way for future research into their evolutionary origins, regulatory functions, and potential post-translational modifications. While initial efforts have

elucidated the properties of the simplest bacterial histones, a vast array of structurally and functionally diverse variants remains uncharacterized. Further investigations will be crucial to unraveling the full spectrum of bacterial histones, their roles in chromatin organization, and their broader implications for gene regulation and cellular processes.

References

1. Lupas, A.N., Ponting, C.P. and Russell, R.B. (2001) On the evolution of protein folds: are similar motifs in different protein folds the result of convergence, insertion, or relics of an ancient peptide world? *J Struct Biol*, **134**, 191-203.
2. Söding, J. and Lupas, A.N. (2003) More than the sum of their parts:: on the evolution of proteins from peptides. *Bioessays*, **25**, 837-846.
3. Alva, V., Söding, J. and Lupas, A.N. (2015) A vocabulary of ancient peptides at the origin of folded proteins. *Elife*, **4**.
4. Linsky, T.W., Noble, K., Tobin, A.R., Crow, R., Carter, L., Urbauer, J.L., Baker, D. and Strauch, E.M. (2022) Sampling of structure and sequence space of small protein folds. *Nat Commun*, **13**, 7151.
5. Durairaj, J., Waterhouse, A.M., Mets, T., Brodiazhenko, T., Abdullah, M., Studer, G., Tauriello, G., Akdel, M., Andreeva, A., Bateman, A. *et al.* (2023) Uncovering new families and folds in the natural protein universe. *Nature*, **622**, 646-653.
6. Doenecke, D., Albig, W., Bode, C., Drabent, B., Franke, K., Gavenis, K. and Witt, O. (1997) Histones: genetic diversity and tissue-specific gene expression. *Histochem Cell Biol*, **107**, 1-10.
7. Kornberg, R.D. and Lorch, Y. (1999) Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell*, **98**, 285-294.
8. Arents, G., Burlingame, R.W., Wang, B.C., Love, W.E. and Moudrianakis, E.N. (1991) The Nucleosomal Core Histone Octamer at 3.1-a Resolution - a Tripartite Protein Assembly and a Left-Handed Superhelix. *P Natl Acad Sci USA*, **88**, 10148-10152.
9. Malik, H.S. and Henikoff, S. (2003) Phylogenomics of the nucleosome. *Nature Structural Biology*, **10**, 882-891.
10. Arents, G. and Moudrianakis, E.N. (1995) The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. *Proc Natl Acad Sci U S A*, **92**, 11170-11174.
11. Alva, V., Ammelburg, M., Soding, J. and Lupas, A.N. (2007) On the origin of the histone fold. *BMC Struct Biol*, **7**, 17.
12. Hadjithomas, M. and Moudrianakis, E.N. (2011) Experimental evidence for the role of domain swapping in the evolution of the histone fold. *Proc Natl Acad Sci U S A*, **108**, 13462-13467.
13. Arents, G. and Moudrianakis, E.N. (1993) Topography of the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA. *Proc Natl Acad Sci U S A*, **90**, 10489-10493.
14. Ammelburg, M. (2006) The C-Domain: A marker for AAA+ proteins and a homolog of the histones.
15. Bowtell, D., Fu, P., Simon, M. and Senior, P. (1992) Identification of Murine Homologs of the Drosophila Son of Sevenless Gene - Potential Activators of Ras. *P Natl Acad Sci USA*, **89**, 6511-6515.

16. Albright, S.R. and Tjian, R. (2000) TAFs revisited: more data reveal new twists and confirm old ideas. *Gene*, **242**, 1-13.
17. Fahrner, R.L., Cascio, D., Lake, J.A. and Slesarev, A. (2001) An ancestral nuclear protein assembly: crystal structure of the Methanopyrus kandleri histone. *Protein Sci*, **10**, 2002-2007.
18. Sandman, K. and Reeve, J.N. (2006) Archaeal histones and the origin of the histone fold. *Curr Opin Microbiol*, **9**, 520-525.
19. Soding, J. (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics*, **21**, 951-960.
20. Alva, V. and Lupas, A.N. (2019) Histones predate the split between bacteria and archaea. *Bioinformatics*, **35**, 2349-2353.
21. Qiu, Y., Tereshko, V., Kim, Y., Zhang, R., Collart, F., Yousef, M., Kossiakoff, A. and Joachimiak, A. (2006) The crystal structure of Aq_328 from the hyperthermophilic bacteria Aquifex aeolicus shows an ancestral histone fold. *Proteins*, **62**, 8-16.
22. Hu, Y., Schwab, S., Deiss, S., Escudeiro, P., van Heesch, T., Joiner, J.D., Vreede, J., Hartmann, M.D., Lupas, A.N., Alvarez, B.H. *et al.* (2024) Bacterial histone HBb from Bdellovibrio bacteriovorus compacts DNA by bending. *Nucleic Acids Res*, **52**, 8193-8204.
23. Kossel, A. (1884) Ueber einen peptonartigen Bestandtheil des Zellkerns.
24. Jones, M.E. (1953) Albrecht Kossel, a biographical sketch. *Yale J Biol Med*, **26**, 80-97.
25. Kornberg, R.D. and Thomas, J.O. (1974) Chromatin structure; oligomers of the histones. *Science*, **184**, 865-868.
26. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 angstrom resolution. *Nature*, **389**, 251-260.
27. Han, M., Chang, M., Kim, U.J. and Grunstein, M. (1987) Histone H2B repression causes cell-cycle-specific arrest in yeast: effects on chromosomal segregation, replication, and transcription. *Cell*, **48**, 589-597.
28. Kim, U.J., Han, M., Kayne, P. and Grunstein, M. (1988) Effects of histone H4 depletion on the cell cycle and transcription of Saccharomyces cerevisiae. *EMBO J*, **7**, 2211-2219.
29. Wyrick, J.J., Holstege, F.C., Jennings, E.G., Causton, H.C., Shore, D., Grunstein, M., Lander, E.S. and Young, R.A. (1999) Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast. *Nature*, **402**, 418-421.
30. Gossett, A.J. and Lieb, J.D. (2012) In vivo effects of histone H3 depletion on nucleosome occupancy and position in Saccharomyces cerevisiae. *PLoS Genet*, **8**, e1002771.
31. Allan, J., Hartman, P.G., Crane-Robinson, C. and Aviles, F.X. (1980) The structure of histone H1 and its location in chromatin. *Nature*, **288**, 675-679.
32. Li, W., Hu, J., Song, F., Yu, J., Peng, X., Zhang, S., Wang, L., Hu, M., Liu, J.-C. and Wei, Y. (2024) Structural basis for linker histone H5–nucleosome binding and chromatin fiber compaction. *Cell Research*, **34**, 707-724.

33. Allfrey, V.G., Faulkner, R. and Mirsky, A. (1964) Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proceedings of the National Academy of Sciences*, **51**, 786-794.
34. Chandrasekharan, M.B., Huang, F. and Sun, Z.W. (2009) Ubiquitination of histone H2B regulates chromatin dynamics by enhancing nucleosome stability. *Proc Natl Acad Sci U S A*, **106**, 16686-16691.
35. Grau-Bove, X., Navarrete, C., Chiva, C., Pribasnig, T., Anto, M., Torruella, G., Galindo, L.J., Lang, B.F., Moreira, D., Lopez-Garcia, P. *et al.* (2022) A phylogenetic and proteomic reconstruction of eukaryotic chromatin evolution. *Nat Ecol Evol*, **6**, 1007-1023.
36. Struhl, K. (1998) Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev*, **12**, 599-606.
37. Garcia-Bassets, I., Kwon, Y.S., Telese, F., Prefontaine, G.G., Hutt, K.R., Cheng, C.S., Ju, B.G., Ohgi, K.A., Wang, J., Escoubet-Lozach, L. *et al.* (2007) Histone methylation-dependent mechanisms impose ligand dependency for gene activation by nuclear receptors. *Cell*, **128**, 505-518.
38. Hublitz, P., Albert, M. and Peters, A.H. (2009) Mechanisms of transcriptional repression by histone lysine methylation. *Int J Dev Biol*, **53**, 335-354.
39. Strahl, B.D. and Allis, C.D. (2000) The language of covalent histone modifications. *Nature*, **403**, 41-45.
40. Wang, Y., Fischle, W., Cheung, W., Jacobs, S., Khorasanizadeh, S. and Allis, C.D. (2004) Beyond the double helix: writing and reading the histone code. *Novartis Found Symp*, **259**, 3-17; discussion 17-21, 163-169.
41. Narlikar, G.J., Fan, H.Y. and Kingston, R.E. (2002) Cooperation between complexes that regulate chromatin structure and transcription. *Cell*, **108**, 475-487.
42. Avvakumov, N., Nourani, A. and Côté, J. (2011) Histone Chaperones: Modulators of Chromatin Marks. *Mol Cell*, **41**, 502-514.
43. Gurard-Levin, Z.A., Quivy, J.P. and Almouzni, G. (2014) Histone chaperones: Assisting histone traffic and nucleosome dynamics. *Annu Rev Biochem*, **83**, 487-+.
44. Sandman, K., Krzycki, J.A., Dobrinski, B., Lurz, R. and Reeve, J.N. (1990) HMf, a DNA-binding protein isolated from the hyperthermophilic archaeon Methanothermus fervidus, is most closely related to histones. *Proc Natl Acad Sci U S A*, **87**, 5788-5791.
45. Decanniere, K., Babu, A.M., Sandman, K., Reeve, J.N. and Heinemann, U. (2000) Crystal structures of recombinant histones HMfA and HMfB from the hyperthermophilic archaeon Methanothermus fervidus. *J Mol Biol*, **303**, 35-47.
46. Bhattacharyya, S., Mattiroli, F. and Luger, K. (2018) Archaeal DNA on the histone merry-go-round. *FEBS J*, **285**, 3168-3174.
47. Henneman, B., van Emmerik, C., van Ingen, H. and Dame, R.T. (2018) Structure and function of archaeal histones. *PLoS Genet*, **14**, e1007582.
48. Mattiroli, F., Bhattacharyya, S., Dyer, P.N., White, A.E., Sandman, K., Burkhart, B.W., Byrne, K.R., Lee, T., Ahn, N.G., Santangelo, T.J. *et al.* (2017) Structure of histone-based chromatin in Archaea. *Science*, **357**, 609-612.

49. Bowerman, S., Wereszczynski, J. and Luger, K. (2021) Archaeal chromatin 'slinkies' are inherently dynamic complexes with deflected DNA wrapping pathways. *Elife*, **10**, e65587.
50. Zhang, Z., Guo, L. and Huang, L. (2012) Archaeal chromatin proteins. *Sci China Life Sci*, **55**, 377-385.
51. Stevens, K.M. and Warnecke, T. (2023), *Semin Cell Dev Biol*. Elsevier, Vol. 135, pp. 50-58.
52. Hocher, A. and Warnecke, T. (2024) Nucleosomes at the Dawn of Eukaryotes. *Genome Biol Evol*, **16**, evae029.
53. Schwab, S., Hu, Y., van Erp, B., Cajili, M.K.M., Hartmann, M.D., Alvarez, B.H., Alva, V., Boyle, A.L. and Dame, R.T. (2024) Histones and histone variant families in prokaryotes. *Nature Communications*, **15**, 7950.
54. Spang, A., Eme, L., Saw, J.H., Caceres, E.F., Zaremba-Niedzwiedzka, K., Lombard, J., Guy, L. and Ettema, T.J.G. (2018) Asgard archaea are the closest prokaryotic relatives of eukaryotes. *PLoS Genet*, **14**, e1007080.
55. Devos, D.P. (2021) Reconciling Asgardarchaeota Phylogenetic Proximity to Eukaryotes and Planctomycetes Cellular Features in the Evolution of Life. *Mol Biol Evol*, **38**, 3531-3542.
56. Rojec, M., Hocher, A., Stevens, K.M., Merkenschlager, M. and Warnecke, T. (2019) Chromatinization of Escherichia coli with archaeal histones. *Elife*, **8**.
57. Muller, J.B., Geyer, P.E., Colaco, A.R., Treit, P.V., Strauss, M.T., Oroshi, M., Doll, S., Virreira Winter, S., Bader, J.M., Kohler, N. *et al.* (2020) The proteome landscape of the kingdoms of life. *Nature*, **582**, 592-596.
58. Dulmage, K.A., Todor, H. and Schmid, A.K. (2015) Growth-Phase-Specific Modulation of Cell Morphology and Gene Expression by an Archaeal Histone Protein. *mBio*, **6**, e00649-00615.
59. Cubonová, L., Katano, M., Kanai, T., Atomi, H., Reeve, J.N. and Santangelo, T.J. (2012) An archaeal histone is required for transformation of Thermococcus kodakarensis *Journal of Bacteriology*, **194**, 6864-6874.
60. Sakrikar, S. and Schmid, A.K. (2021) An archaeal histone-like protein regulates gene expression in response to salt stress. *Nucleic Acids Res*, **49**, 12732-12743.
61. Sanders, T.J., Marshall, C.J. and Santangelo, T.J. (2019) The role of archaeal chromatin in transcription. *J Mol Biol*, **431**, 4103-4115.
62. Hocher, A., Borrel, G., Fadhlaoui, K., Brugere, J.F., Gribaldo, S. and Warnecke, T. (2022) Growth temperature and chromatinization in archaea. *Nat Microbiol*, **7**, 1932-1942.
63. Sakrikar, S., Hackley, R.K., Martinez-Pastor, M., Darnell, C.L., Vreugdenhil, A. and Schmid, A.K. (2023) The hypersaline archaeal histones HpyA and HstA are DNA binding proteins that defy categorization according to commonly used functional criteria. *Mbio*, **14**.

64. Irwin, N.A.T. and Richards, T.A. (2024) Self-assembling viral histones are evolutionary intermediates between archaeal and eukaryotic nucleosomes. *Nature Microbiology*, **9**.
65. Thomas, V., Bertelli, C., Collyn, F., Casson, N., Telenti, A., Goesmann, A., Croxatto, A. and Greub, G. (2011) Lausannevirus, a giant amoebal virus encoding histone doublets. *Environ Microbiol*, **13**, 1454-1466.
66. Liu, Y., Bisio, H., Toner, C.M., Jeudy, S., Philippe, N., Zhou, K., Bowerman, S., White, A., Edwards, G., Abergel, C. *et al.* (2021) Virus-encoded histone doublets are essential and form nucleosome-like structures. *Cell*, **184**, 4237-4250 e4219.
67. Toner, C.M., Hoitsma, N.M., Weerawarana, S. and Luger, K. (2024) Characterization of Medusavirus encoded histones reveals nucleosome-like structures and a unique linker histone. *Nat Commun*, **15**, 9138.
68. Talbert, P.B., Armache, K.-J. and Henikoff, S. (2022) Viral histones: pickpocket's prize or primordial progenitor? *Epigenetics & chromatin*, **15**, 21.
69. Irwin, N.A., Pittis, A.A., Richards, T.A. and Keeling, P.J. (2022) Systematic evaluation of horizontal gene transfer between eukaryotes and viruses. *Nature microbiology*, **7**, 327-336.
70. Cunha, V.D., Gaia, M., Ogata, H., Jaillon, O., Delmont, T.O. and Forterre, P. (2020) Giant viruses encode novel types of actins possibly related to the origin of eukaryotic actin: the viractins. *BioRxiv*, 2020.2006. 2016.150565.
71. Thanbichler, M., Wang, S.C. and Shapiro, L. (2005) The bacterial nucleoid: a highly organized and dynamic structure. *J Cell Biochem*, **96**, 506-521.
72. Dame, R.T. (2005) The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol Microbiol*, **56**, 858-870.
73. Dame, R.T. (2021) Special Issue: Role of Bacterial Chromatin in Environmental Sensing, Adaptation and Evolution. *Microorganisms*, **9**.
74. Drlica, K. and Rouviereyaniv, J. (1987) Histone-Like Proteins of Bacteria. *Microbiol Rev*, **51**, 301-319.
75. Dorman, C.J. and Deighan, P. (2003) Regulation of gene expression by histone-like proteins in bacteria. *Current Opinion in Genetics & Development*, **13**, 179-184.
76. Dillon, S.C. and Dorman, C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nature Reviews Microbiology*, **8**, 185-195.
77. Dorman, C.J., Hinton, J.C.D. and Free, A. (1999) Domain organization and oligomerization among H-NS-like nucleoid-associated proteins in bacteria. *Trends Microbiol*, **7**, 124-128.
78. Rimsky, S. (2004) Structure of the histone-like protein H-NS and its role in regulation and genome superstructure. *Current Opinion in Microbiology*, **7**, 109-114.
79. Smyth, C.P., Lundback, T., Renzoni, D., Siligardi, G., Beavil, R., Layton, M., Sidebotham, J.M., Hinton, J.C., Driscoll, P.C., Higgins, C.F. *et al.* (2000) Oligomerization of the chromatin-structuring protein H-NS. *Mol Microbiol*, **36**, 962-972.

80. Dame, R.T., Wyman, C., Wurm, R., Wagner, R. and Goosen, N. (2002) Structural basis for H-NS-mediated trapping of RNA polymerase in the open initiation complex at the rrnB P1. *J Biol Chem*, **277**, 2146-2150.
81. Claret, L. and RouviereYaniv, J. (1997) Variation in HU composition during growth of Escherichia coli: the heterodimer is required for long term survival. *Journal of Molecular Biology*, **273**, 93-104.
82. Shindo, H., Furubayashi, A., Shimizu, M., Miyake, M. and Imamoto, F. (1992) Preferential binding of E.coli histone-like protein HU alpha to negatively supercoiled DNA. *Nucleic Acids Res*, **20**, 1553-1558.
83. Castaing, B., Zelwer, C., Laval, J. and Boiteux, S. (1995) Hu Protein of Escherichia-Coli Binds Specifically to DNA That Contains Single-Strand Breaks or Gaps. *Journal of Biological Chemistry*, **270**, 10291-10296.
84. Shimizu, M., Miyake, M., Kanke, F., Matsumoto, U. and Shindo, H. (1995) Characterization of the binding of HU and IHF, homologous histone-like proteins of Escherichia coli, to curved and uncurved DNA. *Bba-Gene Struct Expr*, **1264**, 330-336.
85. Kamashev, D. and Rouviere-Yaniv, J. (2000) The histone-like protein HU binds specifically to DNA recombination and repair intermediates. *EMBO J*, **19**, 6527-6535.
86. Bahloul, A., Boubrik, F. and Rouviere-Yaniv, J. (2001) Roles of Escherichia coli histone-like protein HU in DNA replication: HU-beta suppresses the thermosensitivity of dnaA46ts. *Biochimie*, **83**, 219-229.
87. Luijsterburg, M.S., White, M.F., van Driel, R. and Dame, R.T. (2008) The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Crit Rev Biochem Mol Biol*, **43**, 393-418.
88. Craig, N.L. and Nash, H.A. (1984) E. coli integration host factor binds to specific sites in DNA. *Cell*, **39**, 707-716.
89. Yoshua, S.B., Watson, G.D., Howard, J.A.L., Velasco-Berrelleza, V., Leake, M.C. and Noy, A. (2021) Integration host factor bends and bridges DNA in a multiplicity of binding modes with varying specificity. *Nucleic Acids Res*, **49**, 8684-8698.
90. Goosen, N. and van de Putte, P. (1995) The regulation of transcription initiation by integration host factor. *Mol Microbiol*, **16**, 1-7.
91. Ussery, D., Larsen, T.S., Wilkes, K.T., Friis, C., Worning, P., Krogh, A. and Brunak, S. (2001) Genome organisation and chromatin structure in Escherichia coli. *Biochimie*, **83**, 201-212.
92. Swinger, K.K. and Rice, P.A. (2004) IHF and HU: flexible architects of bent DNA. *Curr Opin Struct Biol*, **14**, 28-35.
93. Macchi, R., Montesissa, L., Murakami, K., Ishihama, A., De Lorenzo, V. and Bertoni, G. (2003) Recruitment of σ54-RNA polymerase to the Pu promoter of Pseudomonas putida through integration host factor-mediated positioning switch of α subunit carboxyl-terminal domain on an UP-like element. *Journal of Biological Chemistry*, **278**, 27695-27702.

94. Zhang, A., Rimsky, S., Reaban, M.E., Buc, H. and Belfort, M. (1996) Escherichia coli protein analogs StpA and H-NS: regulatory loops, similar and disparate effects on nucleic acid dynamics. *EMBO J*, **15**, 1340-1349.
95. Chenoweth, M.R. and Wickner, S. (2008) Complex regulation of the DnaJ homolog CbpA by the global regulators σS and Lrp, by the specific inhibitor CbpM, and by the proteolytic degradation of CbpM. *Journal of bacteriology*, **190**, 5153-5161.
96. Niki, H., Jaffe, A., Imamura, R., Ogura, T. and Hiraga, S. (1991) The new gene mukB codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of E. coli. *EMBO J*, **10**, 183-193.
97. Brunetti, R., Prosseda, G., Beghetto, E., Colonna, B. and Micheli, G. (2001) The looped domain organization of the nucleoid in histone-like protein defective Escherichia coli strains. *Biochimie*, **83**, 873-882.
98. Zimmerman, S.B. (2006) Cooperative transitions of isolated nucleoids: Implications for the nucleoid as a cellular phase. *Journal of Structural Biology*, **153**, 160-175.
99. Swinger, K.K., Lemberg, K.M., Zhang, Y. and Rice, P.A. (2003) Flexible DNA bending in HU–DNA cocrystal structures. *The EMBO journal*.
100. Rice, P.A., Yang, S., Mizuuchi, K. and Nash, H.A. (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell*, **87**, 1295-1306.
101. Bailey, K.A., Marc, F., Sandman, K. and Reeve, J.N. (2002) Both DNA and histone fold sequences contribute to archaeal nucleosome stability. *J Biol Chem*, **277**, 9293-9301.
102. Marc, F., Sandman, K., Lurz, R. and Reeve, J.N. (2002) Archaeal histone tetramerization determines DNA affinity and the direction of DNA supercoiling. *J Biol Chem*, **277**, 30879-30886.
103. Hocher, A., Laursen, S.P., Radford, P., Tyson, J., Lambert, C., Stevens, K.M., Montoya, A., Shliaha, P.V., Picardeau, M., Sockett, R.E. *et al.* (2023) Histones with an unconventional DNA-binding mode in vitro are major chromatin constituents in the bacterium Bdellovibrio bacteriovorus. *Nat Microbiol*, **8**, 2006-2019.
104. Driessen, R.P., Meng, H., Suresh, G., Shahapure, R., Lanzani, G., Priyakumar, U.D., White, M.F., Schiessel, H., van Noort, J. and Dame, R.T. (2013) Crenarchaeal chromatin proteins Cren7 and Sul7 compact DNA by inducing rigid bends. *Nucleic Acids Res*, **41**, 196-205.
105. Driessen, R.P., Sitters, G., Laurens, N., Moolenaar, G.F., Wuite, G.J., Goosen, N. and Dame, R.T. (2014) Effect of temperature on the intrinsic flexibility of DNA and its interaction with architectural proteins. *Biochemistry*, **53**, 6430-6438.
106. Erkelens, A.M., Henneman, B., van der Valk, R.A., Kirolos, N.C.S. and Dame, R.T. (2023) Specific DNA binding of archaeal histones HMfA and HMfB. *Front Microbiol*, **14**, 1166608.
107. Hu, Y., Schwab, S., Qiu, K., Zhang, Y., Baer, K., Reichle, H., Panzera, A., Lupas, A.N., Hartmann, M.D., Dame, R.T. *et al.* (2025) DNA Wrapping by a Tetrameric Bacterial Histone. *bioRxiv*, 2025.2005.2008.652872.

108. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S. and Yim, J. (2021) Protein complex prediction with AlphaFold-Multimer. *biorxiv*, 2021.2010. 2004.463034.
109. Henneman, B., Brouwer, T.B., Erkelens, A.M., Kuijntjes, G.J., van Emmerik, C., van der Valk, R.A., Timmer, M., Kirolos, N.C.S., van Ingen, H., van Noort, J. *et al.* (2021) Mechanical and structural properties of archaeal hypernucleosomes. *Nucleic Acids Res*, **49**, 4338-4349.
110. Fuchs, A.C.D., Maldoner, L., Wojtynek, M., Hartmann, M.D. and Martin, J. (2018) Rpn11-mediated ubiquitin processing in an ancestral archaeal ubiquitination system. *Nature Communications*, **9**, 2696.
111. Glazko, V.I., Owen, J.B., ApDewi, I. and Axford, R.F.E. (1997) An association of haemoglobin protein (HBB) with ovulation rate in Cambridge sheep. *Anim Sci*, **64**, 279-282.
112. Mouw, K.W. and Rice, P.A. (2007) Shaping the Borrelia burgdorferi genome: crystal structure and binding properties of the DNA-bending protein Hbb. *Mol Microbiol*, **63**, 1319-1330.
113. Mukherjee, A., DiMario, P.J. and Grove, A. (2009) Mycobacterium smegmatis histone-like protein Hlp is nucleoid associated. *Fems Microbiol Lett*, **291**, 232-240.
114. Tabassum, R., Sandman, K.M. and Reeve, J.N. (1992) HMt, a histone-related protein from Methanobacterium thermoautotrophicum delta H. *J Bacteriol*, **174**, 7890-7895.
115. Zhang, Y., Xue, Y., Shi, J., Ahn, J., Mi, W., Ali, M., Wang, X., Klein, B.J., Wen, H. and Li, W. (2018) The ZZ domain of p300 mediates specificity of the adjacent HAT domain for histone H3. *Nature structural & molecular biology*, **25**, 841-849.
116. Liu, J., Xue, Z., Zhang, Y., Vann, K.R., Shi, X. and Kutateladze, T.G. (2020) Structural Insight into Binding of the ZZ Domain of HERC2 to Histone H3 and SUMO1. *Structure*, **28**, 1225-1230. e1223.
117. Mi, W., Zhang, Y., Lyu, J., Wang, X., Tong, Q., Peng, D., Xue, Y., Tencer, A.H., Wen, H. and Li, W. (2018) The ZZ-type zinc finger of ZZZ3 modulates the ATAC complex-mediated histone acetylation and gene activation. *Nature communications*, **9**, 3759.
118. Karunker, I., Rotem, O., Dori-Bachash, M., Jurkevitch, E. and Sorek, R. (2013) A global transcriptional switch between the attack and growth forms of Bdellovibrio bacteriovorus. *Plos One*, **8**.
119. Lim, H.-W. and Iwafuchi, M. (2022), *DNA-Protein Interactions: Methods and Protocols*. Springer, pp. 59-68.
120. Schmidt, A., Beck, M., Malmström, J., Lam, H., Claassen, M., Campbell, D. and Aebersold, R. (2011) Absolute quantification of microbial proteomes at different states by directed mass spectrometry. *Molecular systems biology*, **7**, 510.
121. Chaves-Sanjuan, A., Gnesutta, N., Gobbini, A., Martignago, D., Bernardini, A., Fornara, F., Mantovani, R. and Nardini, M. (2021) Structural determinants for NF-Y subunit organization and NF-Y/DNA association in plants. *The Plant Journal*, **105**, 49-61.

122. Ceribelli, M., Dolfini, D., Merico, D., Gatta, R., Vigano, A.M., Pavesi, G. and Mantovani, R. (2008) The histone-like NF-Y is a bifunctional transcription factor. *Molecular and cellular biology*, **28**, 2047-2058.
123. Attar, N., Campos, O.A., Vogelauer, M., Cheng, C., Xue, Y., Schmollinger, S., Salwinski, L., Mallipeddi, N.V., Boone, B.A. and Yen, L. (2020) The histone H3-H4 tetramer is a copper reductase enzyme. *Science*, **369**, 59-64.
124. Alpha-Bazin, B., Gorlas, A., Lagorce, A., Joulié, D., Boyer, J.-B., Dutertre, M., Gaillard, J.-C., Lopes, A., Zivanovic, Y. and Dedieu, A. (2021) Lysine-specific acetylated proteome from the archaeon Thermococcus gammatolerans reveals the presence of acetylated histones. *Journal of Proteomics*, **232**, 104044.
125. Eme, L., Tamarit, D., Caceres, E.F., Stairs, C.W., De Anda, V., Schön, M.E., Seitz, K.W., Dombrowski, N., Lewis, W.H. and Homa, F. (2023) Inference and reconstruction of the heimdallarchaeial ancestry of eukaryotes. *Nature*, **618**, 992-999.
126. Azam, T.A. and Ishihama, A. (1999) Twelve species of the nucleoid-associated protein from Escherichia coli: sequence recognition specificity and DNA binding affinity. *Journal of Biological Chemistry*, **274**, 33105-33113.

Acknowledgements

“这个世界充满了各种谜团。即使只是其中的一个不起眼的小小谜团，如果能够靠自己的力量去把它给解释明白，心里的那种愉悦感，就是其他的任何东西都无法取代的。”

——东野圭吾《盛夏的方程式》

These past five years have been a journey filled with challenges, growth, and unforgettable moments, made possible by the support of many incredible people.

First and foremost, I am deeply grateful to Andrei Lupas, for giving me a second chance at the interview and allowing me to start as an intern despite my limited background in protein and biochemistry. His unwavering support, insightful feedback, and kindness over the years have been invaluable. When the first part of this project was scooped, he advised me to step back from the lab and document my work. When I was unsure which conferences to attend, he searched and gave me suggestions. And whenever I made a mistake, he was always there to correct me with patience.

I am especially thankful to Birte Hernandez Alvarez, who took a chance on me despite my inexperience. Without her willingness to supervise me, I would not have had the opportunity to be part of this department. I still remember when she just started working on beta-barrel outer membrane proteins, she humbly assured me that we could learn together. Her guidance during my first weeks and months helped me quickly adapt to the lab and develop good experimental habits. Additionally, I would like to thank her for proofreading not only my thesis but also my progress reports and offering valuable feedback, through which I learned a lot. When I was devastated to learn that another group had published a pre-print on the same protein I was working on, Birte’s support, along with Vikram’s help, allowed us to finish our manuscript soon after. Though it initially felt like a setback, it ultimately marked the beginning of many valuable lessons and opportunities.

I extend my sincere appreciation to Vikram Alva, whose foundational bioinformatic research on histone fold origins and bacterial histones provided a solid basis for this study. Although I was not his student, his dedication and insightful feedback were instrumental in advancing this project.

Marcus Hartmann is the next I would like to mention, who patiently taught me how to determine protein structures through crystallography. Despite his demanding schedule, he always made time to guide me step by step, from performing molecular replacement to structure refinement and deposition. He answered each simple question with patience, helped me bridge knowledge gaps, and recommended workshops and conferences to me.

I would like to thank my thesis advisory committee members, Doron Rapaport and John Weir, for their detailed scientific feedback and constructive comments on this project. Special thanks to Doron for evaluating my dissertation. This project would not have been possible without my collaborators, Remus Dame and Samuel Schwab (Leiden University), Jocelyne Vreede and Thor van Heesch (University of Amsterdam), Yunsen Zhang (University of Illinois Urbana-Champaign), and Erik Schäffer and Laura Muras (University of Tübingen), whose contributions enriched and strengthened this study. A credit also goes to the BioOptics, Genome Center, Electron Microscopy, and NMR Spectroscopy Facilities at MPI Biology Tübingen for their great scientific support.

I am incredibly grateful to my wonderful colleagues and friends, both within the department and across the institute, for creating such a supportive and inspiring environment. In particular, my heartfelt thanks go to Joe, Valeria, Antoine, Matej, Pedro, Kaiyu, Mikel, Timo, Kateryna, Yaning, Jakub, Adrian F., Vera, Alexander, Lukas, Luca, Tobias, Adrian D., Claudio, Felicia, Hadeer, Min, Aurora, Linxuan, Yinan, Yihua, Pengfei, Miao, Kai, Han, Yao, Dingwen, Chuanfu, Ziduan, Wen-Sui, Shanshan, Liping, Josué, Ancilla, Agnes, Martina, and Rebecca for their emotional and scientific support. I also deeply appreciate Silvia, Reinhard, Kerstin, Heidi, Anja, Astrid, Eva,

and Tran, whose guidance and support extended beyond the lab and into everyday life. Without our outstanding technicians, the scientific staff, and Jörg, our lab would not be able to function smoothly. Special thanks to Joe (who proofread the whole thesis draft), Pedro, and Mikel again for proofreading my thesis and providing valuable feedback. A big appreciation also goes to our past and present secretaries, Karin and María, for keeping everything running smoothly. A special mention goes to the undergraduate students I had the pleasure of learning alongside, especially Marco, Annika, and Maxim, whose enthusiasm brought energy to the office and lab. I am also grateful for the experience of supervising two excellent interns, Mozhgan and Vinishaa, which helped me grow both personally and professionally.

I am grateful to the IMPRS program for funding my PhD, and the former and current members of the Researcher Support Team for organizing workshops, courses, and events, as well as the International Office for their assistance in all aspects.

A heartfelt thank you to Shu-Yao, Caroline, and Angela. Though we don't keep in touch very often, I always know I can reach out without hesitation, and they will be there for me. Additionally, I want to acknowledge the joy brought by the institute's wonderful dogs, Taotao, Zoey, Binouze, Jette, and Max, whose kind owners allowed me to destress with their companionship.

Last but not least, I am forever grateful for the unconditional love of my parents and my dog Cream, and for the unwavering support and encouragement from beloved Houming. Finding someone willing to engage in scientific debates (or arguments) with me until 3 a.m. is truly rare. And, I want to thank myself for persevering through the challenges, because without that, none of these would have been possible.

Please allow me to end my dissertation with the lyrics of Wayne Lim Junjie's song "Turn of A Page".

“In the quiet of night
Every shadow and light
Memories overflow

Where the past still lies
Hidden out of sight
The stars lose their glow

Every time and again we see secrets unfold
Every strike of a bell we hear stories retold
Whispers from before
Echoes come and go
Gripping my soul
I’ll let them go, breaking the mold

The sky grows dimmer
But my fire won’t turn cold
From the chains and fears
To dreams I uphold

A turn of a page
I'm leaving the cage
I will step away
From this empty space

Every moment awake
Every word I pray
I'll be less afraid
Of time's silent rage

With the turn of a page
The ink fades away
A new dawn awaits
A new story's made
Turn of a page

Every turn of a page
What we can't erase
What we can't unmake
We’ve got to embrace

Like a ghost on a stage
I just feel my face
Begin to fade

Could this light lead me
Into another place
Above the fray

In the silence of dawn
Every string that I play
Leaves a trace of what
has gone

Coming of a new day
As the night withdraws
I find my way”

- “Turn of A Page”, Wayne Lim Junjie, 2024

Appendix I

Nucleic Acids Research, 2024, 1–12
https://doi.org/10.1093/nar/gkae485
Gene regulation, Chromatin and Epigenetics

Bacterial histone HBb from *Bdellovibrio bacteriovorus* compacts DNA by bending

Yimin Hu[1,†], Samuel Schwab[2,†], Silvia Deiss[1], Pedro Escudeiro[1], Thor van Heesch[3], Joe D. Joiner[1], Jocelyne Vreede[3], Marcus D. Hartmann[1,4], Andrei N. Lupas[1], Birte Hernandez Alvarez[1,*], Vikram Alva[1,*] and Remus T. Dame[2,*]

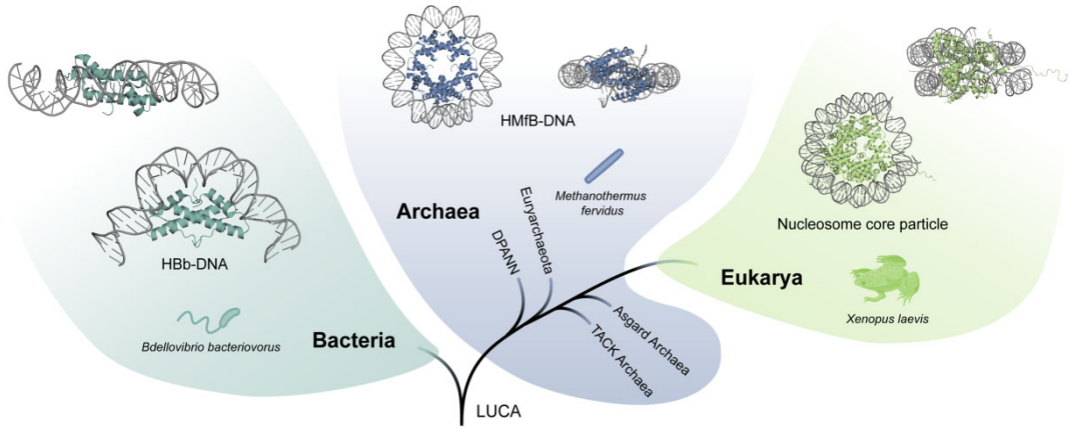
[1]Department of Protein Evolution, Max Planck Institute for Biology Tübingen, Tübingen, Germany
[2]Leiden Institute of Chemistry, Leiden University, Einsteinweg 55, 2333CC Leiden, The Netherlands; Centre for Microbial Cell Biology, Leiden University, Leiden, The Netherlands; Centre for Interdisciplinary Genome Research, Leiden University, Leiden, The Netherlands
[3]Van 't Hoff Institute for Molecular Sciences, University of Amsterdam, The Netherlands
[4]Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany

[*]To whom correspondence should be addressed. Tel: +49 7071 601 356; Fax: +49 7071 601 352; Email: birte.hernandez@tuebingen.mpg.de
Correspondence may also be addressed to Vikram Alva. Tel: +49 7071 601 451; Fax: +49 7071 601 352; Email: vikram.alva@tuebingen.mpg.de
Correspondence may also be addressed to Remus T. Dame. Tel: +31 71 527 5605; Email: rtdame@chem.leidenuniv.nl
[†]The first two authors should be regarded as Joint First Authors.

Abstract

Histones are essential for genome compaction and transcription regulation in eukaryotes, where they assemble into octamers to form the nucleosome core. In contrast, archaeal histones assemble into dimers that form hypernucleosomes upon DNA binding. Although histone homologs have been identified in bacteria recently, their DNA-binding characteristics remain largely unexplored. Our study reveals that the bacterial histone HBb (Bd0055) is indispensable for the survival of *Bdellovibrio bacteriovorus*, suggesting critical roles in DNA organization and gene regulation. By determining crystal structures of free and DNA-bound HBb, we unveil its distinctive dimeric assembly, diverging from those of eukaryotic and archaeal histones, while also elucidating how it binds and bends DNA through interaction interfaces reminiscent of eukaryotic and archaeal histones. Building on this, by employing various biophysical and biochemical approaches, we further substantiated the ability of HBb to bind and compact DNA by bending in a sequence-independent manner. Finally, using DNA affinity purification and sequencing, we reveal that HBb binds along the entire genomic DNA of *B. bacteriovorus* without sequence specificity. These distinct DNA-binding properties of bacterial histones, showcasing remarkable similarities yet significant differences from their archaeal and eukaryotic counterparts, highlight the diverse roles histones play in DNA organization across all domains of life.

Graphical abstract



Introduction

Histones, a class of highly conserved DNA-binding proteins, are predominantly found in eukaryotes and archaea. These proteins share a characteristic 'histone fold', comprising three α-helices connected by two short strap loops (1,2). In eukaryotes, the four core histones—H2A, H2B, H3, and H4—assemble into an octamer wrapped around by approximately 146 base pairs of double-stranded DNA (dsDNA), forming

Received: October 18, 2023. **Revised:** May 1, 2024. **Editorial Decision:** May 22, 2024. **Accepted:** May 24, 2024

the nucleosome, the fundamental unit of chromatin (1–3). A unique feature of eukaryotic histones is their unstructured N-terminal tails, which serve as hotspots for post-translational modifications and play pivotal roles in various cellular processes, including transcriptional regulation, DNA repair, replication, and condensation (4,5).

The first archaeal histone homologs, HMfA and HMfB, were identified in the hyperthermophilic archaeon *Methanothermus fervidus* (6). These homologs assemble into dimers that further oligomerize into elongated, nucleosome-like structures, termed hypernucleosomes, upon DNA binding (7–9). Subsequent research has identified one or more histone variants in the majority of archaeal lineages (10–16), including the Asgard superphylum, which comprises the closest prokaryotic relatives of eukaryotes (8,17). Archaeal histones can assemble into both homo- and hetero-oligomers and exhibit varying sizes, with some even possessing N- or C-terminal tails akin to eukaryotic histones (8,18). Additionally, archaeal histone-DNA complexes display a wide variety of histone-oligomerization states. These states are influenced by lateral dimer-dimer interactions and stacking interactions (19), which are not consistently conserved across species. The widespread occurrence and diversity of archaeal histones suggest distinct roles in genome organization and regulation (14,17,20,21).

Notably, histones are also encoded within viral genomes (22,23). Phylogenetically, viral histones are currently positioned between eukaryotic and archaeal histones (24). While many characterized viral nucleosomal histones structurally resemble eukaryotic nucleosome-like structures (25,26), some stack into larger complexes that lack linker DNA, resembling archaeal hypernucleosomes (27).

Historically, it was widely assumed that bacteria lack histone homologs, and any observed instances were thought to have been acquired through horizontal gene transfer (28,29). Instead of histones, bacteria were believed to rely on a multitude of small architectural proteins termed nucleoid-associated proteins (NAPs) (30). These NAPs are sometimes, albeit misleadingly, referred to as 'histone-like proteins'. However, NAPs are evolutionarily distinct from histones. Their similarities lie in their abundance, basic nature, and DNA-binding properties. While histones are conserved among eukaryotes, no single NAP is found in all prokaryotes. Despite this broad diversity, NAPs in various prokaryotes perform similar architectural functions by bending DNA, bridging DNA segments, and forming filaments on the DNA (30).

To shed light on the origins of histones and nucleosome-based DNA packaging and regulation, we recently conducted a comprehensive search for bacterial histones and discovered about 600 histone homologs in various bacterial classes, including Actinomycetes, Aquificae, Bdellovibrionia, Cyanophyceae, and Spirochaetia (31). Interestingly, these bacterial histones often co-exist with NAPs, whose structural and functional properties have been extensively studied. NAPs contribute to the local and global structuring of genomes and regulate DNA-associated proteins, including those involved in transcription. Bacterial histones can be categorized into pseudodimeric histones, which consist of two consecutive histone domains on a single chain, and single histone fold proteins with only one domain. Importantly, many bacterial histones contain conserved residues essential for DNA binding in eukaryotic and archaeal histones (Figure 1A), suggesting that bacterial histones may also bind DNA (31).

In this study, we present crystal structures and biochemical characterization of the free and DNA-bound forms of the histone homolog HBb from the predatory Gram-negative bacterium *Bdellovibrio bacteriovorus*. While the gene encoding HBb is listed in the NCBI databank under the locus tag *BD_RS00255* (formerly *Bd0055*), and even though it was referred to as Bd0055 in a recent publication (32), we propose using the name HBb (Histone of *B. bacteriovorus*) to maintain consistency with the standard nomenclature of prokaryotic histones in the literature. Our findings demonstrate that the single histone fold protein HBb forms a homodimer that has the ability to bind and bend DNA. Furthermore, we establish that HBb is essential for maintaining bacterial viability. Our results corroborate many of the observations reported in a recent study by Hocher *et al.* (32), while providing strong evidence for different DNA structuring properties of HBb, suggesting a generic role in global genome organization.

Materials and methods

Detailed information is provided in the *SI Appendix*, under *Material and Methods*.

Bacterial strains and cultivation

Genes encoding HBb (*BD_RS00255*, old locus tag: *Bd0055*) and HMfB (GenBank accession number M34778.1) were cloned for recombinant protein expression in *E. coli*. *B. bacteriovorus* HD100 (DSM 50701, DSMZ) was cultured to construct a chromosomal *BD_RS00255* deletion strain using *E. coli* Top10 as prey cells.

Cloning, plasmids, and synthetic DNA

Plasmids pETHis1a (33) and pET28a were used as expression vectors for HBb and HMfB protein production, respectively. For chromosomal deletion of *BD_RS00255* in *B. bacteriovorus*, the suicide vector pT18mobSacB (Addgene plasmid # 72648; http://n2t.net/addgene:72648; RRID:Addgene_72648) was used. Synthetic oligonucleotides (Merck) used in this study are listed in Supplementary Table S1.

Deletion of *BD_RS00255*

Attempts to delete *BD_RS00255* in *B. bacteriovorus* HD100 were conducted as previously described, with slight modifications (34,35).

Protein expression and purification

HBb, HMfB, and FLAG-HBb were recombinantly expressed in *E. coli* Mutant56(DE3) (36) and purified to homogeneity using affinity and size-exclusion chromatography.

Circular dichroism (CD) spectroscopy

CD spectroscopy was performed to analyze the secondary structure and thermal stability of purified HBb protein using a Jasco J-810 spectrometer (JASCO).

Size exclusion chromatography coupled with multi-angle light scattering (SEC-MALS)

SEC-MALS was performed to measure the mass of purified HBb in solution, using a 1260 Infinity II HPLC system

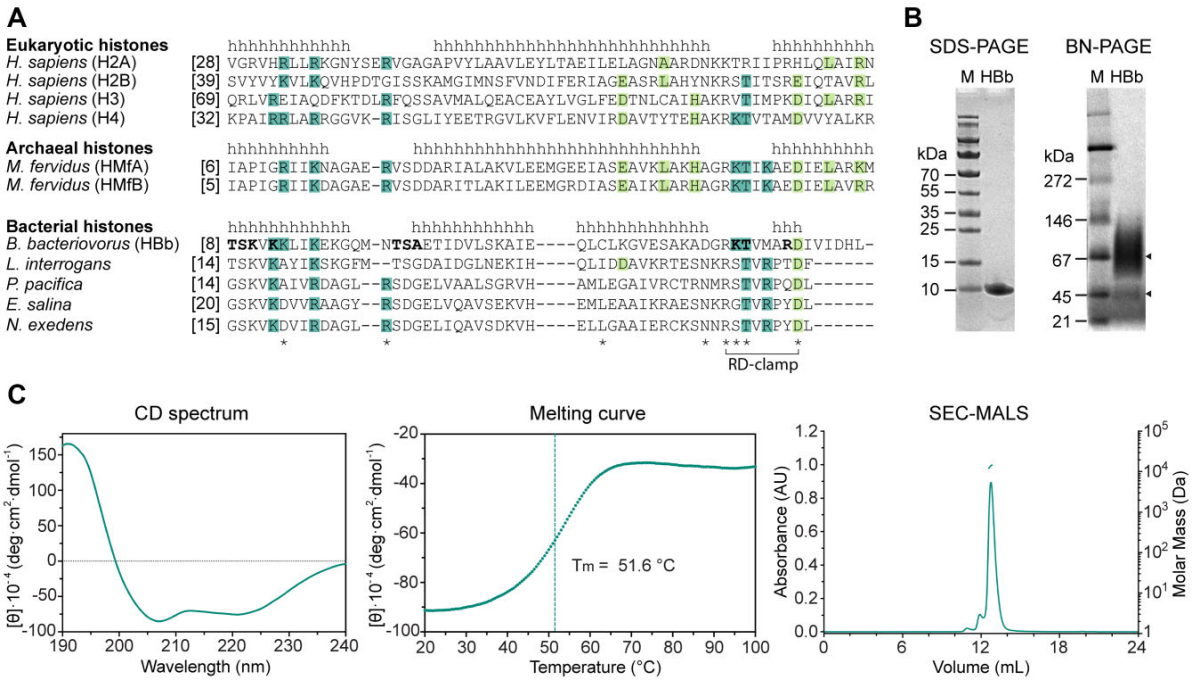


Figure 1. Biophysical characterization of HBb. (**A**) Sequence alignment of representative eukaryotic, archaeal, and bacterial histones from *Homo sapiens* (UniProtKB/Swiss-Prot entry for H2A: P04908; H2B: P62807; H3: P68431; H4: P62805), *Methanothermus fervidus* (HMfA: P48781; HMfB: P19267), *Bdellovibrio bacteriovorus* (Q6MRM1), *Leptospira interrogans* (Q8F3E8), *Plesiocystis pacifica* (A6GF99), *Enhygromyxa salina* (NCBI entry WP_052547863.1), and *Nannocystis exedens* (WP_096326703.1). α-Helices are labeled with 'h' according to the crystal structures of H2A, HMfB and HBb, respectively. Conserved residues involved in DNA binding and tetramerization are highlighted in dark and light green, respectively. HBb residues involved in DNA backbone interaction according to the solved structures HBb-DNA_1 (PDB: 9EZZ) and HBb-DNA_2 (PDB: 9F0E) are in bold. Residues represented as sticks in Figure 3A and B are labeled with *. Residues R51 and D58 forming the RD clamp in HBb are connected by a bracket. (**B**) SDS-PAGE and BN-PAGE showing purified HBb. Triangles mark HBb on BN-PAGE. (**C**) Biophysical characterization of HBb by CD spectroscopy (left: single CD spectrum, middle: thermal melting curve) and SEC-MALS (right) for determination of secondary structure, thermal stability, and oligomeric state.

(Agilent) coupled to a miniDAWN TREOS and Optilab T-rEX refractive index detector (Wyatt Technology).

Crystallization, data collection, and structure determination

Crystallization trials of HBb were performed in presence and absence of DNA. Crystals were obtained and structures solved for the free and two different DNA-bound forms. Detailed methods are described in detail in *SI Appendix*. Data processing and refinement statistics are listed in Supplementary Table S2, and the coordinates and structure factors have been deposited in the PDB under accession codes 8CMP (free HBb), 9EZZ (HBb-DNA_1) and 9F0E (HBb-DNA_2).

Microscale thermophoresis (MST)

The DNA binding affinities of HBb to the 80 bp (9,37) and 80 bp-GC50 oligonucleotides were measured using the Monolith NT.115 instrument (NanoTemper Technologies). MST data were analyzed using the associated MO.Control software.

DNA binding analysis

Binding of HBb and HMfB to DNA was comparatively analyzed using various assays, including electrophoretic mobility shift assay (EMSA), Micrococcal Nuclease (MNase) digestion assay, DNA topology assay, Ligase-mediated circularization assay, and tethered particle motion (TPM) experiments.

DNA affinity purification sequencing (DAP-seq)

Preparation of fragmented DNA libraries from genomic DNA of *B. bacteriovorus* HD100, subsequent DNA affinity purification, and next-generation sequencing were conducted as previously described, with some modifications (38). Data were processed with the nf-core/chipseq (version 2.0.0) pipeline (39), under Nextflow (version 23.04.2).

Modeling of the HBb-DNA complex

Molecular dynamics (MD) simulations of binding of the HBb dimer to DNA were performed using a structural model of an HBb dimer generated based on the two crystal structures of HBb with bound DNA (PDB: 9EZZ and 9F0E) and an AlphaFold2 (AF2) prediction (40). Methods and software used for MD simulations are described in detail in the *SI Appendix, Material and Methods*.

Results

The majority of bacterial histone homologs that we have identified in a recent bioinformatics analysis originate either from sparsely characterized bacteria of marine origin or from environmental samples (31). The predatory bacterium *B. bacteriovorus* represents one of the few organisms identified in the analysis that is both culturable in the lab and amenable to genetic manipulation. For these reasons, we selected the histone

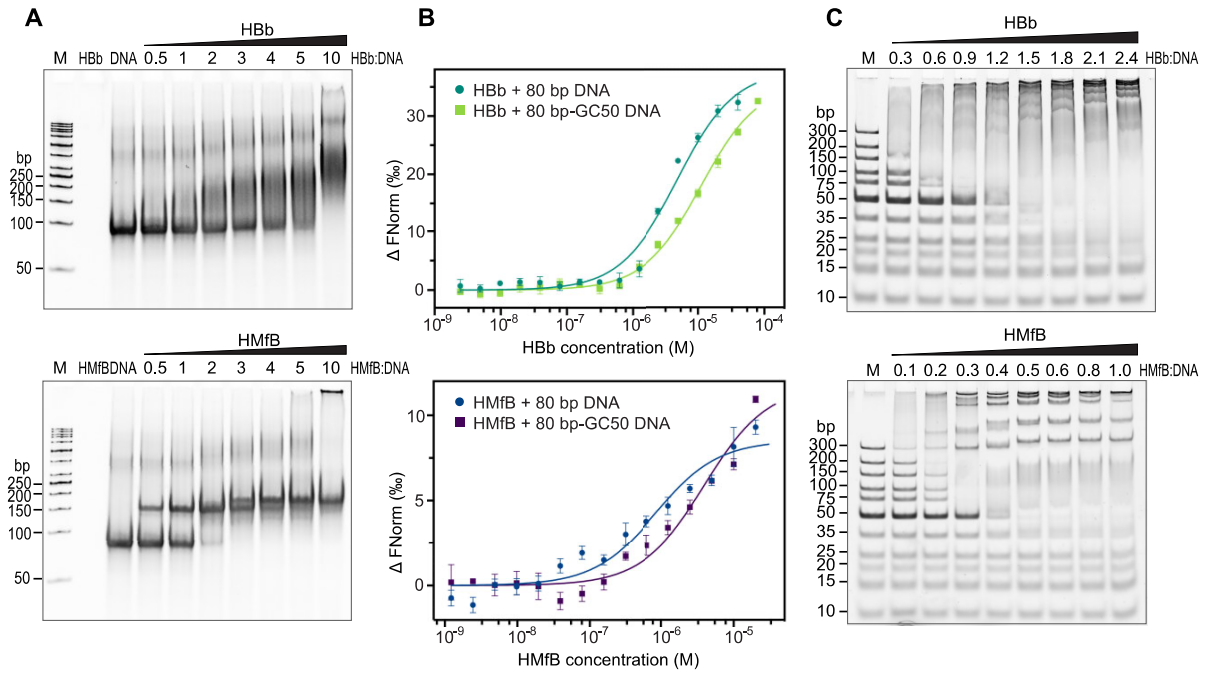


Figure 2. HBb binds DNA. (**A**) EMSA with the 80 bp DNA fragment in the presence of HBb and HMfB as a control. The molar ratios of protein:DNA in lanes 4–10 are indicated. (**B**) MST curves showing binding of HBb to the 80 bp DNA fragment and the 80 bp-GC50 DNA fragment (10 nM), as compared to HMfB. (**C**) EMSA of HBb and HMfB as a control with the GeneRuler Ultra Low Range DNA Ladder (Thermo Fisher Scientific). For lanes 2–9, the protein/DNA mass ratios are labeled.

protein HBb of *B. bacteriovorus* (UniProt entry Q6MRM1) for comprehensive functional and structural characterization.

The histone HBb from *B. bacteriovorus* binds DNA *in vitro*

We overexpressed HBb recombinantly in *E. coli* and purified the protein to homogeneity. Its purity was confirmed through SDS-PAGE, which displayed a single band corresponding to a molecular weight of 10 kDa, in line with the theoretical molecular weight of 7 kDa (Figure 1B). BN-PAGE separation yielded two bands (Figure 1B), suggesting the presence of diverse oligomers. Nonetheless, SEC-MALS analysis showed only a major peak with a molecular weight of 14 ± 0.4 kDa, indicating a dimer (Figure 1C). Circular dichroism (CD) spectroscopy indicated HBb to be an α-helical protein with flexible regions that cooperatively unfolds when heated, exhibiting a melting temperature of 51.6°C (Figure 1C).

We investigated the potential interaction of HBb with DNA using an electrophoretic mobility shift assay (EMSA) with an 80 bp DNA fragment that has a high affinity for HMfB (9,37). We observed a band shift, indicative of HBb-DNA complex formation. This was markedly different from the well-defined protein-DNA bands formed by hypernucleosomal histones, such as HMfB (Figure 2A). The HBb-DNA complexes appeared as smeared bands, suggesting varied conformations, stoichiometries, or potential protein dissociation. We quantified the binding affinity of HBb to the 80 bp DNA fragment and a DNA fragment of random sequence with a GC content of 50% (80 bp-GC50) using microscale thermophoresis (MST) (Figure 2B and Supplementary Table S3). Our results showed that HBb binds nonspecifically to DNA, with an affinity in the lower micromolar range, comparable to HMfB (41–43).

To estimate the size of the DNA binding site of HBb, we performed another EMSA using a DNA ladder (Figure 2C). Stable HBb-DNA complexes were observed with fragments equal to or larger than 35 bp. Although no band shift was observed for the 20 bp and 25 bp fragments, these fragments appear blurred, suggesting that HBb forms unstable complexes with these smaller fragments. In this EMSA ladder assay, HMfB binds to fragments larger than 35 bp, in agreement with previous EMSAs performed on small substrates (44).

HBb crystal structure

Crystallization trials were conducted with purified HBb, and within a few days, we obtained crystals that diffracted to atomic resolution. The best dataset could be processed to a resolution of 1.06 Å and the structure was solved by molecular replacement using the archaeal histone HMk (PDB: 1F1E) as a search model. Except for the first three and last three amino acids, the full chain is clearly resolved in the electron density map. The asymmetric unit contains a single chain of HBb, forming a dimer through crystallographic symmetry. HBb exhibits a characteristic dimeric histone fold, with monomers consisting of three α-helices (α1, α2, α3) connected by two loops (l1 and l2) (Figure 3A). In each monomer, the longer central helix α2 is flanked at both ends by the shorter helices α1 and α3. Dimerization occurs through the antiparallel alignment of the central α2-helices of the two monomers at an angle of approximately 25°. The dimer is stabilized primarily by hydrophobic contacts along the interface, involving residues L31, I35, and L38 in α2, as well as contributions from α1 (e.g. V11) and α3 (e.g. I59).

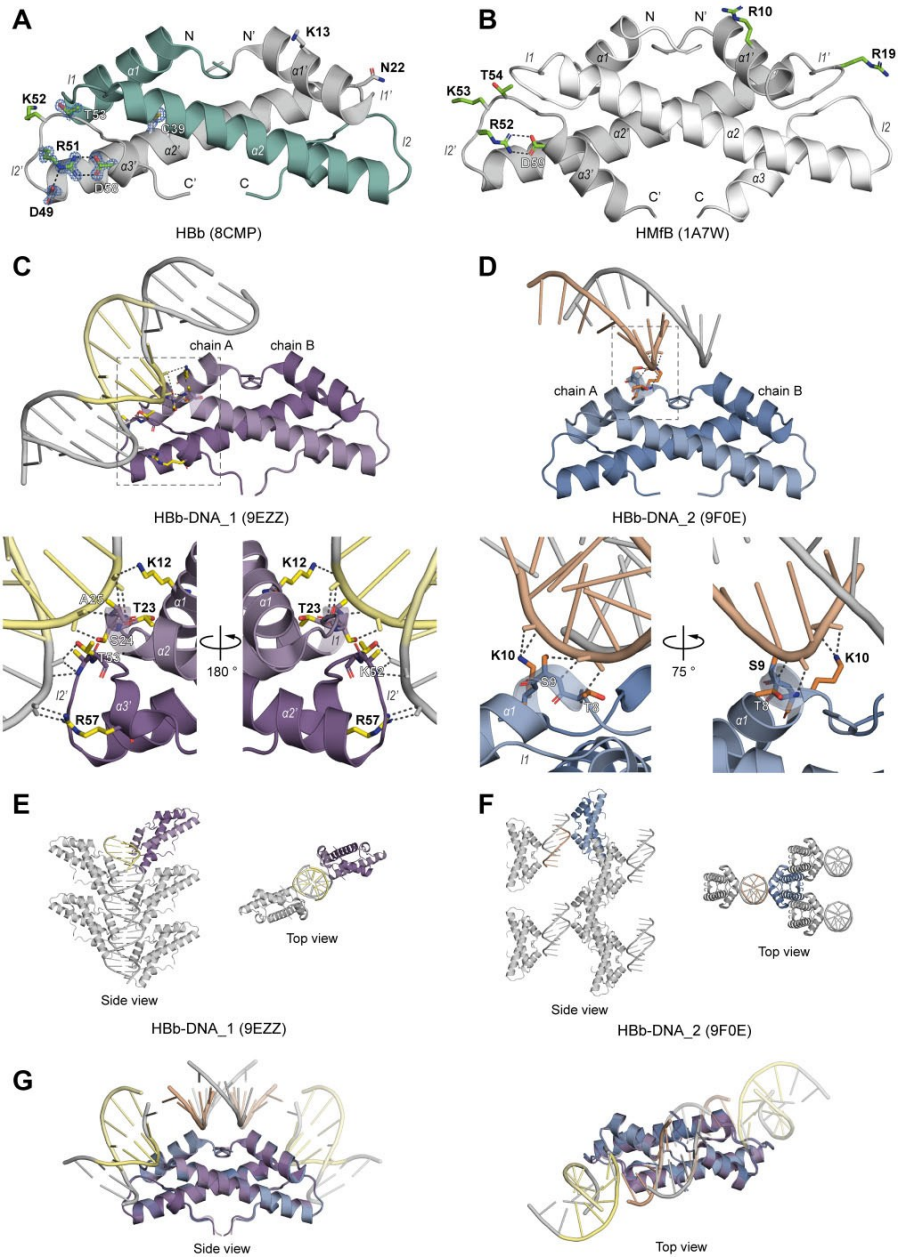


Figure 3. Crystal structures of HBb in its free and DNA-bound forms. (**A**) Crystal structure of the HBb dimer (PDB: 8CMP) in cartoon representation. Selected residues are shown as sticks and conserved residues predicted to be involved in DNA binding or stabilization of the histone fold are in green. The salt bridges formed between D49, R51, and D58 are indicated as dashed lines. A representative 2Fo – Fc electron density map is shown for selected residues at a contour level of at 2.0 σ. (**B**) Crystal structure of the HMfB dimer (PDB: 1A7W) in cartoon representation for comparison. Conserved residues playing a role in DNA binding or stabilization of the histone fold are depicted as sticks. (**C**) Crystal structure of HBb-DNA_1 (PDB: 9EZZ) in cartoon representation with residues involved in DNA binding shown as sticks. The dashed frame marks the enlarged image section. (**D**) Crystal structure of HBb-DNA_2 (PDB: 9F0E) in cartoon representation with DNA- binding residues shown as sticks. The enlarged image section is marked by a dashed square. (**E**) Crystal packing of HBb-DNA_1 with selected symmetry mates generated within 4 Å. (**F**) Crystal packing of HBb-DNA_2 with selected symmetry mates generated within 20 Å. (**G**) Superposition of structures HBb-DNA_1, HBb-DNA_2, and their symmetrical mates, with DNA fragments trimmed for visualization. In all panels, the contents of a single asymmetric unit are shown in color and selected symmetry mates in gray. Hydrogen bonds and salt bridges are indicated as dashed lines.

In contrast to the archaeal histone HMfB (Figure 3B), the α2-helix of HBb is one turn shorter, and α3 comprises only a single helical turn. Like HMfB, HBb features a conserved intramolecular salt bridge, known as the RD clamp, involving residues R51 and D58 (Figures 1A, 3A, B). In HBb, this motif is augmented by D49, which forms an additional salt bridge with R51, potentially stabilizing the conserved KT motif in loop l2. This KT motif in HMfB is known to be involved in DNA interactions. Four of the six DNA-binding residues in HMfB, including K52 and T53, are conserved in HBb (Figure 1A), indicating a potential for DNA binding.

Crystal structure of HBb bound to DNA

Similar to its free form, HBb successfully crystallized in complex with a randomized 20 bp DNA fragment (20 bp-GC50) within a few days, yielding the two distinct crystal forms HBb-DNA_1 and HBb-DNA_2. These could be processed to a resolution of 1.95 Å (HBb-DNA_1) and 1.85 Å (HBb-DNA_2), and solved by molecular replacement using the free form of HBb as a search model. Both structures contain one HBb dimer in the asymmetric unit, with all residues apart from the termini clearly resolved in the electron density. Additionally, clear electron density for short DNA fragments was apparent in both structures, albeit in different positions.

For HBb-DNA_1, the asymmetric unit contains a 5 bp ds-DNA fragment, which forms an infinite dsDNA helix throughout the crystal via crystallographic symmetry (Figure 3C). Both monomers of the HBb dimer interact with the minor groove of this DNA (Figure 3C). Specifically, residues K12 in helix α1, T23 and S24 in loop l1, and A25 in helix α2 of one monomer, as well as K52 and T53 in loop l2 and R57 in helix α3 of the other monomer, form an interaction network with the phosphate groups of the DNA backbone through their side chains and backbones. This structure is largely consistent with the structure recently published by Hocher et al. (PDB: 8FW7), showing the same mode of HBb binding to DNA (Supplementary Figure S1) (32). However, the two structures differ in their crystal packings, notably in the number and arrangement of HBb dimers along the DNA (Figure 3E, Supplementary Figure S1).

In contrast, HBb-DNA_2 reveals a complementary DNA binding interface. Here, the asymmetric unit contains a 9 nucleotide (nt) long single-stranded DNA fragment, which forms dsDNA fragments via crystallographic symmetry (Figure 3D). These fragments are bound by HBb via an additional interface of the minor groove. This involves interactions between residues T8, S9 and K10, all located in the first turn of the helix α1 and conserved in other bacterial histone homologs, and the DNA phosphate backbone (Figure 3D). The crystal packing in HBb-DNA_2 markedly differs from that in HBb-DNA_1 in terms of the arrangement of the HBb dimers and DNA, containing only short DNA segments (Figure 3E, F).

The superposition of the HBb-DNA_1 and HBb-DNA_2 structures illustrates the location of the combined HBb-DNA binding interfaces (Figure 3G), resembling histone-DNA complexes found in eukaryotes and archaea. The engagement of these interfaces suggests that the DNA undergoes bending upon binding with the HBb dimers.

HBb bends DNA *in vitro*

To further substantiate the potential bending of DNA upon HBb binding, as suggested by our crystal structures, we conducted tethered particle motion (TPM) experiments with a linear 685 bp DNA fragment to determine the extent to which HBb alters DNA conformation upon binding. TPM is a single-molecule technique that allows for real-time observation of DNA–protein interactions. In this method, a DNA molecule is tethered at its ends on a bead and a glass surface, respectively, and the motion of the bead serves as a proxy for the conformational state of the DNA. TPM detects conformational changes of the DNA molecule as a function of the protein concentration, measured by the change in root mean square (RMS) displacement of the bound bead. Proteins that compact the DNA reduce the distance between the ends of the DNA, which is measured as a decrease in RMS. In the presence of HBb, the RMS values decrease significantly and reach saturation at ≈ 110 nm (Figure 4A). This value is comparable to the TPM values we measured previously for other DNA-bending proteins, such as *E. coli* HU and *Sulfolobus solfataricus* Sul7 and Cren7 (45,46). However, it is higher than the values reported for the archaeal hypernucleosome-forming histones HMfA and HMfB (19), strongly suggesting that HBb bends DNA rather than wrapping it.

We further characterized the DNA-binding properties of HBb by performing well-established assays, including micrococcal nuclease (MNase) digestion, DNA topology, and ligase-mediated circularization assays (44,47–49). In the MNase digestion assay, MNase is used to digest DNA that is not protected by bound proteins. Unlike the positive control HMfB, which tightly packs DNA and protects multiples of 30 bp fragments from endo-exonuclease digestion (11), the binding of HBb to DNA, similar to NAPs (50,51), offered no such protective effect (Figure 4B). In the DNA topology assay, we analyzed whether the binding of HBb alters the geometry of DNA. Such changes affect the supercoiling state of DNA in the presence of Topoisomerase I and can be visualized by separation of the DNA in an agarose gel. Here, both HBb and HMfB constrain supercoils in a concentration-dependent manner (Figure 4C). Using a ligase-mediated circularization assay, we next examined whether this change in DNA geometry results from DNA bending. DNA-bending proteins effectively render small DNA fragments more flexible, thus promoting the formation of small circular DNA during a ligation reaction. The unambiguous identification of circular DNA is possible by treatment of the ligation product with T5 exonuclease, followed by separation through gel electrophoresis. In the assay, fewer linear fragments and more 240 bp circles formed upon addition of increasing amounts of HBb to 240 bp linear DNA and T4 DNA ligase, indicative of DNA bending induced by HBb (Figure 4D).

MD simulations support the DNA bending model

To further validate our hypothesis that HBb functions as a DNA-bending histone dimer, we turned to Molecular Dynamics (MD) simulations. For the starting structure, we combined the HBb-DNA_1 and HBb-DNA_2 crystal structures to create a complete DNA molecule bent around the HBb dimer (Figure 5A, S2). With this starting structure, we performed two 500 ns MD simulations (Supplementary videos 1 and 2). During both 500 ns simulations, the DNA remains bent by HBb with a bending angle around 84° ± 3.4° across 36 base pairs (Figure 5B). Due to the bending, HBb compacts DNA as illustrated by the lower end-to-end distance compared to the contour length of the DNA (Supplementary Figure S3A). HBb in-

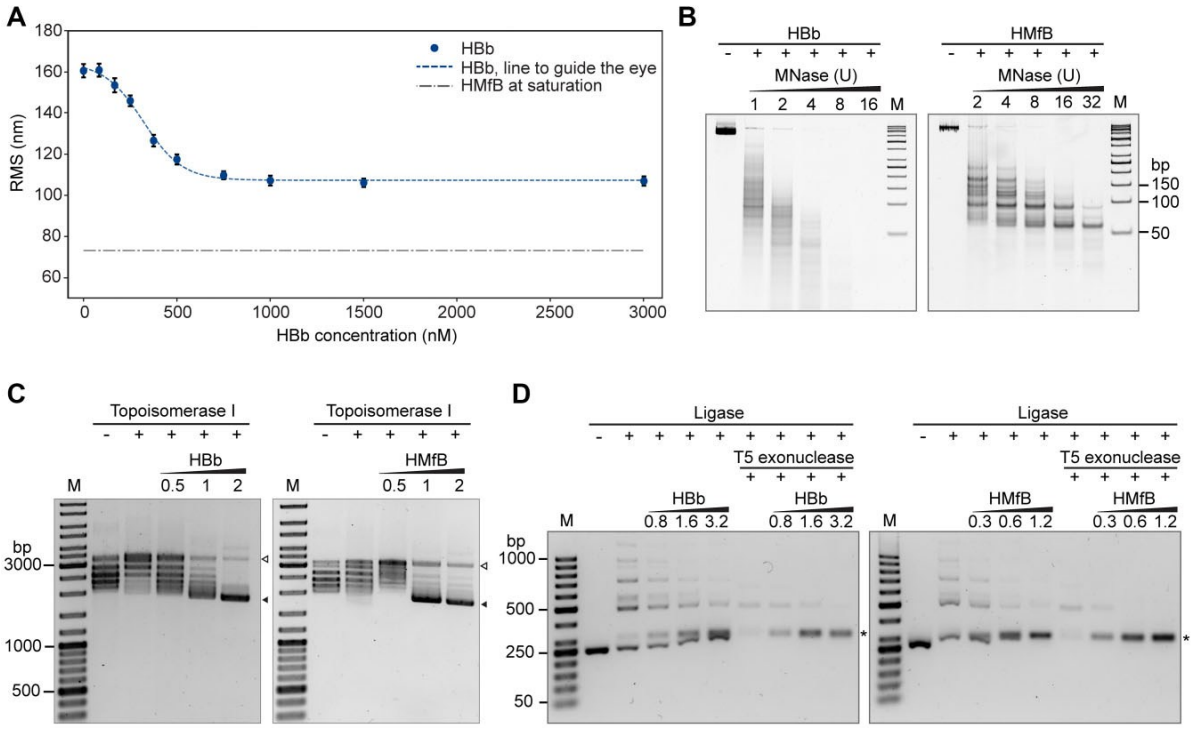


Figure 4. HBb is a DNA bender. (**A**) TPM experiment of HBb with 685 bp DNA. All measurement points are averages of triplicate measurements, with error bars indicating standard deviations. As reference, the RMS value obtained for hypernucleosome structure formed with HMfB is shown as the gray dash-dotted line (19). (**B**) MNase digestion of 685 bp DNA following incubation with HBb and HMfB. Increasing amounts of MNase in the assay are indicated in units. (**C**) DNA topology assay with relaxed pUC19 in the presence of HBb in comparison to HMfB. Protein:DNA mass ratios were labeled in the figure. The bands showing relaxed (△) and supercoiled pUC19 (▲) are marked. (**D**) Ligase-mediated circularization assay of a 240 bp DNA in the presence of HBb and HMfB. Samples are shown before and after T5 exonuclease digestion. Protein:DNA mass ratios are labeled. Monomeric DNA rings are indicated (*).

teracts exclusively with the backbone of the DNA, suggesting nonspecific DNA binding by HBb. Important DNA-binding residues are N22, R57, S9, T53, T8, S24, K10, K12, K52, and K16, ordered from the highest number of DNA contacts (14) to the lowest (5) (Figures 5C, S3B, S3C, S3D). Note that residues with less than five contacts are not considered important. These residues interact exclusively at the minor grooves of the DNA and are conserved in related bacterial histones from Bdellovibrionota, Elusimicrobiota, Spirochaetota, Planctomycetota, Myxococcota, and Chlamydiota (Figure 1A) (52). Bending by HBb causes severe deformations in the minor and major grooves (Supplementary Figure S3E, S3F). The minor grooves where HBb binds with its l1-l2 loop and α1 helix are 0.3 nm narrower, while the major grooves are 0.6 nm wider. In contrast, the major groove between the l1-l2 loop and α1 helix binding sites narrows by 0.5 nm, while the minor groove widens by 0.2 nm. This periodic variation between wide and narrow major and minor groove widths is similar to what is observed in canonical nucleosome histones (53). Overall, the DNA-bending structure of the HBb dimer is highly reminiscent of both the archaeal and eukaryotic histone dimers (Figure 5D, E). Both histones contact the DNA at their l1-l2 loops and α1 helix and interact primarily with the DNA backbone. The bending angle for both HBb and canonical histone HMfB is 84° across 36 bp (9). However, as HBb does not form larger wrapping oligomers like canonical histones, it will provide less DNA compaction than canonical histones, as illustrated by our TPM results (see 'HBb bends DNA *in vitro*').

Nonspecific binding of HBb to the *B. bacteriovorus* chromosome is essential for viability

The bacterium *B. bacteriovorus* is a well-studied model organism with a predatory life cycle. To further understand the functional significance of HBb, we aimed to replace the gene encoding HBb, *BD_RS00255*, with a kanamycin resistance cassette using the suicide vector pT18mobSacB. We achieved successful integration of the plasmid into the chromosome but failed to select for a second crossover event. This suggests that HBb is essential for *B. bacteriovorus* growth and corroborates recent findings (32).

In a subsequent step, we investigated whether HBb has DNA sequence specificity and possibly functions as a transcription factor. To this end, we conducted DAP-seq on a chromosomal scale, a recently developed method to identify genome-wide binding regions of DNA-binding proteins (38). We coupled purified FLAG-HBb to magnetic beads coated with FLAG antibodies and incubated them with a genomic *B. bacteriovorus* DNA library. Protein-bound DNA fragments were isolated, purified, amplified, and subjected to next-generation sequencing. We deviated from the original protocol by expressing FLAG-HBb in *E. coli* rather than in its natural host *B. bacteriovorus*, as this did not yield a sufficient amount of protein. Moreover, we incubated the magnetic beads directly with purified protein instead of cell extract to minimize the risk of nonspecific binding of *E. coli* proteins. DAP-seq analysis was performed for two independently prepared genomic DNA libraries with a fragment size of

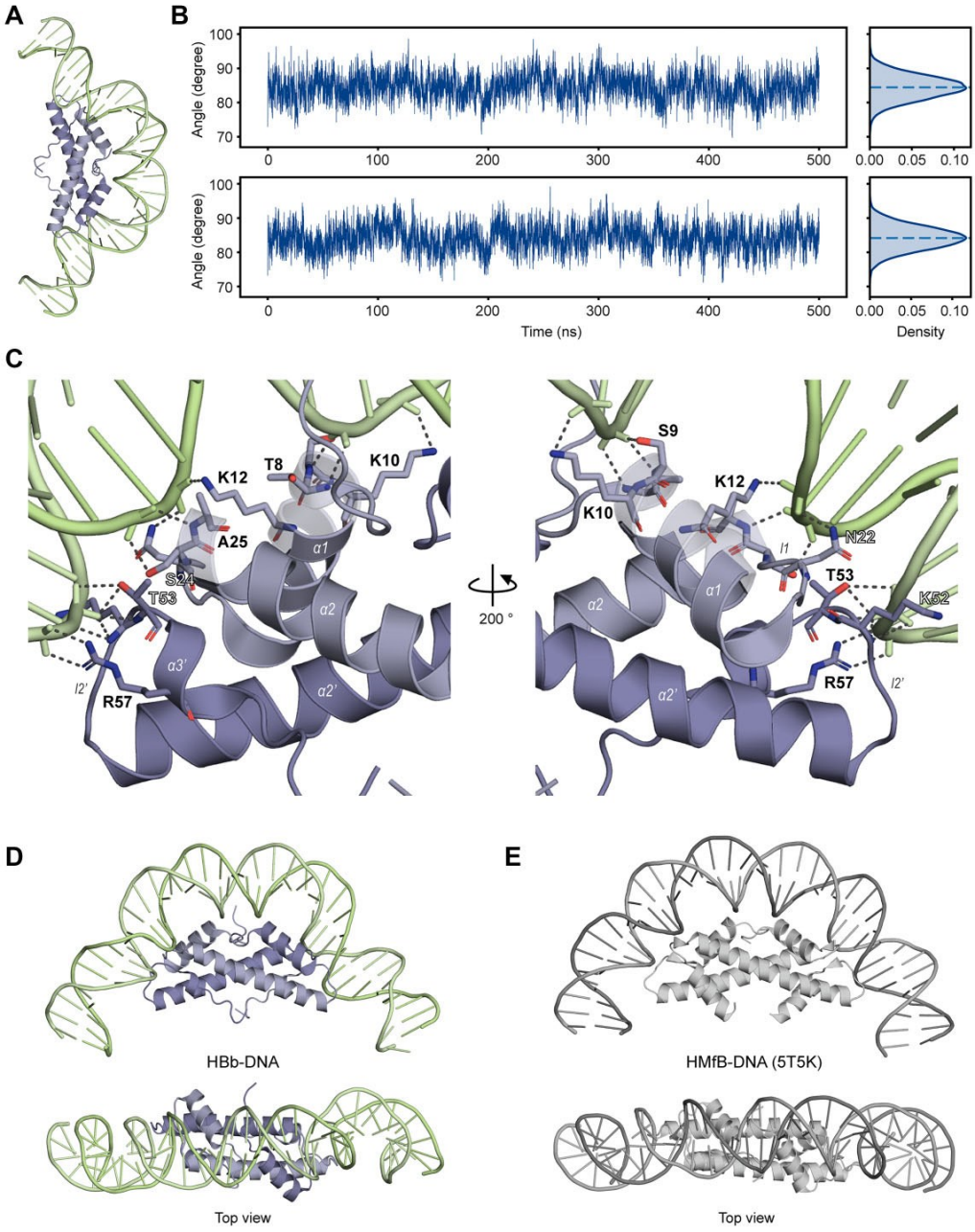


Figure 5. MD simulations supporting DNA bending by HBb. (**A**) The HBb-DNA starting model used for MD simulation. The HBb-DNA model was constructed based on HBb-DNA crystal structures followed by energy minimization, and temperature and pressure equilibration. (**B**) Plots showing the bending angle of the DNA during two independent 500 ns simulations. (**C**) The enlarged image section of the DNA-bound HBb structure obtained from MD simulations. The structure shown is the central structure of the most common structure cluster in the last 100 ns of the simulations. Residues with a large number of contacts to DNA are shown as sticks. (**D**) The side view and top view of the cartoon representation of the central structure of the most common cluster of the simulations. (**E**) The side view and top view of a section of HMfB bound to DNA (PDB: 5T5K). The cartoon representation shows an HMfB dimer with a 38-bp dsDNA wrapping around it.

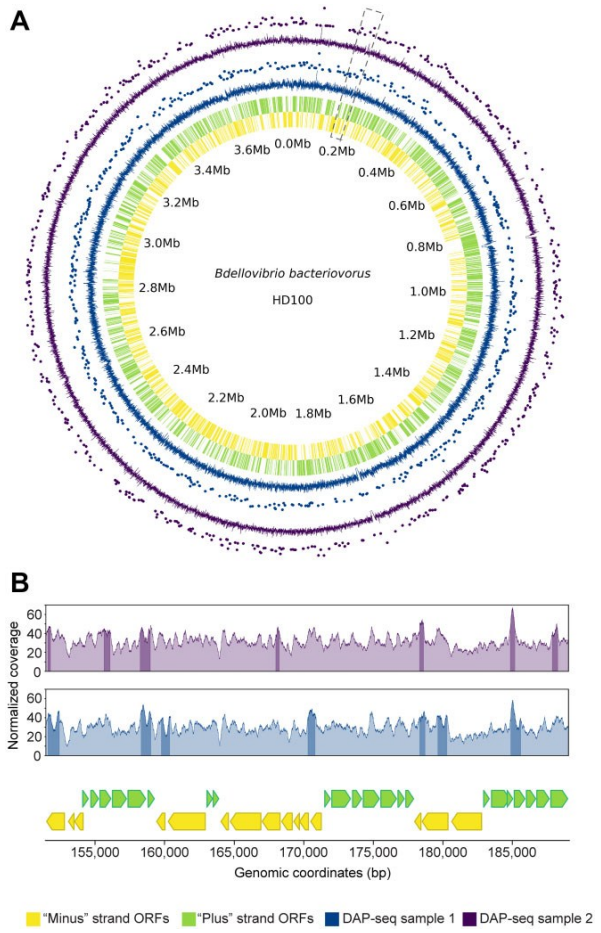


Figure 6. The HBb dimer binds DNA nonspecifically. (**A**) DAP-seq data visualized using a circos plot of the *B. bacteriovorus* HD100 genome. The inner rug plots represent the 'minus' (yellow) and 'plus' (light green) strand ORFs. Each wedge represents the coordinates of a gene. Outer line plots represent the normalized read coverage values, and scatter plots above show the major peaks for the two DAP-seq samples (blue and purple). Peaks are shown as dots, where the strength of the peak signal correlates positively with the distance from the axis. The enlarged section shown in panel E is indicated by gray dashed line. (**B**) Line plots showing the normalized read coverage values for a randomly selected section of the *B. bacteriovorus* HD100 genome encompassing the genomic region from 151 464 bp to 189 033 bp. The bottom plot shows the genomic context, and the top line plots depict the normalized read coverage values for each DAP-seq sample with the major peaks shaded darker. The color code corresponds to panel A.

approximately 200 bp, representing two biological replicates. Samples containing only DNA incubated with anti-FLAG beads in the absence of protein served as controls. For these samples, as determined by Bioanalyzer analysis, the quantity of recovered amplified DNA was less than the amount of primer dimers formed during the amplification. Since this confirms the absence of non-specifically bound DNA, sequencing of these samples was omitted.

Plotting the raw data from DAP-seq analysis on the chromosome map of *B. bacteriovorus* shows HBb binding to both genic and intergenic regions randomly throughout the genome (Figure 6). After peak-calling analysis (as described in *Supplemental Materials and Methods*), we identified a total of 475 statistically significant HBb-binding sites for DAP-seq sample 1 and 455 for DAP-seq sample 2, respectively, distributed across the chromosome (Figure 6A). Most of these peaks are between ~230 bp and ~660 bp in length reaching maximum lengths of 2416 bp and 1747 bp for DAP-seq sample 1 and 2, respectively (Supplementary Figure S4). This indicates that while HBb binds to DNA throughout the entire genome, certain regions are notably enriched with HBb dimers, which are likely spaced along the DNA. Analysis of the peaks in relation to the GC content of the DNA fragments covered yielded a value of approximately 50%, which corresponds to the GC content of 50.5% of the *B. bacteriovorus* HD100 genome (54) (Supplementary Figure S5). This suggests that HBb binds DNA without preferring GC-rich or AT-rich regions. The search for special sequence motifs proved to be infeasible given the broad length distribution and random pattern of binding sites across the chromosome.

Discussion

Building on our recent discovery of bacterial histones (31), this study investigates the DNA binding and structuring properties of the histone protein HBb in *B. bacteriovorus*. We solved the crystal structure of HBb at a resolution of 1.06 Å, which notably, makes it the first structure of a histone protein at atomic resolution; the previous structure with the highest resolution (1.48 Å) was that of histone HMfA from *M. fervidus* (PDB: 1B67) (55). HBb features the characteristic histone fold and forms dimers. Most amino acid residues crucial for DNA binding in archaeal and eukaryotic histones are conserved in HBb. However, residues facilitating tetramerization in archaeal HMfB (17,43) are absent (Figure 1A), suggesting that the dimeric state observed in both the crystal structure and biophysical data likely represent the physiological oligomer.

Employing a comprehensive array of assays, we demonstrate that HBb not only binds to DNA but also alters its topology. Our data indicate that HBb binds as individual dimers along the DNA and induces compaction through DNA bending, akin to the low-density, nonspecific binding mode of individual HU dimers (56). Importantly, MNase digestion assays and TPM measurements reveal that HBb does not compact DNA through wrapping, thereby ruling out nucleosome or hypernucleosome formation, a characteristic of canonical archaeal and eukaryotic histones (8 9,30). This is in agreement with a recent study that showed individual HBb dimers bind sequentially to DNA at increasing protein concentrations without forming hypernucleosomes (32). This binding mode is further supported by the HBb-DNA co-crystal structures presented in this study (32), yet a recently proposed model suggests that high-density binding of HBb would induce the formation of straight HBb-DNA filaments, contrasting our data that indicates compaction through bending. Such a straightening effect has been previously observed in the *E. coli* NAP HU at high density (57,58), but the physiological relevance of these structures remains unclear. It is possible that high-density binding of HU might occur on a local scale, driven by sequence determinants not yet identified. In the aforementioned study, an ensemble Förster Resonance Energy Transfer (FRET) assay was performed with a bi-terminally fluorescently labeled dsDNA molecule (32,59) and it confirmed that HBb does not trigger FRET, unlike the hypernucleosomal histone HTkA, which was used as a reference. This result supports the observation that HBb does not wrap DNA, though

it offers only limited support for the notion that straight HBb-DNA filaments are formed. In fact, their MD simulations, instead, appear to suggest the binding and bending of DNA by individual HBb dimers (32).

Previous studies have reported that the expression of the *BD_RS00255* gene, encoding HBb, is upregulated during the replicative growth phase of *B. bacteriovorus*, a phase closely linked to prey consumption (60). This observation is corroborated by the most recent RNA-seq experiments, which also show that *BD_RS00255* is expressed exclusively during the replicative phase and not during the prey attack phase (Renske van Raaphorst and Géraldine Laloux, personal communication). Multiple independent attempts to disrupt the *BD_RS00255* gene have been unsuccessful, both by us and others (32), underscoring the essential role of HBb in the survival of *B. bacteriovorus*. DAP-seq analyses confirm that HBb binds to DNA across the entire genome, a pattern also observed in bacterial NAPs of the HU/IHF family. These proteins are encoded in 93% of bacterial genomes and are thought to play a role in controlling the dynamics and organization of the bacterial chromosome (61,62).

While our study primarily focuses on the bacterial histone HBb in *B. bacteriovorus*, the evolutionary implications of our findings merit further discussion. The presence of histone-like proteins in bacteria raises compelling questions about the evolutionary trajectory leading to the intricate histone structures and DNA-compacting capabilities observed in eukaryotes. One hypothesis is that the foundational aspects of histones and their DNA-binding abilities were already in place during the time of the Last Universal Common Ancestor (LUCA). These proteins may have initially served as 'simple' NAPs and subsequently underwent a series of evolutionary adaptations, enabling them to form nucleosomes and other higher-order chromatin structures. Such adaptations could have led to a gain of function, attributed to enhanced DNA affinity and (structural) binding specificity, priming the proteins for more efficient DNA packaging, organization, and specific regulatory roles. The presence of histones in a broad spectrum of deeply-branching bacterial lineages, combined with the conservation of key amino acid residues crucial for DNA binding across bacterial, archaeal, and eukaryotic histones, lends support to this notion. However, considering the infrequent occurrence of histones in bacteria and their sequence similarity to archaeal histones, we cannot rule out the possibility that these proteins were acquired by bacteria from archaea through horizontal gene transfer events. The future availability of additional bacterial histone sequences in molecular databases, as well as comparative genomic and functional studies, could illuminate the evolutionary pathways that have shaped histones across the domains of life.

In summary, our study establishes HBb as a unique bacterial histone with distinct DNA-structuring capabilities compared to eukaryotic, archaeal, and viral histones. Although its DNA-binding interfaces are similar to archaeal and eukaryotic histones, its architectural properties closely resemble those of HU/IHF family proteins. While HU/IHF family proteins also exist in *B. bacteriovorus*, they are present in significantly lower quantities compared to HBb (32). This supports the notion of functional redundancy among different bacterial NAPs, particularly in their global effects on chromosome organization. The indispensability of HBb for *B. bacteriovorus* survival suggests that bacterial histones, like their archaeal and eukaryotic counterparts, play critical roles in cellular functions.

Data availability

Coordinates and structure factors of the crystal structures have been deposited in the PDB under accession codes 8CMP (free HBb), 9EZZ (HBb-DNA_1) and 9F0E (HBb-DNA_2).

The TPM, molecular dynamics data, and the supplementary videos are available from the 4TU repository (https://data.4tu.nl) with the DOI 10.4121/6c07e4bb-96f1-4ee0-88ef-6978bda27e5d.

The DAP-seq raw data are deposited at Zenodo (https://zenodo.org/) and accessible with the DOI 10.5281/zenodo.10211306.

The analysis code for the molecular dynamics data is available from figshare (https://figshare.com) with the DOI 10.6084/m9.figshare.25574295.

Supplementary data

Supplementary Data are available at NAR Online.

Acknowledgements

We are grateful to the staff of Beamline X10SA of the Swiss Light Source (PSI, Villigen, Switzerland) and beamline ID23-1 of the European Synchrotron Radiation Facility (Grenoble, France) for excellent technical support. We extend our thanks to Reinhard Albrecht and Antoine Schramm for assistance with crystallization, as well as crystallographic data collection and processing, and Kerstin Bär and Heidi Reichle for assistance with protein purification. We are grateful to Agnes Henschen, Min Zheng, Pengfei Liu (Dept. of Algal Development and Evolution, MPI for Biology, Tübingen, Germany) and Yihua Liu (Dept. of Microbiome Science, MPI for Biology, Tübingen, Germany) for their support with the DAP-seq methodology and data analysis. We thank the staff of the Genome Center, MPI for Biology Tübingen for coordination and help with next-generation DNA sequencing. We thank the ALICE HPC cluster at Leiden University for providing the necessary infrastructure to perform the molecular dynamics simulations.

Author contributions: Conceived and designed the experiments: A.L., B.H.A., V.A., R.T.D.; Performed the experiments: Y.H., S.S., S.D.; Analyzed the data: Y.H., S.S., P.E., T.v.H., J.D.J., J.V., M.D.H., B.H.A., V.A., R.T.D.; Wrote the paper: Y.H., S.S., J.D.J., B.H.A., V.A., R.T.D. with contributions from the other authors

Funding

Max Planck Society; Netherlands Organization for Scientific Research [OCENW.GROOT.2019.012 to R.T.D.]. Funding for open access charge: Max Planck Society.

Conflict of interest statement

None declared.

References

1. Arents,G., Burlingame,R.W., Wang,B.C., Love,W.E. and Moudrianakis,E.N. (1991) The nucleosomal core histone octamer at 3.1 A resolution: A tripartite protein assembly and a left-handed superhelix. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 10148–10152.

2. Luger,K., Mader,A.W., Richmond,R.K., Sargent,D.F. and Richmond,T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 angstrom resolution. *Nature*, **389**, 251–260.
3. Kornberg,R.D. and Thomas,J.O. (1974) Chromatin structure; oligomers of the histones. *Science*, **184**, 865–868.
4. Jenuwein,T. and Allis,C.D. (2001) Translating the histone code. *Science*, **293**, 1074–1080.
5. Luger,K. and Richmond,T.J. (1998) The histone tails of the nucleosome. *Curr. Opin. Genet. Dev.*, **8**, 140–146.
6. Sandman,K., Krzycki,J.A., Dobrinski,B., Lurz,R. and Reeve,J.N. (1990) HMf, a DNA-binding protein isolated from the hyperthermophilic archaeon Methanothermus fervidus, is most closely related to histones. *Proc. Natl. Acad. Sci. U.S.A.*, **87**, 5788–5791.
7. Bhattacharyya,S., Mattiroli,F. and Luger,K. (2018) Archaeal DNA on the histone merry-go-round. *FEBS J.*, **285**, 3168–3174.
8. Henneman,B., van Emmerik,C., van Ingen,H. and Dame,R.T. (2018) Structure and function of archaeal histones. *PLoS Genet.*, **14**, e1007582.
9. Mattiroli,F., Bhattacharyya,S., Dyer,P.N., White,A.E., Sandman,K., Burkhart,B.W., Byrne,K.R., Lee,T., Ahn,N.G., Santangelo,T.J., *et al.* (2017) Structure of histone-based chromatin in Archaea. *Science*, **357**, 609–612.
10. Ammar,R., Torti,D., Tsui,K., Gebbia,M., Durbic,T., Bader,G.D., Giaever,G. and Nislow,C. (2012) Chromatin is an ancient innovation conserved between Archaea and Eukarya. *eLife*, **1**, e00078.
11. Maruyama,H., Harwood,J.C., Moore,K.M., Paszkiewicz,K., Durley,S.C., Fukushima,H., Atomi,H., Takeyasu,K. and Kent,N.A. (2013) An alternative beads-on-a-string chromatin architecture in Thermococcus kodakarensis. *EMBO Rep.*, **14**, 711–717.
12. Pereira,S.L., Grayling,R.A., Lurz,R. and Reeve,J.N. (1997) Archaeal nucleosomes. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 12633–12637.
13. Reeve,J.N., Sandman,K. and Daniels,C.J. (1997) Archaeal histones, nucleosomes, and transcription initiation. *Cell*, **89**, 999–1002.
14. Sanders,T.J., Ullah,F., Gehring,A.M., Burkhart,B.W., Vickerman,R.L., Fernando,S., Gardner,A.F., Ben-Hur,A. and Santangelo,T.J. (2021) Extended archaeal histone-based chromatin structure regulates global gene expression in *Thermococcus kodakarensis*. *Front. Microbiol.*, **12**, 681150.
15. Sandman,K. and Reeve,J.N. (2000) Structure and functional relationships of archaeal and eukaryal histones and nucleosomes. *Arch. Microbiol.*, **173**, 165–169.
16. Sandman,K. and Reeve,J.N. (2006) Archaeal histones and the origin of the histone fold. *Curr. Opin. Microbiol.*, **9**, 520–525.
17. Stevens,K.M., Swadling,J.B., Hocher,A., Bang,C., Gribaldo,S., Schmitz,R.A. and Warnecke,T. (2020) Histone variants in archaea and the evolution of combinatorial chromatin complexity. *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 33384–33395.
18. Stevens,K.M. and Warnecke,T. (2023) Histone variants in archaea—an undiscovered country. *Semin. Cell Dev. Biol.*, **135**, 50–58.
19. Henneman,B., Brouwer,T.B., Erkelens,A.M., Kuijntjes,G.J., van Emmerik,C., van der Valk,R.A., Timmer,M., Kirolos,N.C.S., van Ingen,H., van Noort,J., *et al.* (2021) Mechanical and structural properties of archaeal hypernucleosomes. *Nucleic Acids Res.*, **49**, 4338–4349.
20. Kouzarides,T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.
21. Peeters,E., Driessen,R.P., Werner,F. and Dame,R.T. (2015) The interplay between nucleoid organization and transcription in archaeal genomes. *Nat. Rev. Micro.*, **13**, 333–341.
22. Erives,A.J. (2017) Phylogenetic analysis of the core histone doublet and DNA topo II genes of Marseilleviridae: evidence of proto-eukaryotic provenance. *Epigenetics Chromatin*, **10**, 55.
23. Thomas,V., Bertelli,C., Collyn,F., Casson,N., Telenti,A., Goesmann,A., Croxatto,A. and Greub,G. (2011) Lausannevirus, a giant amoebal virus encoding histone doublets. *Environ. Microbiol.*, **13**, 1454–1466.
24. Talbert,P.B., Armache,K.J. and Henikoff,S. (2022) Viral histones: pickpocket's prize or primordial progenitor? *Epigenetics Chromatin*, **15**, 21.
25. Liu,Y., Bisio,H., Toner,C.M., Jeudy,S., Philippe,N., Zhou,K., Bowerman,S., White,A., Edwards,G., Abergel,C., *et al.* (2021) Virus-encoded histone doublets are essential and form nucleosome-like structures. *Cell*, **184**, 4237–4250.
26. Valencia-Sanchez,M.I., Abini-Agbomson,S., Wang,M., Lee,R., Vasilyev,N., Zhang,J., De Ioannes,P., La Scola,B., Talbert,P., Henikoff,S., *et al.* (2021) The structure of a virus-encoded nucleosome. *Nat. Struct. Mol. Biol.*, **28**, 413–417.
27. Irwin,N.A.T. and Richards,T.A. (2023) Self-assembling viral histones unravel early nucleosome evolution. bioRxiv doi: https://doi.org/10.1101/2023.09.20.558576, 22 September 2023, preprint: not peer reviewed.
28. Alva,V., Ammelburg,M., Soding,J. and Lupas,A.N. (2007) On the origin of the histone fold. *BMC Struct. Biol.*, **7**, 17.
29. Qiu,Y., Tereshko,V., Kim,Y., Zhang,R., Collart,F., Yousef,M., Kossiakoff,A. and Joachimiak,A. (2006) The crystal structure of Aq_328 from the hyperthermophilic bacteria *Aquifex aeolicus* shows an ancestral histone fold. *Proteins*, **62**, 8–16.
30. Dame,R.T., Rashid,F.M. and Grainger,D.C. (2020) Chromosome organization in bacteria: mechanistic insights into genome structure and function. *Nat. Rev. Genet.*, **21**, 227–242.
31. Alva,V. and Lupas,A.N. (2019) Histones predate the split between bacteria and archaea. *Bioinformatics*, **35**, 2349–2353.
32. Hocher,A., Laursen,S.P., Radford,P., Tyson,J., Lambert,C., Stevens,K.M., Montoya,A., Shliaha,P.V., Picardeau,M., Sockett,R.E., *et al.* (2023) Histones with an unconventional DNA-binding mode in vitro are major chromatin constituents in the bacterium Bdellovibrio bacteriovorus. *Nat. Microbiol.*, **8**, 2006–2019.
33. Bogomolovas,J., Simon,B., Sattler,M. and Stier,G. (2009) Screening of fusion partners for high yield expression and purification of bioactive viscotoxins. *Protein Expr. Purif.*, **64**, 16–23.
34. Hobley,L., Fung,R.K., Lambert,C., Harris,M.A., Dabhi,J.M., King,S.S., Basford,S.M., Uchida,K., Till,R., Ahmad,R., *et al.* (2012) Discrete cyclic di-GMP-dependent control of bacterial predation versus axenic growth in Bdellovibrio bacteriovorus. *PLoS Pathog.*, **8**, e1002493.
35. Roschanski,N., Klages,S., Reinhardt,R., Linscheid,M. and Strauch,E. (2011) Identification of genes essential for prey-independent growth of Bdellovibrio bacteriovorus HD100. *J. Bacteriol.*, **193**, 1745–1756.
36. Baumgarten,T., Schlegel,S., Wagner,S., Low,M., Eriksson,J., Bonde,I., Herrgard,M.J., Heipieper,H.J., Norholm,M.H., Slotboom,D.J., *et al.* (2017) Isolation and characterization of the E. coli membrane protein production strain Mutant56(DE3). *Sci. Rep.*, **7**, 45089.
37. Bailey,K.A., Pereira,S.L., Widom,J. and Reeve,J.N. (2000) Archaeal histone selection of nucleosome positioning sequences and the procaryotic origin of histone-dependent genome evolution. *J. Mol. Biol.*, **303**, 25–34.
38. Bartlett,A., O'Malley,R.C., Huang,S.C., Galli,M., Nery,J.R., Gallavotti,A. and Ecker,J.R. (2017) Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc.*, **12**, 1659–1672.
39. Ewels,P.A., Peltzer,A., Fillinger,S., Patel,H., Alneberg,J., Wilm,A., Garcia,M.U., Di Tommaso,P. and Nahnsen,S. (2020) The nf-core framework for community-curated bioinformatics pipelines. *Nat. Biotechnol.*, **38**, 276–278.
40. Evans,R., O'Neill,M., Pritzel,A., Antropova,N., Senior,A., Green,T., Žídek,A., Bates,R., Blackwell,S., Yim,J., *et al.* (2022) Protein complex prediction with AlphaFold-Multimer. bioRxiv doi: https://doi.org/10.1101/2021.10.04.463034, 10 March 2022, preprint: not peer reviewed.

41. Bailey,K.A., Marc,F., Sandman,K. and Reeve,J.N. (2002) Both DNA and histone fold sequences contribute to archaeal nucleosome stability. *J. Biol. Chem.*, **277**, 9293–9301.
42. Erkelens,A.M., Henneman,B., van der Valk,R.A., Kirolos,N.C.S. and Dame,R.T. (2023) Specific DNA binding of archaeal histones HMfA and HMfB. *Front. Microbiol.*, **14**, 1166608.
43. Marc,F., Sandman,K., Lurz,R. and Reeve,J.N. (2002) Archaeal histone tetramerization determines DNA affinity and the direction of DNA supercoiling. *J. Biol. Chem.*, **277**, 30879–30886.
44. Bailey,K.A., Chow,C.S. and Reeve,J.N. (1999) Histone stoichiometry and DNA circularization in archaeal nucleosomes. *Nucleic Acids Res.*, **27**, 532–536.
45. Driessen,R.P., Meng,H., Suresh,G., Shahapure,R., Lanzani,G., Priyakumar,U.D., White,M.F., Schiessel,H., van Noort,J. and Dame,R.T. (2013) Crenarchaeal chromatin proteins Cren7 and Sul7 compact DNA by inducing rigid bends. *Nucleic Acids Res.*, **41**, 196–205.
46. Driessen,R.P., Sitters,G., Laurens,N., Moolenaar,G.F., Wuite,G.J., Goosen,N. and Dame,R.T. (2014) Effect of temperature on the intrinsic flexibility of DNA and its interaction with architectural proteins. *Biochemistry*, **53**, 6430–6438.
47. Broyles,S.S. and Pettijohn,D.E. (1986) Interaction of the Escherichia-Coli Hu Protein with DNA - Evidence for Formation of Nucleosome-Like Structures with Altered DNA Helical Pitch. *J. Mol. Biol.*, **187**, 47–60.
48. Musgrave,D., Forterre,P. and Slesarev,A. (2000) Negative constrained DNA supercoiling in archaeal nucleosomes. *Mol. Microbiol.*, **35**, 341–349.
49. Hodges-Garcia,Y., Hagerman,P.J. and Pettijohn,D. (1989) DNA ring closure mediated by protein HU. *J. Biol. Chem.*, **264**, 14621–14623.
50. Cajili,M.K.M. and Prieto,E.I. (2022) Interplay between Alba and Cren7 regulates chromatin compaction in *Sulfolobus solfataricus*. *Biomolecules*, **12**, 481.
51. Maruyama,H., Prieto,E.I., Nambu,T., Mashimo,C., Kashiwagi,K., Okinaga,T., Atomi,H. and Takeyasu,K. (2020) Different proteins mediate step-wise chromosome architectures in *Thermoplasma acidophilum* and *Pyrobaculum calidifontis*. *Front. Microbiol.*, **11**, 1247.
52. Schwab,S., Hu,Y., van Erp,B., Cajili,M.K.M., Hartmann,M.D., Alvarez,B.H., Alva,V., Boyle,A.L. and Dame,R.T. (2023) Novel histones and histone variant families in prokaryotes. bioRxiv doi: https://doi.org/10.1101/2023.06.01.543357, 24 May 2024, preprint: not peer reviewed.
53. Nash,J.A., Singh,A., Li,N.K. and Yingling,Y.G. (2015) Characterization of nucleic acid compaction with histone-mimic nanoparticles through all-atom molecular dynamics. *Acs Nano*, **9**, 12374–12382.
54. Rendulic,S., Jagtap,P., Rosinus,A., Eppinger,M., Baar,C., Lanz,C., Keller,H., Lambert,C., Evans,K.J., Goesmann,A., *et al.* (2004) A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a genomic perspective. *Science*, **303**, 689–692.
55. Decanniere,K., Babu,A.M., Sandman,K., Reeve,J.N. and Heinemann,U. (2000) Crystal structures of recombinant histones HMfA and HMfB from the hyperthermophilic archaeon *Methanothermus fervidus*. *J. Mol. Biol.*, **303**, 35–47.
56. Pinson,V., Takahashi,M. and Rouviere-Yaniv,J. (1999) Differential binding of the Escherichia coli HU, homodimeric forms and heterodimeric form to linear, gapped and cruciform DNA. *J. Mol. Biol.*, **287**, 485–497.
57. Skoko,D., Wong,B., Johnson,R.C. and Marko,J.F. (2004) Micromechanical analysis of the binding of DNA-bending proteins HMGB1, NHP6A, and HU reveals their ability to form highly stable DNA— protein complexes. *Biochemistry*, **43**, 13867–13874.
58. van Noort,J., Verbrugge,S., Goosen,N., Dekker,C. and Dame,R.T. (2004) Dual architectural roles of HU: formation of flexible hinges and rigid filaments. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 6969–6974.
59. Lorenz,M., Hillisch,A., Goodman,S.D. and Diekmann,S. (1999) Global structure similarities of intact and nicked DNA complexed with IHF measured in solution by fluorescence resonance energy transfer. *Nucleic Acids Res.*, **27**, 4619–4625.
60. Karunker,I., Rotem,O., Dori-Bachash,M., Jurkevitch,E. and Sorek,R. (2013) A global transcriptional switch between the attack and growth forms of Bdellovibrio bacteriovorus. *PLoS One*, **8**, e61850.
61. Holowka,J., Trojanowski,D., Ginda,K., Wojtas,B., Gielniewski,B., Jakimowicz,D. and Zakrzewska-Czerwinska,J. (2017) HupB is a bacterial nucleoid-associated protein with an indispensable eukaryotic-like tail. *mBio*, **8**, e01272.
62. Prieto,A.I., Kahramanoglou,C., Ali,R.M., Fraser,G.M., Seshasayee,A.S. and Luscombe,N.M. (2012) Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in Escherichia coli K12. *Nucleic Acids Res.*, **40**, 3524–3537.

Received: October 18, 2023. **Revised:** May 1, 2024. **Editorial Decision:** May 22, 2024. **Accepted:** May 24, 2024

Supplementary Information Appendix

Bacterial histone HBb from *Bdellovibrio bacteriovorus* compacts DNA by bending

Yimin Hu[#], Samuel Schwab[#], Silvia Deiss, Pedro Escudeiro, Thor van Heesch, Joe D. Joiner, Jocelyne Vreede, Marcus D. Hartmann, Andrei N. Lupas, Birte Hernandez Alvarez*, Vikram Alva*, Remus T. Dame*

*To whom correspondence should be addressed
E-mail: birte.hernandez@tuebingen.mpg.de
or vikram.alva@tuebingen.mpg.de
or rtdame@chem.leidenuniv.nl

This file includes:

SI Appendix, Figure S1 to S5
SI Appendix, Material and Methods
SI Appendix, Table S1 to S3
SI Appendix, References

SI Appendix, Figures

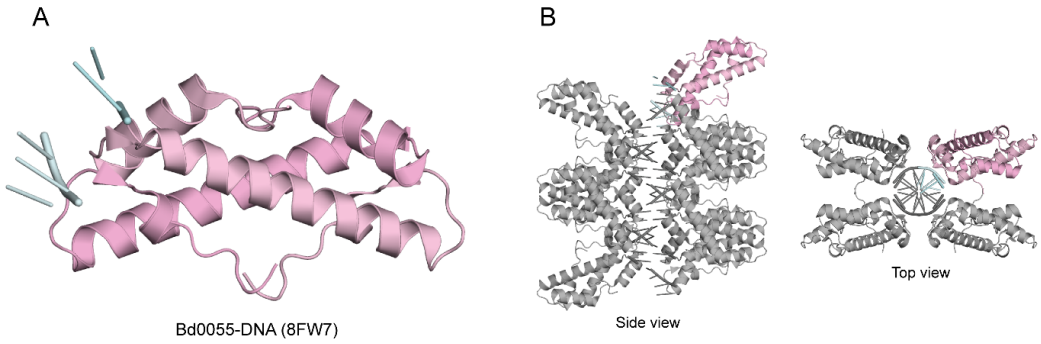


Fig. S1. Crystal structure and crystal packing of DNA-bound HBb (Bd0055) structure (PDB: 8FW7) published by Hocher *et al.* (1). A. The protein is shown in cartoon representation (in light pink) bound to DNA (in light cyan). **B.** The related crystal packing, visualized by selected symmetry mates (in gray) generated within 4 Å.

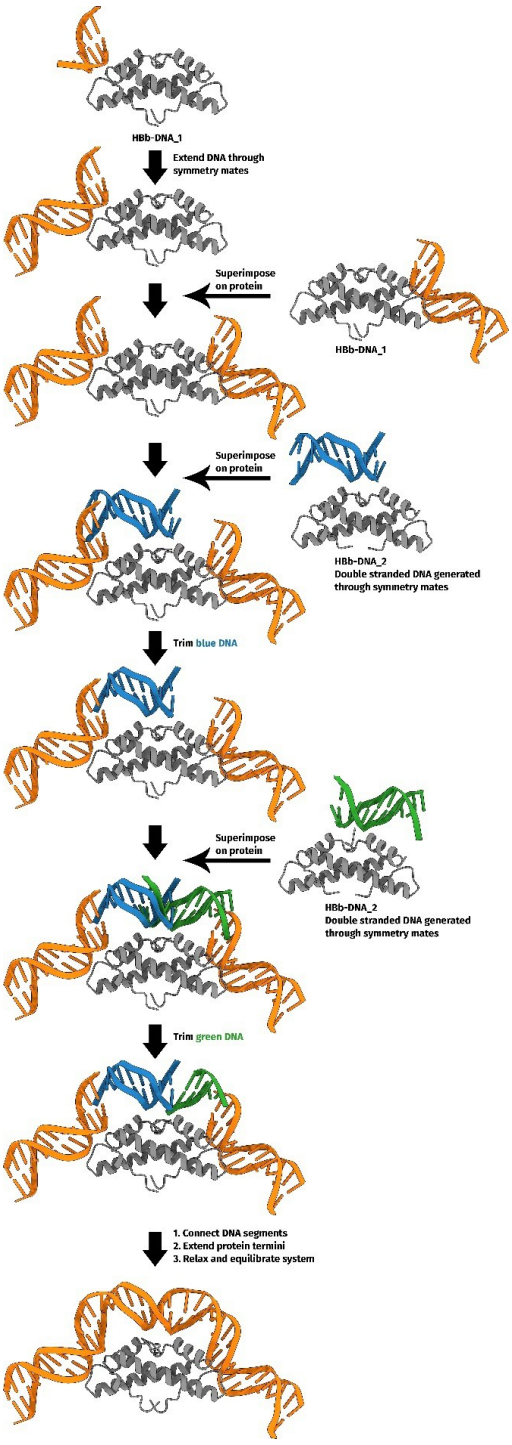


Fig. S2. Method of generating a complete DNA molecule bent around HBb.

The starting structure for the molecular dynamics simulations was constructed from the HBb-DNA_1 and 2 crystal structures. The DNA from HBb-DNA_1 is colored in orange. The DNA from HBb-DNA_2 is colored in blue and green. Protein is colored in gray.

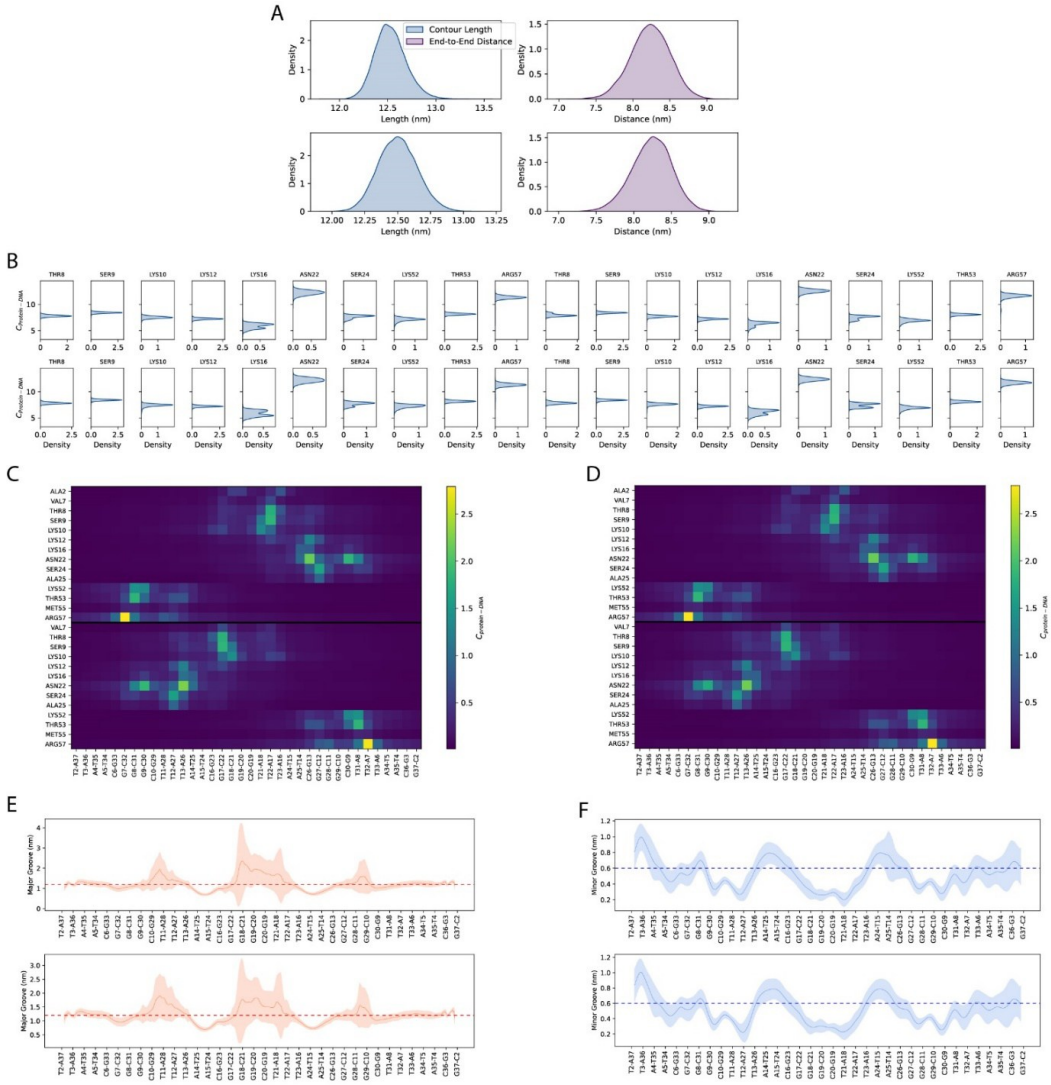


Fig. S3. DNA structure and protein-DNA interaction analysis of the two 500 ns MD simulations. A. Density estimates of the contour length and end-to-end distance for the DNA molecules of run 1 (top) and run 2 (bottom). **B.** Decomposed density estimates of the number of contacts of each protein residue with respect to the complete DNA of run 1 (top) and run 2 (bottom). Only residues with an average number of contacts above 5 are shown. The data for each monomer are concatenated. **C, D.** Heat map of the mean number of contacts of each protein residue with each base pair for run 1 (**C**) and run 2 (**D**). The data for each protein monomer is concatenated with a black dividing line separating the data of each monomer. **E, F.** Major (**E**) and minor (**F**) groove widths of the DNA molecules of run 1 (top) and run 2 (bottom). The shaded area represents one standard deviation. The normal groove width of B-DNA is indicated by the striped line.

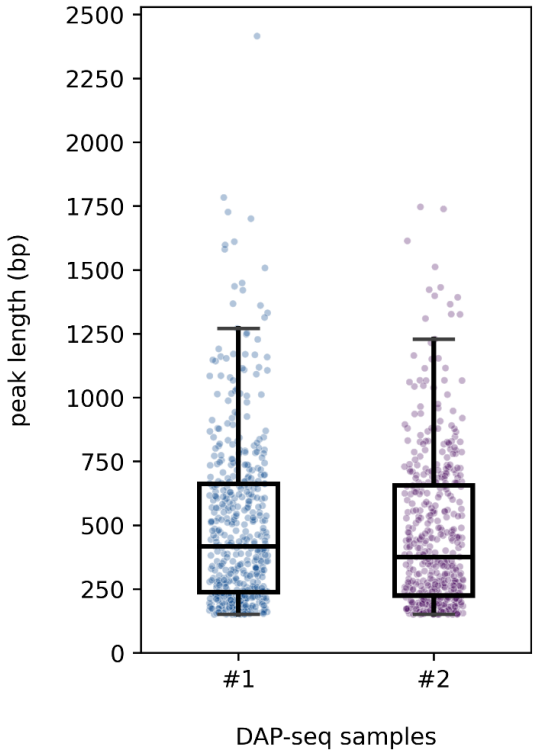


Fig. S4. Peak length distribution for DAP-seq samples 1 and 2.

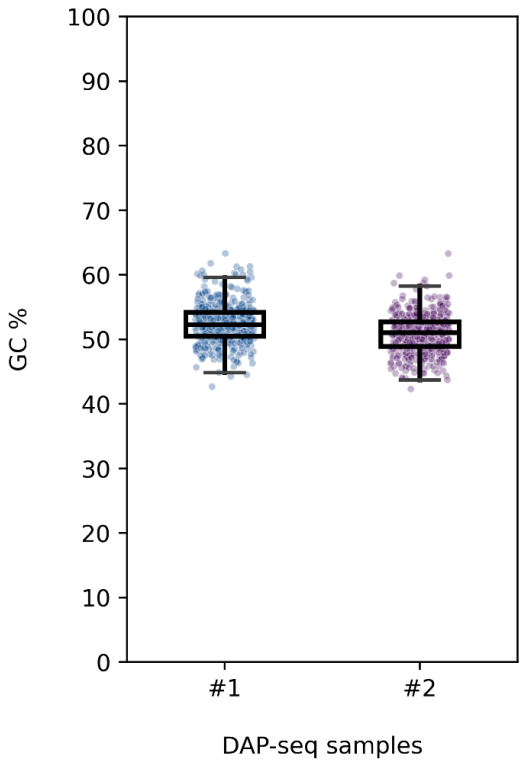


Fig. S5. GC content distribution of the genomic segments located between the coordinates of a given peak for DAP-seq samples 1 and 2.

SI Appendix, Material and Methods

Bacterial strains and cultivation

Cloning procedures were performed in *Escherichia coli* Top10 cells. *E. coli* strain Mutant56(DE3) (2) was used as the expression host. *E. coli* cells were grown on Luria-Bertani (3) medium.

B. bacteriovorus HD100 (DSM 50701, DSMZ) was cultured at 30°C using *E. coli* Top10 as prey cells in PYE broth. Media was supplemented with antibiotics as required.

Cloning, plasmids, and synthetic DNA

The genes encoding HBb (*BD_RS00255,* old locus tag: *Bd0055*) and HMfB (GenBank accession number M34778.1) were codon-optimized and synthesized (Synbio Technologies; BioCat GmbH). The *hbb* gene was cloned into the expression vector pETHis1a (4) for recombinant expression of HBb (UniProt entry Q6MRM1) fused to an N-terminal $(histidine)_6$ tag and FLAG tag, respectively. The HMfB-encoding gene (UniProt entry P19267) was cloned in pET28a for expression with an N-terminal $(histidine)_6$ tag.

For chromosomal deletion of *BD_RS00255* in *B. bacteriovorus*, the suicide vector pT18mobSacB, kindly gifted by Brian Kvitko (Addgene plasmid # 72648; http://n2t.net/addgene:72648; RRID: Addgene 72648), was used. The kanamycin resistance (Km^r) cassette was amplified from plasmid pT2SK, provided as a gift from Shelley Copley (Addgene plasmid # 59383; http://n2t.net/addgene:59383; RRID:Addgene_59383). Vector pUC18 was used for subcloning. Synthetic oligonucleotides (Merck) used in this study are listed in Table S1.

Deletion of *BD_RS00255*

For deletion of *BD_RS00255* in *B. bacteriovorus* HD100, the genomic regions upstream and downstream of *BD_RS00255* (nucleotides 47964-49345 and 49567-50958 of the complete genome of *B. bacteriovorus* strain HD100; segment 1/11, Sequence ID: BX842646.1) were amplified by PCR using specific primers (Table S1) and cloned into pUC18. The Km^r cassette was amplified by PCR from vector pT2SK and inserted into pUC18 between both fragments. A fragment comprising the Km^r cassette flanked by the *BD_RS00255* upstream and downstream genomic region was amplified by PCR and subcloned into the suicide vector pT18mobSacB. The resulting plasmid was introduced into *B. bacteriovorus* via conjugation transformation. Chromosomal integration of the plasmid by homologous recombination was verified by PCR and sequencing. To select for a second crossover event resulting in the replacement of *BD_RS00255* by the Km^r cassette, cells were grown on medium containing 5-10% sucrose.

Protein expression and purification

For expression of FLAG-HBb, His-HBb, and HMfB in *E. coli* Mutant56(DE3), cells that were transformed with the respective plasmids encoding were cultivated in LB broth, supplemented with kanamycin at 37°C. At an OD_{600} of 0.5, IPTG was added at a final concentration of 1 mM to induce protein expression. Following further cultivation at 25°C for 16 h for FLAG-HBb, and at 37°C for 4 hours for His-HBb and HMfB, respectively, cells were harvested by centrifugation. All purification steps were performed at 4°C. For HBb proteins, all buffers were supplemented with 2-mercaptoethanol (2.5 mM) to prevent cysteine disulfide bridge formation. Cell pellets were resuspended in a buffer containing 20 mM Tris, pH 8.0, 300 mM NaCl and 10 mM imidazole, supplemented with protease inhibitor mix (cOmplete™, EDTA-free Protease Inhibitor Cocktail, Roche), 0.1 mM PMSF, 3 mM $MgCl_2$ and DNase. Cells were lysed by sonication for 15 min. Cell debris and membranes were pelleted at 95,000 x g for 45 min. The supernatant was filtered with a 0.45 µm filter and applied on a 5 mL HisTrap column (Cytiva). The bound proteins were eluted with a linear gradient of 10-500 mM imidazole in the aforementioned buffer. For His-HBb, the histidine tag was cleaved with TEV protease. Following a second HisTrap column to remove the histidine-tagged TEV, HBb was purified to homogeneity by gel filtration chromatography on a Superdex 75 column (Cytiva) equilibrated with SEC buffer (20 mM Tris, pH 8.0, 150 mM NaCl). Protein purity was assessed by SDS-PAGE (15%) and BN-PAGE (4-16%, ThermoFisher Scientific). The protein concentration was determined using the BCA protein assay (ThermoFisher Scientific). Purified HBb was dialyzed against 50 mM Tris, pH 7.0, 75 mM KCl and 10% Glycerol for further assays.

Circular dichroism (CD) spectroscopy

HBb sample was dialyzed in a dialysis buffer (10 mM Tris, pH 8.0, 100 mM KF and 1 mM 2-mercaptoethanol) overnight, and then diluted with the dialysis buffer to a final concentration of 25 µM. CD spectra were measured with a Jasco J-810 spectrometer (JASCO) in a wavelength range of 190-240 nm, using a cuvette with a path length of 1 mm and a read speed of 100 nm/min. A total of ten single spectra were recorded and averaged. Thermal stability of HBb was analyzed measuring the ellipticity (θ) at 220 nm over a temperature gradient of 20-100°C applying a ramp of 1°C/min, a data pitch of 0.5, and a response time of 1 s. Data analysis, including blank subtraction and curve smoothing, was performed using the Spectra Manager software suite (JASCO). CD spectrum and the melting curve were plotted using Prism 9 (GraphPad).

Size exclusion chromatography coupled with multi-angle light scattering (SEC-MALS)

SEC-MALS of HBb was performed in SEC buffer. HBb at a concentration of 2.7 mg/mL was incubated on ice for 20 min before loading onto a Superdex 75 Increase 10/300 GL column

(Cytiva), equilibrated with the corresponding buffer. The run was performed at a flow rate of 0.5 mL/min on a 1260 Infinity II HPLC system (Agilent) coupled to a miniDAWN TREOS and Optilab T-rEX refractive index detector (Wyatt Technology). HBb was detected at 215 nm and measured in triplicate. Molar mass distributions were calculated using the ASTRA v.7.3.0.18 software suite (Wyatt Technology).

Crystallization, data collection, and structure determination

Crystallization trials were performed by mixing 300 nL of protein with 300 nL of reservoir solution in 96-well sitting-drop vapor-diffusion plates using commercially available screens with reservoir volumes of 100 µL. In free form, HBb was set up at a concentration of 2.7 mg/ml in 20 mM Tris, pH 8.0, 150 mM NaCl, 1 M Urea, and best crystals were obtained with a reservoir solution containing 0.1 M CH_3COONa, pH 4.5 and 25% (w/v) PEG 3350. For co-crystallization, the protein was incubated with the randomized double-stranded DNA fragment 20 bp-GC50 (Table S1) at a volume ratio of 1:1 at 37 °C for 10 min, and aggregates removed by centrifugation after cooling to room temperature. Best crystals were obtained with a protein solution containing 800 µM HBb and 200 µM DNA and reservoir containing 0.1 M HEPES, pH 7.0, 0.2 M LiCl and 20% (w/v) PEG 6k (HBb-DNA_1) and protein solution containing 600 µM HBb and 150 µM DNA and a reservoir of 0.2 M $(NH_4)_2SO_4$, 0.1 M CH_3COONa, 22% (w/v) PEG 4k (HBb-DNA_2). For cryo protection, crystals were transferred to droplets of their reservoir solution spiked with 10% (v/v) PEG 400 (free HBb), 20% (v/v) PEG 400 (HBb-DNA_1) or 15% (v/v) PEG 200 (HBb-DNA_2), loop-mounted, and flash-cooled in liquid nitrogen. Data were collected at beamline X10SA of the Swiss Light Source (Villigen, Switzerland) (free HBb) and beamline ID23-1 of the European Synchrotron Radiation Facility (Grenoble, France) (HBb-DNA_1, HBb-DNA_2) at 100 K, using EIGER X 16M hybrid pixel detectors (Dectris Ltd., Switzerland). Data were reduced, processed and scaled using XDS (5).

The structure of free HBb was solved by molecular replacement (MR) using Phaser (6) and HMk (PDB: 1F1E) as a search model, locating one monomer in the asymmetric unit (AU). The structure was rebuilt and refined in cycles with manual modelling in Coot (7) and refinement with REFMAC5 (8). DNA-bound structures were solved by MR using MOLREP (9) and the refined free HBb coordinates as a model, locating one dimer in the AU for both structures. After initial rigid body refinement with REFMAC5 (8), electron density for double-stranded DNA fragments became apparent in both structures in different locations, which were modeled manually based on the DNA sequence. All structures were finalized in cycles of manual modelling in Coot (7) and refinement in REFMAC5 (8). Data processing and refinement statistics are listed in Table S3, and the coordinates and structure factors deposited in the PDB under accession codes 8CMP (free HBb), 9EZZ (HBb-DNA_1) and 9F0E (HBb-DNA_2).

Microscale thermophoresis (MST)

To determine the DNA binding affinity of HBb, Cy5-labelled 80 bp DNA (10,11) and 80 bp-GC50 oligonucleotides were annealed as described. A dilution series of label-free HBb was prepared in 20 mM Tris, pH 8.0, 150 mM NaCl and 2.5 mM 2-mercaptoethanol and titrated against the annealed DNA fragments at a concentration of 10 nM. The HBb-DNA mixtures were incubated at 37°C for 10 min, centrifuged to remove precipitates, and loaded into Monolith NT premium capillaries (MO-K025, NanoTemper Technologies) for measurement. HMfB was used as a control. Similarly, a dilution series of the label-free HMfB was titrated against both DNA fragments at the concentration of 20 nM. HMfB-DNA mixtures were incubated at RT for 5 min, centrifuged and loaded into Monolith NT capillaries (MO-K022). All measurements were repeated in triplicate and performed at 25°C using the Monolith NT.115 instrument with a Nano RED Detector and MST power set to medium. MST data were analyzed by fitting them to a K_d model using MO.Control (NanoTemper Technologies).

DNA binding analysis

Electrophoretic mobility shift assay (EMSA)

For *in vitro* DNA-binding tests, an 80 bp DNA fragment (10,11) was used. For annealing, complementary oligonucleotides were incubated at 95°C for 5 min, followed by slow cooling to room temperature. The annealed 80 bp DNA was mixed with protein at different ratios in 25 mM Tris, pH 8.0, 50 mM NaCl, 50 mM KCl and incubated at 37°C for 10 min. Samples were separated on a 6% DNA retardation gel (ThermoFisher Scientific) at 100 V and DNA was visualized by SYBR Gold (ThermoFisher Scientific) staining.

To approximate the length of DNA fragments required for protein binding, GeneRuler Ultra Low Range DNA Ladder (ThermoFisher Scientific) was mixed with HBb or HMfB in 50 mM Tris, pH 7.0, 75 mM KCl at indicated ratios and incubated at room temperature for 30 min. Glycerol was added to a final concentration of 10%, and samples were separated on a self-made 10% polyacrylamide gel in TAE buffer at 120 V. DNA was stained with GelRed (Biotium) and visualized using Gel Doc XR+ imaging system (Bio-Rad Laboratories).

Micrococcal nuclease (MNase) digestion assay

For HBb-DNA complex formation, 1440 ng HBb was mixed with 900 ng 685 bp DNA in 50 mM Tris, pH 7.0, 75 mM KCl, and incubated at room temperature for 30 min. Then, indicated amounts of MNase (New England BioLabs) were added and incubation was continued in digestion buffer, provided by the manufacturer, for 12 min at 37°C in a total reaction volume of 80 µL. The reaction was stopped with EDTA and SDS at a final concentration of 95 mM and 0.5%, respectively. The protected DNA was recovered by phenol/chloroform extraction, followed by ethanol precipitation. The purified DNA was dissolved in 5 µL nuclease-free H_2O

and separated on a self-made 10% polyacrylamide gel in TAE buffer at 120 V for 1 h. To visualize the DNA, the gel was stained with GelRed (Biotium) and imaged using Gel Doc XR+ imaging system (Bio-Rad Laboratories). For the positive control HMfB, 540 ng protein was incubated with 900 ng 685 bp DNA and the assay was performed as described.

DNA topology assay

E. coli DH5α cells were freshly transformed with plasmid pUC19 and grown to an OD_{600} of 2 to isolate the plasmid using GeneJET Plasmid Miniprep Kit (ThermoFisher Scientific). Topologically relaxed pUC19 was obtained by nicking with endonuclease Nb.BsrDI (New England BioLabs) followed by covalent closure with T4 DNA ligase (ThermoFisher Scientific). Successful relaxation was confirmed by agarose gel electrophoresis. In the assay, 200 ng of relaxed pUC19 was mixed with 100 ng, 200 ng, and 400 ng HBb and HMfB, respectively, in 50 mM Tris, pH 7.0, 75 mM KCl. After 30 min incubation at room temperature, 7.5 U of topoisomerase I (type IB) (ThermoFisher Scientific) was added and incubation continued for 60 min at 37°C in a total reaction volume of 50 µL. Pure DNA was obtained by phenol/chloroform extraction followed by ethanol precipitation. Samples were separated on a 1% TAE agarose gel. To visualize different topological states of the plasmid, DNA was stained with GelRed (Biotium) and imaged using Gel Doc XR+ imaging system (Bio-Rad Laboratories).

Ligase-mediated circularization assay

The ligase-mediated circularization assay was performed with a linear DNA fragment of 240 bp comprising the first 240 bp of the DNA substrate used in the TPM experiments. 380 ng DNA was mixed with 108 ng, 216 ng and 432 ng HMfB, and 288 ng, 576 ng and 1152 ng HBb in 50 mM Tris, pH 7.0, 75 mM KCl. Following incubation at RT for 30 min, $MgCl_2$, ATP, DTT, and T4 DNA ligase were added to final concentrations of 10 mM, 1 mM, 10 mM, and 0.2 U/µL, respectively, in a total volume of 100 µL and incubated for 24 hours at room temperature. The DNA was purified by phenol/chloroform extraction and ethanol precipitation. One third of purified DNA was subjected to T5 exonuclease (New England BioLabs) treatment in CutSmart™ buffer (New England BioLabs) for 60 min at 37°C. DNA samples were separated on a 2% TAE agarose gel; stained with GelRed (Biotium) and visualized using Gel Doc XR+ imaging system (Bio-Rad Laboratories).

Tethered Particle Motion (TPM) experiments

TPM experiments were performed as previously described (12) using 50 mM Tris, pH 7.0, 75 mM KCl as buffer. A standard deviation cutoff of 8% and an anisotropic ratio cutoff of 1.3 were used to select single-tethered beads. Measurements at each HBb concentration were done in triplicate. Means and standard deviations of the individual measurement series for each HBb

concentration were calculated by maximum likelihood estimation assuming a normal distribution. Outliers with a Z-score >3 or <-3 were not considered for fitting. The “line to guide the eye” was generated by fitting the means to a logistic function. A custom Python script was used for fitting and plotting the TPM data. For plotting, the means of the three individual measurements were averaged for each measured concentration and the standard deviations were error-propagated ($Std(\bar{X}) = \sqrt[2]{\frac{\sum_{i=1}^{n} Var(X_i)}{n^2}}$).

DAP-seq

Library construction and DNA precipitation

Genomic DNA (gDNA) libraries of *B. bacteriovorus* were constructed with an average fragment length of 200 bp as described (13). For DNA precipitation, 5 µg of purified FLAG-HBb was diluted in Binding buffer (PBS supplemented with 1 mM 2-mercaptoethanol) to a final volume of 400 µL and incubated with 25 µL of pre-washed Anti-FLAG M2 magnetic beads (Sigma-Aldrich) for 60 min at room temperature in an orbital shaker. Subsequently, the beads were washed 5 times with Binding buffer supplemented with 0.005% NP-40, followed by three wash steps with non-supplemented Binding buffer. Washed beads were resuspended in 80 µL of Binding buffer containing 1 µg of adaptor-ligated gDNA library and incubated for 60 min at room temperature in an orbital shaker. The beads were washed a total of eight times with 200 µL of Binding buffer, with the first 5 washing steps conducted in the presence of 0.005% NP-40. Bound DNA-protein complexes were eluted from the beads by incubation with 30 µL 3xFLAG peptide (100 ng/µL, Sigma-Aldrich) at room temperature for 30 min in an orbital shaker. Eluted DNA was amplified using different pairs of indexed primers for each sample. Removal of primer dimers and size selection of PCR products was performed using AMPure XP beads (Beckman Coulter) according to the instruction and subsequent elution in 20 µL 0.1x TE buffer. The concentration of each library was measured using Qubit dsDNA HS assay kit (Invitrogen) and the average size of each library was determined by Bioanalyzer DNA HS kit (Agilent). Libraries were diluted to a concentration of 2 nM and pooled for sequencing. The negative control experiment was conducted by incubating the gDNA library with beads in the absence of protein. All DAP-seq experiments were repeated twice.

Next-generation Sequencing

Sequencing was performed using an Illumina NextSeq 2000 instrument. An average of 2 million pair-end 151-bp reads per sample were generated.

Data analysis

The paired-end FASTQ files were processed with the nf-core/chipseq (version 2.0.0) pipeline (14), under Nextflow (version 23.04.2). The *B. bacteriovorus* HD100 genome (assembly ASM19617v1, RefSeq accession GCF_000196175.1) was used as reference, and the respective nucleotide FASTA, and GTF annotation files were provided as inputs to the pipeline. The chosen configuration profile was that of Singularity, and the effective genome size parameter was set to 3,782,950. All other parameters were left to default. The BAM files generated by the nf-core/chipseq pipeline (".mLb.clN.sorted.bam" suffix) were provided as input to MACS (version 3.0.0b3) (15). MACS was run without building a model, with a shift-size of 151 bp, the "broad" flag set, and the effective genome size parameter set to 3,782,950. All other parameters were left to default. The BED files generated by MACS ("peaks.gappedPeak" suffix), were analyzed with pandas (version 1.5.3) (16). The GC content of the genomic segments located between the coordinates of a given peak was calculated with BioPython (version 1.81) (17), using these BED files, and the nucleotide FASTA for the genome described above as inputs. The bigWig files generated by the nf-core/chipseq pipeline were analyzed with Bioframe (version 0.4.1) (18), in conjunction with the genomic features table file, and GenBank flatfile pertaining to the genome described above. Normalized read coverage values, peak coordinates, peak signal values, peak length, GC content, and genomic coordinates were visualized with pyCircos (github.com/ponnhide/pyCircos) and Matplotlib (version 3.7.1) (19). All Python packages were installed and used under Python 3.10.12.

Molecular dynamics (MD) simulations

For MD simulations, we used GROMACS (20) with the AMBER ff14sb-ParmBSC1 force field (21-23). All systems were solvated in a dodecahedron box with a distance of at least 1.0 nm to the box boundary and filled with SPC/E water molecules (24). Water molecules were replaced at random with Na^+ and Cl^- ions to charge-neutralize the system. 75 mM of Na^+ and Cl^- ions were additionally added. Energy minimization was performed using the steepest descent method for 5000,000 steps until the largest force was below 1000.0 kJ/mol/nm. To equilibrate the solvent and ions, heavy atoms were position constrained for 100 ps at 310 K and 1 bar. The cutoff for van der Waals interactions was 1.1 nm. Electrostatic interactions beyond a cutoff of 1.1 nm were treated with the particle-mesh Ewald method using a grid spacing of 0.16 nm. Temperature and pressure were kept constant with the V-rescale thermostat (Bussi-Donadio-Parrinello thermostat) (25) and the Parrinello-Rahman barostat (26), respectively. Bonds were constrained with LINCS (27) and simulations were conducted in time steps of 2 fs.

The starting structure for the MD, with DNA bent around the HBb dimer, was constructed from the HBb-DNA_1 and HBb-DNA_2 crystal structures (Fig. S2). The DNA at the l1-l2 loops

of the dimer comes from the HBb-DNA_1 crystal structure. The DNA from HBb-DNA_1 was extended through crystal symmetry mates. The DNA between the two l1-l2 loops was filled in with the DNA from the HBb-DNA_2 crystal structure by superimposition on the protein, followed by trimming of the DNA. All DNA segments were connected together and the sequence was changed to the sequence of the 20 bp-GC50 fragment used for crystallization. For the final structure, the protein structure of HBb-DNA_2 was used and the missing terminal protein residues were filled in based on an AlphaFold prediction of the HBb dimer, generated with ColabFold (28,29). All protein superimpositions, trimming of the DNA, and the changes to the DNA sequence were conducted in ChimeraX v1.7.1 (30). Crystal symmetry mates were generated in Open-Source PyMOL v2.5.0 (31,32).

Snapshots were collected every 40 ps from the last 100 ns of both runs and clustered on RMSD of all protein and DNA atoms with a cutoff value of 0.2 Å in GROMACS with the gromos algorithm (33). Clustering resulted in 60 clusters, the largest of which contained 1873 snapshots (37% of all snapshots analysed during clustering). From the most abundant cluster, the central structure, i.e., the structure with the smallest average RMSD among all other structures of the cluster, was used as the representative structure of the MD in Figure 5C and 5D.

The atoms involved in protein-DNA contacts were identified as described in a previous study by van Heesch et al. (34).

The bending angle of the DNA was computed by first converting the DNA base pairs into a rigid body model (35), which defines mean reference frames for each base pair. Then, after defining vectors through the origins of the mean reference frames of the second base pair and center base pair, and the second to last base pair and center base pair, the bending angle was computed as 180 ° minus the inverse cosine of the dot product of these vectors. For the bending angle of HMfB, the same procedure was performed on the crystal structure of HMfB bound to DNA (PDB: 5T5K) (11).

Data was plotted in Python with the Matplotlib package (19,32).

SI Appendix, Tables

Table S1

Oligonucleotides used in this work

Oligo name	Sequence (5'-3')	Application
hbb up fwd (SalI)	5'-CTAGTCGACCATGTGTGAAGACCCCATCAATC-3'	*hbb* deletion, amplification of *BD_RS00255* upstream region
hbb up rev (Acc65I)	5'-CTAGGTACCCCATGTTAATCGAACAGATAAAAG-3'	*hbb* deletion, amplification of *BD_RS00255* upstream region
hbb dwn fwd (Acc65I)	5'-CTAGGTACCGAAAATTCGTTCAGCATTGGTTC-3'	*hbb* deletion, amplification of *BD_RS00255* downstream region
hbb dwn rev (SacI)	5'-CTAGAGCTCCTTGGGTGTGAAGATCATCTCTTG-3'	*hbb* deletion, amplification of *BD_RS00255* downstream region
Km fwd (Acc65I)	5'-CTAGGTACCCTCTGATGTTACATTGCACAAGATAAAA-3'	*hbb* deletion, amplification of Km^r cassette
Km rev (Acc65II)	5'-CTAGGTACCTCCTGCGTTACGCCCCGCCCTGC-3'	*hbb* deletion, amplification of Km^r cassette
pUC/M13 fwd (XbaI)	5'-GCTATCTAGAGTAAAACGACGGCCAGTGCC-3'	*hbb* deletion, amplification of Km^r cassette flanked by *BD_RS00255* up- and downstream regions
pUC/M13 rev (XbaI)	5'-GCTATCTAGACAGGAAACAGCTATGACCATG-3'	of Km^r cassette flanked by *BD_RS00255* up- and downstream regions
20 bp-GC50-F	5'-TTAAAGCCCGTTAAAGCCCG-3'	Co-crystallization of HBb and DNA
20 bp-GC50-R	5'-CGGGCTTTAACGGGCTTTAA-3'	Co-crystallization of HBb and DNA
80 bp-DNA-Fwd [Cy5]	5'-[cyanine5]CCGTACTGTCGTCTGCGGCCTTTGATTATCAATTAAAGCGTTCTACGGCGTTTTTGATCGCTCAACGTGCGGAGCTAGAT-3'	MST
80 bp-DNA-Fwd	5'-CCGTACTGTCGTCTGCGGCCTTTGATTATCAATTAAAGCGTTCTACGGCGTTTTTGATCGCTCAACGTGCGGAGCTAGAT-3'	EMSA
80 bp-DNA-Rev	5'-ATCTAGCTCCGCACGTTGAGCGATCAAAAACGCCGTAGAACGCTTTAATTGATAATCAAAGGCCGCAGACGACAGTACGG-3'	EMSA, MST

80 bp-GC50-Fwd [Cy5]	5'-[cyanine5]TTAAAGCCCGTTAAAGCCCGTTAAAGCCCGTTAAAGCCCGTTAAAGCCCGTTAAAGCCCGTTAAAGCCCG-3'	MST
80 bp-GC50-Rev	5'-CGGGCTTTAACGGGCTTTAACGGGCTTTAACGGGCTTTAACGGGCTTTAACGGGCTTTAACGGGCTTTAA-3	MST
240 bp-DNA-Fwd	5'-[Phos]TTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG-3'	Ligase-mediated circularization assay
240 bp-DNA-Rev	5'-[Phos]TGGTTTCTTAGACGTCAGGTGGCACTTTTCGG-3'	Ligase-mediated circularization assay
685 bp-DNA-Fwd [Biotin]	5'-[Biotin]TTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG-3'	TPM
685 bp-DNA-Rev [DIG]	5'-[DIG]CCAAGTAGCGAAGCGAGCAGGACTGGGCGG-3'	TPM
685 bp-DNA-Fwd	5'-TTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG-3'	MNase assay
685 bp-DNA-Rev	5'-CCAAGTAGCGAAGCGAGCAGGACTGGGCGG-3'	MNase assay

Table S2

Binding affinities (K_d) of HBb and HMfB to DNA substrates.

Protein	DNA substrate	K_d ± SD (∝M)
HBb	80 bp DNA	4.84 ± 0.97
	80 bp-GC50 DNA	11.1 ± 1.07
HMfB	80 bp DNA	0.83 ± 0.27
	80 bp-GC50 DNA	3.92 ± 1.16

Table S3

Data collection and refinement statistics of free HBb, HBb-DNA_1 and HBb-DNA_2

Values for the outer shell are given in parentheses.

	HBb	HBb-DNA_1	HBb-DNA_2
Data collection			
Space group	$P2_12_12$	$P2_12_12_1$	C2
Cell dimensions			
a, *b*, *c* (Å)	34.47, 56.35, 26.70	32.32, 34.22, 151.83	108.70, 34.76, 56.43
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 100, 90
Resolution range (Å)	28.17–1.06 (1.13–1.06)	31.63-1.95 (2.07-1.95)	42.49-1.85 (1.96-1.85)
Completeness (%)	96.3 (79.4)	99.9 (99.8)	93.0 (95.2)
Redundancy	10.8 (3.88)	7.89 (8.33)	4.68 (4.70)
$\langle I/\sigma(I)\rangle$	23.80 (1.66)	11.3 (0.86)	9.59 (2.18)
R_{meas}	0.047 (0.921)	0.104 (2.16)	0.110 (0.637)
Refinement			
No. of reflections, working set	20904	11629	15089
No. of reflections, test set	1161	645	843
Final R_{cryst}	0.142	0.221	0.249
Final R_{free}	0.169	0.268	0.267
R.m.s. deviations			
Bonds (Å)	0.019	0.005	0.008
Angles (°)	1.421	1.271	1.297

References

1. Hocher, A., Laursen, S.P., Radford, P., Tyson, J., Lambert, C., Stevens, K.M., Montoya, A., Shliaha, P.V., Picardeau, M., Sockett, R.E. *et al.* (2023) Histones with an unconventional DNA-binding mode in vitro are major chromatin constituents in the bacterium Bdellovibrio bacteriovorus. *Nat Microbiol*, **8**, 2006-2019.
2. Baumgarten, T., Schlegel, S., Wagner, S., Low, M., Eriksson, J., Bonde, I., Herrgard, M.J., Heipieper, H.J., Norholm, M.H., Slotboom, D.J. *et al.* (2017) Isolation and characterization of the E. coli membrane protein production strain Mutant56(DE3). *Sci Rep*, **7**, 45089.
3. Alva, V., Ammelburg, M., Soding, J. and Lupas, A.N. (2007) On the origin of the histone fold. *BMC Struct Biol*, **7**, 17.
4. Bogomolovas, J., Simon, B., Sattler, M. and Stier, G. (2009) Screening of fusion partners for high yield expression and purification of bioactive viscotoxins. *Protein Expr Purif*, **64**, 16-23.
5. Kabsch, W. (2010) Xds. *Acta Crystallogr D Biol Crystallogr*, **66**, 125-132.
6. McCoy, A.J. (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr*, **63**, 32-41.
7. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60**, 2126-2132.
8. Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F. and Vagin, A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr*, **67**, 355-367.
9. Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J Appl Crystallogr*, **30**, 1022-1025.
10. Bailey, K.A., Pereira, S.L., Widom, J. and Reeve, J.N. (2000) Archaeal histone selection of nucleosome positioning sequences and the procaryotic origin of histone-dependent genome evolution. *J Mol Biol*, **303**, 25-34.
11. Mattiroli, F., Bhattacharyya, S., Dyer, P.N., White, A.E., Sandman, K., Burkhart, B.W., Byrne, K.R., Lee, T., Ahn, N.G., Santangelo, T.J. *et al.* (2017) Structure of histone-based chromatin in Archaea. *Science*, **357**, 609-612.
12. Henneman, B., Heinsman, J., Battjes, J. and Dame, R.T. (2018) Quantitation of DNA-Binding Affinity Using Tethered Particle Motion. *Methods Mol Biol*, **1837**, 257-275.
13. Bartlett, A., O'Malley, R.C., Huang, S.C., Galli, M., Nery, J.R., Gallavotti, A. and Ecker, J.R. (2017) Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat Protoc*, **12**, 1659-1672.
14. Ewels, P.A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M.U., Di Tommaso, P. and Nahnsen, S. (2020) The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol*, **38**, 276-278.
15. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol*, **9**, R137.
16. Team, T. (2023) pandas-dev/pandas: Pandas, v2. 1.1. *Zenodo, doi*, **10**, 5281.
17. Cock, P.J., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F. and Wilczynski, B. (2009) Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, **25**, 1422.
18. Open2C, Abdennur, N., Fudenberg, G., Flyamer, I., Galitsyna, A.A., Goloborodko, A., Imakaev, M. and Venev, S.V. (2022) Bioframe: Operations on Genomic Intervals in Pandas Dataframes. *bioRxiv*, 2022.2002.2016.480748.
19. Hunter, J.D. (2007) Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, **9**, 90-95.
20. Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B. and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, **1**, 19-25.
21. Ivani, I., Dans, P.D., Noy, A., Perez, A., Faustino, I., Hospital, A., Walther, J., Andrio, P., Goni, R., Balaceanu, A. *et al.* (2016) Parmbsc1: a refined force field for DNA simulations. *Nature Methods*, **13**, 55-58.
22. Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E. and Simmerling, C. (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput*, **11**, 3696-3713.
23. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E. and Berendsen, H.J. (2005) GROMACS: fast, flexible, and free. *J Comput Chem*, **26**, 1701-1718.

24. Berendsen, H.-J.-C., Grigera, J.-R. and Straatsma, T.P. (1987) The missing term in effective pair potentials. *Journal of Physical Chemistry*, **91**, 6269-6271.
25. Bussi, G., Zykova-Timan, T. and Parrinello, M. (2009) Isothermal-isobaric molecular dynamics using stochastic velocity rescaling. *The Journal of chemical physics*, **130**.
26. Parrinello, M. and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, **52**, 7182-7190.
27. Hess, B., Bekker, H., Berendsen, H.J. and Fraaije, J.G. (1997) LINCS: A linear constraint solver for molecular simulations. *Journal of computational chemistry*, **18**, 1463-1472.
28. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S. and Steinegger, M. (2022) ColabFold: making protein folding accessible to all. *Nature methods*, **19**, 679-682.
29. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S. and Yim, J. (2021) Protein complex prediction with AlphaFold-Multimer. *biorxiv*, 2021.2010. 2004.463034.
30. Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris, J.H. and Ferrin, T.E. (2023) UCSF ChimeraX: Tools for structure building and analysis. *Protein Science*, **32**, e4792.
31. DeLano, W.L. (2002) Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr*, **40**, 82-92.
32. Van Rossum, G. and Drake, F.L. (2009). CreateSpace.
33. Daura, X., Gademann, K., Jaun, B., Seebach, D., Van Gunsteren, W.F. and Mark, A.E. (1999) Peptide folding: when simulation meets experiment. *Angewandte Chemie International Edition*, **38**, 236-240.
34. van Heesch, T., Bolhuis, P.G. and Vreede, J. (2023) Decoding dissociation of sequence-specific protein–DNA complexes with non-equilibrium simulations. *Nucleic Acids Research*, **51**, 12150-12160.
35. Lavery, R., Moakher, M., Maddocks, J.H., Petkeviciute, D. and Zakrzewska, K. (2009) Conformational analysis of nucleic acids revisited: Curves+. *Nucleic acids research*, **37**, 5917-5929.

Appendix II

DNA Wrapping by a Tetrameric Bacterial Histone

Yimin Hu[a], Samuel Schwab[b,#], Kaiyu Qiu[a,#], Yunsen Zhang[c], Kerstin Bär[a], Heidi Reichle[a], Aurora Panzera[d], Andrei N. Lupas[a], Marcus D. Hartmann[a,e], Remus T. Dame[b], Vikram Alva[a,*], Birte Hernandez Alvarez[a,*]

[a] Max Planck Institute for Biology Tübingen, Department of Protein Evolution, Tübingen, Germany

[b] Leiden Institute of Chemistry, Leiden University, Leiden, The Netherlands; Centre for Microbial Cell Biology, Leiden University, Leiden, The Netherlands; Centre for Interdisciplinary Genome Research, Leiden University, Leiden, The Netherlands

[c] Theoretical and Computational Biophysics Group, Beckman Institute for Advanced Science and Technology, Center of Biophysics and Quantitative Biology, University of Illinois Urbana-Champaign, Urbana, United States

[d] Max Planck Institute for Biology Tübingen, BioOptics Facility, Tübingen, Germany

[e] Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany

[#] Joint second authors
*To whom correspondence should be addressed
E-mail: vikram.alva@tuebingen.mpg.de or birte.hernandez@tuebingen.mpg.de

Classification: Biological sciences: Biochemistry; Physical sciences: Biophysics and computational biology

Keywords: Histone; Crystal structure; Nucleosome; Chromatin; Nucleoid

Abstract

Histones are conserved DNA-packaging proteins found across all domains of life. In eukaryotes, canonical histones form octamers that wrap ~147 base pairs of DNA into nucleosomes — the fundamental units of chromatin. In archaea, histones form dimers that further multimerize into extended hypernucleosomes along DNA. Although bacteria were long thought to lack histones, recent studies have uncovered histone homologs in diverse bacterial lineages, many of which possess key DNA-binding features. We previously characterized HBb, a bacterial histone from *Bdellovibrio bacteriovorus*, which binds DNA as a dimer and induces bending. Here, we describe HLp from *Leptospira perolatii*, a representative of a distinct bacterial histone group. Crystallographic and biophysical analyses reveal that HLp forms stable tetramers. Like HBb, HLp binds DNA non-specifically; however, it adopts a different mode of interaction — wrapping ~60 base pairs of DNA around its tetrameric core. This wrapping mode, supported by molecular dynamics simulations and DNA-binding assays, promotes DNA compaction and alters its topology. When expressed heterologously in *Escherichia coli*, HLp reorganizes nucleoid morphology, consistent with a role in chromatin organization. These findings expand the known repertoire of histone-DNA interaction in bacteria and underscore the structural and functional diversity of histone-based genome organization across the tree of life.

Introduction

Histones are DNA-organizing proteins conserved across all domains of life. They share a characteristic histone fold — comprising three α-helices (α1, α2, α3) connected by two loops — that mediates histone-histone and histone-DNA interactions[1, 2]. In eukaryotes, histones assemble into nucleosomes, which serve as the basic units of chromatin. Nucleosomes consist of an octamer, composed of two copies each of histones H2A, H2B, H3, and H4, around which ~147 base pairs (bp) of DNA wraps[1, 2, 3]. These nucleosomes are arranged like beads on a string and are stabilized by linker histones (H1, H5), which bind DNA between adjacent nucleosomes[4, 5]. A hallmark of eukaryotic histones is their unstructured N-terminal tails, which undergo extensive post-translational modifications to regulate chromatin structure and gene expression[6, 7].

Histones were long believed to be unique to eukaryotes, although they play a central role in chromatin packaging and regulation. However, homologs have since been identified in archaea and, more recently, in bacteria — revealing that histone-based DNA organization is an ancient and widespread strategy for genome compaction. Despite this shared ancestry, key differences exist across domains. While eukaryotic histones are highly conserved and structurally uniform, forming nucleosomes with defined stoichiometry, prokaryotic histones exhibit far greater structural and functional diversity. Notably, N-terminal tails are generally absent in prokaryotes; however, histones from Asgard archaea — the closest known prokaryotic relatives of eukaryotes — do possess such tails[8, 9, 10, 11]. This suggests that chromatin regulation via tail modifications may have emerged prior to the evolution of the eukaryotic nucleus[9]. In both archaea and bacteria, genome organization has traditionally been attributed to nucleoid-associated proteins (NAPs), such as HU, H-NS, and IHF, which lack the histone fold but perform analogous roles in DNA compaction and transcriptional regulation[12, 13, 14, 15]. Based on sequence analysis and structural predictions, prokaryotic histones are broadly categorized into two major families: (i) nucleosomal histones, found exclusively in archaea, and (ii) α3 histones, present in both archaea and bacteria[9]. The α3 histones are defined by a shorter α2 helix and a truncated α3 helix, and exhibit extensive variation in their quaternary structures and domain organizations, leading to classification into multiple subfamilies[9].

Archaeal histones are more abundant and better characterized than their bacterial counterparts, particularly the nucleosomal variants. A well-established example is HMfB from *Methanothermus fervidus*, which forms homodimers that tetramerize upon DNA binding, assembling into continuous helices where each dimer binds

approximately 30 bp of DNA[16, 17]. Similarly, HTkA from *Thermococcus kodakarensis* has been shown to influence chromatin compaction and significantly modulate transcription initiation and elongation[18]. Beyond the nucleosomal type, non-nucleosomal archaeal histones exhibit even greater structural diversity. Members of the face-to-face (FtF) histone subfamily — part of the α3 histone group — form homotetramers that are predicted to wrap DNA[9]. This architecture is fundamentally distinct from the spiraling assembly seen in nucleosomal histones, as demonstrated by the crystal structure of HTkC from *T. kodakarensis*[9]. Another example, MJ1647 from *Methanocaldococcus jannaschii*, belongs to a histone subgroup found exclusively in *Methanococcales*. Unlike canonical histones, MJ1647 tetramerizes through its C-terminal helices rather than the histone fold and binds DNA in a bridging rather than wrapping mode[19].

Bacterial histones are rare, present in fewer than 2% of sequenced bacterial genomes, in stark contrast to the NAP HU, which is found in over 90% of genomes[20]. Their limited distribution, relatively recent discovery, and frequent occurrence in uncultured or metagenomically characterized bacteria have left most bacterial histones uncharacterized at the functional level[21]. Among those identified, two main families dominate, α3 histones and DUF1931 pseudodimeric histones, both of which also occur in archaea[9]. The DUF1931 proteins, exemplified by structures from *Thermus thermophilus* (PDB: 1WWI) and *Aquifex aeolicus* (PDB: 1R4V), form pseudodimers but lack the conserved residues required for DNA binding and nucleosome-like assembly[9, 22, 23]. In contrast, α3 histones possess DNA-binding residues and are subdivided into five distinct subfamilies: (i) FtF histones, (ii) bacterial dimer histones, (iii) ZZ histones, (iv) phage histones, and (v) Rab GTPase histones[9]. The recent characterization of HBb from *Bdellovibrio bacteriovorus*, a member of the bacterial dimer subfamily, provided the first experimental evidence of a functional bacterial histone. HBb is essential for viability and binds genomic DNA non-specifically as a dimer[20, 24]. Although initially thought to bind without affecting DNA topology, structural and biophysical analyses later revealed that HBb induces local DNA bending, employing an interaction mode reminiscent of eukaryotic histones[24].

To broaden our understanding of bacterial histone diversity, we investigated HLp from *Leptospira perolatii*, a member of the FtF subfamily of α3 histones. Crystallographic and biophysical analyses show that HLp forms stable tetramers. Structures of HLp in both free and DNA-bound forms reveal extensive protein-DNA contacts and a tetrameric architecture resembling that of archaeal HTkC[9]. Remarkably, HLp wraps ~60 bp of DNA around its core, representing the first example of DNA wrapping by a bacterial histone. This mode of interaction, corroborated by molecular

dynamics simulations and DNA-binding assays, drives DNA compaction and topological change, extending the known mechanisms of histone-DNA association in bacteria.

Materials and methods

Bioinformatic analyses

For cluster analysis, we used the histone sequences HMfB from *M. fervidus* (UniProtKB ID: P19267), HBb from *B. bacteriovorus* (Q6MRM1), the *A. aeolicus* pseudodimeric histone (O66665), the FtF histone from *L. interrogans* (Q8F3E8), and HTkC from *T. kodakarensis* (Q5JDW7) as queries in BLAST searches to retrieve archaeal and bacterial histone homologs. Searches were performed via the NCBI BLAST webserver with the 'Max target sequences' parameter set to 5000[25, 26]. Sequences shorter than 50 residues, longer than 200, or annotated as 'partial' or 'fragment' were excluded to eliminate truncated entries and multidomain proteins. The full-length hits were pooled and filtered using MMseqs2 to reduce redundancy at 80% pairwise sequence identity over 80% length coverage[27]. The resulting non-redundant dataset of 3341 sequences was clustered using CLANS based on all-against-all pairwise sequence similarities computed via BLAST[28, 29]. Clustering was performed to equilibrium using an E-value cut-off of 1e-6. Genome neighborhood analysis was carried out using the EFI Genome Neighborhood Tool, and structural predictions were generated using the AlphaFold2 (AF2)[30, 31].

Bacterial strains and cultivation

Cloning and plasmid amplification were conducted using *Escherichia coli* Top10. *E. coli* BL21(DE3) and *E. coli* Mutant56(DE3) were used as hosts for protein expression. Bacteria were grown in LB media supplemented with appropriate antibiotics.

Cloning, plasmids, and synthetic DNA

The nucleotide sequences of the HLp-encoding gene CH373_10155 [GenBank: NPDY01000005.1, 141608 – 141799 (+)] from *L. perolatii strain FH1-B-C1* and the *hmfB* gene (GenBank: M34778.1) were codon-optimized and synthesized (BioCat GmbH) for protein expression in *E. coli* (Fig. S1A).

HLp was expressed fused to an N-terminal (histidine)$_6$ tag using the vector pET-30a(+), and HMfB was expressed using pET-28a(+). Ubiquitin from *Caldiarchaeum subterraneum* (*Cs*Ub), used as a control for light microscopy imaging, was expressed from pET-30b(+)[32].

Synthetic oligonucleotides (Merck) used in this study are listed in Table S1. For DNA binding analyses, complementary single-stranded (ss) oligonucleotides were mixed, heated to 95 ºC for 5 min, and slowly cooled to room temperature to generate double-stranded (ds) DNA fragments. The 600-bp-GC40 and 240-bp-GC40 DNA fragments used in DNA binding assays were amplified by PCR using the vector pETHis1a as the template[33].

Protein expression and purification

HMfB and HLp were expressed in *E. coli* Mutant56(DE3) and *E. coli* BL21(DE3), respectively. Bacterial cultures were maintained in LB broth supplemented with kanamycin at 37 ºC. At an optical density of OD_{600} = 0.6, isopropyl-β-d-thiogalactoside (IPTG) was added at a final concentration of 1 mM to induce protein expression. Cells were agitated at 37 ºC for an additional 4 h. Cells were pelleted and resuspended in lysis buffer containing 20 mM Tris, pH 8.0, 300 mM NaCl, and 10 mM imidazole, supplemented with protease inhibitor mix (cOmplete™, EDTA-free Protease Inhibitor Cocktail, Roche), 0.1 mM PMSF, 3 mM $MgCl_2$ and DNase I (AppliChem GmbH). Cells were lysed using a French press, and cell debris and insoluble material were pelleted by centrifugation at 95,000 x g for 45 min. The supernatant was filtered through a 0.45 µm filter and applied to a 5 mL HisTrap column (Cytiva). Bound proteins were eluted using a linear imidazole gradient, ranging from 10-500 mM, in the aforementioned buffer. For HLp, the $(histidine)_6$ tag was cleaved with *Tobacco Etch Virus* (TEV) protease. A second purification step using a HisTrap column was performed using the aforementioned buffers containing 150 mM NaCl to separate the protein from the cleaved tag. Purified proteins were dialyzed against 20 mM Tris buffer at the indicated pH with 150 mM NaCl and concentrated using an Amicon Ultra Centrifugal Filter Device (3 kDa MWCO, Millipore). Protein purity was confirmed by SDS-PAGE (mPAGE 4-12% Bis-Tris Precast Gel, Millipore), and protein concentration was determined spectrophotometrically or using a BCA protein assay (ThermoFisher Scientific). For the tethered particle motion (TPM) assay, purified HLp was dialyzed against 50 mM Tris, pH 7.0, 75 mM KCl, and 10% glycerol.

Circular dichroism (CD) spectroscopy

HLp was dialyzed against 10 mM phosphate buffer, pH 8.0, 75 mM KF, and diluted to a final concentration of 15 µM. CD spectra were recorded on a Jasco J-810 spectrometer (JASCO) in the wavelength range of 190-250 nm using a cuvette with a 1 mm path length and a 100 nm/min reading speed. A total of ten single spectra were recorded and averaged. To analyze the thermal stability of HLp, the ellipticity (θ) was

measured at 222 nm over a temperature gradient of 10-100 ºC, using a ramp of 1 ºC/min, a data pitch of 0.5, and a response time of 1 s. Data analysis, blank subtraction, and curve smoothing were performed using the Spectra Manager software suite (JASCO). CD spectra and thermal melting curves were plotted using GraphPad Prism 9 (GraphPad Software, Inc.).

Size exclusion chromatography coupled with multi-angle light scattering (SEC-MALS)

The HLp-DNA complex was analyzed by SEC-MALS following incubation of HLp with the 30-bp-GC40 DNA fragment at a protein-to-DNA molar ratio of 2:1. All samples were incubated at 37 ºC for 10 min, and aggregates were removed by centrifugation. Samples were applied onto a Superdex 75 Increase 10/300 GL column (Cytiva), pre-equilibrated with SEC buffer (25 mM Tris, 50 mM NaCl, 50 mM KCl, pH 7.5). The run was performed at a 0.5 mL/min flow rate on a 1260 Infinity II HPLC system (Agilent) coupled to a miniDAWN TREOS and Optilab T-rEX refractive index detector (Wyatt Technology). HLp and the 30 bp-GC40 DNA fragment were detected at 215 nm and 260 nm, respectively. Measurements were performed in triplicate, and molecular mass distributions were calculated using the ASTRA v.7.3.0.18 software suite (Wyatt Technology).

Crystallization, data collection, and structure determination

Crystallization trials were set up by mixing 300 nL of protein with 300 nL of reservoir solution in 96-well sitting-drop vapor-diffusion plates using commercially available screens with reservoir volumes of 100 µL. Without DNA, HLp was prepared at 11 mg/mL in 20 mM Tris, pH 8.0, and 150 mM NaCl. The best diffracting crystals were obtained with a reservoir solution containing 0.1 M citric acid, pH 5.0, and 3.15 M ammonium sulfate. For co-crystallization of HLp and DNA, 500 µM HLp was mixed with 500 µM 30-bp-GC40 dsDNA (Table S1) and incubated at 37 °C for 10 min, after which aggregates were removed by centrifugation. The conditions under which the crystals used for structure determination grew were 0.1 M HEPES, pH 7.5 and 25% PEG 6000 for HLp-DNA_1, and the Morpheus condition G12 [0.1 M carboxylic acids, 0.1 M Morpheus Buffer System 3, pH 8.5 and 50% (V/V) Morpheus Precipitant Mix 4] for HLp-DNA_2. For cryo-protection, the crystals were transferred to droplets of their reservoir solution spiked with 30% glycerol (free HLp) or 15% PEG 400 (HLp-DNA_1), loop-mounted, and flash frozen in liquid nitrogen. Data were collected at beamline X10SA of the Swiss Light Source (Villigen, Switzerland) at 100 K, using an EIGER X 16M hybrid pixel detector (Dectris, Ltd.). Data were reduced, processed, and

scaled using XDS[34]. Due to pronounced anisotropy, diffraction data for HLp-DNA_2 were submitted to the STARANISO server for ellipsoidal truncation and anisotropic scaling, following the unmerged data protocol[35].

The structure of free HLp was solved by molecular replacement (MR) using MOLREP and an AlphaFold2 prediction as a search model, locating one HLp dimer in the asymmetric unit (ASU)[31, 36, 37]. The DNA-bound structures were solved by MR using MOLREP and the refined free HLp coordinates as a search model, locating an HLp dimer and 16 bp of dsDNA in the ASU for HLp-DNA_1, and a monomer and 15 nucleotides of ssDNA in the ASU for HLp-DNA_2. The three structures were modeled, refined, and finalized in cycles of manual modeling in Coot and refinement with REFMAC5[38, 39]. Data processing and refinement statistics are given in Table S3. The coordinates and structure factors have been deposited in the PDB under the accession numbers 9QT0 (free HLp), 9QT1 (HLp-DNA_1), and 9QT2 (HLp-DNA_2). All structures were visualized using PyMOL (The PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.).

Molecular dynamics (MD) simulation

We performed all-atom molecular dynamics (MD) simulations in GROMACS (version 2023.2) to study two possible topologies of the HLp-DNA complexes formed by bridging or wrapping[40]. The DNAs in the HLp-DNA_1 and HLp-DNA_2 structures were extended through crystal symmetry. The obtained structures were superimposed based on the central HLp, the excess DNA fragments were trimmed up to the overlap point, and then the remaining DNA fragments were connected to obtain the two initial structures of HLp-DNA complexes (Fig. 5 and Fig. S6). Two initial models were constructed for each binding mode (wrapping and bridging) based on the HLp-DNA_1 and HLp-DNA_2 structures. One representative model for each binding mode was pursued further. The constructed models were then preprocessed using PDBFixer (version 1.10, https://github.com/openmm/pdbfixer) to add missing hydrogen atoms and adjust chain configurations as necessary. The CHARMM all-atom force field (CHARMM36m) was applied to parameterize and describe both intra- and intermolecular interactions between the protein and DNA[41]. Ion parameters and the TIP3P water model were selected following the CHARMM36m recommendations[42]. Each system was solvated in a sufficiently large cubic water box to ensure a minimum water layer of 1.5 nm surrounding the protein-DNA complex in all directions. Sodium and chloride ions were added to neutralize the system. Periodic boundary conditions (PBC) were imposed to eliminate edge effects, enabling the simulation to be conducted under an infinite solvent environment.

Each system underwent a multistage process of energy minimization and equilibration before production simulation. First, a harmonic positional restraint of 1000 kJ/(mol·nm²) was applied to the backbones of both proteins and DNAs to relax the distributions of water molecules and ions. The system was then subjected to 5000 steps of conjugate gradient and steepest descent minimization until the maximum force on any atom dropped below a predefined threshold of 1000 kJ/(mol·nm²). Following energy minimization, each complex system was equilibrated in multiple stages using the canonical (NVT) and isothermal-isobaric (NPT) ensembles. During the NVT phase, the system was gradually heated to 303.15 K with a V-rescale thermostat[43]. In the NPT phase, density equilibration was achieved at a constant pressure of 1 atm using a Parrinello-Rahman barostat[44]. Once equilibration was complete, a 1 μs production simulation was conducted on the processed protein-DNA complex, with the integration carried out using the leap-frog algorithm at a timestep of 2 fs. Meanwhile, a 1.2 nm cutoff was set for van der Waals (VDW) and electrostatic potentials to calculate the short-range interaction, and the Particle Mesh Ewald (PME) method was employed to compute the long-range interaction[45]. To enhance computational efficiency, the LINCS algorithm was employed to constrain hydrogen bond lengths in water molecules and the backbones of both proteins and DNAs[46]. Two independent 1 μs simulations were performed for each system to ensure the robustness of the results (SI movie 1 and 2). Both simulations exhibited highly similar behaviors in either the wrapping or the bridging model; therefore, the analysis was conducted on one representative trajectory for each system. To assess global structural stability, we calculated the root mean square deviation (RMSD) of the DNA backbone relative to the initial structure. Binding fluctuations were evaluated by determining the difference between the contact number of each frame and the average contact number over all captured frames. Protein-DNA contacts were defined by the presence of any protein-heavy atom within 0.4 nm of a DNA-heavy atom. Hydrogen bonds were identified with a distance cutoff of 0.3 nm for the donor-to-acceptor (D-A) distance and an angular cutoff of 150° for the donor-hydrogen-acceptor (D-H-A) angle. To further study the binding dynamics of the local interaction sites, we calculated the binding contact number for each residue throughout the simulation and manually selected residues to define four binding sites (Fig. 5C and Table S4).

To further characterize and understand the binding dynamics of the proposed wrapping model, an unwrapping simulation was performed in the framework of a two-step steered molecular dynamics (SMD)[47], with the constructed wrapping model (Fig. S6) as the starting structure. The first SMD simulation targeted the dissociation of the B-site 2, followed by a second SMD simulation that used the last snapshot from the B-

site dissociation as the starting structure to investigate the unbinding of the A-site. In the SMD setup, a reaction coordinate was defined by connecting the center of mass of the relevant DNA regions (residues 1 to 12) to that of the protein, ensuring that the applied force was directed along an axis expected to promote the unwrapping process. A pulling velocity of 0.002 nm/ps and a harmonic spring constant of 500 kJ/(mol·nm²) were selected to balance the computational feasibility with the need to capture quasi-equilibrium behavior. During the production simulation of SMD, force and extension data were recorded at regular intervals to allow integration of the force over distance and estimation of the work required for unwrapping. Simultaneously, the temporal evolution of protein-DNA contacts was monitored using a 0.4 nm cutoff to pinpoint specific events where the DNA disengaged from the binding interface. Once both B-site 2 and A-site 2 were dissociated, we integrated the force-distance curves obtained from the SMD simulation to generate dissociation energy profiles. All simulations were analyzed using VMD (version 1.9.4) and the MDAnalysis package (version 2.8.0), with Matplotlib and Seaborn employed for visualization[48, 49].

DNA binding assays

All DNA binding assays were performed as described previously, with minor modifications[24].

Electrophoretic mobility shift assay (EMSA)

The 80-bp dsDNA fragment, 30-bp dsDNA fragments of various GC content (30%, 40%, 50%, and 60%), and the GeneRuler 1 kb Ladder (ThermoFisher Scientific) were used for *in vitro* DNA binding tests. For comparative DNA binding analysis, the proteins were mixed with the annealed dsDNA fragments at the indicated molar ratios in a binding buffer containing 25 mM Tris, pH 8.0, 50 mM NaCl, and 50 mM KCl. Following incubation at 37 ºC for 10 min, the samples were separated on a 6% DNA retardation gel (ThermoFisher Scientific). To analyze the binding of HLp and HMfB to the GeneRuler 1 kb Ladder, the proteins and the DNA were mixed at the indicated mass ratios in the binding buffer and incubated at 37 ºC for 10 min. Samples were separated on a 1% agarose gel, and the DNA was visualized using SYBR Gold nucleic acid gel stain (ThermoFisher Scientific) and imaged using the Fusion SL imaging system (Vilber).

Micrococcal nuclease (MNase) digestion assay

The 600-bp-GC40 DNA fragment was amplified using primers pET-600-bp-GC40-F and pET-600-bp-GC40-R with pETHis1a as a template, followed by purification using the QIAquick PCR Purification Kit (QIAGEN). For HLp-DNA complex formation, 720 ng

of HLp was mixed with 900 ng of 600-bp-GC40 DNA in MNase buffer (10 mM Tris, pH 8.0, 75 mM NaCl) and incubated at 37 ºC for 10 min. Indicated amounts of MNase (New England Biolabs) were added, and digestion was performed in a reaction volume of 80 µL at 37 ºC for 10 min. As a positive control, HMfB (540 ng) was preincubated with 600-bp-GC40 DNA (900 ng) and treated with the indicated amounts of MNase at 37 ºC for 15 min. EDTA and SDS were added to final concentrations of 95 mM and 0.5%, respectively, to stop the reaction. DNA fragments protected by bound proteins were purified by phenol/chloroform extraction followed by ethanol precipitation. The purified DNA fraction was dissolved in 5 µL of water and separated on a 10% Novex TBE gel (ThermoFisher Scientific). The gel was stained with SYBR Gold nucleic acid gel stain (ThermoFisher Scientific) and imaged using the Gel Doc XR+ imaging system (Bio-Rad Laboratories). All experiments were performed at least in triplicate, and the length of the selected DNA bands was calculated using Image Lab 6.1 (Bio-Rad Laboratories). For better visualization, the image background was subtracted using the rolling-ball algorithm with a radius of 50 pixels in FIJI (version 2.3.0, https://fiji.sc/)[50].

Tethered particle motion (TPM) assay

TPM experiments were performed as previously described in 50 mM Tris, pH 7.0, and 75 mM KCl [51]. A standard deviation cutoff of 8% and an anisotropic ratio cutoff of 1.3 were used to select single-tethered beads[51]. Measurements at each HLp concentration were done in triplicate. Means and standard deviations of the individual measurement series for each HLp concentration were calculated by maximum likelihood estimation, assuming a normal distribution. Outliers with a robust Z-score >3 or <-3 were not considered for fitting. The “line to guide the eye” was generated by fitting the means to a logistic function. A custom Python script was used to fit and plot the TPM data. For plotting, the means of the three individual measurements were averaged for each measured concentration, and the standard deviations were error-propagated ($Std(\overline{X}) = \sqrt[2]{\frac{\sum_{i=1}^{n} Var(X_i)}{n^2}}$).

DNA topology assay

Plasmid pUC19 was purified from *E. coli* Top10 using NucleoSpin Plasmid QuickPure Kit (Macherey-Nagel). For topological relaxation, the plasmid pUC19 was treated with the nicking endonuclease Nb.BsrDI (New England BioLabs), followed by ligation with T4 DNA ligase (ThermoFisher Scientific). Plasmid relaxation was verified by agarose gel electrophoresis. 200 ng of relaxed pUC19 was mixed with HLp or HMfB at indicated protein-to-DNA mass ratios in assay buffer containing 10 mM Tris, pH 7.0, and 75 mM

NaCl) and incubated at room temperature for 30 min. Then, 1 U of Topoisomerase I (ThermoFisher Scientific) was added, and the reaction mixture was incubated at 37 ºC for 30 min in a final volume of 50 µL. DNA was purified by phenol/chloroform extraction followed by ethanol precipitation. Purified DNA samples were separated on a 0.8% TAE agarose gel. Plasmids were visualized by staining with SYBR Gold nucleic acid gel stain (ThermoFisher Scientific) and imaged using the Gel Doc XR+ imaging system (Bio-Rad Laboratories).

Ligase-mediated circularization assay

The 240-bp-GC40 DNA fragment was amplified from plasmid pETHis1a using primers pET-240-bp-GC40-Fp and pET-240-bp-GC40-Rp (Table S1) and purified using QIAquick PCR Purification Kit (QIAGEN)[33]. 400 ng of DNA was incubated with HMfB and HLp at the indicated protein-to-DNA mass ratios in LMC buffer (10 mM Tris, pH 7.0, 75 mM KCl, and 5% glycerol) in a total reaction volume of 20 µL at room temperature for 30 min. T4 DNA ligase buffer and T4 ligase at a final concentration of 1 U/µL were added in a final reaction volume of 100 µL. A sample not treated with T4 DNA ligase was used as a control. The sample was incubated for 24 h at room temperature, and the DNA was purified by phenol/chloroform extraction and ethanol precipitation. Half of each sample was treated with 1 U of T5 exonuclease (New England BioLabs) in the appropriate buffer at 37 ºC for 1 h. DNA samples were separated on a 2% TAE agarose gel, stained with SYBR Gold nucleic acid gel stain (ThermoFisher Scientific), and imaged using the Gel Doc XR+ imaging system (Bio-Rad Laboratories).

Microscale thermophoresis (MST)

To determine the DNA binding affinity of HLp to the Cy5 labeled 80-bp and 80-bp-GC40 dsDNAs, a dilution series of HLp was prepared in 20 mM Tris, pH 8.0, 150 mM NaCl, and titrated against the dsDNA fragments at a concentration of 10 nM. The HLp-DNA mixtures were incubated at 37 ºC for 10 min and loaded into Monolith NT premium capillaries (MO-K025, NanoTemper Technologies) after centrifugation to remove precipitates. HMfB was used as a positive control and titrated against both DNA fragments at a concentration of 20 nM. HMfB-DNA mixtures were incubated at room temperature for 5 min, centrifuged, and loaded into Monolith NT capillaries (MO-K022, NanoTemper Technologies). All measurements were performed in triplicate at 25 ºC using the Monolith NT 115 instrument with a Nano RED detector and MST power set to medium. MST data were analyzed by fitting them to a K_d model using MO Control (NanoTemper Technologies).

Light microscopy imaging and data processing

Vectors pET-30a(+) and pET-30a(+) encoding *C. subterraneum* ubiquitin (*Cs*Ub) were transformed in *E. coli* BL21(DE3)[32]. The transformed strains, as well as non-transformed *E. coli* BL21(DE3) and *E. coli* Mutant56(DE3), served as negative controls.

Sample preparation

Overnight cultures of each strain were diluted 30-fold (V/V) in LB medium supplemented with the appropriate antibiotic and incubated at 37 ºC, 170 rpm. Except for the negative controls, protein expression was induced with IPTG at a final concentration of 1 mM at an OD_{600} = 0.4-0.6. Following cultivation for another 4 h, 500 µL bacterial culture was pelleted. The cells were washed with PBS and fixed using 4% (V/V) formaldehyde. Fixation was stopped with glycine at a final concentration of 150 mM. Following two wash steps with PBS, the cells were incubated with 1 µg/mL DAPI solution (Invitrogen), washed, and resuspended in PBS. The treated cells were dropped onto a 2% agarose pad, which was placed upside down in a 35 mm imaging dish (High Glass Bottom, Ibidi) for imaging. Triplicates of each sample were analyzed.

Imaging and image processing

Light microscopy imaging was conducted using a Zeiss LSM 780 inverted confocal microscope (Carl Zeiss AG, Oberkochen, Germany) equipped with a 63 x oil/1.4NA oil-immersion objective. The 405 nm diode laser line was used for DAPI excitation and brightfield imaging. Airyscan datasets were processed with Airyscan software to generate 32-bit images using pixel reassignment and Wiener filter-based deconvolution[52]. Brightfield images were used to segment individual bacterial cells via a custom-developed Macro in FIJI, which included background subtraction and auto-thresholding[50]. Segmentation results were manually curated to eliminate artefacts. For fluorescence analysis, a maximum intensity projection was first applied to the Airyscan z-stacks. The segmented bacterial outlines were overlaid onto the fluorescence channel, followed by a second segmentation step on the fluorescence channel to identify nucleoid regions using auto-thresholding.

Results

Bioinformatic analysis

To explore the diversity and evolutionary relationships of bacterial FtF histones and identify candidates for experimental characterization, we performed a comparative sequence analysis based on a curated dataset comprising α3 histones from the FtF

and bacterial dimer subfamilies, archaeal nucleosomal histones, and pseudodimeric DUF1931-family histones. Only proteins consisting solely of the histone fold, without additional domains, were included. All-against-all pairwise sequence similarities were analyzed using CLANS, enabling visualization of the sequence space occupied by FtF histones and assessment of their relationships to other prokaryotic histone families. The resulting cluster map reveals a clear separation among major prokaryotic histone families (Fig. 1A). DUF1931 pseudodimeric histones form a distinct, compact cluster, well separated from canonical histones. Archaeal nucleosomal histones group into a large, coherent cluster, with a small subgroup of closely related sequences extending from its core, including the *Haloferax volcanii* pseudodimer HstA (HVO_0520, UniProt D4GS56). Archaeal FtF histones, bacterial FtF histones, and bacterial dimer histones form separate, well-defined clusters. The adjacency of the bacterial FtF cluster to the archaeal FtF cluster suggests potential shared structural features, such as tetramerization.

While FtF histones are widespread across archaeal phyla, in bacteria they are predominantly found in specific lineages, including Spirochaetota, Planctomycetota, Bdellovibrionota, and Myxococcota. The majority of these bacterial representatives are derived from uncultured organisms or metagenomic assemblies, limiting opportunities for functional characterization. Among the few exceptions, we identified two species with complete genomes from culturable organisms within the bacterial FtF cluster: *L. interrogans*, a pathogenic species and the primary causative agent of leptospirosis, and *L. perolatii*, a species occasionally associated with disease in humans and animals. The FtF histone from *L. interrogans* is known to be highly abundant and essential for viability, underscoring the functional relevance of this histone group[20]. Notably, neither *L. interrogans* nor *L. perolatii* encodes classical bacterial NAPs such as HU or Dps, suggesting that FtF histones may serve as the principal DNA organizers in these lineages. Consistent with this, we found only one histone homolog in each of the two *Leptospira* species. Interestingly, in many *Leptospira* species, the *hlp* gene is located adjacent to genes encoding the chromosome segregation protein SMC (WP_100713470.1), methionine aminopeptidase (MAP; WP_100713471.1), and an uncharacterized DUF350 domain-containing transmembrane protein (WP_100713472.1). This conserved genomic neighborhood hints at potential functional links between histone-mediated DNA organization and core aspects of cellular physiology. We focused our efforts on the histone homolog HLp from *L. perolatii*, thereby enabling the characterization of a previously unexplored bacterial FtF histone and advancing our understanding of histone-based chromatin organization in Leptospira, beyond the limited insights available from *L. interrogans*[53]. Moreover, AF

structure predictions indicated that HLp forms a homotetramer, resembling that of the archaeal FtF histone HTkC from *T. kodakarensis*, further supporting its relevance as a representative of this bacterial subgroup[9].

Structure of HLp from *L. perolatii*

Based on the above bioinformatic analysis, HLp from *L. perolatii* was selected as a representative of the bacterial FtF histone group for structural and functional characterization. In sequence databases, two variants of the HLp protein are documented: a shorter form comprising 63 amino acid residues and a longer form with 77 residues, differing in their N-terminal regions (Fig. S1A). To determine the correct variant, we examined the genomic context of the *hlp* gene, particularly analyzing the position of the Shine-Dalgarno (SD) sequence — a ribosomal binding site crucial for initiating bacterial translation[54]. Our analysis revealed that the SD sequence aligns appropriately with the start codon of the shorter, 63-residue variant, suggesting that this form is the authentic translation product (Fig. S1A). HLp was recombinantly overexpressed in *E. coli* and purified to homogeneity, as confirmed by SDS-PAGE, which showed a single band at ~10 kDa, consistent with its theoretical molecular weight of 7.1 kDa (Fig. S2A). SEC-MALS revealed a predominant species of 26 ± 0.3 kDa, indicating that HLp forms tetramers in solution (Fig. S2B). CD spectroscopy showed that HLp is a predominantly α-helical protein and unfolds upon heating, with a melting temperature of 57.1 °C (Fig. S2C and D).

HLp crystallized under multiple conditions within a few days. The best data set was processed to a resolution of 1.30 Å, and the structure was solved by molecular replacement using an AF model as the search template (Fig. 2)[31, 36]. The entire chain, except for the first seven residues, was well resolved in the electron density map. The asymmetric unit contained an HLp dimer, which assembles into a homotetramer by crystallographic symmetry. This tetramer closely matches the AF model, with an RMSD value of 0.907 Å (Fig. S3A).

The HLp monomer exhibits the characteristic histone fold, comprising three α-helices (α1, α2, and α3) connected by two loops (l1 and l2) (Fig. 2A). The central α2 helix is one turn shorter than in the archaeal histone HMfB (Fig. S4), while the C-terminal α3 helix is truncated and forms a single helical turn composed of the final four residues — similar to what is observed in the bacterial dimeric histone HBb from *B. bacteriovorus*[24]. Within the dimer, the monomers are arranged in a head-to-tail orientation, with the interface formed primarily by the antiparallel crossing of the α2 helices at an angle of ~40°. It is stabilized by hydrophobic interactions involving residues V16 and I20 (α1 helix), a leucine-rich stretch in the α2 helix (L35, L39, L42,

and L47), and F63 (α3 helix), along with polar contacts between loop regions. Notably, M26 and S28 in l1 of one monomer interact with T56 and V58 in l2 of the opposing monomer (Fig. 2B).

Tetramerization is mediated by contacts between the C-terminal end of α2, loop l2, and α3 of opposing dimers. This includes a branched hydrogen-bonding network formed between monomers from opposing dimers: specifically, between R49, S52, and N53 of one, and F63, T61, and R55 of the other monomer, respectively (Fig. 2C). Notably, R49, N53, and R55 are conserved across predicted tetrameric bacterial FtF histones, suggesting a common mechanism of tetramer assembly (Fig. 1B). This mode of oligomerization is fundamentally distinct from that of archaeal nucleosomal histones such as HMfB, which assemble into extended spirals (Fig. S4B)[16].

HLp also harbors a conserved RD-clamp — an intramolecular salt bridge between R55 and D62 — that stabilizes loop l2, a feature shared with both HMfB and HBb (Fig. 1B, Fig. 2A, and Fig. S4)[24, 55]. Compared to canonical histones, HLp shows partial conservation of residues implicated in DNA binding (Fig. 1B). Nevertheless, electrostatic surface potential analysis of the HLp tetramer reveals a continuous band of positive charge encircling the structure, consistent with a DNA-wrapping mode of interaction. (Fig. S3B)[56].

HLp binds non-specifically to DNA *in vitro*

We first assessed the DNA-binding ability of HLp using EMSAs. The 80-bp DNA fragment — previously shown to bind HMfB with high affinity — exhibited reduced mobility in polyacrylamide gels upon incubation with HLp (Fig. 3A), indicating the formation of HLp-DNA complexes[16, 57]. Complex formation was concentration-dependent, and the appearance of smeared bands suggested that the HLp-DNA complexes were less stable or more heterogeneous than those formed with HMfB.

To further investigate sequence preferences, we repeated the EMSA with shorter 30-bp DNA fragments varying in GC content. These assays yielded sharper band patterns, and HLp showed the highest affinity for DNA with 40% GC content (Fig. 3B and Fig. S5). Binding was confirmed by microscale thermophoresis (MST), which demonstrated that HLp interacts with DNA non-specifically, with the strongest binding to the DNA of 40% GC content and a dissociation constant in the low micromolar range (Fig. 3C and Table S2).

We next used SEC-MALS to determine the stoichiometry of the HLp-DNA complex. HLp alone and the 30-bp-GC40 DNA fragment eluted as single peaks, with molecular weights matching their theoretical values (Fig. 3D). Upon incubation of HLp with 30-bp-GC40 DNA, a new peak appeared, corresponding to a molecular weight of 69.2 ±

5.7 kDa, consistent with a complex consisting of an HLp tetramer bound to two 30-bp DNA duplexes.

Crystal structures of HLp bound to DNA

To gain structural insight into the interaction between HLp and DNA, we co-crystallized HLp and the 30-bp-GC40 fragment. Diffraction data were collected for two different crystal forms, HLp-DNA_1 and HLp-DNA_2, processed to resolutions of 2.10 Å and 1.90 Å, respectively (Table S3). Both structures were solved by molecular replacement using the DNA-free HLp structure as the search model. In both HLp–DNA crystal forms, the DNA fragments appear to form continuous double helices winding throughout the crystal lattice, each revealing a distinct interface between HLp and the DNA (Fig. 4A and B). In both structures, the seemingly continuous DNA bases within central segments (16-bp dsDNA for HLp-DNA_1 and 15-nt ssDNA for HLp-DNA_2) were well resolved in the electron density maps, whereas the first six N-terminal residues of HLp were not visible.

In HLp-DNA_1, the ASU contains an HLp dimer, which assembles into tetramers by crystallographic symmetry, and a 16-bp dsDNA segment of the seemingly infinite DNA helices winding throughout the crystal (Fig. 4A). As observed in other histone-DNA complexes, binding is mediated primarily through interactions between basic or polar side chains of HLp and the phosphate backbone of the DNA. On HLp, the interface spans the two monomers in the ASU, across the dimer. In each monomer, this interface involves residues A13, S14, and K15, which correspond to the "paired end of helices" motif in eukaryotic histones, as well as K21 in helix α1, and T27 and G29 in loop l1 (Fig. 4C and E)[58].

In HLp-DNA_2, the ASU comprises a single HLp monomer, which forms tetramers via crystallographic symmetry, and a 15-nt ssDNA segment of the apparently endless DNA (Fig. 4B). Although the overall crystal packing is similar to HLp-DNA_1, it exhibits a complementary DNA binding mode, with the dsDNA engaged along the HLp tetramerization interface. Therein, in both HLp dimers, the monomers bind to DNA via two sets of interactions. The first corresponds to the “β-bridge” motif in eukaryotic histones and involves T27 and G29 in loop l1, as well as T56 and T57 in loop l2, while the second set involves K17 and K21 from helix α1 (Fig. 4D and F)[58].

The two DNA-bound crystal structures thereby reveal complementary, but overlapping DNA-binding interfaces distributed along the orbicular surface of the HLp tetramer, corresponding to the canonical interaction motifs of archaeal and eukaryotic histones[16, 58]. Taken together, these findings suggest that HLp is able to bind DNA through a wrapping mechanism.

Molecular dynamics (MD) simulation

The two crystal structures suggested a bridging mode of DNA binding by HLp. However, given the distinct yet overlapping interfaces observed, we hypothesized that HLp might also support a wrapping mode, in which the tetramer could engage a DNA segment across the entire DNA-binding interface. To investigate whether HLp indeed exhibits both bridging and wrapping modes, and to explore their relative stability and potential interconversion, we performed all-atom MD simulations. Using the HLp-DNA_1 and HLp-DNA_2 structures as templates, we generated two starting models: a wrapping model, consisting of an HLp tetramer wrapped by a 67-bp dsDNA fragment, and a bridging model, in which an HLp tetramer engages two separate 32-bp dsDNA fragments (Fig. 5A and Fig. S6). For each model, we performed two independent 1 μs all-atom simulations, which yielded highly consistent trajectories (SI movie 1 and 2). For subsequent analyses, one representative trajectory per model was examined in detail.

The wrapping model showed high stability, with a DNA backbone RMSD maintained at approximately 6 Å throughout the simulation (Fig. 5B). The final frame of the simulation shows the HLp tetramer fully wrapped by ~60 bp of dsDNA (Fig. 5A). In contrast, the bridging model exhibited substantial conformational rearrangements, characterized by a progressive increase in protein-DNA contacts and a spontaneous shift toward the wrapping configuration (Fig. 5A and B). Consistently, the DNA potential energy of the wrapping model is lower than the bridging model (Fig. S7A), suggesting that the wrapping interaction provides greater thermodynamic stability through an expanded protein-DNA interface.

To quantify HLp-DNA contacts, we calculated atom-atom interactions between protein and DNA heavy atoms for both the bridging and wrapping models (Fig. S8 and S9). Most contacts involved the DNA phosphate backbone and sugar moieties (Fig. S7B), consistent with the non-sequence-specific binding observed in crystal structures. HLp side chains contributed slightly more to DNA interactions than backbone atoms (Fig. S6C). Across both models, residues located in α1 and loops l1 and l2 exhibited an average of more than three DNA contacts (Fig. S8 and S9). In particular, K54 and R59 in loop l2 formed persistent hydrogen bonds with the DNA throughout the simulations (Fig. S10).

To dissect the spatial dynamics of binding, we defined four major DNA-binding regions on the HLp tetramer: A-sites 1 and 2, each comprising two “β-bridge” motifs with their flanking DNA-binding residues; and B-sites 1 and 2, each centered on a "paired end of helices" motif along with adjacent DNA-interacting residues (Fig. 5C and

Table S4). In the bridging model, DNA remained stably associated with B-sites 1 and 2, but progressively engaged A-site 2, and to a lesser extent A-site 1, as it transitioned into the wrapping mode (Fig. 5D and Fig. S11). In the wrapping model, local interactions at all four sites were stably maintained or even enhanced over time (Fig. 5D and Fig. S12). Notably, fluctuations in contact number at A-site 1 in both models appear to result from steric interference at the DNA termini.

To probe the unwrapping dynamics of the HLp-DNA complex, we performed steered molecular dynamics (SMD) simulations (Fig. S13). Unwrapping occurred in two distinct phases: initial disruption at B-site 2, followed by disengagement at A-site 2, yielding a partially unwrapped intermediate ensemble. The unwrapping energy landscape revealed two main barriers: ~500 kJ/mol at B-site 2 and ~400 kJ/mol at A-site 2. Energy decomposition analysis identified five residues — R59, K21, K15, S14, and K17 — as major contributors to DNA binding, collectively accounting for 52% of the total binding energy. These findings are consistent with the hydrogen bond and contact analyses from the unbiased simulations and highlight the cooperative, multivalent nature of HLp-DNA interactions. The high energetic cost of unwrapping suggests that dissociation is a regulated, stepwise process, reinforcing the stability of the wrapped state.

Together, our simulations support both wrapping and bridging as viable DNA-binding modes for HLp. However, the wrapping mode is energetically favored and yields more extensive, stable interactions — pointing to its likely predominance under physiological conditions.

HLp is wrapped by DNA *in vitro*

To investigate the mechanism of DNA binding by HLp *in vitro*, we employed a series of well-established biochemical and biophysical assays. We first used a MNase digestion assay, which measures the extent to which protein-bound DNA is protected from nucleolytic cleavage. The 600-bp dsDNA fragment was incubated with HLp and subsequently digested with increasing concentrations of MNase. The resulting digestion pattern revealed a characteristic ladder of fragments ranging from 35 to 72 bp (Fig. 6A), indicating that HLp protects bound DNA from enzymatic degradation. This protective effect is reminiscent of that observed for the nucleosomal archaeal histone HMfB, which served as a positive control. However, the HLp-protected fragments differ in both size and regularity: while HMfB consistently generates ~30-bp increments, the fragment sizes protected by HLp are more heterogeneous. These observations suggest that although both proteins compact DNA, they do so through distinct binding modes.

To directly monitor HLp-induced DNA compaction, we performed TPM experiments using the 685-bp linear DNA fragment. TPM enables real-time monitoring of protein-DNA interactions by measuring the Brownian motion of a DNA-tethered polystyrene bead; changes in DNA flexibility or contour length, such as those caused by protein-induced compaction, result in measurable shifts in root mean square (RMS) displacement of the bead. Upon addition of HLp, we observed a progressive, concentration-dependent decrease in RMS values, indicating compaction of the DNA (Fig. 6B). Saturation was reached at ~6000 nM HLp with a final RMS value of ~80 nm, closely matching the values observed for the archaeal histones HMfA and HMfB, suggesting that HLp wraps DNA rather than bridging it[59].

To investigate whether HLp induces DNA condensation across a range of DNA sizes, we performed ladder EMSA assays with linear dsDNA fragments of increasing length. HLp altered the electrophoretic mobility of DNA on agarose gels in a size-dependent manner, with fragments longer than 2000 bp migrating faster, while those shorter than 1000 bp showed reduced mobility (Fig. 6C). These results are consistent with protein-induced changes in DNA conformation and are qualitatively similar to those observed with HMfB, although HMfB required lower protein concentrations and produced more pronounced mobility shifts[17].

We next examined whether HLp binding affects DNA topology using a topoisomerase I relaxation assay. In the presence of HLp, the relaxed plasmid DNA became progressively supercoiled in a protein concentration-dependent manner, indicating the introduction of topological strain upon HLp binding (Fig. S14A). This behavior mirrors that of HMfB, albeit with a less pronounced effect.

Finally, we tested whether HLp promotes DNA end-joining using a ligase-mediated circularization assay. In this assay, short DNA fragments are circularized by T4 DNA ligase and non-circularized linear species are digested by T5 exonuclease. While HMfB efficiently promotes the circularization of DNA monomers due to its pronounced DNA-bending activity, HLp had only modest effects, even at high concentrations (Fig. S14B). In contrast, HLp favored the formation of linear DNA multimers as protein concentration increased, indicative of open-ended protein-DNA complexes formed upon binding (Fig. S14B).

Taken together with the crystal structures and molecular dynamics simulations, our *in vitro* data support a model in which HLp wraps and compacts DNA through a mechanism that is distinct from the bending mode employed by the bacterial histone HBb and from the nucleosome assembly of eukaryotic histones, differing in both its dynamics and topological effects.

HLp binds to genomic DNA *in vivo*

Due to the lack of facilities for handling *Leptospira* strains, we used *E. coli* as a heterologous model to investigate the effects of HLp on genomic DNA *in vivo* using light microscopy. HLp was expressed in *E. coli* BL21(DE3), while the archaeal histone HMfB was expressed in *E. coli* Mutant56(DE3) as a positive control. Ubiquitin from *Caldiarchaeum subterraneum* (*Cs*Ub), which lacks DNA-binding activity, served as a negative control[32]. Because expression levels varied significantly between constructs, the strain showing the highest expression level for each protein was selected for further analysis.

Growth curves were recorded to determine the onset of the stationary phase and to assess potential effects of protein expression on cell proliferation (Fig. S15). HLp and *Cs*Ub expression in *E. coli* BL21(DE3) led to a modest increase in doubling time and a reduction in final cell density compared to uninduced controls. However, similar effects were observed in cells carrying empty vectors, suggesting that these changes were not specific to HLp or *Cs*Ub overexpression. In contrast, HMfB expression had no measurable impact on cell growth. To minimize variability due to nucleoid dynamics during the cell cycle, microscopy samples were collected in early stationary phase, when cells had exited active division.

Microscopic analysis revealed that HLp expression markedly alters nucleoid organization. DAPI staining of HLp-expressing cells demonstrated clear elongation and a substantial increase in nucleoid volume relative to controls (Fig. 7), consistent with impaired chromosome condensation and/or interference with cell division. These effects were already evident in the absence of IPTG, likely due to basal expression from the leaky T7 promoter. Upon overexpression of HLp, the cytoplasm of HLp-expressing cells was predominantly occupied by decondensed genomic DNA, a phenotype also observed upon HMfB overexpression, suggesting that HLp disrupts the chromatin organization of the *E. coli* cells through its DNA-binding ability. Segmentation of the DAPI fluorescence signals in HMfB-expressing cells exhibited single, well-defined regions with clear contours. In contrast, segmentation of HLp-expression cells resulted in fluorescence signals dispersed across multiple regions with diffuse contours, indicating a more heterogeneous distribution of these nucleoid regions. Combined with TPM assay results, which showed HLp-induced DNA compaction only at high protein concentrations, these findings suggest that HLp induces localized DNA condensation and reveal subtle differences in chromatin organization between the two histones.

Discussion

In this study, we present a comprehensive analysis of the bacterial histone HLp from *L. perolatii*, integrating structural, biophysical, and functional data to elucidate its role in DNA binding and compaction. Our findings reveal that HLp assembles into a unique, DNA-wrapped tetramer, expanding the known repertoire of bacterial histone architectures, and providing new insights into prokaryotic chromatin organization. This represents the first experimental characterization of a bacterial histone belonging to the FtF subgroup of α3 histones, previously identified bioinformatically as one of the largest subfamilies of prokaryotic histones[9].

Structurally, HLp closely resembles the archaeal FtF histone HTkC from *T. kodakarensis*, consistent with bioinformatic predictions based on CLANS clustering, where the two histones are positioned adjacent to each other. Both proteins adopt a characteristic histone fold comprising three α-helices connected by two loops, yet differ notably from nucleosomal histones in their truncated α2 and α3 helices. This truncation is also found in HBb, the only other experimentally studied bacterial histone, highlighting a common structural adaptation within α3 histones[20, 24]. Unlike HBb, which strictly forms dimers, HLp and HTkC assemble into tetramers even in the absence of DNA, mediated by highly conserved residues situated at the C-terminal region of α2, loop l2, and helix α3. This unique oligomerization interface appears to be a defining feature of FtF histones and sets them apart from other prokaryotic histone families, including archaeal nucleosomal histones, which typically assemble into higher-order oligomers only upon DNA binding. Given their structural similarity and conserved tetramerization interface, we anticipate that HTkC may also wrap DNA similarly to HLp.

The HLp tetramer displays a continuous band of positive surface charge encircling the complex, consistent with its role in non-specific DNA binding. EMSA and MST assays confirm that HLp binds dsDNA without sequence specificity, with strongest affinity for DNA of ~40% GC content, which is similar to the native genomic composition of *L. perolatii*. Crystal structures of HLp-DNA complexes reveal the HLp tetramer to be decorated with distinct DNA-binding interfaces along its perimeter, including motifs corresponding to the canonical "paired end of helices" and “β-bridge” motifs[58]. Thereby, their spatial arrangement in HLp resembles that of eukaryotic and archaeal histones.

While each of the two individual DNA-bound crystal structures might hint at a DNA-bridging mode, their structural superposition and the location of the individual DNA-binding motifs suggested a potential DNA-wrapping mode for HLp. Of these two modes, MD simulations favor the wrapping mode as the energetically preferred configuration. Even simulations initiated in the bridging conformation exhibited spontaneous

progression toward wrapping. Notably, DNA wrapping involves ~60 bp encircling the tetramer, consistent with structural predictions. However, we cannot rule out that HLp also bridges DNA *in vivo*, potentially in concert with other factors.

Biochemical assays further support the wrapping-based binding mode. In MNase digestion assays, HLp protects DNA from nucleolytic cleavage, producing irregularly spaced fragments (35-72 bp) distinct from the regular 30-bp pattern generated by HMfB and the ~147-bp protection seen in eukaryotic nucleosomes. These fragment sizes correspond approximately to half and full turns of DNA around the HLp tetramer, indicating protection via wrapping but with flexible nucleosome-like spacing. Additional *in vitro* assays, including TPM, topoisomerase relaxation and ladder EMSA, demonstrate that HLp compacts DNA and alters its topology, albeit less efficiently than HMfB. HLp does not promote the formation of circular DNA monomers in the circularization assay, suggesting that it does not bring DNA ends into close proximity for ring closure under the tested conditions.

Expression of HLp in *E. coli* leads to dramatic reorganization of the nucleoid. DAPI staining reveals marked nucleoid decondensation and increased cellular length, indicative of disrupted DNA compaction and/or interference with cell division. These phenotypes closely mirror those observed upon HMfB expression, although the nucleoid appears more homogeneous with HMfB than with HLp, suggesting differences in chromatin architecture. These *in vivo* findings support the hypothesis that HLp functions as a global chromatin organizer, similar to its homolog in *L. interrogans*, where the FtF histone is among the most highly expressed and essential proteins[20]. Notably, neither *L. interrogans* nor *L. perolatii* encodes other histone homologs or canonical bacterial NAPs, further indicating that HLp serves as the principal chromatin component in these species. In contrast, species such as *B. bacteriovorus* encode two distinct histones (HBb and Bd3044) as well as NAPs such as HU, underscoring lineage-specific diversity in bacterial chromatin organization strategies[20, 24]. Consistent with HLp's proposed central role, genomic analysis reveals that in *L. perolatii*, the *hlp* gene resides within a conserved locus adjacent to genes encoding SMC, MAP, and a putative transmembrane protein (Fig. S1B). This synteny is preserved in *L. interrogans* and other *Leptospira* species, suggesting a functionally co-evolved genomic module (Fig. S1B). Such an arrangement raises the possibility that HLp cooperates closely with SMC to maintain proper chromosome structure and segregation, while the neighboring MAP enzyme, responsible for cleaving the initiating methionine from nascent proteins, may support efficient maturation of HLp, SMC, and other critical factors involved in nucleoid dynamics.

From an evolutionary perspective, HLp is only the second bacterial histone

characterized structurally and functionally, after the dimeric histone HBb. While bacterial dimeric histones like HBb appear restricted to bacteria and nucleosomal histones are found exclusively in archaea and eukaryotes, FtF histones uniquely span both archaeal and bacterial domains. Our bioinformatic analysis demonstrates that bacterial and archaeal FtF histones form distinct and clearly separated clusters, supporting the hypothesis that FtF histones may trace their ancestry back to the last universal common ancestor (LUCA). Alternatively, the broad distribution of FtF histones across archaeal phyla, including Asgard archaea, and their sparse, patchy occurrence in bacteria, suggest a scenario involving multiple independent horizontal gene transfer events from archaea into specific bacterial lineages[9]. Thus, the evolutionary history of FtF histones likely reflects a combination of ancestral origin, lineage-specific adaptations, and recurrent gene-transfer events.

In summary, our work expands the known diversity of bacterial histones and establishes HLp as a DNA-wrapping histone that assembles into a stable homotetramer. Rather than being simplified ancestors of eukaryotic histones, prokaryotic histones represent structurally diverse, functionally adaptable proteins that contribute to genome organization through distinct DNA-binding modes. Beyond structural and mechanistic insight, future studies will need to explore how these histones function in their native cellular contexts — how they interact with other histone variants, cooperate with or substitute for NAPs, and participate in processes such as transcription, replication, and genome repair.

Data availability

Coordinates and structure factors of the crystal structures have been deposited in the PDB under entry numbers 9QT0 (DNA-free HLp), 9QT1 (HLp-DNA_1), and 9QT2 (HLp-DNA_2).

For the MD analysis, all configuration files, trajectory files, movies, and analysis scripts are deposited at Zenodo (http://zenodo.org/) and accessible with the DOI 10.5281/zenodo.15234989.

The TPM data is deposited in the 4TU repository (https://data.4tu.nl) and accessible with the DOI 10.4121/5b604dd5-5498-46aa-b437-775e957f93e3.

Acknowledgements

We thank the staff of Beamline X10SA of the Swiss Light Source (PSI, Villigen, Switzerland) for excellent technical support. We are grateful to Reinhard Albrecht for assistance with crystallization and crystallographic data collection. We extend our

thanks to Linxuan Li (Dept. of Integrative Evolutionary Biology, MPI for Biology Tübingen, Germany) and Agnes Henschen (Dept. of Algal Development and Evolution, MPI for Biology Tübingen, Germany) for their support with light microscopy sample preparation. We thank Pedro Escudeiro (Dept. of Protein Evolution, MPI for Biology Tübingen, Germany) for proofreading the manuscript. We acknowledge the HPC system Raven at the Max Planck Computing and Data Facility for the performed computational work.

Funding

This work was supported by institutional funds from the Max Planck Society and funding from the Netherlands Organization for Scientific Research [OCENW.GROOT.2019.012].

Author Contributions

Conceived and designed the experiments: A.N.L., B.H.A., V.A., Y.H.
Performed the experiments: A.P., H.R., K.B., S.S., Y.H.
Performed MD simulations: K.Q., Y.Z.
Performed bioinformatic analyses: V.A.
Analyzed the data: A.P., B.H.A., M.D.H., S.S., R.T.D, V.A., Y.H.
Wrote the paper: B.H.A., V.A., Y.H., with contributions from the other authors.

Competing Interest Statement

The authors declare no competing interests.

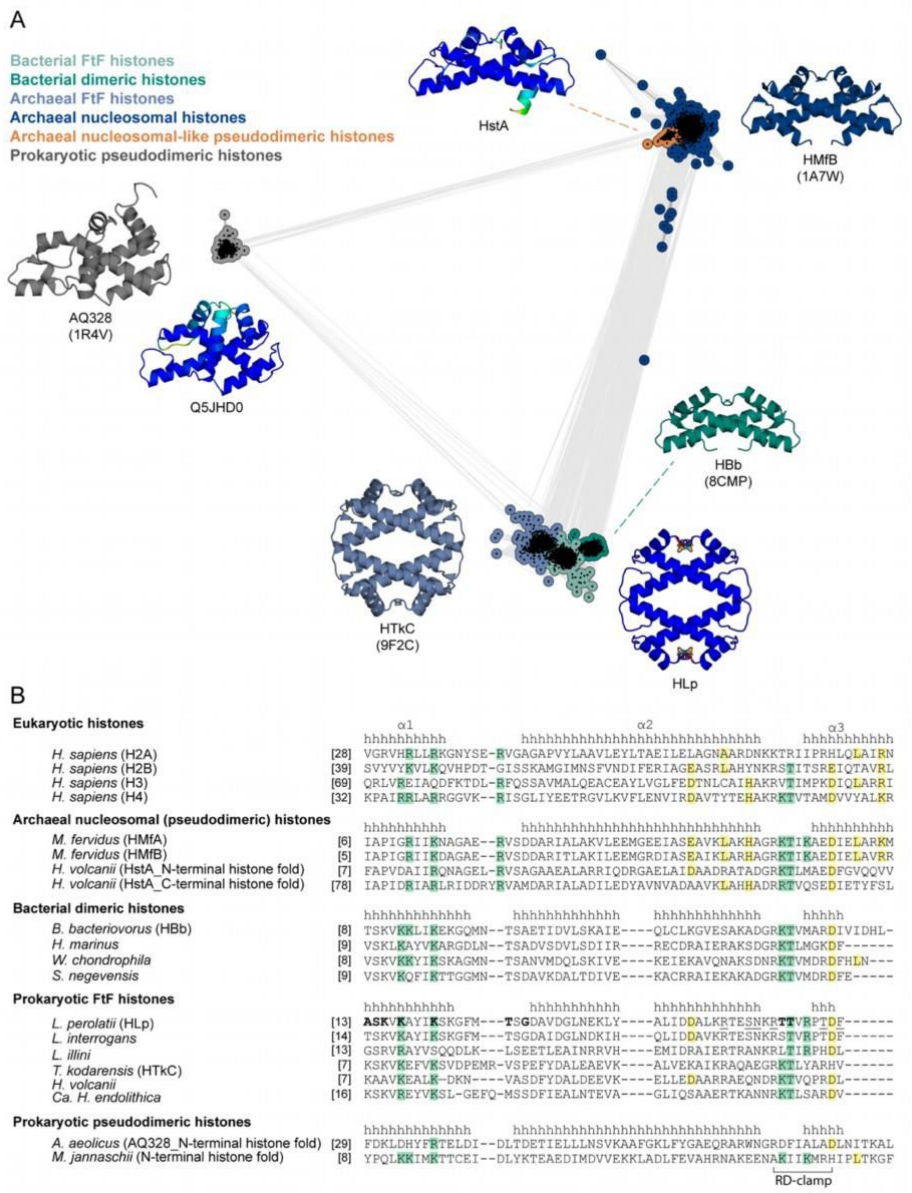


Fig. 1 Comparative analysis of prokaryotic histones composed exclusively of histone fold domains.

A. Cluster map of ~3,300 prokaryotic histones containing only the histone fold domain. Each dot represents a single protein sequence, edges reflect pairwise sequence similarity. Structures of representative proteins — either crystal structures (PDB IDs in parentheses) or AlphaFold models — are shown in cartoon representation. AlphaFold models are colored by pLDDT confidence scores. Representative proteins include: HstA from *Haloferax volcanii* (D4GS56), an archaeal nucleosomal-like pseudodimeric histone; HMfB from *Methanothermus fervidus* (P19267), an archaeal nucleosomal histone; HBb from *Bdellovibrio bacteriovorus* (Q6MRM1), a bacterial dimeric histone; HLp from *Leptospira perolatii* (A0A2M9ZN55), a bacterial FtF histone; HTkC from *Thermococcus kodakarensis* (Q5JDW7), an archaeal FtF histone; and Q5JHD0 from *T. kodakarensis* (Q5JHD0) and AQ328

from *Aquifex aeolicus* (O66665), both prokaryotic pseudodimeric histones. **B.** Multiple sequence alignment of representative histones from both eukaryotes and prokaryotes, including *Homo sapiens* (H2A: P04908; H2B: P62807; H3: P68431; H4: P62805), *M. fervidus* (HMfA: P48781; HMfB: P19267), *H. volcanii* (HstA: D4GS56), *B. bacteriovorus* (HBb: Q6MRM1), *Hymenobacter marinus* (NCBI accession No.: ATH09486), *Waddlia chondrophila* (D6YWW1), *Simkania negevensis* (F8L7X8), *L. perolatii* (HLp: A0A2M9ZN55), *L. interrogans* (Q8F3E8), *Leptonema illini* (H2CFS2), *T. kodakarensis* (HTkC: Q5JDW7), *H. volcanii* (D4GZE0), Candidatus *Heimdallarchaeum endolithica* (A0A9Y1FPJ9), *A. aeolicus* (O66665), and *Methanocaldococcus jannaschii* (A0A832T4V6). α-helices are annotated with “h” based on crystal structures of H2A, HMfA, HBb, HLp, and AQ328. Conserved residues associated with DNA binding and oligomerization are highlighted in green and yellow, respectively. HLp residues contributing to tetramerization in the HLp crystal structure (PDB: 9QT0) are underlined, while those interacting with the DNA backbone in HLp-DNA complexes (PDB: 9QT1 and 9QT2) are shown in bold. Unless otherwise stated, accession numbers in parentheses refer to UniProtKB.

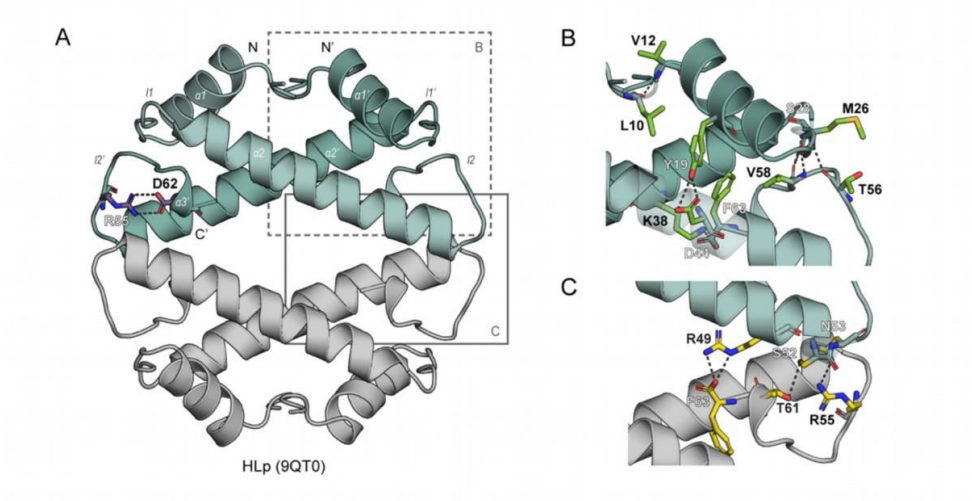


Fig. 2 Crystal structure of the HLp tetramer.

A. Crystal structure of the HLp tetramer (PDB: 9QT0) shown in cartoon representation. Residues R55 and D62, which form the RD clamp, are shown as sticks. Salt bridges are indicated as dashed lines. **B.** Close-up view of the HLp dimerization interface. Residues involved in dimerization are shown as green sticks, with hydrogen bonds depicted as dashed lines. **C.** Close-up view of the HLp tetramerization interface. Tetramerization-associated residues are shown as yellow sticks. Salt bridges and hydrogen bonds are indicated by dashed lines. In all panels, the contents of the asymmetric unit are shown in color, with selected symmetry mates in gray.

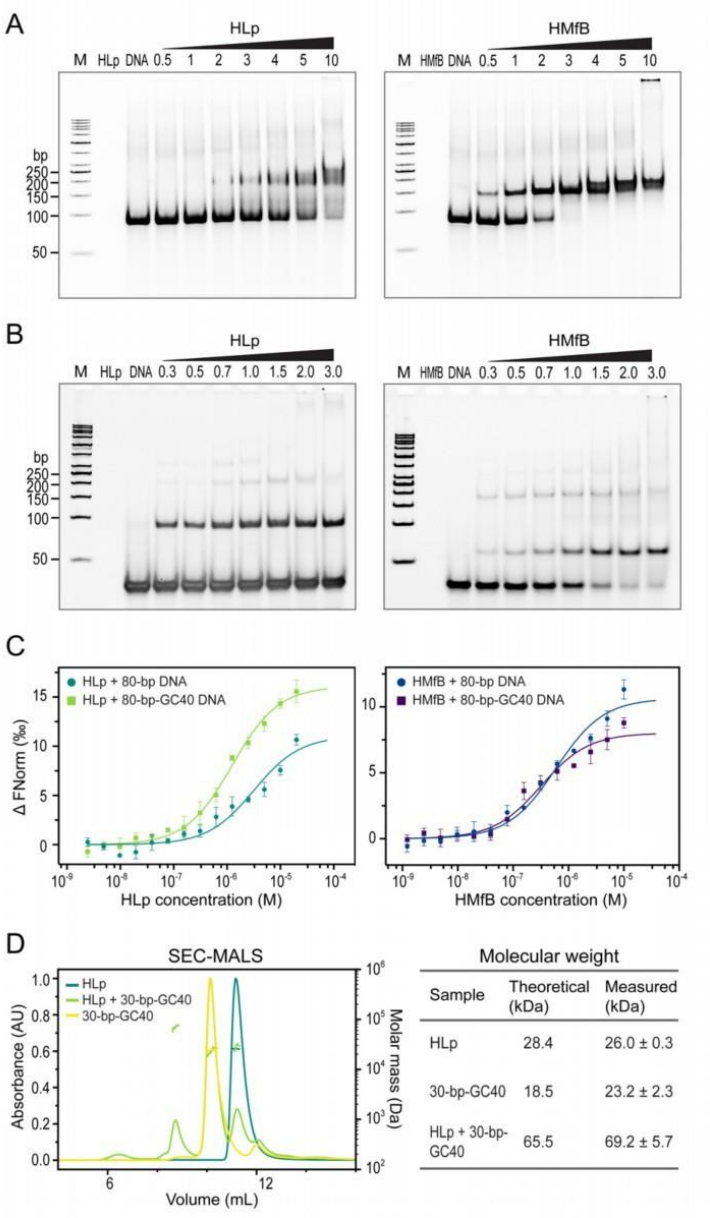


Fig. 3 HLp binds DNA *in vitro*.

EMSAs showing the binding of HLp and HMfB (control) to the 80-bp DNA fragment (**A**) and the 30-bp-GC40 DNA fragment (**B**). Increasing protein concentrations (lanes 4-10), indicated as molar protein-to-DNA ratios, were incubated with the corresponding DNA and analyzed on a 6% polyacrylamide gel. **C.** MST measurements comparing the binding of HLp to the 80-bp and the 80-bp-GC40 DNA fragments, alongside HMfB. **D**. Profiles of SEC-MALS runs of HLp, the 30-bp-GC40 DNA, and a mixture of both; the table summarizes both theoretical and experimentally determined molecular weights.

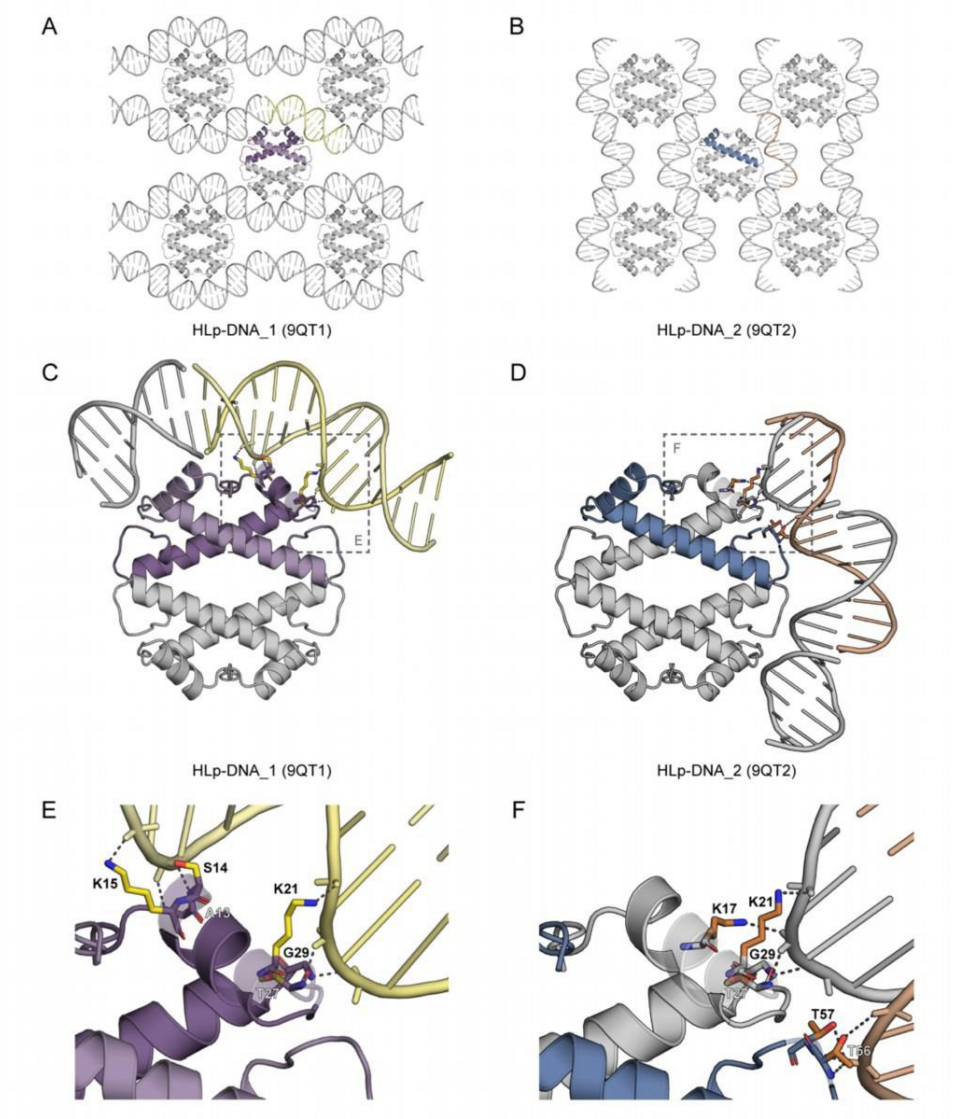


Fig. 4 Crystal structures of DNA-bound HLp.

A, B. Crystal packing of HLp-DNA_1 (**A**) and HLp-DNA_2 (**B**) showing selected symmetry mates within 20 Å. **C, D.** Crystal structures of HLp-DNA_1 (PDB: 9QT1) (**C**) and HLp-DNA_2 (PDB: 9QT2) (**D**) shown in cartoon representation. The framed regions in panels **C** and **D** correspond to the magnified views shown in **E** and **F**, respectively. **E, F.** Residues involved in DNA binding are shown as sticks, and protein-DNA interactions are depicted as dashed lines. In all panels, the contents of the asymmetric unit are shown in color, and selected symmetry mates in gray.

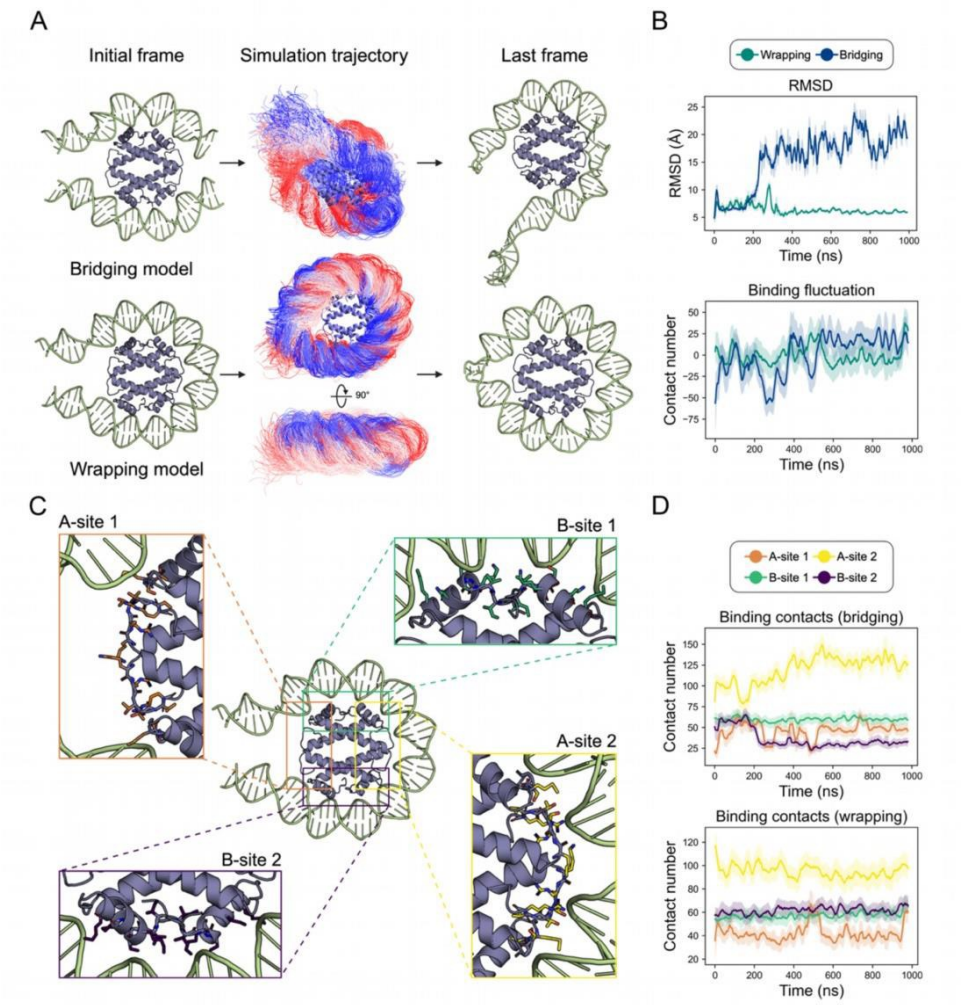
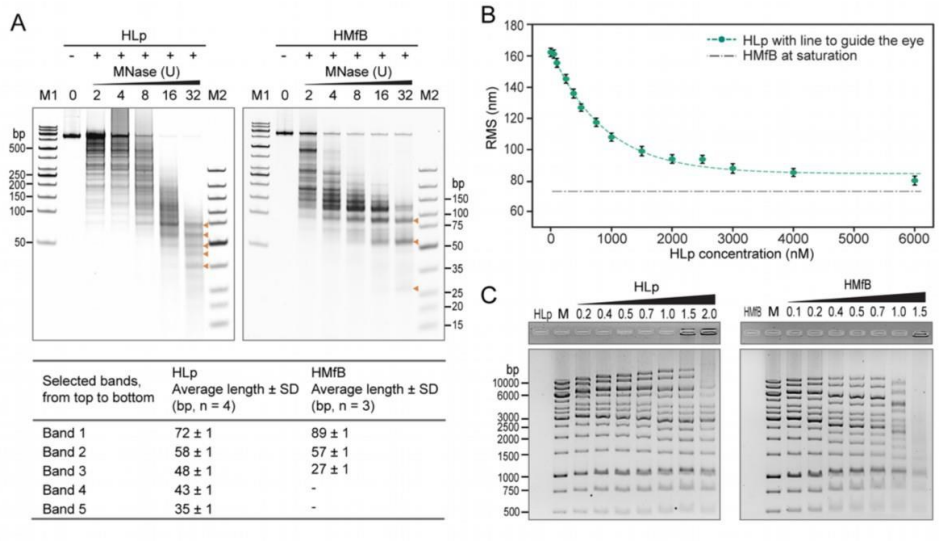


Fig. 5 Molecular dynamics simulations of the wrapping and bridging modes.

A. Visualization of the starting model (left), the conformational ensemble derived from the simulation (middle), and the structure of the last frame (right). **B.** Overall characterization of the simulations of the two proposed binding modes, showing the RMSD of the DNA backbones (top) and binding fluctuations (bottom). Binding fluctuation is defined as the deviation of the contact number in each frame from the average contact number across all frames. Each plot is smoothed with a window size of 100, with green and blue lines represent the wrapping and bridging models, respectively. **C.** Structures of the four defined binding sites: A-site 1 (orange), A-site 2 (yellow), B-site 1 (green), and B-site 2 (purple). Residues involved in each site are shown as sticks and listed in Table S4. **D.** Binding contacts for the four defined binding sites throughout the simulation for the wrapping (top) and bridging (bottom) models. Line colors correspond to those in panel C. Binding contacts are defined as any pair of heavy atoms within 4 Å.



Selected bands, from top to bottom	HLp Average length ± SD (bp, n = 4)	HMfB Average length ± SD (bp, n = 3)
Band 1	72 ± 1	89 ± 1
Band 2	58 ± 1	57 ± 1
Band 3	48 ± 1	27 ± 1
Band 4	43 ± 1	-
Band 5	35 ± 1	-

Fig.6 DNA wraps around HLp.

A. MNase digestion assay analyzing the protection of the 600-bp-GC40 DNA fragment by HLp in comparison to HMfB. Increasing MNase concentrations used in the assay are indicated. The lengths of DNA fragments obtained at the highest MNase concentration were determined by densitometric analysis, using the molecular weight marker M2 (GeneRuler Ultra Low Range DNA Ladder, ThermoFisher Scientific) as the standard. **B.** TPM experiment with HLp and the 685-bp DNA. Each measurement point represents the average of triplicate measurements, with error bars indicating standard deviations. The line connecting the points was generated by fitting the means to a logistic function. **C.** EMSA analyzing the binding of HLp and HMfB (control) to the DNA fragments of the GeneRuler 1 kb Ladder (ThermoFisher Scientific). Increasing protein concentrations (lanes 3-9), indicated as protein-to-DNA mass ratios, were incubated with the corresponding DNA and analyzed on a 1% agarose gel.

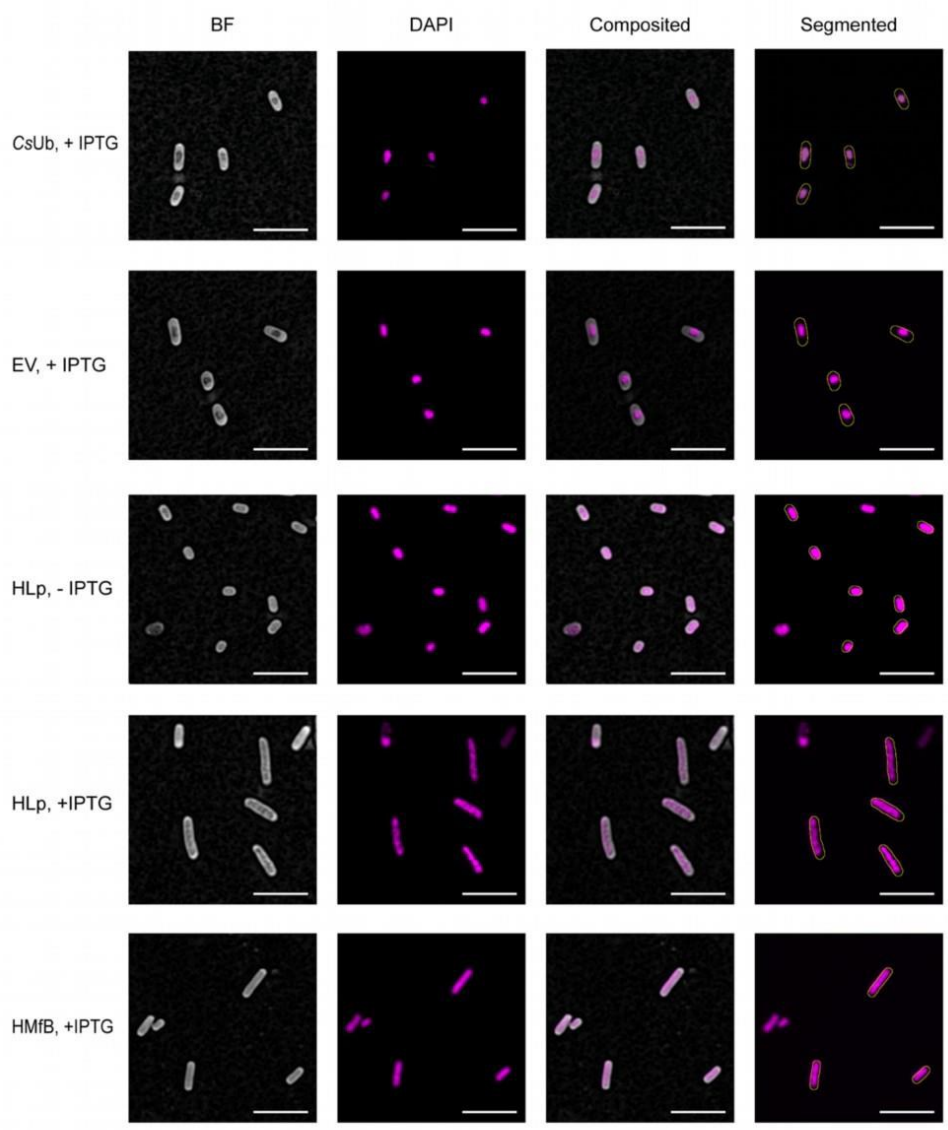


Fig. 7 HLp binds to genomic DNA *in vivo*.

Airyscan imaging of *E. coli* cells expressing HLp and HMfB (positive control) with DAPI dye, compared to negative controls transformed with the empty vector (EV), the vector expressing *Cs*Ub, and the uninduced HLp sample (-IPTG). The left column presents the brightfield image with subtracted background. The second column displays the maximum intensity projection of fluorescence. The third one is the overlap, and the last column shows the segmented bacteria outline on top of the fluorescence channel. Scale bar: 5 µm.

References

1. Arents G, Burlingame RW, Wang BC, Love WE, Moudrianakis EN. The Nucleosomal Core Histone Octamer at 3.1-a Resolution - a Tripartite Protein Assembly and a Left-Handed Superhelix. *P Natl Acad Sci USA* **88**, 10148-10152 (1991).

2. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 angstrom resolution. *Nature* **389**, 251-260 (1997).

3. Kornberg RD, Thomas JO. Chromatin structure; oligomers of the histones. *Science* **184**, 865-868 (1974).

4. Allan J, Hartman PG, Crane-Robinson C, Aviles FX. The structure of histone H1 and its location in chromatin. *Nature* **288**, 675-679 (1980).

5. Li W*, et al.* Structural basis for linker histone H5–nucleosome binding and chromatin fiber compaction. *Cell Research* **34**, 707-724 (2024).

6. Luger K, Richmond TJ. The histone tails of the nucleosome. *Curr Opin Genet Dev* **8**, 140-146 (1998).

7. Jenuwein T, Allis CD. Translating the histone code. *Science* **293**, 1074-1080 (2001).

8. Hocher A, Warnecke T. Nucleosomes at the Dawn of Eukaryotes. *Genome Biol Evol* **16**, evae029 (2024).

9. Schwab S*, et al.* Histones and histone variant families in prokaryotes. *Nature Communications* **15**, 7950 (2024).

10. Henneman B, van Emmerik C, van Ingen H, Dame RT. Structure and function of archaeal histones. *PLoS Genet* **14**, e1007582 (2018).

11. Stevens KM*, et al.* Histone variants in archaea and the evolution of combinatorial chromatin complexity. *P Natl Acad Sci USA* **117**, 33384-33395 (2020).

12. Dame RT. The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol Microbiol* **56**, 858-870 (2005).

13. Dorman CJ, Deighan P. Regulation of gene expression by histone-like proteins in bacteria. *Current Opinion in Genetics & Development* **13**, 179-184 (2003).

14. Dillon SC, Dorman CJ. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nature Reviews Microbiology* **8**, 185-195 (2010).

15. Dame RT. Special Issue: Role of Bacterial Chromatin in Environmental Sensing, Adaptation and Evolution. *Microorganisms* **9**, (2021).

16. Mattiroli F*, et al.* Structure of histone-based chromatin in Archaea. *Science* **357**, 609-612 (2017).

17. Sandman K, Krzycki JA, Dobrinski B, Lurz R, Reeve JN. HMf, a DNA-binding protein isolated from the hyperthermophilic archaeon Methanothermus fervidus, is most closely related to histones. *Proc Natl Acad Sci U S A* **87**, 5788-5791 (1990).

18. Sanders TJ*, et al.* Extended Archaeal Histone-Based Chromatin Structure Regulates Global Gene Expression in Thermococcus kodakarensis. *Front Microbiol* **12**, 681150 (2021).

19. Ofer S*, et al.* DNA-bridging by an archaeal histone variant via a unique tetramerisation interface. *Communications Biology* **6**, 968 (2023).

20. Hocher A*, et al.* Histones with an unconventional DNA-binding mode in vitro are major chromatin constituents in the bacterium Bdellovibrio bacteriovorus. *Nat Microbiol* **8**, 2006-2019 (2023).

21. Alva V, Lupas AN. Histones predate the split between bacteria and archaea. *Bioinformatics* **35**, 2349-2353 (2019).

22. Qiu Y*, et al.* The crystal structure of Aq_328 from the hyperthermophilic bacteria Aquifex aeolicus shows an ancestral histone fold. *Proteins* **62**, 8-16 (2006).

23. Paratsaphan S*, et al.* Characterization of a novel peptide from pathogenic leptospira and its cytotoxic effect. *Pathogens* **9**, 906 (2020).

24. Hu Y*, et al.* Bacterial histone HBb from Bdellovibrio bacteriovorus compacts DNA by bending. *Nucleic Acids Res* **52**, 8193-8204 (2024).

25. Camacho C*, et al.* BLAST+: architecture and applications. *BMC bioinformatics* **10**, 1-9 (2009).

26. Sayers EW*, et al.* Database resources of the National Center for Biotechnology Information in 2025. *Nucleic acids research* **53**, D20-D29 (2025).

27. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature biotechnology* **35**, 1026-1028 (2017).

28. Frickey T, Lupas A. CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics* **20**, 3702-3704 (2004).

29. Zimmermann L*, et al.* A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *Journal of molecular biology* **430**, 2237-2243 (2018).

30. Oberg N, Zallot R, Gerlt JA. EFI-EST, EFI-GNT, and EFI-CGFP: enzyme function initiative (EFI) web resource for genomic enzymology tools. *Journal of molecular biology* **435**, 168018 (2023).

31. Jumper J*, et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).

32. Fuchs ACD, Maldoner L, Wojtynek M, Hartmann MD, Martin J. Rpn11-mediated ubiquitin processing in an ancestral archaeal ubiquitination system. *Nature Communications* **9**, 2696 (2018).

33. Bogomolovas J, Simon B, Sattler M, Stier G. Screening of fusion partners for high yield expression and purification of bioactive viscotoxins. *Protein Expr Purif* **64**, 16-23 (2009).

34. Kabsch W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132 (2010).

35. Tickle I*, et al.* Staraniso. *Cambridge, United Kingdom: Global Phasing Ltd* **923**, (2018).

36. Evans R*, et al.* Protein complex prediction with AlphaFold-Multimer. *biorxiv*, 2021.2010. 2004.463034 (2021).

37. Vagin A, Teplyakov A. MOLREP: an automated program for molecular replacement. *J Appl Crystallogr* **30**, 1022-1025 (1997).

38. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).

39. Murshudov GN*, et al.* REFMAC5 for the refinement of macromolecular crystal

structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-367 (2011).

40. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. *Journal of computational chemistry* **26**, 1701-1718 (2005).

41. Huang J*, et al.* CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature methods* **14**, 71-73 (2017).

42. Mark P, Nilsson L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *The Journal of Physical Chemistry A* **105**, 9954-9960 (2001).

43. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics* **52**, 7182-7190 (1981).

44. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *The Journal of chemical physics* **126**, (2007).

45. Darden T, York D, Pedersen L. Particle mesh Ewald: An N log (N) method for Ewald sums in large systems. *Journal of chemical physics* **98**, 10089-10089 (1993).

46. Hess B, Bekker H, Berendsen HJ, Fraaije JG. LINCS: A linear constraint solver for molecular simulations. *Journal of computational chemistry* **18**, 1463-1472 (1997).

47. Park S, Schulten K. Calculating potentials of mean force from steered molecular dynamics simulations. *The Journal of chemical physics* **120**, 5946-5961 (2004).

48. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *Journal of molecular graphics* **14**, 33-38 (1996).

49. Michaud-Agrawal N, Denning EJ, Woolf TB, Beckstein O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *Journal of computational chemistry* **32**, 2319-2327 (2011).

50. Schindelin J*, et al.* Fiji: an open-source platform for biological-image analysis. *Nature methods* **9**, 676-682 (2012).

51. Henneman B, Heinsman J, Battjes J, Dame RT. Quantitation of DNA-Binding Affinity Using Tethered Particle Motion. *Methods Mol Biol* **1837**, 257-275 (2018).

52. Huff J. The Airyscan detector from ZEISS: confocal imaging with improved signal-to-noise ratio and super-resolution. (eds). Nature Publishing Group US New York (2015).

53. Thibeaux R*, et al.* Deciphering the unexplored Leptospira diversity from soils uncovers genomic evolution to virulence. *Microbial genomics* **4**, e000144 (2018).

54. Shine J, Dalgarno L. The 3′-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences* **71**, 1342-1346 (1974).

55. Decanniere K, Babu AM, Sandman K, Reeve JN, Heinemann U. Crystal structures of recombinant histones HMfA and HMfB from the hyperthermophilic archaeon Methanothermus fervidus. *J Mol Biol* **303**, 35-47 (2000).

56. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences* **98**, 10037-10041 (2001).

57. Bailey KA, Pereira SL, Widom J, Reeve JN. Archaeal histone selection of nucleosome positioning sequences and the procaryotic origin of histone-dependent genome evolution. *J Mol Biol* **303**, 25-34 (2000).

58. Arents G, Moudrianakis EN. Topography of the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA. *Proc Natl Acad Sci U S A* **90**, 10489-10493 (1993).

59. Henneman B*, et al.* Mechanical and structural properties of archaeal hypernucleosomes. *Nucleic Acids Res* **49**, 4338-4349 (2021).

Supplementary information of

DNA Wrapping by a Tetrameric Bacterial Histone

Yimin Hu[a], Samuel Schwab[b,#], Kaiyu Qiu[a,#], Yunsen Zhang[c], Kerstin Bär[a], Heidi Reichle[a], Aurora Panzera[d], Andrei N. Lupas[a], Marcus D. Hartmann[a,e], Remus T. Dame[b], Vikram Alva[a,*], Birte Hernandez Alvarez[a,*]

[a] Max Planck Institute for Biology Tübingen, Department of Protein Evolution, Tübingen, Germany

[b] Leiden Institute of Chemistry, Leiden University, Leiden, The Netherlands; Centre for Microbial Cell Biology, Leiden University, Leiden, The Netherlands; Centre for Interdisciplinary Genome Research, Leiden University, Leiden, The Netherlands

[c] Theoretical and Computational Biophysics Group, Beckman Institute for Advanced Science and Technology, Center of Biophysics and Quantitative Biology, University of Illinois Urbana-Champaign, Urbana, United States

[d] Max Planck Institute for Biology Tübingen, BioOptics Facility, Tübingen, Germany

[e] Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany

[#] Joint second authors
*To whom correspondence should be addressed
E-mail: vikram.alva@tuebingen.mpg.de or birte.hernandez@tuebingen.mpg.de

Supplementary Tables

Supplementary Table S1
Oligonucleotides used in this work.

Oligo name	Sequence (5' to 3')	Application
30-bp-GC30-F	TTTAAAACGCTTTAAAACGCTTTAAAACGC	EMSA
30-bp-GC30-R	GCGTTTTAAAGCGTTTTAAAGCGTTTTAAA	EMSA
30-bp-GC40-F	TTTAAAGCCGTTTAAAGCCGTTTAAAGCCG	SEC-MALS, EMSA, Crystallization
30-bp-GC40-R	CGGCTTTAAACGGCTTTAAACGGCTTTAAA	SEC-MALS, EMSA, Crystallization
30-bp-GC50-F	TTAAAGCCCGTTAAAGCCCGTTAAAGCCCG	EMSA
30-bp-GC50-R	CGGGCTTTAACGGGCTTTAACGGGCTTTAA	EMSA
30-bp-GC60-F	TTAAGCCCCGTTAAGCCCCGTTAAGCCCCG	EMSA
30-bp-GC60-R	CGGGGCTTAACGGGGCTTAACGGGGCTTAA	EMSA
80-bp-DNA-F	CCGTACTGTCGTCTGCGGCCTTTGATTATCA ATTAAAGCGTTCTACGGCGTTTTTGATCGCT CAACGTGCGGAGCTAGAT	EMSA
80-bp-DNA-F[Cy5]	[Cyanine5]CCGTACTGTCGTCTGCGGCCTTTG ATTATCAATTAAAGCGTTCTACGGCGTTTTTG ATCGCTCAACGTGCGGAGCTAGAT	MST
80-bp-DNA-R	ATCTAGCTCCGCACGTTGAGCGATCAAAAAC GCCGTAGAACGCTTTAATTGATAATCAAAGG CCGCAGACGACAGTACGG	EMSA, MST
pET-600-bp-GC40-F	CGCGAATTTTAACAAAATATTAACGTTTACA	MNase digestion
pET-600-bp-GC40-R	ATTCAGGTGAAAATATTGTTGATGCG	MNase digestion
80-bp-GC40-F[Cy5]	[Cyanine5]TTTAAAGCCGTTTAAAGCCGTTTAA AGCCGTTTAAAGCCGTTTAAAGCCGTTTAAA GCCGTTTAAAGCCGTTTAAAGCCG	MST
80-bp-GC40-R	CGGCTTTAAACGGCTTTAAACGGCTTTAAAC GGCTTTAAACGGCTTTAAACGGCTTTAAACG GCTTTAAACGGCTTTAAA	MST
685-bp-DNA-F[Biotin]	[Biotin]TTACTTTCACCAGCGTTTCTGGGTGAG CAAAAACAG	TPM

685-bp-DNA-R[DIG]	[DIG]CCAAGTAGCGAAGCGAGCAGGACTGGGCGG	TPM
pET-240-bp-GC40-Fp	[Phos]TGCAATTTATTCATATCAGGATTATCA	Ligase-mediated circularization assay
pET-240-bp-GC40-Rp	[Phos]GCATAAACTTTTGCCATTCTCACC	Ligase-mediated circularization assay

Supplementary Table S2
Binding affinities (K_d) of HLp and HMfB to DNA substrates measured with MST.

Protein	DNA substrate	K_d ± SD* (μM)
HLp	80-bp DNA	3.18 ± 1.06
	80-bp-GC40 DNA	1.24 ± 0.14
HMfB	80-bp DNA	0.58 ± 0.13
	80-bp-GC40 DNA	0.31 ± 0.08

* Average K_d value determined from three independent measurements with standard deviation (SD)

Supplementary Table S3

Data collection and refinement statistics of DNA-free HLp, HLp-DNA_1 and HLp-DNA_2.

Values for the outer shell are given in parentheses. Values for the ellipsoidal completeness are given in square brackets.

	HLp	HLp-DNA_1	HLp-DNA_2
Data collection			
Space group	$P3_121$	$P4_12_12$	F222
Cell dimensions			
a, *b*, *c* (Å)	55.70, 55.70, 58.46	67.93, 67.93, 97.90	48.77, 91.61, 100.33
α, β, γ (°)	90, 90, 120	90, 90, 90	90, 90, 90
Resolution range (Å)	48.25-1.30 (1.38-1.30)	43.16-2.10 (2.23-2.10)	39.56-1.90 (2.05-1.90)
Completeness (%)	99.7 (98.3)	99.7 (99.5)	68.4 (19.9) [85.4 (48.4)]
Redundancy	9.43 (5.34)	5.62 (5.53)	7.51 (6.44)
$\langle I/\sigma(I)\rangle$	19.87 (1.14)	11.09 (1.96)	13.86 (1.34)
R_{meas}	0.052 (1.40)	0.081 (0.743)	0.060 (1.534)
Refinement			
No. of reflections, working set	24883	13229	5261
No. of reflections, test set	1303	696	918
Final R_{cryst}	0.175	0.250	0.235
Final R_{free}	0.211	0.280	0.288
R.m.s. deviations			
Bonds (Å)	0.0046	0.0068	0.0042
Angles (°)	1.164	1.297	1.173

Supplementary Table S4
Amino acid residues of the four defined binding sites.

Binding site	Residues (Chain ID-Residue ID)
A-site 1	A-MET26, A-THR27, A-SER28, A-GLY29, B-LYS54, B-ARG55, B-THR56, B-THR57, B-VAL58, B-ARG59, C-MET26, C-THR27, C-SER28, C-GLY29, D-LYS54, D-ARG55, D-THR56, D-THR57, D-VAL58, D-ARG59
A-site 2	A-LYS54, A-ARG55, A-THR56, A-THR57, A-VAL58, A-ARG59, B-MET26, B-THR27, B-SER28, B-GLY29, C-LYS54, C-ARG55, C-THR56, C-THR57, C-VAL58, C-ARG59, D-MET26, D-THR27, D-SER28, D-GLY29
B-site 1	C-ILE11, C-VAL12, C-ALA-13, C-SER14, C-LYS15, C-LYS17, C-LYS21, D-ILE11, D-VAL12, D-ALA-13, D-SER14, D-LYS15, D-LYS17, D-LYS21
B-site 2	A-ILE11, A-VAL12, A-ALA-13, A-SER14, A-LYS15, A-LYS17, A-LYS21, B-ILE11, B-VAL12, B-ALA-13, B-SER14, B-LYS15, B-LYS17, B-LYS21

Supplementary figures

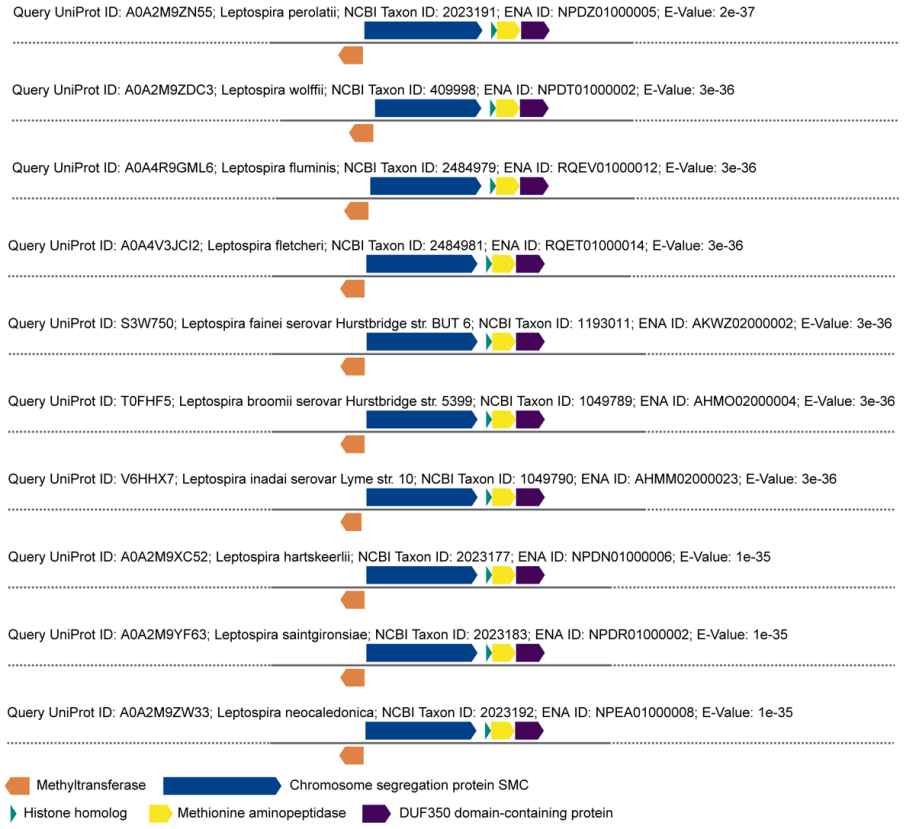
A

>NPDY01000005.1:141450-141799 Leptospira perolatii strain FH1-B-C1 contig_5, whole genome shotgun sequence
CTTCTGAAAATGATTCCTTTTTTGAGGGTATTCTGGAATGGTATCTTTCTAAGTACTGAGTACCTTCTATAAGAACTTACAGAAAAGGAGAAAGAGCTCCCCGGCAA
AAATTGGATATGGTACAAAAGAAAAAAACTACAGTTAAGAGGAAGAGTTCCATGGCTCAAAACGCTGAAAAGGATACCCTCATCGTCGCAAGTAAGGTGAAAGCCTA
TATCAAATCCAAAGGGTTCATGACTTCTGGGGATGCAGTCGATGGTTTGAATGAAAAGTTGTACGCATTAATCGATGATGCGTTAAAGCGCACTGAGTCCAACAAAC
GGACTACGGTTCGCCCAACCGACTTCTAA

>WP_207761458.1 hypothetical protein [Leptospira perolatii]
MVQKKTTVKRKSMAQNAEKDTLIVASKVKAYIKSKGFMTSGDAVDGLNEKLYALIDDALKRTESNKRTTVRPTDF

>PJZ70190.1 hypothetical protein CH360_07900 [Leptospira perolatii]
MAQNAEKDTLIVASKVKAYIKSKGFMTSGDAVDGLNEKLYALIDDALKRTESNKRTTVRPTDF

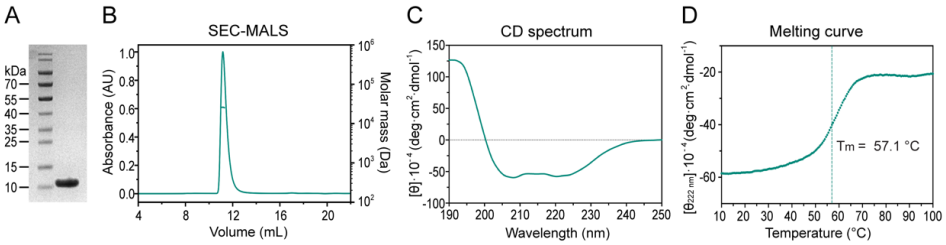
B



Supplementary Figure S1

DNA and protein sequences of HLp and conserved genomic context across Leptospira strains.

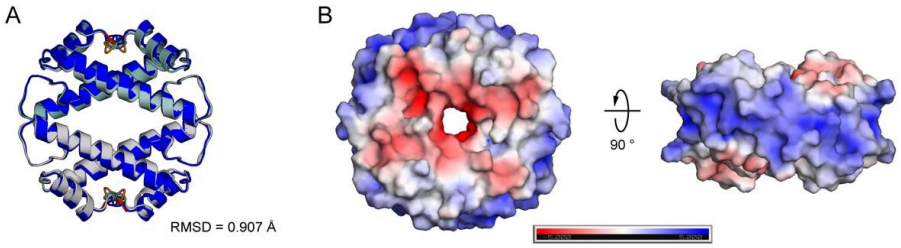
A. Nucleotide and amino acid sequences of HLp. The start codon of the longer *hlp* gene and its corresponding methionine are colored green, while those of the shorter gene variant are highlighted in orange. The Shine-Dalgarno sequence is highlighted in yellow. **B.** Genomic neighborhood diagrams showing HLp and its homologs in various Leptospira species, illustrating conserved synteny and gene context.



Supplementary Figure S2

Biophysical characterization of HLp in terms of purity, stability, and oligomeric state.

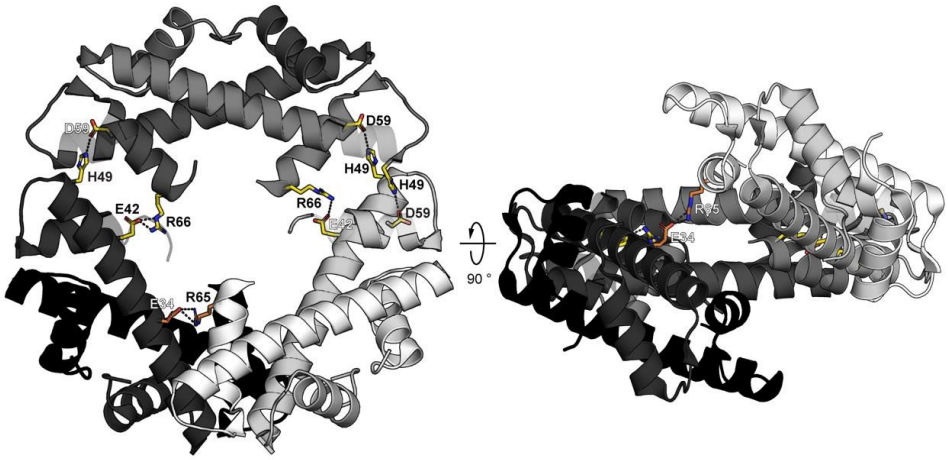
A. SDS-PAGE showing purified HLp. **B.** SEC-MALS analysis of HLp showing its tetrameric state. **C.** Single CD spectrum of HLp. **D.** Thermal melting curve of HLp measured with CD spectroscopy at a wavelength of 222 nm.



Supplementary Figure S3

Analysis of the HLp crystal structure.

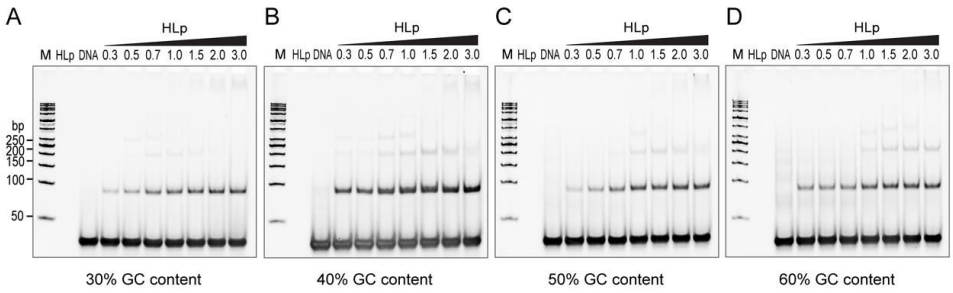
A. Superposition of the crystal structure of the HLp tetramer with its AlphaFold prediction, showing an RMSD of 0.907 Å. **B.** APBS (Advanced Poisson-Boltzmann Solver) electrostatic analysis of the HLp tetrameric structure reveals a continuous, positively charged surface encircling the entire tetramer.



Supplementary Figure S4

Structure of three assembled HMfB dimers.

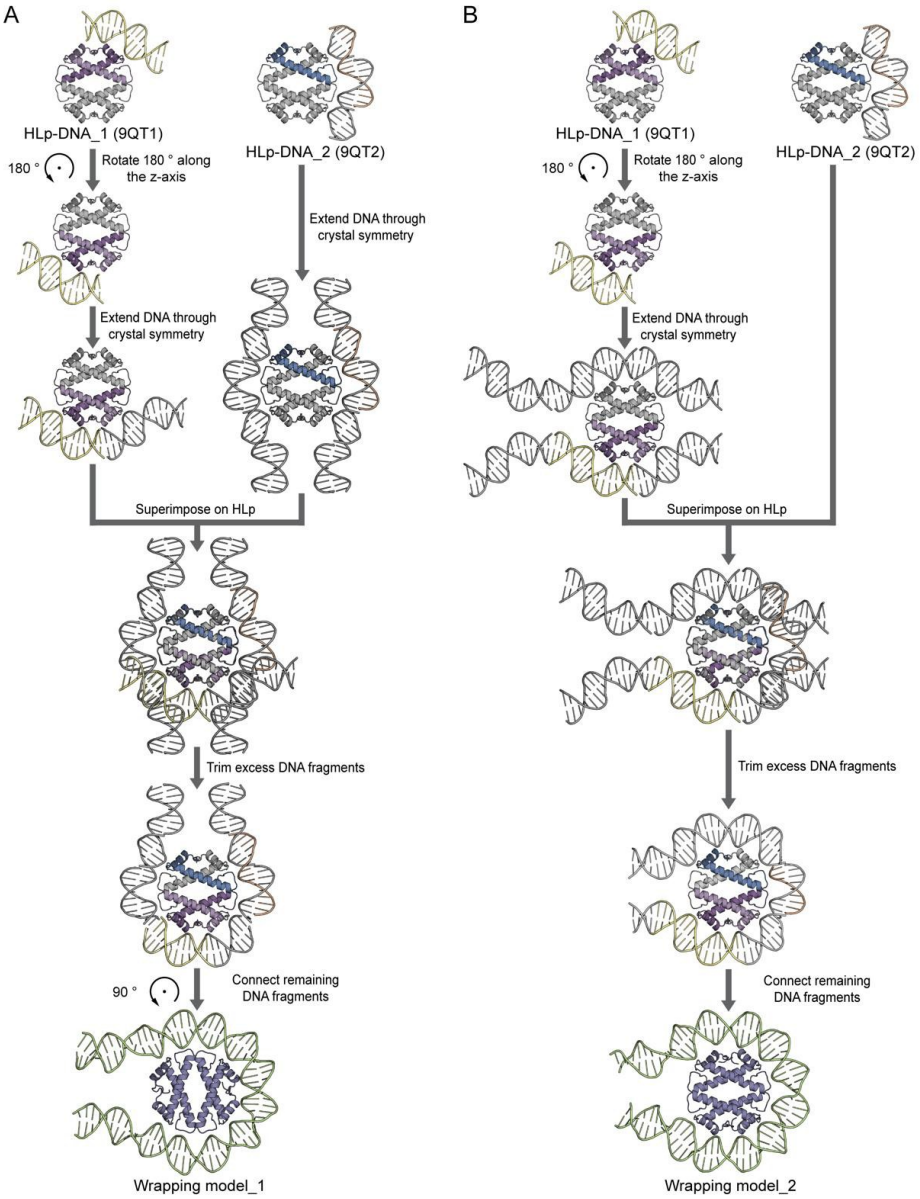
Crystal structure of three HMfB dimers in spiral arrangement (PDB: 5T5K) in cartoon representation. Residues involved in oligomerization are shown as sticks and the salt bridges formed between HMfB dimers are indicated as dashed lines.



Supplementary Figure S5

Binding of HLp to DNA fragments of variable GC content.

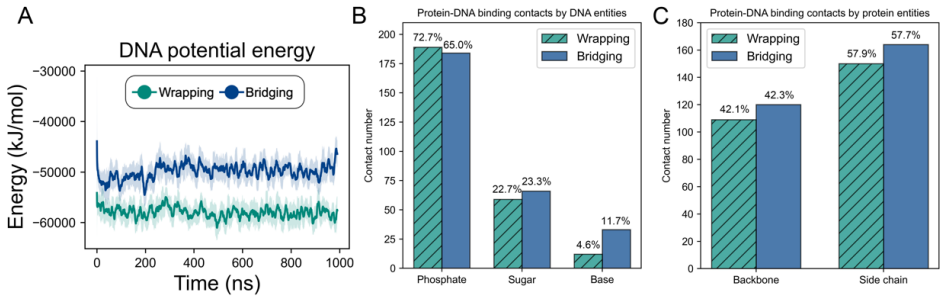
EMSAs showing binding of HLp to DNA fragments 30-bp-GC30 (**A**), 30-bp-GC40 (**B**), 30-bp-GC50 (**C**), and 30-bp-GC60 (**D**). The molar protein to DNA ratios loaded in lanes 4-10 are indicated.



Supplementary Figure S6

The workflow to generate starting models for wrapping mode.

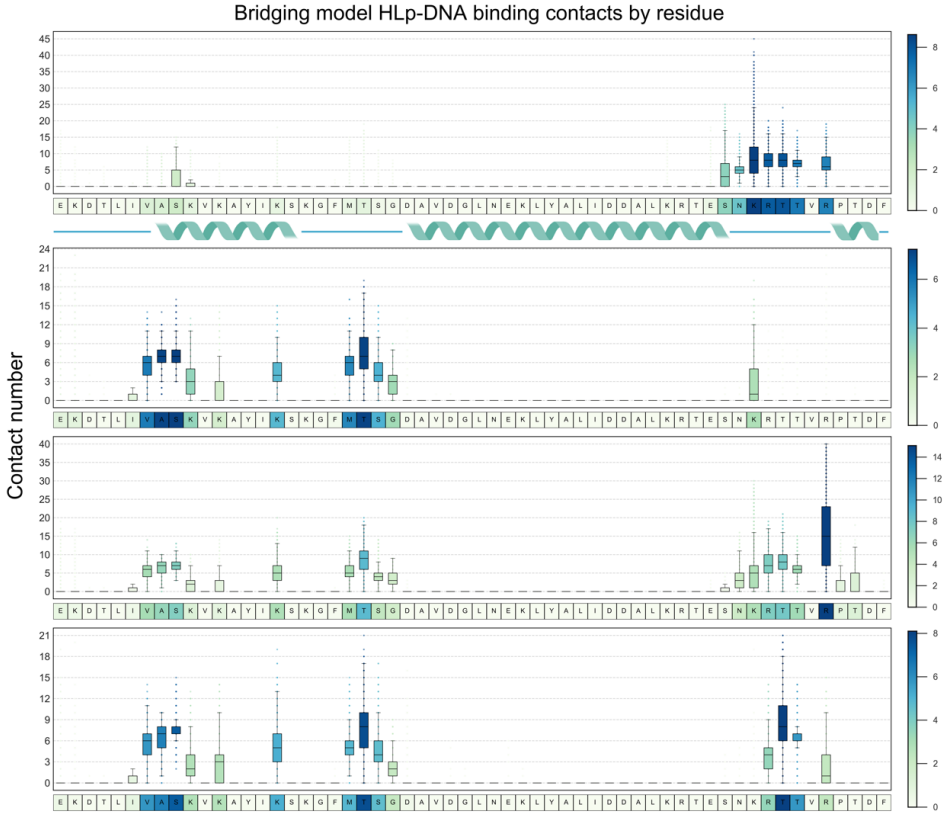
Workflows, demonstrating how the starting models — wrapping model_1 (A) and wrapping model_2 (B) — were constructed. The crystal structures of HLp-DNA_1 (PDB: 9QT1) and HLp-DNA_2 (PDB: 9QT2) depict the contents of a single asymmetric unit in color, with selected symmetry mates shown in gray. To enhance visualization, the HLp-DNA_1 structure was rotated 180° along the z-axis before extending the DNA fragments through crystal symmetry.



Supplementary Figure S7

DNA potential energy and protein-DNA interactions in the wrapping and the bridging Model.

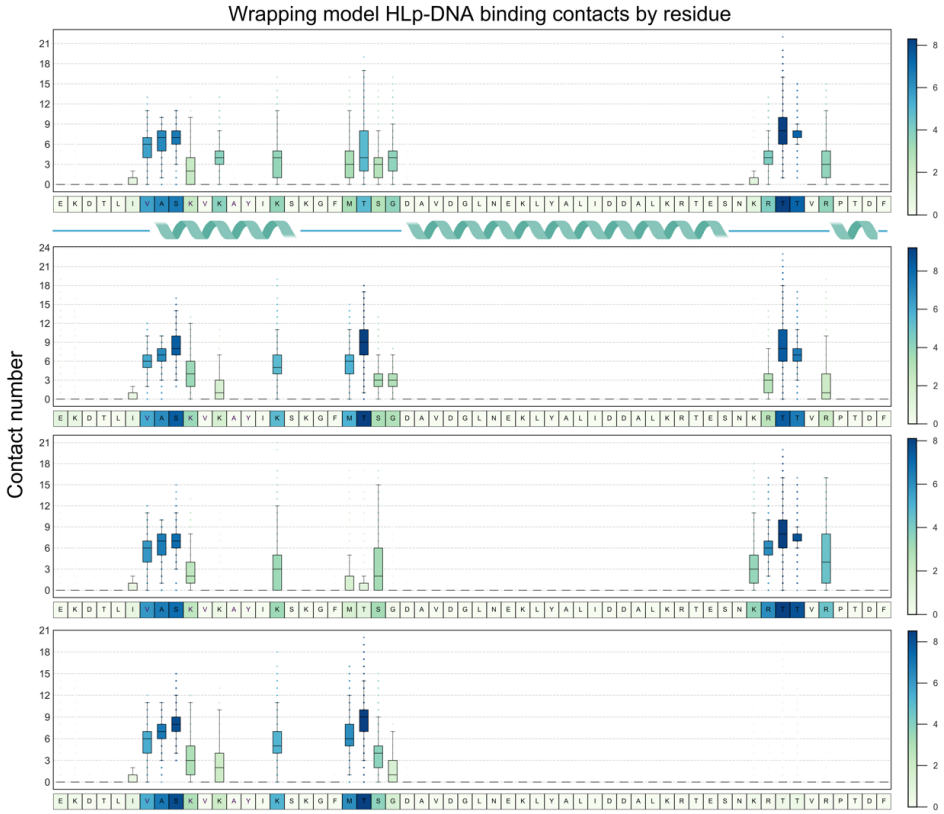
A. DNA potential energy plotted throughout the simulations. **B.** Average number of protein-DNA atom-atom contacts classified by DNA interaction entities (phosphate groups, sugars, and bases). **C.** Average number of protein-DNA atom-atom contacts classified by protein interaction entities (backbones and side chains). In all panels, the colors green and blue correspond to the wrapping model and bridging model, respectively.



Supplementary Figure S8

Residue-wise binding contact analysis in the bridging model.

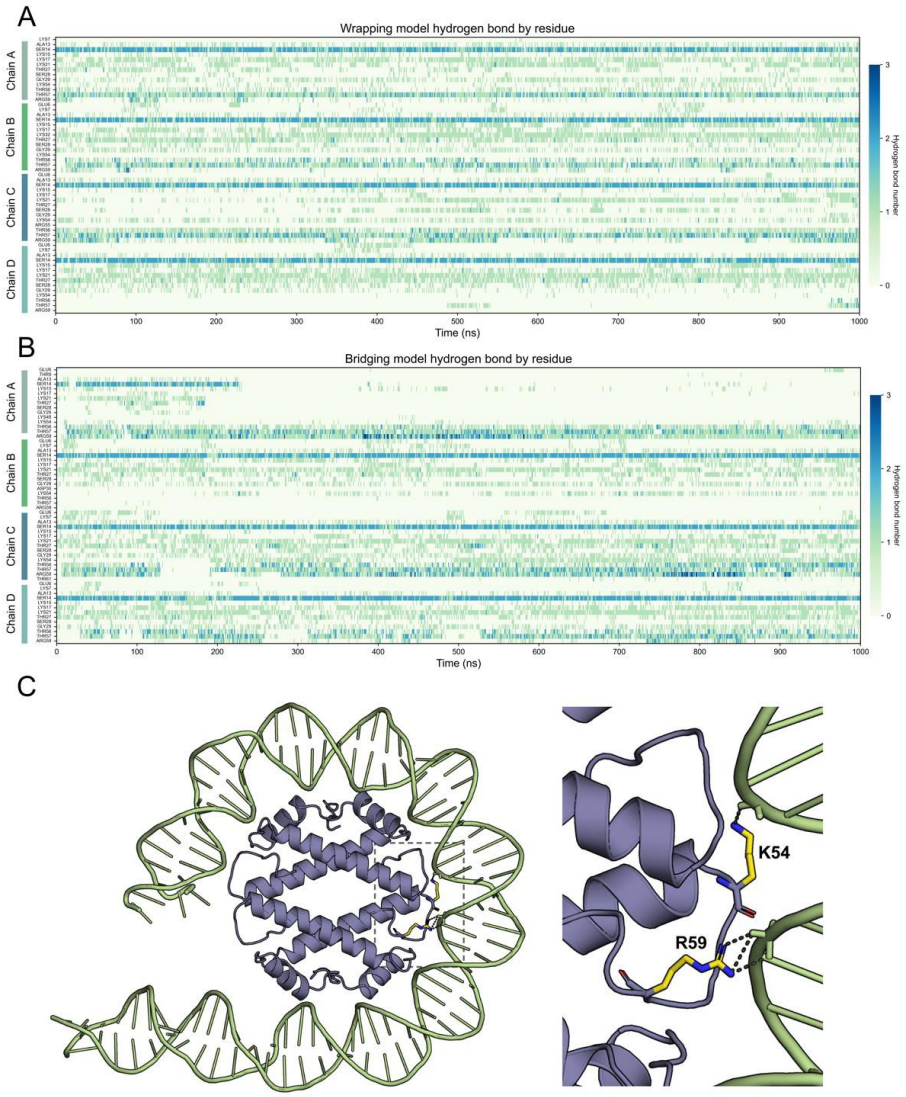
Residue-wise boxplots depicting binding contact numbers for the bridging model. Chains A, B, C, and D are presented sequentially from top to bottom. Darker blue shades indicate higher contact numbers. Below each boxplot, the corresponding amino acid sequence is shown, with a secondary structure diagram provided beneath the first sequence for reference.



Supplementary Figure S9

Residue-wise binding contact analysis in the wrapping model.

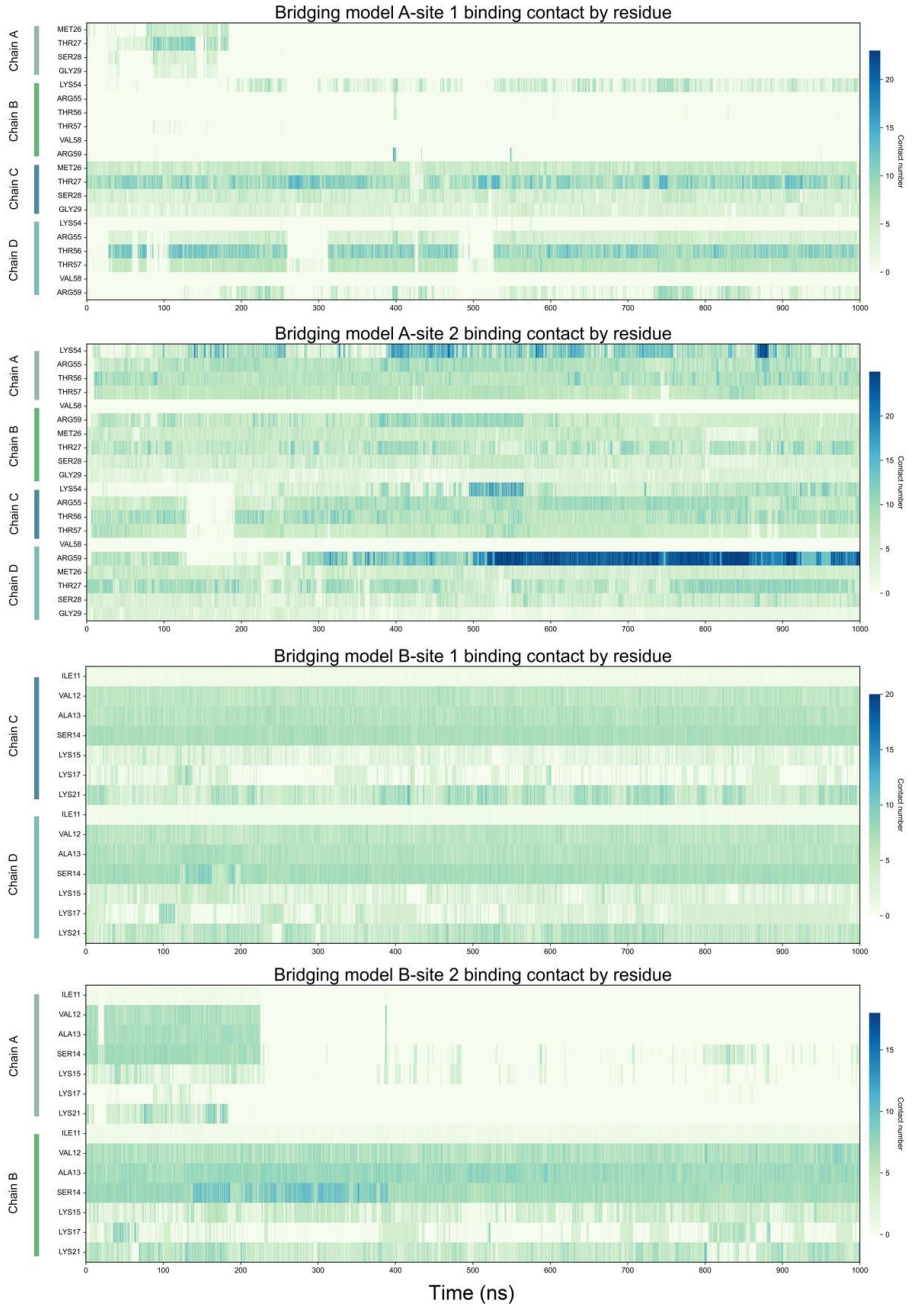
Residue-wise boxplots depicting binding contact numbers for the wrapping model. Chains A, B, C, and D are presented sequentially from top to bottom. Darker blue shades indicate higher contact numbers. Below each boxplot, the corresponding amino acid sequence is shown, with a secondary structure diagram provided beneath the first sequence for reference.



Supplementary Figure S10

Protein-DNA hydrogen bond analysis of the wrapping and the bridging models.

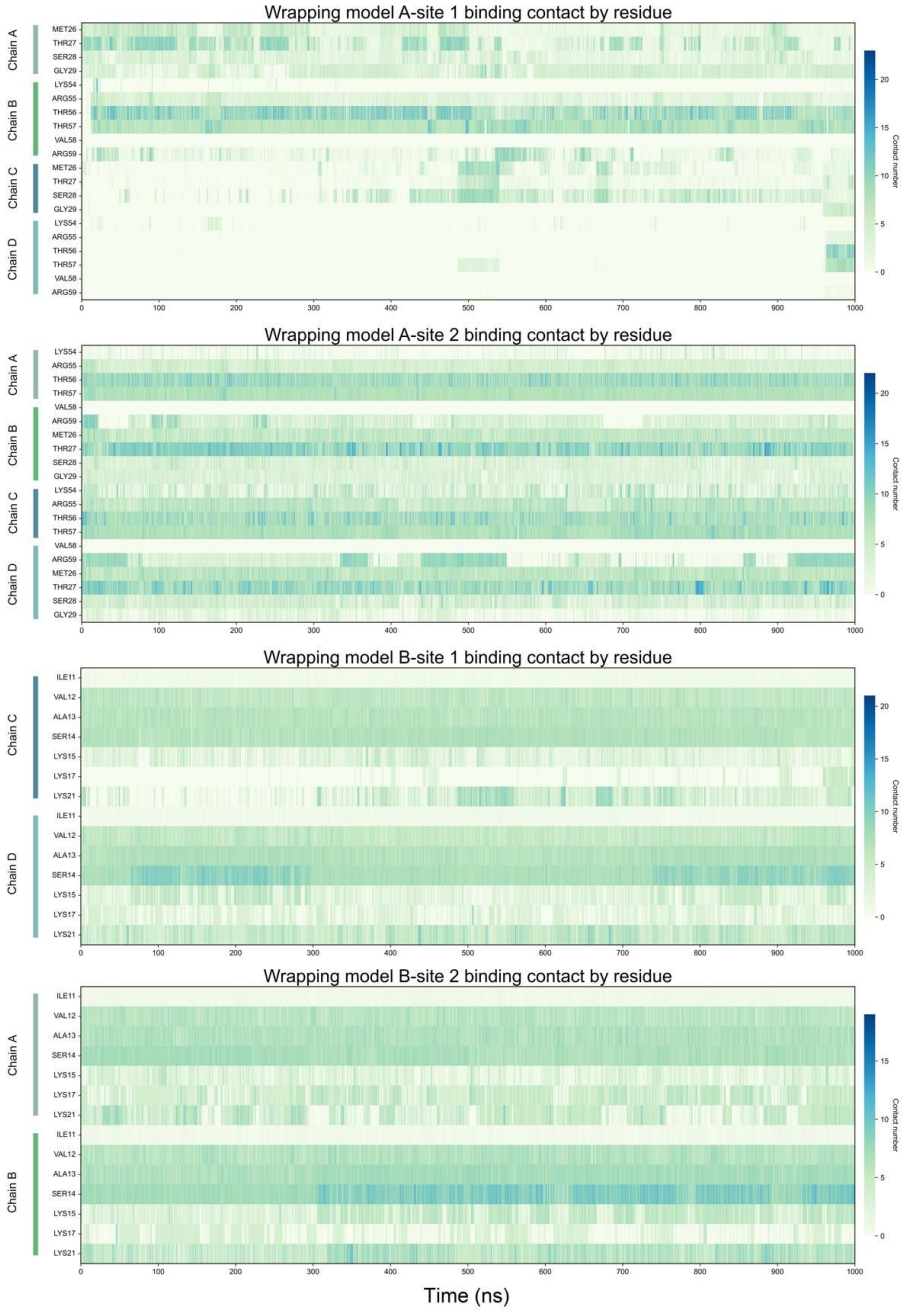
Heatmaps depicting the number of protein-DNA hydrogen bonds for each residue in the wrapping (**A**) and bridging model (**B**) throughout the simulation. The chain identifier of each residue is labeled on the left, with darker blue colors indicating a higher number of hydrogen bonds. **C**. Cartoon representation highlighting structural details of residues K54 and R59 in one monomer of the wrapping model, captured at the 150 ns snapshot of the HLp-DNA simulation. Both residues are shown as sticks with yellow-colored side chains, and hydrogen bonds between residues and DNA are represented by dashed lines.



Supplementary Figure S11

Protein-DNA contact analysis across binding sites in the bridging model.

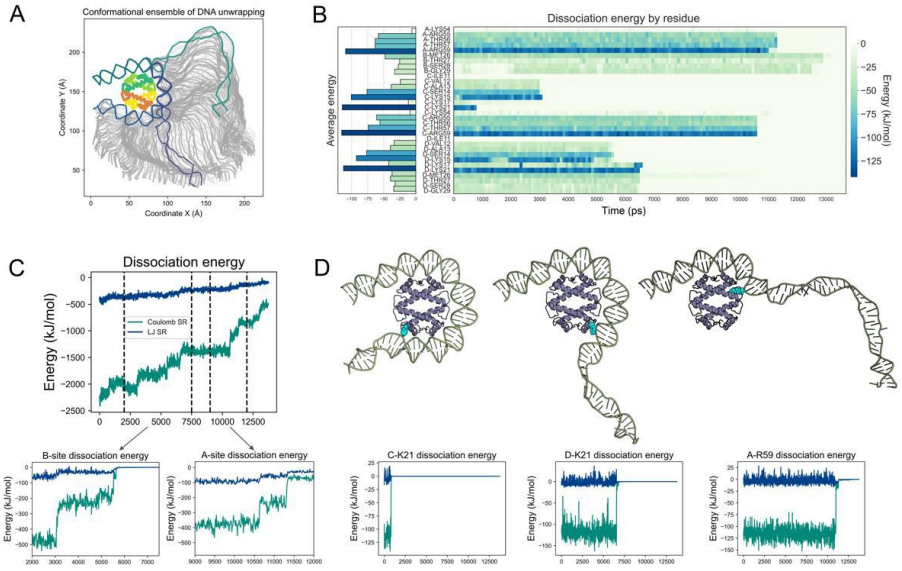
Heatmap illustrating the protein-DNA contact number for each residue across the four binding sites in the bridging model throughout the simulation. Darker blue colors indicate higher contact numbers. The four binding sites — A-site 1, A-site 2, B-site 1, and B-site 2 — are arranged sequentially from top to bottom, with the chain identifier of each residue labeled on the left.



Supplementary Figure S12

Protein-DNA contact analysis across binding sites in the wrapping model.

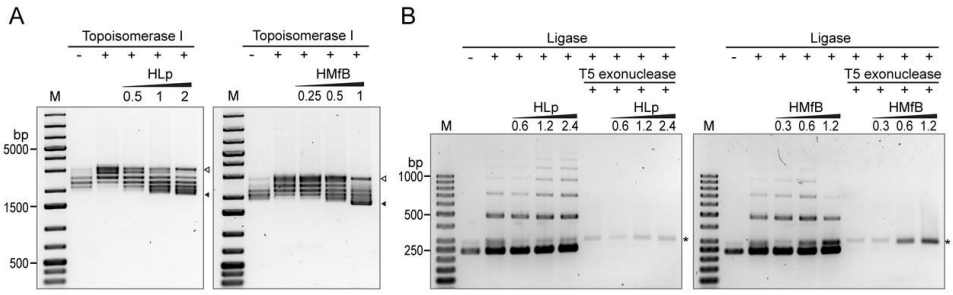
Heatmap illustrating the protein-DNA contact number for each residue across the four binding sites in the bridging model throughout the simulation. Darker blue colors indicate higher contact numbers. The four binding sites — A-site 1, A-site 2, B-site 1, and B-site 2 — are arranged sequentially from top to bottom, with the chain identifier of each residue labeled on the left.



Supplementary Figure S13

DNA unwrapping simulation.

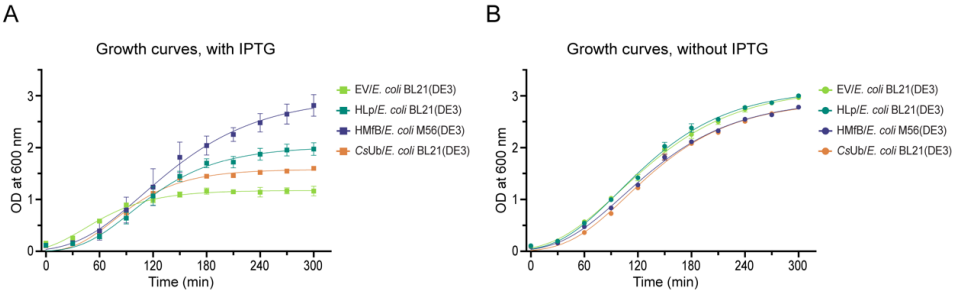
A. This panel shows the conformational ensemble of DNA throughout the simulation. The two ssDNA strands are colored light gray and dark gray, respectively. The HLp tetramer is colored by chains. **B.** Heatmap illustrating the energy profile of each residue involved in A-site 2 and B-site 2 during the unwrapping simulation. Darker blue shades indicate lower energy. The average energy of each residue, excluding frames where the residue is dissociated (with an energy of 0 kJ/mol), is shown as a bar plot on the left. **C.** Energy profile of the HLp-DNA system throughout the simulation, with blue and green lines representing short-range Lennard-Jones (LJ SR) potential energy and short-range Coulomb energy (Coulomb SR), respectively. The dissociation energy of the two binding sites involved in the simulation, B-site 2 and A-site 2, is presented below. **D.** Dissociation energy profiles of three residues with low energy: Chain D-K15, Chain C-R59, and Chain B-K17. These residues are shown as spheres and colored in cyan.



Supplementary Figure S14

HLp binding changes DNA topology.

A. DNA topology assay with relaxed pUC19 plasmid DNA in the presence of HLp and HMfB. Protein to DNA mass ratios are indicated. The bands corresponding to relaxed (△) and supercoiled pUC19 (▲) are labelled. **B.** Ligase-mediated circularization assay with the 240-bp-GC40 DNA and HLp or HMfB. Samples are shown before and after T5 exonuclease digestion. The ratio of protein to DNA mass is labelled. Circularized monomeric DNA is marked (*).



Supplementary Figure S15

Fig. S13 Growth curves of *E. coli* expression strains.

E. coli strains were transformed with empty pET-30a(+) (EV), HLp in pET-30a(+), *Cs*Ub in pET-30a(+), and HMfB in pET-28a(+). Growth curves were recorded in triplicate, with protein expression induced (**A**) and uninduced (**B**) by IPTG.