

1 **Histone-organized chromatin in bacteria**

2

3 Antoine Hocher<sup>1,2#\*</sup>, Shawn P. Laursen<sup>3#</sup>, Paul Radford<sup>4</sup>, Jess Tyson<sup>4</sup>, Carey Lambert<sup>4</sup>,

4 Kathryn M Stevens<sup>1,2</sup>, Mathieu Picardeau<sup>5</sup>, R. Elizabeth Sockett<sup>4</sup>, Karolin Luger<sup>6,7\*</sup>, Tobias

5 Warnecke<sup>1,2\*</sup>

6

7 <sup>1</sup>Medical Research Council London Institute of Medical Sciences, London, United Kingdom

8 <sup>2</sup>Institute of Clinical Sciences, Faculty of Medicine, Imperial College London, London,

9 United Kingdom

10 <sup>3</sup>Department of Molecular, Cellular, and Developmental Biology, University of Colorado

11 Boulder, Boulder, United States

12 <sup>4</sup>School of Life Sciences, Medical School, Queen's Medical Centre, University of

13 Nottingham, Nottingham, United Kingdom

14 <sup>5</sup>Institut Pasteur, Université Paris Cité, CNRS UMR 6047, Biology of Spirochetes Unit, Paris,

15 France

16 <sup>6</sup>Department of Biochemistry, University of Colorado Boulder, Boulder, United States

17 <sup>7</sup>Howard Hughes Medical Institute, Chevy Chase, United States

18

19 #joint first authors

20 \*corresponding authors ([a.hocher@lms.mrc.ac.uk](mailto:a.hocher@lms.mrc.ac.uk), [karolin.luger@colorado.edu](mailto:karolin.luger@colorado.edu),

21 [tobias.warnecke@lms.mrc.ac.uk](mailto:tobias.warnecke@lms.mrc.ac.uk))

22

1 **ABSTRACT**

2 Histones are the principal constituents of chromatin in eukaryotes and most archaea, while  
3 bacteria generally rely on an orthogonal set of proteins to organize their chromosomes.  
4 However, several bacterial genomes encode proteins with putative histone fold domains.  
5 Whether these proteins are structurally and functionally equivalent to archaeal and eukaryotic  
6 histones is unknown. Here, we demonstrate that histones are essential and are major  
7 components of chromatin in the bacteria *Bdellovibrio bacteriovorus* and *Leptospira*  
8 *interrogans*. Patterns of sequence evolution suggest important roles in several additional  
9 bacterial clades. Structural analysis of the *B. bacteriovorus* histone (Bd0055) dimer shows  
10 that histone fold topology is conserved between bacteria, archaea, and eukaryotes. Yet,  
11 unexpectedly, Bd0055 binds DNA end-on and forms a sheath of tightly packed histone  
12 dimers to encase straight DNA. This binding mode is in stark contrast to archaeal, eukaryotic,  
13 and viral histones, which invariably bend and wrap DNA around their outer surface. Our  
14 results demonstrate that histones are integral chromatin components across the tree of life and  
15 highlight organizational innovation in the domain Bacteria.

16

17

18

## 1 INTRODUCTION

2 The genomes of all eukaryotes are organized by nucleosomes, which are composed of  
3 four core histones that assemble into an octamer to wrap DNA in two tight superhelical turns  
4 (Luger *et al.* 1997). This arrangement severely restricts access to the genome. Consequently,  
5 eukaryotes rely on a complex system of access control to coordinate DNA-templated  
6 processes such as transcription, replication, and DNA repair (Kornberg and Lorch 2020).

7 Histones are among the most conserved and abundant proteins across eukaryotes  
8 (Postberg *et al.* 2010; Patwal *et al.* 2021; Soo and Warnecke 2021). The ‘histone fold’ is  
9 composed of three alpha helices (connected by two short loops) and dimerizes in a head-to-  
10 tail ‘handshake motif’ (Arents *et al.* 1991). Additional structural elements, including  
11 divergent N-terminal tails, distinguish the four eukaryotic core histones from each other  
12 (Luger and Richmond 1998a; b). Deletion or depletion of individual histones leads to  
13 transcriptional dysregulation, cell cycle arrest, and, ultimately, cell death (Han *et al.* 1987;  
14 Kim *et al.* 1988; Wyrick *et al.* 1999; Gossett and Lieb 2012).

15 Minimalist histones consisting only of the histone fold are pervasive in archaea,  
16 where they can play a major role in chromatin organization alongside other nucleoid-  
17 associated proteins (NAPs) (Henneman *et al.* 2018; Talbert *et al.* 2019; Laursen *et al.* 2021;  
18 Hocher *et al.* 2022). In some archaea, histones rank among the most abundant proteins in the  
19 cell (Hocher *et al.* 2022) and their presence is required for viability in the model archaeon  
20 *Thermococcus kodakarensis* (Čuboňová *et al.* 2012). Like their eukaryotic counterparts,  
21 archaeal histone dimers bend DNA around their outer surface by contacting three consecutive  
22 minor grooves of DNA. Three independent DNA interaction interfaces are formed by main  
23 and side chains of paired L1-L2 loops and the central  $\alpha 1$ - $\alpha 1$  interface (Luger and Richmond  
24 1998a; Mattioli *et al.* 2017). Whereas eukaryotic histones are obligate heterodimers, model  
25 archaeal histones such as HTkA/B from *T. kodakarensis* and HMfA/B from *Methanothermus*  
26 *fervidus* (henceforth HMf/HTk histones) can form homo- as well as heterodimers (Sandman  
27 *et al.* 1994; Stevens *et al.* 2020). These dimers then oligomerize into a histone stack of  
28 variable size that wraps DNA to form slinky-like ‘hypernucleosomes’ (Maruyama *et al.*  
29 2013; Mattioli *et al.* 2017; Bowerman *et al.* 2021).

30 Bacteria use a variety of small, basic NAPs to organize their DNA. The most  
31 widespread NAP is HU, but other abundant, lineage-restricted proteins are commonplace  
32 (Luijsterburg *et al.* 2008; Dillon and Dorman 2010). Deleting individual NAPs is often not

1 lethal to bacterial cells. This is true even for NAPs that, like HU in *E. coli*, are major  
2 constituents of the nucleoid (Baba *et al.* 2006). Histones are generally considered to be absent  
3 from bacteria. However, homology searches have revealed the presence of proteins with  
4 putative histone fold domains in a small and eclectic assortment of bacterial genomes (Alva  
5 and Lupas 2018). The role of these proteins in organizing bacterial chromatin is unknown.  
6 Here, we show that histone proteins are major and essential nucleoid components in the  
7 bacterium *Bdellovibrio bacteriovorus*. In vitro, these histones completely coat linear DNA by  
8 binding end-on, forming a nucleohistone filament. In all other known histone-DNA  
9 complexes, DNA is wrapped around the outside of a helical ramp of histones. As such, this  
10 bacterial ‘inside-out’ arrangement presents a fundamental divergence from the established  
11 histone binding mode.

12

## 13 **RESULTS**

### 14 ***Histones are phylogenetically entrenched in several bacterial clades***

15 We conducted a systematic homology search for histone fold proteins in bacteria.  
16 Across a phylogenetically balanced database of 18,343 bacterial genomes, we found 416  
17 proteins that contain a histone fold domain (Table S1). 1.86% of genomes encode at least one  
18 histone fold domain, compared to 92.8% of genomes that encode HU. In agreement with  
19 prior work (Alva and Lupas 2018), we identified two major histone size classes, containing  
20 either a singlet or a doublet histone fold. Both classes are typically devoid of other  
21 recognized domains (Fig. 1a, Table S1). Like their archaeal homologs, most bacterial singlets  
22 also lack the long, disordered N-terminal tails that are characteristic of eukaryotic histones  
23 (Henneman *et al.* 2018; Stevens and Warnecke 2022). Amino acid conservation across the  
24 three domains of life is particularly high in the L2 loop (RKTV). In archaeal and eukaryotic  
25 histones, this region contacts DNA. The highly conserved ‘RD clamp’ stabilizes this  
26 arrangement. Bacterial singlets are six residues shorter than HMf/HTk histones, on average.  
27 This is mostly due to a shorter  $\alpha 2$  helix, which is diagnostic of bacterial singlet histones (Fig.  
28 1b, c). Several residues that are normally present in archaeal and eukaryotic histones,  
29 including the ‘sprocket arginine’ R19 (Hodges *et al.* 2015), are not found in bacterial histones  
30 (Fig. 1c). The N-terminus as a whole exhibits considerable divergence, is more hydrophobic  
31 in nature, and includes a conserved serine-lysine motif (S9-K10) that is not found in archaea.

32

## 1 ***Bdellovibrio bacteriovorus* encodes an abundant histone that localizes to the nucleoid**

2 The phyletic distribution of histones across the bacterial domain is patchy (Fig. 1d,  
3 Table S1). In some instances, this might indicate assembly contaminants or recent/transient  
4 horizontal gene transfer. However, we identified clades where histone fold proteins are  
5 present in several closely related sister lineages. Such phylogenetic persistence is particularly  
6 evident in the phylum Bdellovibrionota (Figs. 2a, S1) and the order Leptospirales (Fig. S2).

7 *Bdellovibrio bacteriovorus* HD100, the model organism of the Bdellovibrionota, is a  
8 bacterial predator with a biphasic, bacterially-invasive life cycle (Sockett 2009). Small,  
9 motile attack phase cells breach, enter through, and subsequently re-seal the outer membrane  
10 of other gram-negative prey bacteria (e.g. *E. coli*). From the periplasm, *B. bacteriovorus* then  
11 consumes its prey by secreting proteases and nucleases into the prey cytoplasm, culminating  
12 in cycles of replication and coordinated non-binary division to yield new attack phase cells,  
13 which are released from the husk of the ravaged prey following induced rupture of its wall  
14 and outer membrane (Negus *et al.* 2017; Harding *et al.* 2020).

15 *B. bacteriovorus* encodes two predicted singlet histones (Table S1): Bd0055 is a  
16 bacterial singlet (Fig. 1c) with a detectable homolog in 67% of Bdellovibrionota genomes  
17 (Figs. 2a, S1). It is highly conserved at the amino acid level (Fig. S3). Bd3044 is a longer,  
18 less well conserved protein that is present in only 37% of Bdellovibrionota genomes (Figs.  
19 S1,3). Prior transcriptomic data from across different stages of the *B. bacteriovorus* life cycle  
20 indicate high expression of Bd0055, especially during active replication in the host (Fig. S4).  
21 We therefore focus on Bd0055 as a candidate global organizer of the nucleoid.

22 We used quantitative label-free proteomics in attack phase cells to confirm high  
23 abundance of Bd0055 at the protein level (Fig. 2b). Of note, protein abundance in attack  
24 phase cells is better correlated with RNA abundance in growth phase, consistent with a time  
25 lag in protein production where transcript levels during growth phase foreshadow protein  
26 levels in attack phase cells (Fig. S4).

27 To investigate cellular localization of Bd0055 and to monitor expression throughout  
28 the *B. bacteriovorus* life cycle, we generated *B. bacteriovorus* HD100 strains in which the  
29 native copy of *bd0055* was replaced with a version that was C-terminally tagged with either  
30 mCherry or mCitrine. Both tagged strains exhibit no gross morphological defects and carry  
31 out predation efficiently. We detect strong fluorescence from tagged Bd0055 throughout the  
32 life cycle, including in free-swimming attack phase cells and following invasion of the *E. coli*

1 prey (Fig. 2c). Compared to a previously characterized mCherry-tagged control protein with  
2 known cytoplasmic localization (Willis *et al.* 2016; Makowski *et al.* 2019) the fluorescent  
3 signal emanating from Bd0055-mCherry is absent from the cell poles, consistent with  
4 localization to the nucleoid as captured by Hoechst staining (Fig. 2d-f). In line with the  
5 absence of a signal peptide, we find no evidence for secretion into the *E. coli* prey, suggesting  
6 that Bd0055 is unlikely to be used for manipulation of prey chromatin.

7 To obtain orthogonal *in vivo* confirmation that Bd0055 associates with the *B.*  
8 *bacteriovorus* nucleoid, we carried out sucrose gradient-based nucleoid enrichment  
9 experiments coupled to quantitative proteomics. We analyzed two fractions from sucrose  
10 gradients of lysed attack phase cells: a lower density fraction that is enriched for soluble  
11 cytosolic proteins and a nucleoid fraction, enriched for DNA and DNA-binding proteins.  
12 Subunits of the RNA polymerase are, as in other bacteria, very abundant (Fig. 2g). Bd0055 is  
13 highly represented and enriched 29-fold in the nucleoid fraction. Homologs of classic  
14 bacterial NAPs encoded in the *B. bacteriovorus* genome, including HU and Dps, are present  
15 but less abundant and less enriched in the nucleoid fraction (Fig. S5).

16

### 17 ***Bd0055 folds into a histone fold dimer***

18 We determined the crystal structure of Bd0055 at 1.8 Å resolution (Table S2).  
19 Bd0055 forms a crystallographic dimer where one monomer is related to its partner through  
20 two-fold symmetry (Fig. 3a). It exhibits the overall topology of a histone fold dimer, as well  
21 as a ‘basic ridge’ along its ‘top’ surface that is also found in archaeal and eukaryotic histones,  
22 where it is used for DNA binding (Luger *et al.* 1997; Mattioli *et al.* 2017) (Fig. 3b). The  
23 most pronounced structural difference between bacterial and archaeal/eukaryotic histones is  
24 the 7 Å shorter  $\alpha 2$  helix (1 helical turn shorter compared to archaeal HMfB, Fig. 1c),  
25 reducing the overall size of the dimer (Fig. 3a). In addition, the equivalent of the  $\alpha 3$  helix in  
26 Bd0055 forms only one helical turn and packs against  $\alpha 2$  of the dimerization partner. This is  
27 unlike the three-turn  $\alpha 3$  helix in other histones, which is engaged in tetramerization (Mattioli  
28 *et al.* 2017) by contributing to the four-helix bundle interface, interacting with a conserved  
29 histidine in  $\alpha 2$  (H49) that is conspicuously absent from all bacterial histones (Fig. 1c).  
30 Finally, the surface opposite the basic ridge of the Bd0055 dimer is more acidic than that of  
31 HMf/HTk histones (Fig. 3b).

32

## 1 ***Bd0055 binding to DNA differs from that of archaeal histones***

2 We showed that Bd0055 interacts with DNA *in vitro* using fluorescence polarization  
3 (FP) (Fig. 4a). Using gel electromobility shift assays, we previously demonstrated that the  
4 archaeal histone HTkA shifts 147 bp DNA to a single discrete band (Bowerman *et al.* 2021).  
5 HTkA saturates this DNA fragment at a ratio of ~10 to 1 and is unable to produce higher  
6 shifts with added protein. In contrast, Bd0055 shifts DNA to several regularly spaced bands  
7 indicative of multiple binding events (Fig. 4b). Increasing the protein/DNA ratio of Bd0055  
8 continues to decrease electrophoretic mobility, suggesting a different binding mode (Fig. 4b).

9 To test whether Bd0055 bends DNA into a nucleosome-like geometry, we assembled  
10 histones on 147 bp DNA fragments end-labeled with a Förster resonance energy transfer  
11 (FRET) donor-acceptor pair. Titration of archaeal HTkA onto this DNA brings the ends into  
12 FRET proximity by forming nucleosome-like structures, as observed previously (Mattioli *et al.*  
13 *al.* 2017). In contrast, no signal was observed upon titrating even a large excess of Bd0055  
14 (Fig. 4c). By monitoring fluorescence polarization of the same samples, we verified that both  
15 proteins bind DNA under these conditions (Fig. 4a). To further investigate this different  
16 behavior of Bd0055, we superimposed the Bd0055 histone structure onto an archaeal  
17 hypernucleosome consisting of four HMfB dimers and 118 bp of DNA (based on PDB  
18 5T5K) and carried out all-atom molecular dynamics simulations. While archaeal histone  
19 dimers remain stably stacked throughout the simulation, the modeled Bd0055  
20 hypernucleosome unfolds within a few nanoseconds (Fig. 4d, Movie S1). Although Bd0055  
21 dimers remain bound to DNA during the simulation, they no longer contact other dimers  
22 through protein-protein interactions, suggesting a failure to form stable tetramers on DNA. In  
23 archaeal and eukaryotic histones, a tetramer is the smallest unit capable of wrapping DNA.

24 To understand how Bd0055 binds DNA, we determined the crystal structure of  
25 Bd0055 in complex with a 35 bp DNA fragment (Table S2). This structure shows that  
26 Bd0055 contacts the DNA with only one L1-L2 binding interface across the minor groove  
27 (Fig. 5a). The molecular details of this edge-on interaction are very similar to those observed  
28 for other known histones. It employs the L1-L2 motif that is conserved across the domains of  
29 life, although the sprocket arginine that reaches into the minor groove in nearly all other  
30 histone-DNA complexes is missing (Fig. 1c). Surprisingly, rather than bending the DNA  
31 around its positively charged surface, the DNA maintains a straight trajectory. Additional  
32 histone dimers bind the next two phosphates through L1-L2 interfaces engaged in identical  
33 interactions, but flipped by 180 degrees (Fig. 5b). The filament is stabilized through protein-

1 protein interactions between histone dimers that involve the basic DNA binding ridge and the  
2 acidic ‘underside’ of histone dimers 1 and 5, and through exchanging N-terminal tails  
3 between dimers 1 and 4 (Fig. 5c). This nucleohistone filament, which requires a histone to  
4 DNA ratio of 1 histone dimer per 2.5 bp, reverses known histone convention by completely  
5 protecting straight DNA from the solvent. All other known histone-DNA complexes have  
6 histones on the inside and wrap DNA around them in a stoichiometry of 1 histone dimer per  
7 30 bp (Fig. 5b).

8

9 ***Why does Bd0055 bind DNA edge-on rather than wrapping it around its surface?***

10 Three main structural differences between Bd0055 and other histones could be  
11 responsible for this binding mode. First, the N-terminus promotes interactions between  
12 dimers on the same strand of DNA, potentially stabilizing the fiber structure. Second, the  
13 shorter  $\alpha 2$  helix of Bd0055 might force the DNA to bend too severely to wrap around the  
14 outside of the shortened dimer. Third, key residues responsible for tetramerization in archaeal  
15 histones are absent, most notably a histidine at the equivalent position of A48 (Figs. 1c, 5c).

16 To test if one of these differences is responsible for the inability of Bd0055 to wrap  
17 DNA, we made the compensating mutants *in vitro*, and assayed their ability to wrap 147 bp  
18 DNA using FRET. Neither deletion of the four N-terminal amino acids nor insertion of four  
19 amino acids into  $\alpha 2$  of Bd0055 produce a FRET signal between DNA ends. In contrast,  
20 mutations to mimic the archaeal tetramerization domain (A48H, S45F, I61L) enable Bd0055  
21 to bring the ends of the DNA within FRET distance (Fig. 5d). None of these amino acids are  
22 implicated in DNA binding in either binding mode and their mutation does not impair the  
23 ability of Bd0055 to bind DNA (Fig. S6). The magnitude of the FP signal observed for the  
24 tetramerization mutant binding 147 bp DNA closely matches that of HTkA, and the shift to  
25 the left indicates the higher affinity of Bd0055 for DNA compared to HTkA (Fig. 5d). These  
26 data suggest that the inability of wildtype Bd0055 to form tetramers is responsible for its  
27 inability to wrap DNA around its outer perimeter as is observed in archaeal and eukaryotic  
28 histones. This is consistent with prior findings from HMfB mutagenesis and analysis of a  
29 small set of archaeal histones that also have a non-histidine residue at position 49 and a  
30 congruent inability to form stable tetramer interfaces (Marc *et al.* 2002; Stevens *et al.* 2020).

31

32



## 1 ***Bd0055 is essential for predatory and prey-independent growth***

2 We sought to delete *bd0055* using an established silent gene deletion approach to  
3 investigate its physiological role (Harding *et al.* 2020; Banks *et al.* 2022). Attempts to delete  
4 *bd0055* in the wild type HD100 background resulted in no deletion strains and 133 reversions  
5 to wild type. We also attempted to delete *bd0055* in a host-independent strain, HID13  
6 (Lambert *et al.* 2010), to establish whether impaired predation was the reason that no  
7 deletions were obtained in HD100. These attempts resulted in no successful deletions and  
8 150 reversions to wild type. This suggests that *bd0055* is essential during both predatory and  
9 prey-independent growth.

10

## 11 ***Bacterial histones have essential roles outside the Bdellovibrionota***

12 Is *B. bacteriovorus* unique among bacteria in using histones to organize chromatin?  
13 We examined publicly available gene expression data and generated additional  
14 transcriptomic and proteomic data. At both the transcript and protein level, we find high  
15 expression of histone-fold proteins in several distantly related bacteria (Fig. 6a). We also  
16 observe notable phylogenetic persistence in clades where gene expression profiles are not  
17 available, including in the Planctomycetota and Elusimicrobia (Figs. S7-10).

18 We became particularly curious about *Leptospira interrogans*, the causative agent of  
19 leptospirosis, whose histone (Uniprot ID: A0A2H1XGH2 gene ID: LA\_2458) is  
20 exceptionally abundant (ranked 4 of 1502 quantified proteins; Figs. 6a, S11). Echoing  
21 findings for *Bd0055*, we find this histone to be highly conserved at the amino acid level with  
22 homologs present in all members of the order Leptospirales, including saprophytic and  
23 pathogenic species (Figs. 6b, S2,12). Nucleoid enrichment experiments followed by mass  
24 spectrometry confirmed high abundance and strong nucleoid enrichment (Fig. 6c). Finally,  
25 several attempts to delete the gene encoding this histone from the *L. interrogans* genome  
26 failed, indicating that histones are also essential in this species.

27

28

## 29 **DISCUSSION**

30 Our results demonstrate that histone-based chromatin organization is not exclusive to  
31 eukaryotes and archaea. The bacterial histone *Bd0055* from *B. bacteriovorus* HD100 is  
32 structurally similar to its archaeal and eukaryotic homologs, but binds DNA edge-on and

1 oligomerizes *in vitro* to form a dense nucleohistone fiber that completely encases DNA. This  
2 is a striking departure from all histones known to date, which wrap DNA around their  
3 exterior to form (hyper)nucleosomes, leaving DNA partially accessible. Our data show that  
4 the inability of Bd0055 to form tetramers is responsible for its inability to wrap DNA around  
5 its outside. The fact that the tetramerization domain is also absent from all other bacterial  
6 singlets (Fig. 1c) further suggests that the inability to wrap might be a general characteristic  
7 of bacterial histones.

8 Filamentation of Bd0055 on DNA is further promoted by complementary charge  
9 interactions and hydrogen bonds between individual DNA-bound dimers. Nucleohistone  
10 filament formation observed here demands a histone to DNA ratio of 1 histone dimer per  $\sim 3$   
11 bp of DNA, 10-fold higher than the 1 histone dimer per  $\sim 30$  bp of DNA observed in  
12 canonical nucleosomes and hypernucleosomes. Because crystals were set up at  
13 concentrations assuming the much lower stoichiometry, it is unlikely that the nucleohistone  
14 filament is an artifact of excessive histones in solution.

15 The crystal lattice indicates the ability of histones to generate higher order fiber  
16 packing through the interaction of histones on adjacent nucleohistone filaments (Movie S2).  
17 Although interactions between filaments do not appear to be very strong, it is worth pointing  
18 out that one of the three DNA binding regions in the Bd0055 dimer is available for  
19 interaction with DNA. Tomograms from *B. bacteriovorus* and *L. interrogans* suggest  
20 occasional close packing of DNA fibers *in vivo* (Butan *et al.* 2011; Raddi *et al.* 2012), but  
21 whether these are nucleohistone filaments remains to be shown.

22 The degree of compaction of the *B. bacteriovorus* nucleoid varies throughout its life  
23 cycle. Attack phase cells have highly compacted nucleoids that cannot be penetrated even by  
24 small fluorescent proteins (Kaljević *et al.* 2021). The nucleoid decondenses shortly after  
25 entry into host cells, and a less compact nucleoid state persists until the *B. bacteriovorus* cell  
26 is ready to septate. At this time, the newly replicated genomes of each daughter cell again  
27 become discretely packaged into highly compacted nucleoids (Borgnia *et al.* 2008). Bd0055  
28 nucleohistone filaments could contribute to nascent genome segregation and the management  
29 of extreme compaction during the attack phase. This process could be controlled by changing  
30 protein levels, variations in ionic strength, or post-translational modifications.

31 Histones likely play a role in many other bacteria. Our work shows that histone  
32 sequences exist in diverse organisms across the domain Bacteria, are often highly abundant,  
33 and are essential in at least one other organism, *L. interrogans*. Like other spirochaetes, *L.*

1 *interrogans* has a long, slender filamentous shapes, not dissimilar to the shape of *B.*  
2 *bacteriovorus* during predatory growth inside prey, prior to septation. Leptospira histones  
3 share key features with Bd0055, such as the shorter  $\alpha 2$  helix, the conserved L2 loop, and the  
4 absence of the histidine responsible for tetramerization in archaea. More work is needed to  
5 discover whether other bacterial histones bind DNA in in the same manner as Bd0055, when  
6 and how these sequences were acquired, and what function other bacterial histones that  
7 encode doublets or attached domains serve.

8

## 9 **METHODS**

10

### 11 **Homology survey of histone fold proteins**

12

13 PFAM HMM models of known histone fold domains (Histone (PF00125),  
14 CBF<sub>D</sub>\_NFYB\_HMF (PF00808) and DUF1931 (PF09123)) were searched against a  
15 phylogenetically diverse database of 18,343 bacterial genomes (Table S3) using hmmsearch  
16 (v 3.1b2) with the `-cut_ga` option to ensure reproducibility. In addition, a list of eight  
17 prokaryotic histone fold seed sequences were obtained from (Alva and Lupas 2018) and used  
18 to direct homology search with Jackhmmer (v 3.1b2). Jackhmmer results with p-values  $<1e-3$   
19 were kept. Hits were combined and proteins longer than 200 amino acids discarded. All  
20 proteomes in the database were obtained from NCBI (accessed on 20 May 2020). When  
21 GenBank proteome files were not available, proteomes were predicted using Prodigal  
22 (v2.6.3), with default settings.

23

24 For the analysis of  $\alpha 2$  length (Fig. 1b), we considered curated sets of histone fold  
25 proteins to facilitate the systematic identification of  $\alpha 2$ . Bacterial singlets were filtered from  
26 the bulk of histone fold proteins by removal of DUF1931 HMM hits and further removal of  
27 proteins that were longer than 65 amino acids. HMf/HTk-type archaeal histones were pruned  
28 from a phylogenetic tree based on the alignment of bacterial and archaeal histone-fold  
29 proteins (K. Stevens). Secondary structures were then predicted for each histones using  
30 Jpred4 (Drozdetskiy *et al.* 2015) and the length of  $\alpha 2$  calculated based on those predictions.  
31 The  $\alpha 2$  was taken to be the helix that overlapped the peptide region L28-L32 (in HMfB  
32 coordinates, see Fig. 1c) and was identified as the longest helix in the protein in 1,439 out of  
33 1,458 cases. Outliers in  $\alpha 2$  length (see Fig. 1b) were manually scrutinized and are owing to

1 secondary structure mispredictions, often splitting  $\alpha 2$  up into two smaller helices. Removal  
2 of these outliers does not affect conclusions.

3

#### 4 **Protein alignment and phylogenetic trees**

5

6 All species trees were obtained from GTDB (<https://gtdb.ecogenomic.org/>). Trees were  
7 rendered by iTol and finalized in Adobe Illustrator.

8

#### 9 **Best reciprocal blasts**

10

11 Best reciprocal blast hits, used to compute amino acid conservation (Figs. S3,12) were  
12 obtained using the `blast_best_reciprocal_hit` function from the R library `metablastr` (v 0.3),  
13 requiring a minimal E-value of 0.001.

14

#### 15 **Bacterial culture**

16

17 *B. bacteriovorus* HD100 was cultured on double-layer YPSC plates to initialize growth from  
18 frozen stocks and subsequently grown in liquid calcium/HEPES buffer containing *E. coli*  
19 S17-1 as described previously (Lambert and Sockett 2008). The kanamycin-resistant *E. coli*  
20 S17-1 (pZMR100) strain was used as prey for the culture of kanamycin-resistant *B.*

21 *bacteriovorus* strains. Media were supplemented with kanamycin (50  $\mu\text{g/ml}$ ) when required.

22 The host-independent *B. bacteriovorus* strain HID13 was grown in YP medium at 30°C as  
23 described previously (Lambert *et al.* 2010).

24

25 *L. interrogans* was cultured at 30°C in EMJH medium to a density of  $10^7$  bacteria/ml. Cells  
26 were harvested via centrifugation at  $4,000 \times g$  for 20 min and the pellets were washed with  
27 Phosphate Buffered Saline (PBS). The resulting pellets were frozen at -80°C until use.

28

#### 29 **Whole cell extract preparation**

30

31 Total proteins were extracted from ~5 mg of -80°C frozen, unfractionated pellet using the iST  
32 proteomic kit (see below). Experiments were carried in biological duplicates at the protein

1 extraction step, and each biological replicate was itself technically duplicated at the  
2 proteomic step.

3

#### 4 **Nucleoid enrichment and protein purification**

5

6 Nucleoid enrichment was carried out as described in (Ohniwa *et al.* 2011) with the following  
7 modifications: Lysozyme concentration in buffer B was doubled for *L. interrogans*;  
8 Lysozyme incubation was carried out on ice for 10 min for *B. bacteriovorus* and at room  
9 temperature for 10 min for *L. interrogans*. 10 to 60% sucrose gradients were poured  
10 manually by 2 mL of 10% increment in Ultra-Clear (14x89 mm) Beckman-Coulter tubes.  
11 Gradients were allowed to cool down to 4°C for before the experiment. Lysed cells were  
12 deposited on gradients and centrifuged at 10,000 rpm on a SW41-Ti Beckman-Coulter rotor  
13 by a Beckman Optima centrifuge pre-cooled at 4°C. Acceleration was set to their minimum  
14 value for both the start and the end of the run. Soluble cytosolic proteins settle at low density  
15 towards the top of the tube (“top fraction”) whereas the nucleoid, along with membrane  
16 proteins (Portalier and Worcel 1976; Murphy and Zimmerman 1997), will settle at a higher  
17 density and can be identified as an opaque, viscous band (“nucleoid fraction”). Proteins from  
18 both fractions were concentrated and purified following a methanol chloroform treatment as  
19 described in (Hocher *et al.* 2022). Samples were then prepared for mass spectrometry using  
20 the iST preomics kit, as described below. Experiments were carried out in biological  
21 triplicates at the protein extraction step, and each biological replicate was itself technically  
22 duplicated at the proteomic step.

#### 23 **Protein preparation for mass spectroscopy**

24 Proteins from whole cell extract and from different nucleoid enrichment fractions were all  
25 processed using the iST 8x Preomics kit. The sonication step was carried out as  
26 recommended by the manufacturer on a Bioruptor Plus (high intensity setting). Following  
27 heat denaturation at 95°C on an Eppendorf Thermomixer C and sonication, total protein  
28 amount was estimated using 205 nm absorbance by Nanodrop (method scope 31). A total of  
29 100 µg for WCE and 37.5 (36.5) µg for *B. bacteriovorus* (*L. interrogans*) sucrose fractions  
30 were used subsequently. Enzymatic digestion (LysC/Trypsin) was carried out on a Eppendorf  
31 Thermomixer C for 90 min at 37°C 500 rpm. For each experiment, all samples were  
32 processed using the same kit on the same day until storage at -80°C in LC-load buffer.

## 1 **Liquid chromatography-tandem mass spectrometry (LC-MS/MS)**

2

3 Chromatographic separation was performed using an Ultimate 3000 RSLC nano liquid  
4 chromatography system (Thermo Scientific) coupled to a Thermo Scientific LTQ Orbitrap  
5 Velos (*B. bacteriovorus* nucleoid enrichment samples) or an Orbitrap Q-Exactive mass  
6 spectrometer (all other proteomics samples) via an EASY-Spray source.

7

8 For sample analysis on the Velos, peptide solutions were injected and loaded onto a trapping  
9 column (Acclaim PepMap 100 C18, 100  $\mu\text{m} \times 2\text{ cm}$ ) for desalting and concentration at 8  
10  $\mu\text{L}/\text{min}$  in 2% acetonitrile, 0.1% TFA. Peptides were then eluted on-line to an analytical  
11 column (EASY-Spray PepMap RSLC C18, 75 $\mu\text{m} \times 50\text{cm}$ ) at a flow rate of 250nL/min.  
12 Peptides were separated using a 120 minute stepped gradient, 1-22% of buffer B for 90  
13 minutes followed by 22-42% buffer B for another 30 minutes (composition of buffer A –  
14 95/5%: H<sub>2</sub>O/DMSO + 0.1% FA, buffer B – 75/20/5% MeCN/H<sub>2</sub>O/DMSO + 0.1% FA) and  
15 subsequent column conditioning and equilibration. Eluted peptides were analysed by the  
16 mass spectrometer operating in positive polarity using a data-dependent acquisition mode.  
17 Ions for fragmentation were determined from an initial MS1 survey scan at 30,000 resolution,  
18 followed by CID (Collision Induced Dissociation) of the top 10 most abundant ions. MS1 and  
19 MS2 scan AGC targets were set to 1e6 and 3e4 for maximum injection times of 500 ms and  
20 100 ms respectively. A survey scan m/z range of 350 – 1,500 was used, with normalised  
21 collision energy set to 35%, charge state screening enabled with +1 charge states rejected and  
22 a minimal fragmentation trigger signal threshold of 500 counts.

23

24 For sample analysis on the Q-Exactive, chromatographic separation was performed using an  
25 Ultimate 3000 RSLC nano liquid chromatography system (Thermo Scientific) coupled to an  
26 Orbitrap Q-Exactive mass spectrometer (Thermo Scientific) via an EASY-Spray source.  
27 Peptide solutions were injected and loaded onto a trapping column (Acclaim PepMap 100  
28 C18, 100  $\mu\text{m} \times 2\text{ cm}$ ) for desalting and concentration at 8  $\mu\text{L}/\text{min}$  in 2% acetonitrile, 0.1%  
29 TFA. Peptides were then eluted on-line to an analytical column (EASY-Spray PepMap RSLC  
30 C18, 75  $\mu\text{m} \times 75\text{ cm}$ ) at a flow rate of 200 nL/min. See below for further details.

31

32

33

34

## 1 **Specific downstream settings – *B. bacteriovorus* whole cell extract proteomics**

2

3 Peptides were separated using a 120 minute gradient, 4-25% of buffer B for 90 minutes  
4 followed by 25-45% buffer B for another 30 minutes (composition of buffer B – 80%  
5 acetonitrile, 0.1% FA) and subsequent column conditioning and equilibration. Eluted  
6 peptides were analysed by the mass spectrometer in positive polarity using a data-dependent  
7 acquisition mode. Ions for fragmentation were determined from an initial MS1 survey scan at  
8 70,000 resolution, followed by HCD (Higher-energy Collision Induced Dissociation) of the  
9 top 12 most abundant ions at 17,500 resolution. MS1 and MS2 scan AGC targets were set to  
10 3e6 and 5e4 for maximum injection times of 50 ms and 50 ms respectively. A survey scan  
11 m/z range of 400–1800 was used, normalised collision energy set to 27 and charge exclusion  
12 enabled for unassigned and +1 ions. Dynamic exclusion was set to 45 seconds.

13

## 14 **Specific downstream settings – *L. interrogans* proteomics**

15

16 Peptides were separated using 90-minute gradient, 4-25% of buffer B for 60 minutes  
17 followed by 25-45% buffer B for another 30 minutes (composition of buffer B – 80%  
18 acetonitrile, 0.1% FA) and subsequent column conditioning and equilibration. Eluted  
19 peptides were analysed by the mass spectrometer in positive polarity using a data-dependent  
20 acquisition mode. Ions for fragmentation were determined from an initial MS1 survey scan at  
21 70,000 resolution, followed by HCD (Higher-energy Collision Induced Dissociation) of the  
22 top 10 most abundant ions at 17,500 resolution. MS1 and MS2 scan AGC targets were set to  
23 3e6 and 5e4 for maximum injection times of 50 ms and 100 ms respectively. A survey scan  
24 m/z range of 350–1,800 was used, normalised collision energy set to 27 and charge exclusion  
25 enabled for unassigned and +1 ions. Dynamic exclusion was set to 45 seconds.

26

## 27 **Proteomics data processing**

28

29 Data were processed using the MaxQuant software platform (v1.6.10.43) (Tyanova *et al.*  
30 2016), with database searches carried out by the in-built Andromeda search engine against  
31 the GenBank proteome of each organism. A reverse decoy database approach was used at a  
32 1% false discovery rate (FDR) for peptide spectrum matches. Search parameters were as  
33 follows: maximum missed cleavages set to 2, fixed modification of cysteine  
34 carbamidomethylation and variable modifications of methionine oxidation, protein N-

1 terminal acetylation, asparagine deamidation and cyclisation of glutamine to pyro-glutamate.  
2 Label-free quantification was enabled with an LFQ minimum ratio count of 1. The ‘match  
3 between runs’ function was used with match and alignment time limits of 0.7 and 20 minutes  
4 respectively.

## 6 **Nucleoid enrichment analysis**

7  
8 Nucleoid enrichment analysis was carried out as in (Hocher *et al.* 2022), comparing soluble  
9 and nucleoid enriched fraction using the R package DEP (v1.8.0). Globally, we observe  
10 robust enrichment of proteins with predicted membrane localization, providing validation for  
11 the approach (Fig. S13).

## 13 **RNA extraction and sequencing**

14  
15 Total RNA was extracted from exponentially growing host-independent *B. bacteriovorus*  
16 HID13, and from -80°C frozen *Aquifex aeolicus* pellets (purchased from the  
17 Archaeenzentrum Regensburg, Germany) using the RNEasy kit (Qiagen), including DNase I  
18 treatment. RNA quality was assessed using an Agilent 2100 Bioanalyser.

19  
20 For RNA extraction from *B. bacteriovorus* HID13, cells were diluted in 10 mL of fresh YP  
21 medium at an OD of 0.1 and grown until reaching a density of 0.6 (20 hours at 30°C with  
22 shaking). For each replicate, 2 mL of culture were spun at maximum speed, snap frozen and  
23 kept at -80°C until RNA extraction.

24  
25 For the *B. bacteriovorus* samples (all RIN scores > 9.5), ribosomal RNA was depleted using  
26 the NEBNext rRNA Depletion Kit (Bacteria), and NGS libraries made using the NEBNext  
27 Ultra II Directional RNALibrary Prep Kit for Illumina according to manufacturer’s  
28 instructions. Paired-end 55 bp reads were generated on a NextSeq 2000 with dual 8 bp  
29 indexing. For the *A. aeolicus* samples (all RIN scores > 5), rRNA was depleted using  
30 Illumina Ribozero Kit (Bacteria), and NGS libraries made using the TruSeq Stranded  
31 Total RNA LT Kit according to manufacturer’s instructions. Single End 50 bp reads were  
32 generated on a MiSeq with single 6 bp indexing.

33



## 1 **RNA-seq analysis**

2

3 Untrimmed reads were mapped using Bowtie2 (Langmead and Salzberg 2012) for *A. aeolicus*  
4 (single end reads) and using BWA for *B. bacteriovorus* (paired-end). Read counts of all  
5 genes were estimated using the python package HTSeq (Anders *et al.* 2015).

6

## 7 **Fluorescence microscopy**

8

9 The fluorophore mCherry/mCitrine was fused to the C-terminus of Bd0055 by PCR  
10 amplification of the gene without its stop codon and amplification of the fluorophore gene,  
11 followed by Gibson cloning using the Geneart assembly kit (Invitrogen) according to  
12 instructions into the mobilisable broad host range vector pK18*mobsacB* and this was  
13 conjugated into *B. bacteriovorus* HD100 as described previously (Lambert *et al.* 2003). PCR  
14 amplification was carried out with phusion polymerase (New England Biolabs) according to  
15 instructions, using primers: Bd0055mC\_mCit\_F 5'

16 cgttgtaaaacgacggccagtgccACACTCATCATACCCGCC 3'; Bd0055mC\_Link\_R

17 5' TGCTCACCATAACCACCGCTGCCACCGCCGCCGCTACCGCCAAGGTGGTTCG 3';

18 0055mCherry\_Link\_F

19 5' CGACCACCTTGCGGGTAGCGGCGGCGGTGGCAGCGGTGGTATGGTGAGCA 3';

20 0055mCherry\_R 5' ggaaacagctatgaccatgattagTTACTTGTACAGCTCGTCCATG 3'.

21

22 Cells were imaged on a Nikon Ti-E epifluorescence microscope equipped with a Apo x100  
23 Ph3 oil objective lens (NA: 1.45) and images were acquired on an Andor Neo sCMOS  
24 camera using Nikon NIS software. The following filters were used for fluorescence images:  
25 mCherry (excitation: 555 nm, emission: 620/60 nm), DAPI (Hoechst 33342): (excitation: 395  
26 nm, emission: 435-485 nm). DNA was stained by the addition of Hoechst 33342 at a final  
27 concentration of 5  $\mu$ g/ml. Images were analysed with FIJI software with the MicrobeJ plugin  
28 (Ducret *et al.* 2016). mCitrine expressing cells were imaged on a Leica DMRB with phase  
29 contrast and DIC for transmitted light illumination. To image mCitrine during predation, cells  
30 were fixed (1% paraformaldehyde, 5 min, quenched using 150 mM glycine) and DNA  
31 stained by DAPI in 1x PSB (5min of staining removal by centrifugation).

32

33

## 1 **Protein purification from *E.coli***

2

3 The Bd0055 (or mutant) ORF was codon optimized for expression in *E. coli* and synthesized  
4 into a dsDNA gBlock (IDT) with 35 bp and 22 bp overhangs on the 5' and 3' sides,  
5 respectively. The gBlock was cloned into a *lac*-inducible, ampicillin-resistant pET expression  
6 vector through restriction-free cloning. The protein was expressed in BD *E. coli* cells from  
7 overnight cultures using 0.4 mM IPTG induction after cells reached an OD of ~1.0. Cells  
8 were harvested after 2 hours and centrifuged at 6,000 rpm for 20 min at 4°C. Media was  
9 decanted and the cells were flash frozen in liquid nitrogen and stored at -80°C. Cells were  
10 thawed on ice for 30 min and resuspended in lysis buffer (50 mM Tris pH 7.5, 5 mM EDTA,  
11 0.1% Triton X-100, 5 mM BME, 1 mM AEBSF, and 1 Pierce Complete protease inhibitor  
12 tablet per 50 ml). The cells were then lysed by sonication (3 rounds of 1s on/off for 1min)  
13 and centrifuged at 16,000 rpm for 20 min at 4°C. The lysate was filtered through a 0.45 µm  
14 filter and run over a 5 ml SP HP column (Cytiva) on an AKTA Pure FPLC using a linear  
15 gradient starting at buffer A (0 M NaCl, 50 mM Tris-HCl pH 7.5, 1 mM TCEP, 1 mM  
16 AEBSF) to 1 M NaCl (1 M NaCl, 50 mM Tris-HCl pH 7.5, 1 mM TCEP, 1 mM AEBSF)  
17 over 40 column volumes. Protein typically eluted around 300 mM NaCl. Fractions were  
18 sampled uniformly and analyzed by SDS-PAGE, as Bd0055 has very little absorption at 280  
19 nm. Fractions containing Bd0055 were pooled and diluted in buffer A to ~100 mM NaCl  
20 before being applied to a 5ml Heparin HP column (Cytiva). Protein was eluted with a linear  
21 gradient from buffer A to buffer B over 40 CV, with the protein typically eluting at 450 mM  
22 NaCl. Fractions were pooled and concentrated to a volume of 1 ml before being loaded onto a  
23 120 ml S75 (Cytiva) ran in buffer B containing 10% glycerol, with the protein having a  
24 typical retention volume of 74-82 ml. Fractions were pooled, concentrated, aliquoted, flash  
25 frozen in liquid nitrogen, and stored at -80°C.

26

## 27 **Apo protein crystallization**

28

29 Recombinant Bd0055 in storage buffer (1 M NaCl, 1 mM EDTA, 50 mM HEPES pH 7.5,  
30 10% glycerol) was crystallized through hanging drop diffusion in a mother liquor (3.2 M  
31 (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 100 mM BICINE pH 9) by gently mixing 1 µl of protein to 1 µl of mother liquor  
32 on a coverslip and sealing it onto a pre-greased well of a 24 well crystal tray containing 1 ml  
33 of mother liquor. Crystals formed within 24 hrs and matured within 3 days. Crystals were

1 looped and frozen in liquid nitrogen. X-ray diffraction data were collected on a Rigaku  
2 XtaLAB MM003 and indexed in DIALS. Molecular replacement was done using Phenix and  
3 structures were refined in Coot. Short, idealized alpha-helices were used as search models,  
4 most of the remaining structure was built automatically using AutoBuild. Final residues were  
5 placed by hand in Coot.

## 6 7 **DNA binding (FP)**

8  
9 50  $\mu$ M protein was titrated using an Opentrons OT-2 liquid handling robot into 10 nM  
10 Alexa488 end-labeled 147 bp DNA by serial dilution down to a protein concentration less  
11 than the DNA probe in reaction buffer (10 mM NaCl, 1 mM EDTA, 50 mM HEPES pH 6).  
12 After incubation at RT for > 15 min (overnight incubation did not change overall results),  
13 Fluorescence polarization data were collected using a BMG Labtech CLARIOstar. Data were  
14 analyzed in Prism and fit to an [Inhibitor] vs. response – four parameter nonlinear regression  
15 curve.

## 16 17 **DNA binding (FRET)**

18  
19 FRET experiments were conducted simultaneously with the FP experiments, as the 147 bp  
20 DNA used was labeled on the opposite end of the Alexa488 with an Alexa647 fluorophore.  
21 Data were collected on a BMG Labtech CLARIOstar through excitation of 488 nm light  
22 while recording the emission at 647 nm. Data were analyzed in Prism and fit to an [Inhibitor]  
23 vs. response – four parameter nonlinear regression curve.

## 24 25 **Electrophoretic mobility shift assay (EMSA)**

26  
27 1  $\mu$ M 147 bp DNA was mixed with varying concentrations of protein in reaction buffer (10  
28 mM NaCl, 1 mM EDTA, 50 mM HEPES pH 7.5). Reactions were mixed 1:1 with 80%  
29 glycerol, incubated at RT for > 15 min, and ran on 10% native PAGE (0.2x TBE running  
30 buffer, 150 V, 90 min). Gels stained with ethidium bromide and visualized on a Typhoon,  
31 using the appropriate laser and filter.

32

33

## 1 **Molecular dynamics simulations**

2

3 The HMfB hypernucleosome structure was built in ChimeraX and Chimera, by extending  
4 PDB 5T5K to include four HMfB dimers and 118 bp of dsDNA. The Bd0055  
5 hypernucleosome structure was built by docking the Bd0055 apo structure into the HMfB  
6 hypernucleosome. All-atom molecular dynamics simulations using explicit solvent were  
7 carried out using AMBER18 using the ff14SB, bsc1, and tip3p forcefields (for protein, DNA,  
8 and water respectively). Structures were protonated and hydrogen mass repartitioned (as  
9 implemented through parm9 in AMBER). Structures were placed in cubic boxes  
10 surrounding the structures by at least 25 Å, charge neutralized using potassium ions, and  
11 hydrated with water molecules. The structures were energy minimized in two, 5,000 step  
12 cycles: the first restraining the protein and DNA molecules to allow the solvent to relax, and  
13 the second to allow the whole system to relax. Minimized structures were then heated to 300  
14 K and slowly brought to 1.01325 atm. These systems were then simulated for 500 ns in 4 fs  
15 steps. Simulations were carried out on NVIDIA GPUs (RTX6000s or A100s) using CU  
16 Boulder's Blanca Condo cluster. RMSD analysis was carried out using cpptraj through  
17 AMBER18.

18

## 19 **DNA/protein crystallization**

20

21 DNA fragments ranging between 35 bp and 45 bp in length, with randomized sequence and  
22 ~50% GC content, were screened for their ability to form discrete complexes with Bd0055,  
23 using EMSA. The 35 bp DNA  
24 (sequence: 5'TCTTGCACTAAGAGCTACTGGAGTGCATCAGATGT3')

25 was selected as it formed a discrete DNA shift at roughly 4:1 protein to DNA (it shifted to  
26 higher smears upon addition of more protein). 75 μM DNA and 1200 μM protein (16:1) were  
27 mixed in crystallization buffer (10 mM NaCl, 0.1 mM EDTA, 50 mM HEPES pH 7.5) and  
28 dialyzed at RT against crystallization buffer. Hanging drop crystals were set in 4 μl drops at  
29 1:1 complex to mother liquor (15% PEG 550 MME, 50 mM HEPES pH 8.0) and incubated at  
30 20°C. Crystals formed overnight and matured within a week. Large cubic crystals were  
31 looped, cryoprotected by washing in 30% glycerol, and flash frozen in liquid nitrogen. X-ray  
32 diffraction data were collected at the ALS synchrotron (12,397.9 eV, 225 mm detector, ΔΦ =  
33 0.25) and indexed in DIALS. Molecular replacement was done using Phenix and structures

1 were refined in Coot. The apo Bd0055 structure was used as a search model along with 2  
2 ssDNA fragments (2 and 3 bp).

3

#### 4 **Gene deletion**

5

6 Attempts to generate a markerless deletion mutant of the *bd0055* open reading frame in  
7 *B. bacteriovorus* HD100 were by PCR amplification of the flanking 1000 bp upstream region  
8 with the first 2 codons of the *bd0055* open reading frame and final 3 codons with the  
9 downstream 1000 bp flanking region and fusing these by Gibson assembly into the  
10 mobilisable broad host range vector pK18*mobsacB*. The primers used were: Bd0055\_UP\_F  
11 5' cgttgtaaaacgacggccagtccaATTGATTGAACACGGCAATC 3'; Bd0055\_UP\_R 5'  
12 cgaattaaaggtgTGCCATGATATACCTCTCC 3'; Bd0055\_DN\_F 5'  
13 gtatatcatggcaCACCTTTAATTCGGTTCGAACG 3'; Bd0055\_DN\_R 5'  
14 ggaaacagctatgaccatgattacgTCGCTGGATGTCTCCGGC 3'

15 This construct was conjugated into *B. bacteriovorus* HD100 as described previously  
16 (Lambert *et al.* 2003) and resulting exconjugants screened for kanamycin sensitivity and  
17 screened by PCR to test for either gene deletion or revertant to wild-type. Screening primers  
18 were: Bd0055KO\_S\_F 5' atctggagcttcacttcccg 3' and Bd0055KO\_S\_R 5'  
19 ggtgatgatccgggctctaa 3'.

20

21 For targeted mutagenesis in *L. interrogans*, a kanamycin resistance cassette replacing the  
22 coding sequence of LA\_2458 and 0.8-0.9 kb sequences homologous to the sequences  
23 flanking the target gene was synthesized by GeneArt (Life Technologies, Grand Island, NY,  
24 USA) and cloned in an *E. coli* vector which is not able to replicate in *L. interrogans*. Plasmid  
25 DNA was then introduced in *L. interrogans* serovar Manilae by electroporation as previously  
26 described (Picardeau *et al.* 2001) with a Biorad Gene Pulser Xcell™. Electroporated cells  
27 were plated on EMJH agar plates supplemented with 50 µg/ml kanamycin. Plates were  
28 incubated for 4 weeks at 30°C in sealed plastic bags or wrapped in foil to avoid desiccation.

29

#### 30 **Dataset availability**

31

32 All datasets are publicly available. Proteomics data have been deposited to PRIDE  
33 (PXD039405) and RNA-seq data to GEO under accession GSE220534. All genomes used

1 were publicly available with no usage restriction (Table S3). Structure data have been  
2 deposited at the Protein Data Bank (PDB 8FVX for apo Bd0055 and 8FW7 for DNA bound  
3 Bd0055).

#### 4 5 **Script availability**

6  
7 Analysis scripts were run using R version 3.6.2. Scripts required to reproduce figures are  
8 available at <https://github.com/hocherantoinne/BHF/>.

9  
10

#### 11 **ACKNOWLEDGMENTS**

12 We thank the LMS Proteomics Facility for generating mass spectrometry data, the LMS  
13 Genomics Facility for RNA sequencing, and Annette Erbse at CU Boulder for help with  
14 crystallography. PR, JT, CL, and RES are funded by a Wellcome Trust Investigator Award in  
15 Science (209437/Z/17/Z). Work by AH, KMS, and TW is funded by a UKRI MRC core grant  
16 to TW (MC-A658-5TY40). KL and SL are funded by the Howard Hughes Medical Institute.

17

#### 18 **COMPETING INTERESTS**

19 The authors declare no competing interest

20

#### 21 **OPEN ACCESS**

22 For the purpose of open access, the author has applied a Creative Commons Attribution (CC  
23 BY) licence (where permitted by UKRI, ‘Open Government Licence’ or ‘Creative Commons  
24 Attribution No-derivatives (CC BY-ND) licence may be stated instead) to any Author  
25 Accepted Manuscript version arising’.

26

27 This article is subject to HHMI’s Open Access to Publications policy. HHMI lab heads have  
28 previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable  
29 license to HHMI in their research articles. Pursuant to those licenses, the author-accepted  
30 manuscript of this article can be made freely available under a CC BY 4.0 license  
31 immediately upon publication.

32

33

## 1 MATERIALS & CORRESPONDENCE

2 Correspondence and requests for materials can be addressed to KL  
3 (karolin.luger@colorado.edu) or TW (tobias.warnecke@lms.mrc.ac.uk).

4  
5

## 6 CONTRIBUTIONS

7 AH initiated the project and carried out all bioinformatic analyses, sequencing experiments  
8 and proteomics (including upstream bacterial culture and biochemistry). SL performed all *in*  
9 *vitro* work on *E. coli*-expressed Bd0055 (biophysical characterization, crystallography,  
10 structure analysis) as well as MD simulations. KMS assisted in histone homolog curation  
11 and pan-bacterial analysis of histone proteins. PR built Bd0055-mCherry and -mCitrine  
12 fusion strains of *B. bacteriovorus* which were phenotyped microscopically by JT. CL and PR  
13 attempted gene deletions of *bd0055* in predatory and prey-independent cultures. MP provided  
14 biomass for *L. interrogans* biochemistry and attempted gene deletions of LA\_2458. RES, KL  
15 and TW supervised the project. AH, SL, KL, and TW led data analysis and interpretation and  
16 wrote the manuscript with input from all authors.

17  
18  
19

## 20 SUPPLEMENTARY TABLES

21

22 **Table S1. Bacterial histone-fold proteins identified via homolog search.**

23

24 **Tables S2. X-ray diffraction statistics.** Data collection and refinement statistics from  
25 Bd0055 apo crystal structures (pdb 8FVX) and Bd0055 bound to 35 bp of dsDNA (pdb  
26 8FW7).

27

28 **Table S3. Bacterial genomes searched for histone fold proteins.**

29

## 30 SUPPLEMENTARY FILES

31

32 **Movie S1.** 500 ns simulations of HMfB and modeled Bd0055 hypernucleosomes.

33

1 **Movie S2.** Animation of nucleohistone filament and higher-order lattice formation. In the  
2 crystal lattice, histone-bound DNA fibers assemble in a parallel manner with a DNA spacing  
3 of ~60 Å to the nearest strand. The protein-protein contacts stabilizing the lattice engage the  
4 second L1-L2 interface of a given Bd0055 dimer. Arginine 57 is sandwiched between E36  
5 and D29 of a histone dimer binding to another DNA, and C-terminal tails cross over between  
6 the two dimers (Fig. 5c). Additionally, the amino terminal NH<sub>3</sub> hydrogen bonds with D49 in  
7 the L2 loop of the neighboring dimer.

8

9

## 10 REFERENCES

11

12

13 Alva V., and A. N. Lupas, 2018 Histones predate the split between bacteria and archaea.  
14 *Bioinformatics* 35: 2349–2353. <https://doi.org/10.1093/bioinformatics/bty1000>

15 Anders S., P. T. Pyl, and W. Huber, 2015 HTSeq—a Python framework to work with high-  
16 throughput sequencing data. *Bioinformatics* 31: 166–169.  
17 <https://doi.org/10.1093/bioinformatics/btu638>

18 Arents G., R. W. Burlingame, B. C. Wang, W. E. Love, and E. N. Moudrianakis, 1991 The  
19 nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a  
20 left-handed superhelix. *Proc National Acad Sci* 88: 10148–10152.  
21 <https://doi.org/10.1073/pnas.88.22.10148>

22 Baba T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, *et al.*, 2006 Construction of  
23 *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol.*  
24 *Syst. Biol.* 2. <https://doi.org/10.1038/msb4100050>

25 Beder T., and H. P. Saluz, 2018 Virulence-related comparative transcriptomics of infectious  
26 and non-infectious chlamydial particles. *Bmc Genomics* 19: 575.  
27 <https://doi.org/10.1186/s12864-018-4961-x>

28 Borgnia M. J., S. Subramaniam, and J. L. S. Milne, 2008 Three-Dimensional Imaging of the  
29 Highly Bent Architecture of *Bdellovibrio bacteriovorus* by Using Cryo-Electron  
30 Tomography. *J Bacteriol* 190: 2588–2596. <https://doi.org/10.1128/jb.01538-07>

31 Bowerman S., J. Wereszczynski, and K. Luger, 2021 Archaeal chromatin ‘slinkies’ are  
32 inherently dynamic complexes with deflected DNA wrapping pathways. *Elife* 10: e65587.  
33 <https://doi.org/10.7554/elife.65587>

34 Čuboňová L., M. Katano, T. Kanai, H. Atomi, J. N. Reeve, *et al.*, 2012 An Archaeal Histone  
35 Is Required for Transformation of *Thermococcus kodakarensis*. *J Bacteriol* 194: 6864–  
36 6874. <https://doi.org/10.1128/jb.01523-12>



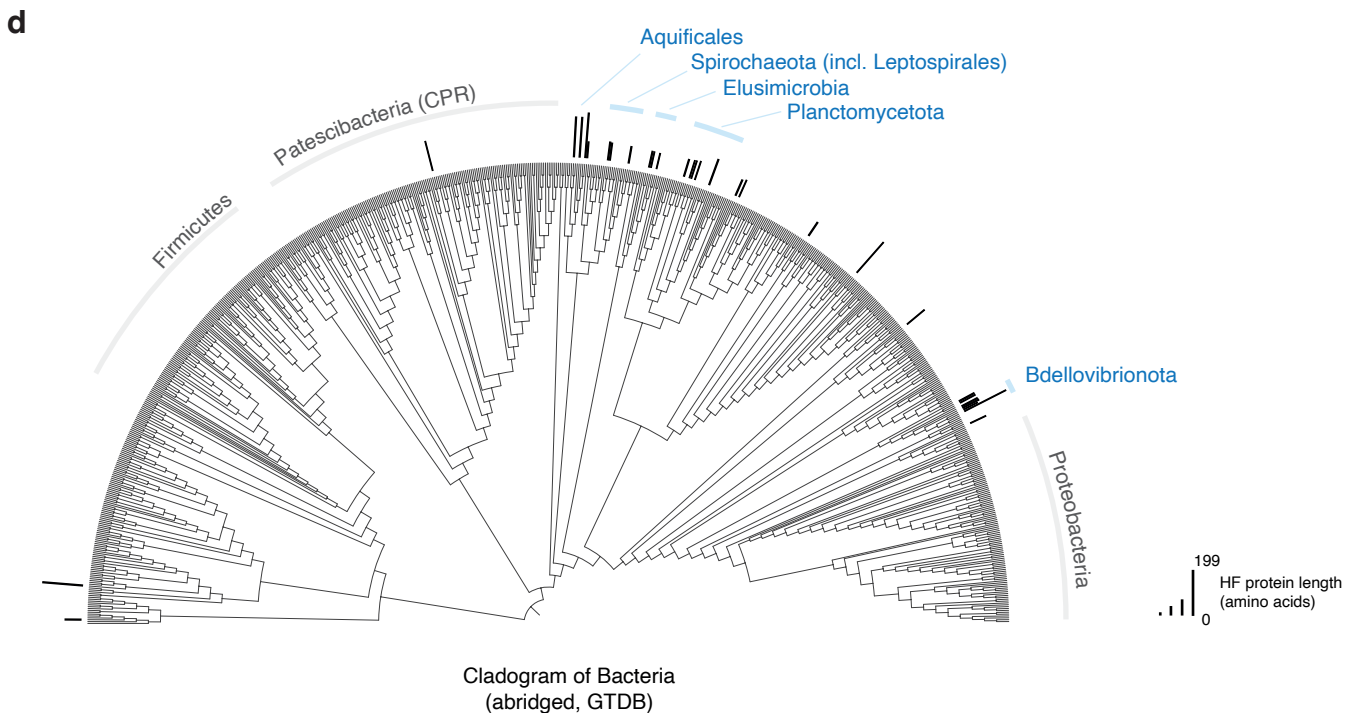
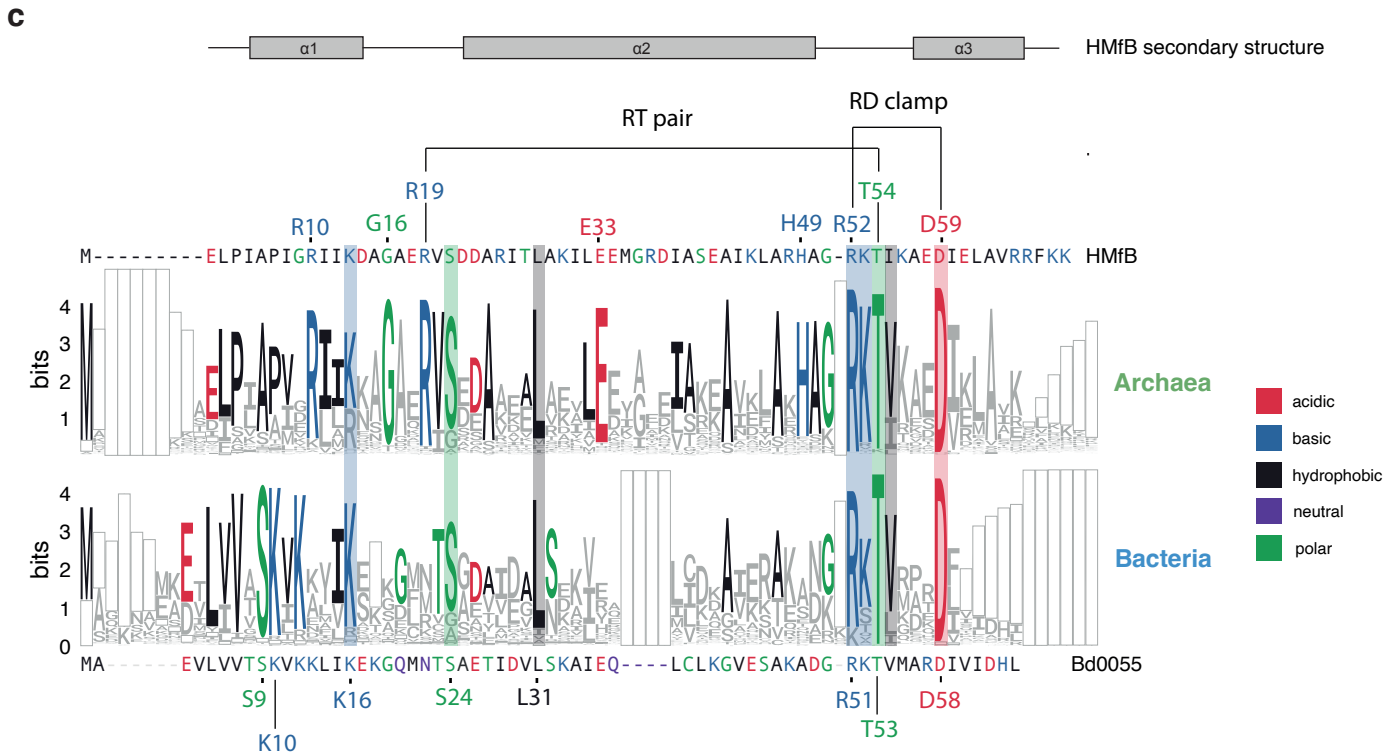
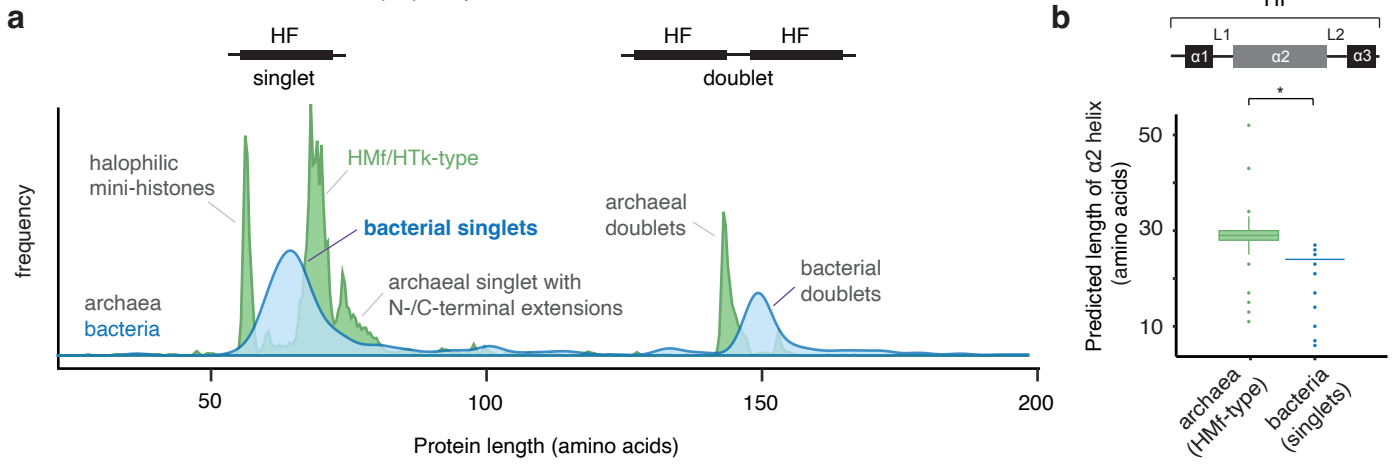
- 1 Drozdetskiy A., C. Cole, J. Procter, and G. J. Barton, 2015 JPred4: a protein secondary  
2 structure prediction server. *Nucleic Acids Res* 43: W389–W394.  
3 <https://doi.org/10.1093/nar/gkv332>
- 4 Ducret A., E. M. Quardokus, and Y. V. Brun, 2016 MicrobeJ, a tool for high throughput  
5 bacterial cell detection and quantitative analysis. *Nat Microbiol* 1: 16077.  
6 <https://doi.org/10.1038/nmicrobiol.2016.77>
- 7 Gossett A. J., and J. D. Lieb, 2012 In Vivo Effects of Histone H3 Depletion on Nucleosome  
8 Occupancy and Position in *Saccharomyces cerevisiae*. *Plos Genet* 8: e1002771.  
9 <https://doi.org/10.1371/journal.pgen.1002771>
- 10 Han M., M. Chang, U.-J. Kim, and M. Grunstein, 1987 Histone H2B repression causes cell-  
11 cycle-specific arrest in yeast: Effects on chromosomal segregation, replication, and  
12 transcription. *Cell* 48: 589–597. [https://doi.org/10.1016/0092-8674\(87\)90237-6](https://doi.org/10.1016/0092-8674(87)90237-6)
- 13 Henneman B., C. van Emmerik, H. van Ingen, and R. T. Dame, 2018 Structure and function  
14 of archaeal histones. *Plos Genet* 14: e1007582.  
15 <https://doi.org/10.1371/journal.pgen.1007582>
- 16 Hocher A., G. Borrel, K. Fadhlou, J.-F. Brugère, S. Gribaldo, *et al.*, 2022 Growth  
17 temperature and chromatinization in archaea. *Nat Microbiol* 7: 1932–1942.  
18 <https://doi.org/10.1038/s41564-022-01245-2>
- 19 Hodges A. J., I. J. Gallegos, M. F. Laughery, R. Meas, L. Tran, *et al.*, 2015 Histone Sprocket  
20 Arginine Residues Are Important for Gene Expression, DNA Repair, and Cell Viability in  
21 *Saccharomyces cerevisiae*. *Genetics* 200: 795–806.  
22 <https://doi.org/10.1534/genetics.115.175885>
- 23 Kaljević J., T. N. V. Saaki, S. K. Govers, O. Remy, R. van Raaphorst, *et al.*, 2021  
24 Chromosome choreography during the non-binary cell cycle of a predatory bacterium.  
25 *Curr Biol* 31: 3707-3720.e5. <https://doi.org/10.1016/j.cub.2021.06.024>
- 26 Kim U. J., M. Han, P. Kayne, and M. Grunstein, 1988 Effects of histone H4 depletion on the  
27 cell cycle and transcription of *Saccharomyces cerevisiae*. *Embo J* 7: 2211–2219.  
28 <https://doi.org/10.1002/j.1460-2075.1988.tb03060.x>
- 29 Kornberg R. D., and Y. Lorch, 2020 Primary Role of the Nucleosome. *Mol Cell* 79: 371–375.  
30 <https://doi.org/10.1016/j.molcel.2020.07.020>
- 31 Lambert C., M. C. M. Smith, and R. E. Sockett, 2003 A novel assay to monitor predator–prey  
32 interactions for *Bdellovibrio bacteriovorus* 109 J reveals a role for methyl-accepting  
33 chemotaxis proteins in predation. *Environ Microbiol* 5: 127–132.  
34 <https://doi.org/10.1046/j.1462-2920.2003.00385.x>
- 35 Lambert C., and R. E. Sockett, 2008 Laboratory Maintenance of *Bdellovibrio*. *Curr Protoc*  
36 *Microbiol* 9: 7B.2.1-7B.2.13. <https://doi.org/10.1002/9780471729259.mc07b02s9>

- 1 Lambert C., C.-Y. Chang, M. J. Capeness, and R. E. Sockett, 2010 The First Bite— Profiling  
2 the Predatosome in the Bacterial Pathogen *Bdellovibrio*. *Plos One* 5: e8599.  
3 <https://doi.org/10.1371/journal.pone.0008599>
- 4 Langmead B., and S. L. Salzberg, 2012 Fast gapped-read alignment with Bowtie 2. *Nat*  
5 *Methods* 9: 357–359. <https://doi.org/10.1038/nmeth.1923>
- 6 Laursen S. P., S. Bowerman, and K. Luger, 2021 Archaea: The Final Frontier of Chromatin. *J*  
7 *Mol Biol* 433: 166791. <https://doi.org/10.1016/j.jmb.2020.166791>
- 8 Luger K., A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond, 1997 Crystal  
9 structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389: 251–260.  
10 <https://doi.org/10.1038/38444>
- 11 Luger K., and T. J. Richmond, 1998a DNA binding within the nucleosome core. *Curr Opin*  
12 *Struc Biol* 8: 33–40. [https://doi.org/10.1016/s0959-440x\(98\)80007-9](https://doi.org/10.1016/s0959-440x(98)80007-9)
- 13 Luger K., and T. J. Richmond, 1998b The histone tails of the nucleosome. *Curr Opin Genet*  
14 *Dev* 8: 140–146. [https://doi.org/10.1016/s0959-437x\(98\)80134-2](https://doi.org/10.1016/s0959-437x(98)80134-2)
- 15 Makowski Ł., D. Trojanowski, R. Till, C. Lambert, R. Lowry, *et al.*, 2019 Dynamics of  
16 Chromosome Replication and Its Relationship to Predatory Attack Lifestyles in  
17 *Bdellovibrio bacteriovorus*. *Appl Environ Microb* 85: e00730-19.  
18 <https://doi.org/10.1128/aem.00730-19>
- 19 Marc F., K. Sandman, R. Lurz, and J. N. Reeve, 2002 Archaeal histone tetramerization  
20 determines DNA affinity and the direction of DNA supercoiling. *The Journal of*  
21 *Biological Chemistry* 277: 30879–30886. <https://doi.org/10.1074/jbc.m203674200>
- 22 Maruyama H., J. C. Harwood, K. M. Moore, K. Paszkiewicz, S. C. Durley, *et al.*, 2013 An  
23 alternative beads-on-a-string chromatin architecture in *Thermococcus kodakarensis*.  
24 *EMBO Rep* 14: 711–717. <https://doi.org/10.1038/embor.2013.94>
- 25 Mattioli F., S. Bhattacharyya, P. N. Dyer, A. E. White, K. Sandman, *et al.*, 2017 Structure of  
26 histone-based chromatin in Archaea. *Science* 357: 609–612.  
27 <https://doi.org/10.1126/science.aaj1849>
- 28 Murphy L. D., and S. B. Zimmerman, 1997 Isolation and Characterization of Spermidine  
29 Nucleoids from *Escherichia coli*. *J Struct Biol* 119: 321–335.  
30 <https://doi.org/10.1006/jsbi.1997.3883>
- 31 Ohniwa R. L., Y. Ushijima, S. Saito, and K. Morikawa, 2011 Proteomic Analyses of  
32 Nucleoid-Associated Proteins in *Escherichia coli*, *Pseudomonas aeruginosa*, *Bacillus*  
33 *subtilis*, and *Staphylococcus aureus*. *Plos One* 6: e19172.  
34 <https://doi.org/10.1371/journal.pone.0019172>
- 35 Patwal I., H. Trinh, A. Golden, and A. Flaus, 2021 Histone sequence variation in divergent  
36 eukaryotes facilitates diversity in chromatin packaging. *Biorxiv* 2021.05.12.443918.  
37 <https://doi.org/10.1101/2021.05.12.443918>

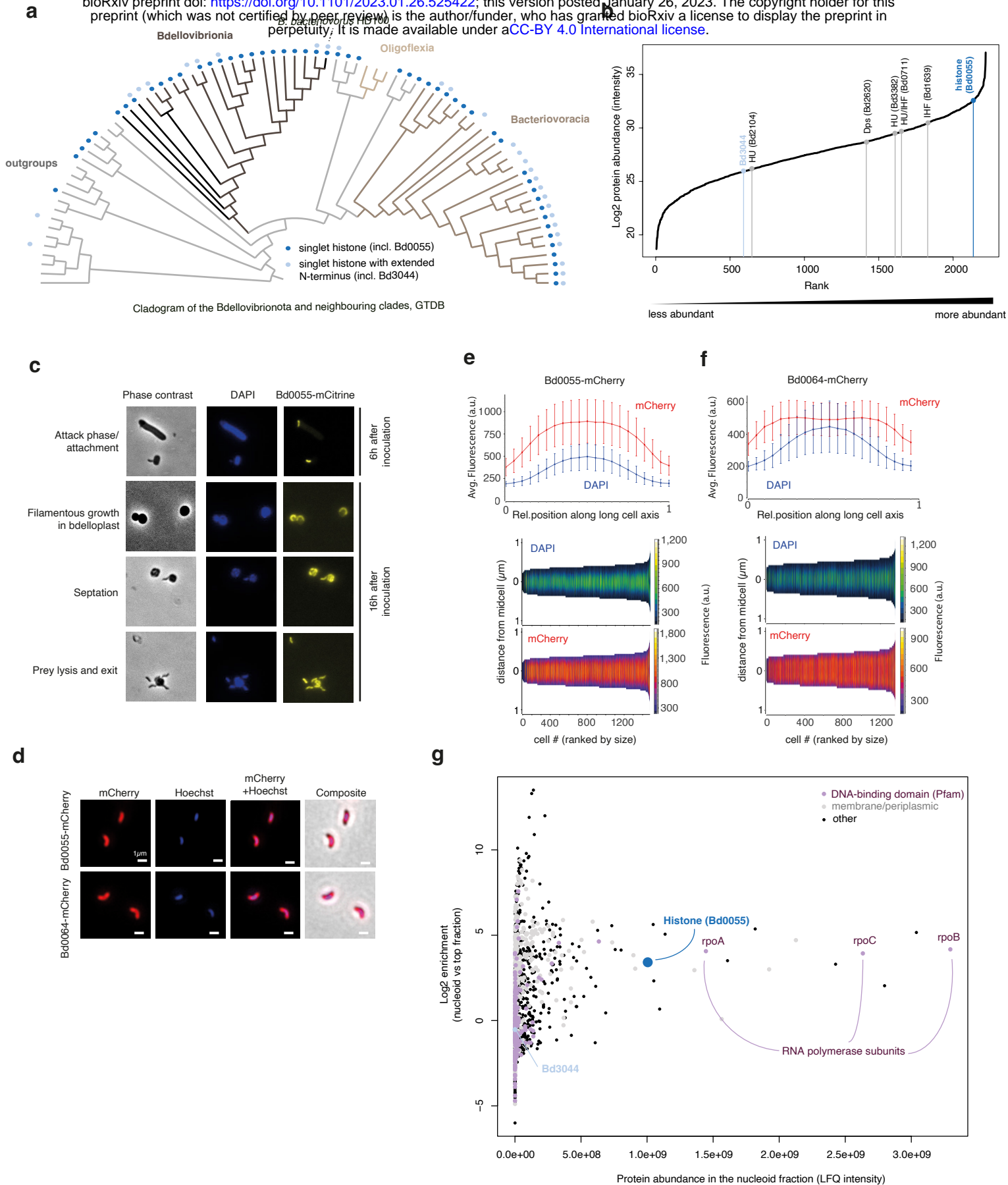
- 1 Picardeau M., A. Brenot, and I. S. Girons, 2001 First evidence for gene replacement in  
2 *Leptospira* spp. Inactivation of *L. biflexa* *flaB* results in non-motile mutants deficient in  
3 endoflagella. *Mol Microbiol* 40: 189–199. [https://doi.org/10.1046/j.1365-](https://doi.org/10.1046/j.1365-2958.2001.02374.x)  
4 [2958.2001.02374.x](https://doi.org/10.1046/j.1365-2958.2001.02374.x)
- 5 Portalier R., and A. Worcel, 1976 Association of the folded chromosome with the cell  
6 envelope of *E. coli*: Characterization of the proteins at the DNA-membrane attachment  
7 site. *Cell* 8: 245–255. [https://doi.org/10.1016/0092-8674\(76\)90008-8](https://doi.org/10.1016/0092-8674(76)90008-8)
- 8 Postberg J., S. Forcob, W.-J. Chang, and H. J. Lipps, 2010 The evolutionary history of  
9 histone H3 suggests a deep eukaryotic root of chromatin modifying mechanisms. *BMC*  
10 *Evol. Biol.* 10: 259. <https://doi.org/10.1186/1471-2148-10-259>
- 11 Sandman K., R. A. Grayling, B. Dobrinski, R. Lurz, and J. N. Reeve, 1994 Growth-phase-  
12 dependent synthesis of histones in the archaeon *Methanothermus fervidus*. *proceedings of*  
13 *the National Academy of Sciences of the United States of America* 91: 12624–12628.
- 14 Schmidt A., M. Beck, J. Malmström, H. Lam, M. Claassen, *et al.*, 2011 Absolute  
15 quantification of microbial proteomes at different states by directed mass spectrometry.  
16 *Mol. Syst. Biol.* 7. <https://doi.org/10.1038/msb.2011.37>
- 17 Sockett R. E., 2009 Predatory Lifestyle of *Bdellovibrio bacteriovorus*. *Microbiology+* 63:  
18 523–539. <https://doi.org/10.1146/annurev.micro.091208.073346>
- 19 Soo V. W. C., and T. Warnecke, 2021 Slaying the last unicorn: discovery of histones in the  
20 microalga *Nanochlorum eucaryotum*. *Roy Soc Open Sci* 8: 202023.  
21 <https://doi.org/10.1098/rsos.202023>
- 22 Stevens K. M., J. B. Swadling, A. Hocher, C. Bang, S. Gribaldo, *et al.*, 2020 Histone variants  
23 in archaea and the evolution of combinatorial chromatin complexity. *P Natl Acad Sci Usa*  
24 117: 33384–33395. <https://doi.org/10.1073/pnas.2007056117>
- 25 Stevens K. M., and T. Warnecke, 2022 Histone variants in archaea – An undiscovered  
26 country. *Semin Cell Dev Biol.* <https://doi.org/10.1016/j.semcdb.2022.02.016>
- 27 Swarts D. C., J. J. Koehorst, E. R. Westra, P. J. Schaap, and J. van der Oost, 2015 Effects of  
28 Argonaute on Gene Expression in *Thermus thermophilus*. *Plos One* 10: e0124880.  
29 <https://doi.org/10.1371/journal.pone.0124880>
- 30 Talbert P. B., M. P. Meers, and S. Henikoff, 2019 Old cogs, new tricks: the evolution of gene  
31 expression in a chromatin context. *Nat Rev Genet* 20: 283–297.  
32 <https://doi.org/10.1038/s41576-019-0105-7>
- 33 Tyanova S., T. Temu, and J. Cox, 2016 The MaxQuant computational platform for mass  
34 spectrometry-based shotgun proteomics. *Nat Protoc* 11: 2301–2319.  
35 <https://doi.org/10.1038/nprot.2016.136>
- 36 Willis A. R., C. Moore, M. Mazon-Moya, S. Krokowski, C. Lambert, *et al.*, 2016 Injections  
37 of Predatory Bacteria Work Alongside Host Immune Cells to Treat *Shigella* Infection in  
38 Zebrafish Larvae. *Curr Biol* 26: 3343–3351. <https://doi.org/10.1016/j.cub.2016.09.067>

- 1 Wyrick J. J., F. C. P. Holstege, E. G. Jennings, H. C. Causton, D. Shore, *et al.*, 1999
- 2 Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast.
- 3 Nature 402: 418–421. <https://doi.org/10.1038/46567>

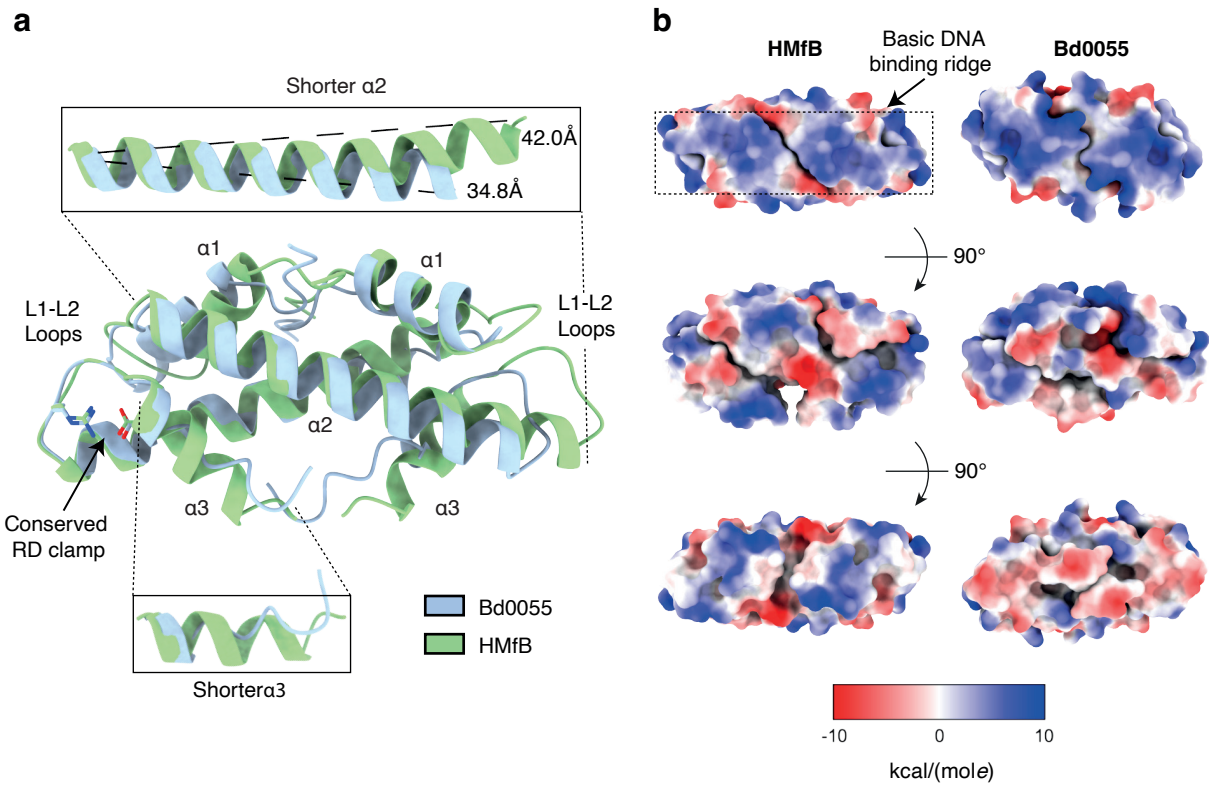
4



**Figure 1. Histone-fold proteins in bacteria.** (a) Length distribution of proteins (<200 amino acids) encoding predicted histone fold domains in bacteria and archaea. Relative frequencies are shown. (b) Length of the  $\alpha 2$  helix in bacterial versus HMf/HTk archaeal singlet histones ( $P < 2.2 \times 10^{-16}$ , Mood test). (c) Weblogo representation for bacterial versus archaeal singlet histones. Bacterial and archaeal histones were aligned separately followed by profile-profile alignment to allow comparison across kingdoms. Alignment gaps were coded as a separate character so as to retain their information value and are visualized as empty boxes. Residues that are notably conserved across kingdoms are highlighted by shaded boxes. (d) Abridged (see Methods) bacterial species tree illustrating the phyletic distribution of histone fold-containing proteins across the kingdom.

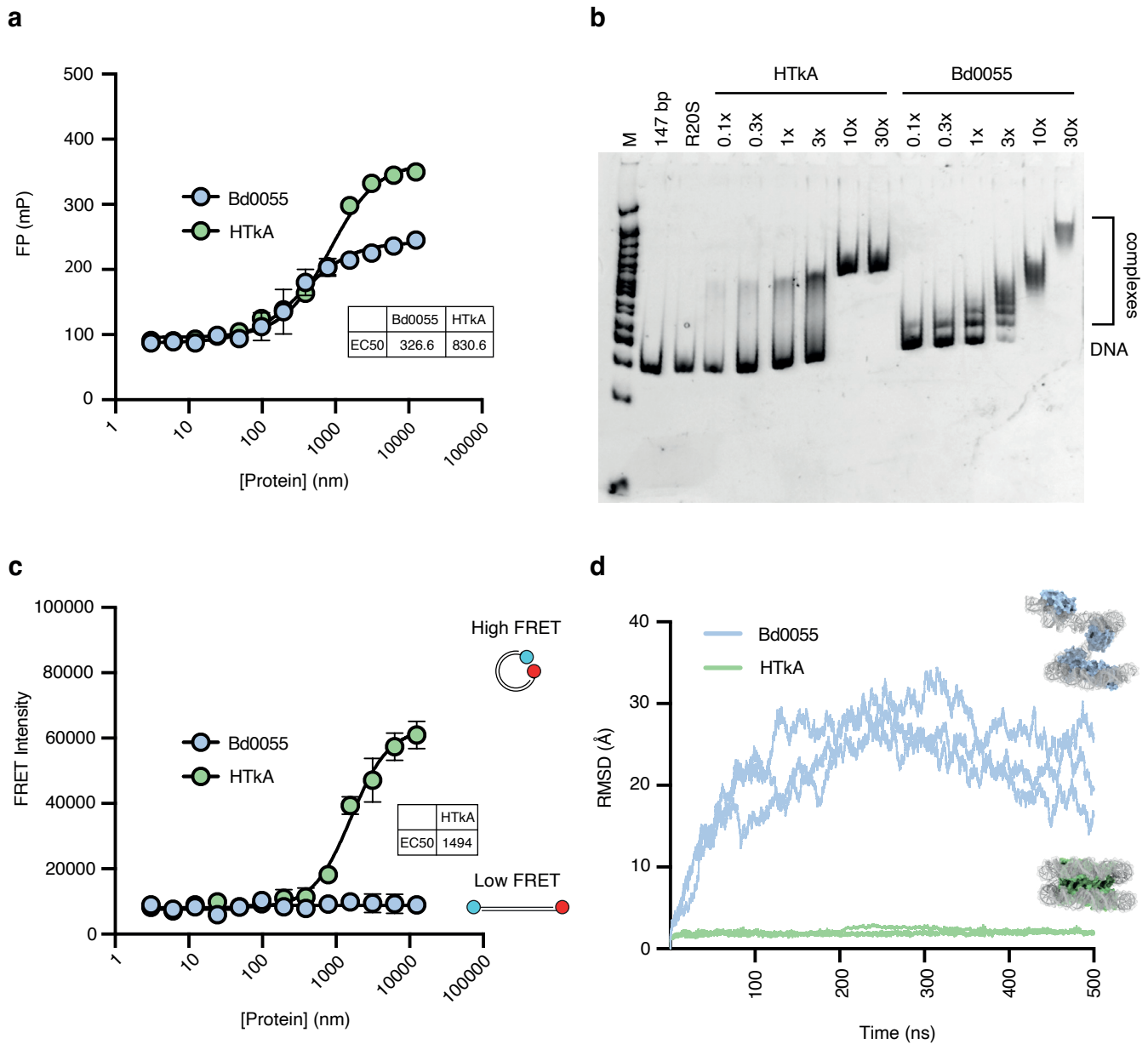


**Figure 2. Histones in Bdellovibrio.** (a) Phyletic distribution of singlet histones across the Bdellovibrionota. (b) Ranked protein abundance in *B. bacteriovorus* attack phase cells, based on quantitative label-free proteomics. All quantified proteins are plotted. (c) Representative images of different phases of the *B. bacteriovorus* life cycle from strains expressing a Bd0055-mCitrine fusion protein. (d-f) Representative images and quantification describing the localization of Bd0055-mCherry and Bd0064-mCherry, a protein with previously established cytosolic localization, in *B. bacteriovorus*. (g) High abundance and prominent nucleoid enrichment of Bd0055 in *B. bacteriovorus* attack phase cells (also see Fig. S5). Note that the nucleoid fraction is also enriched for membrane components that have previously been reported to co-sediment with the nucleoid (Portalier and Worcel 1976; Murphy and Zimmerman 1997; Hoher *et al.* 2022).

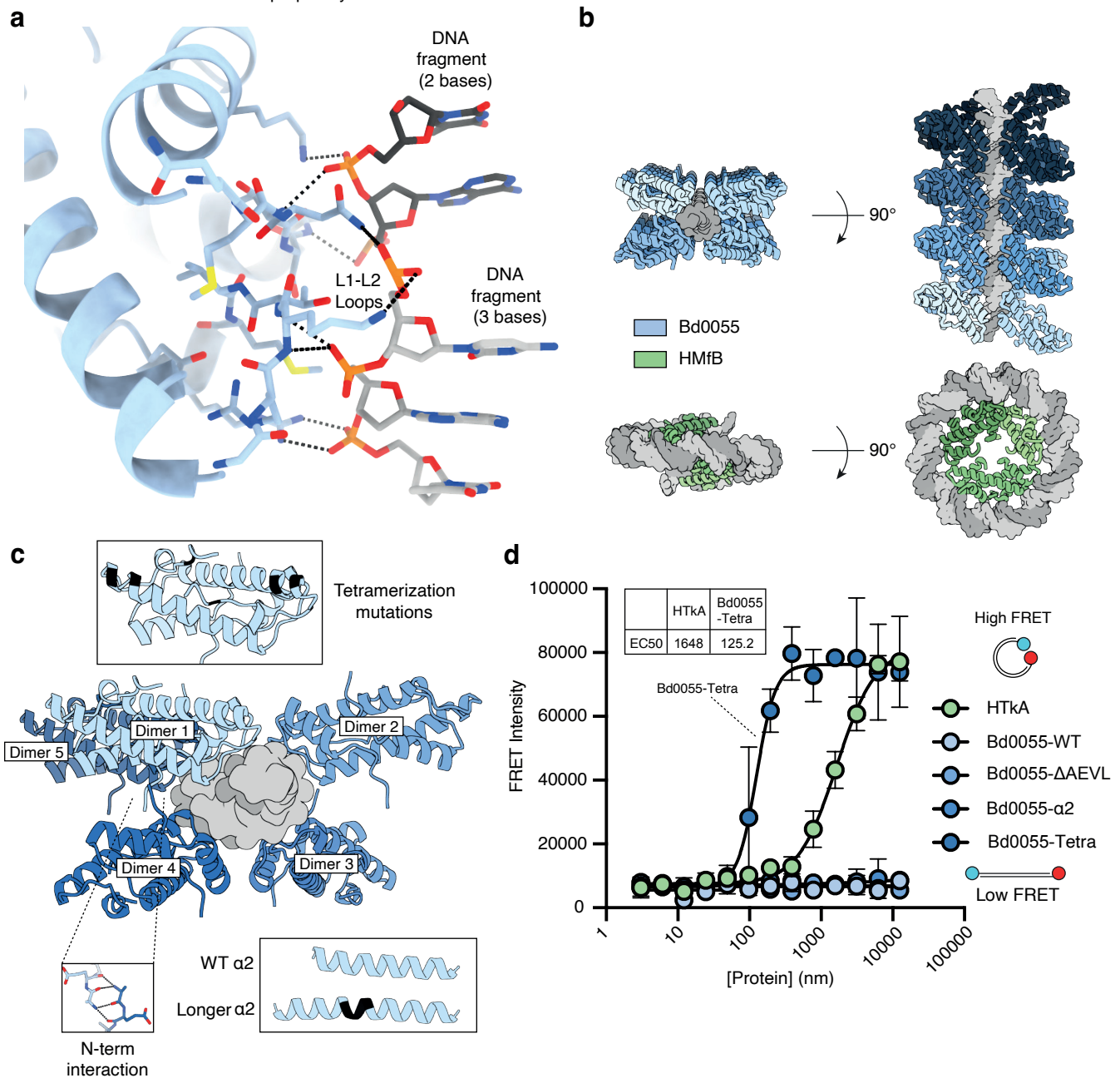


**Figure 3. Bd0055 forms a histone fold dimer. (a)** Crystal structure of Bd0055 superimposed onto archaeal histone HMfB (PDB 1A7W). Bd0055 maintains the overall topology of a histone fold, with shorter  $\alpha 2$  and  $\alpha 3$  helices. Bd0055 conserves the signature RD-clamp and L1/L2 loops. **(b)** Columbic surface charge calculation of HMfB and Bd0055, shown in three orientations. Bd0055 maintains the basic ridge of residues important for binding DNA in other histones (top), but has a net acidic charge on its opposite face (bottom).



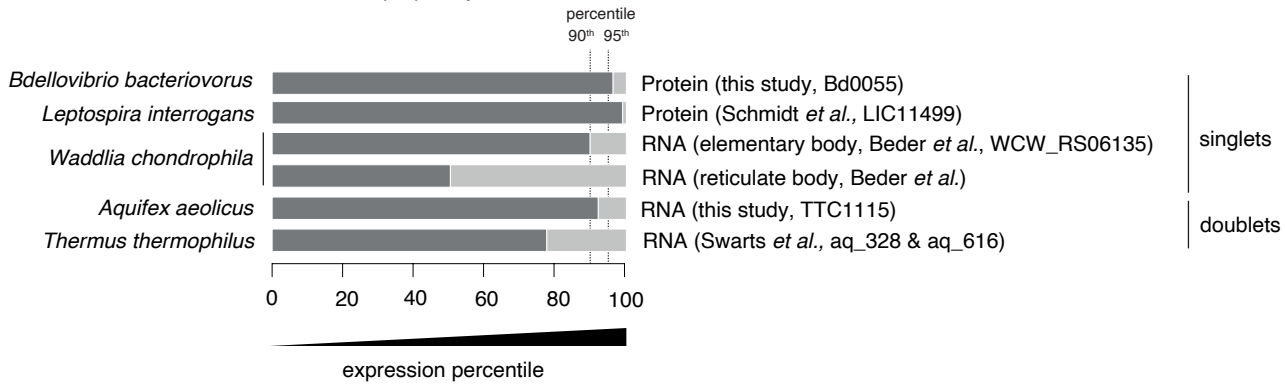


**Figure 4. Bd0055 binds DNA differently than HTkA.** (a) Binding of Bd0055 to an Alexa488-labeled piece of 147 bp DNA (Widom 601 sequence), measured by fluorescence polarization (FP) ( $n=3$ ). Error bars represent standard deviation of the mean. (b) Bd0055 - DNA binding monitored by EMSA (using unlabeled 147 bp DNA). Bd0055 binds with similar affinity as HTkA, but produces a different gel shift pattern. (c) Bd0055 was titrated into dual labeled 147 bp DNA (Alexa488 and Alexa647, same DNA as in FP experiment). At concentrations of HTkA high enough to bind DNA, histones wrap the DNA and produce a high FRET signal. Bd0055 does not cause FRET between DNA ends, even with a large excess of protein ( $n=3$ ). Error bars represent standard deviation of the mean. (d) Plot of root mean square deviations (RMSDs) from simulating a hypothetical Bd0055 hypernucleosome modeled onto PDB 5T5K, using four histone dimers and 118 bp of DNA (see Movie S1). After 500 ns of simulation, hypernucleosomes with HMfB remain stable, whereas hypothetical Bd0055 hypernucleosomes fall apart after as little as 10 ns.

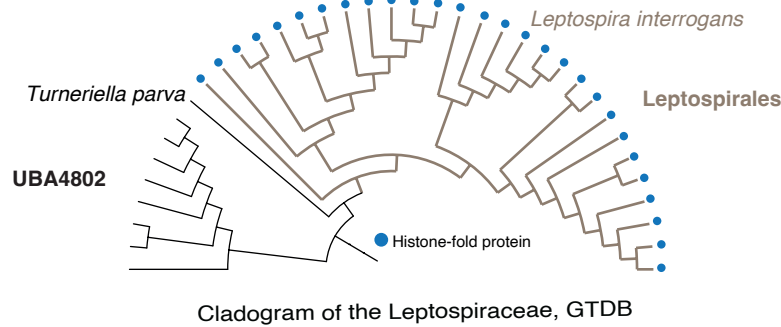


**Figure 5. Bd0055 binds DNA end-on and encases straight DNA.** (a) Crystal structure of Bd0055 in complex with 35 bp DNA, showing interactions between the L1-L2 loop across and the phosphate backbone across the minor groove of DNA. Asymmetric unit is shown. (b) Top: Bd0055 dimers encase dsDNA by binding to five phosphates (2-3 on each strand) and interacting with neighboring dimers through electrostatic interactions. Consecutive dimers as they filament on the DNA are shown in decreasing shades of blue. Bottom panel shows how archaeal HMfB wraps DNA around a core of histone dimers that are linked through a four-helix bundle structure (PDB 5T5K). (c) Details of histone-histone interactions in the Bd0055 nucleohistone filament. Sites of mutagenesis to switch the binding mode are indicated. (d) Mutation of A48H, S45F, I61L (Bd0055-tetra) in Bd0055 enables it to bring DNA ends into FRET proximity, while extending  $\alpha 2$  by inserting YAIE (Bd0055- $\alpha 2$ ) or deleting the N-terminus (Bd0055- $\Delta$ AEVL) has no effect, even though all mutants still bind DNA (see Fig. S6). Error bars represent standard deviation of the mean.

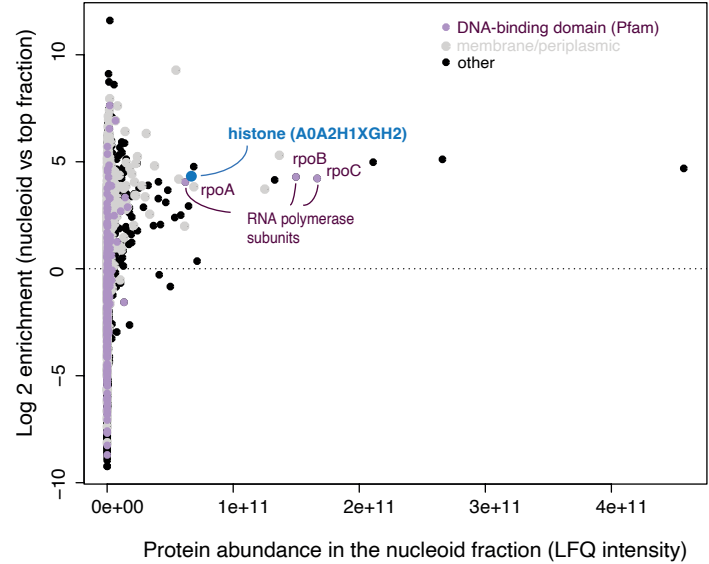
**a**



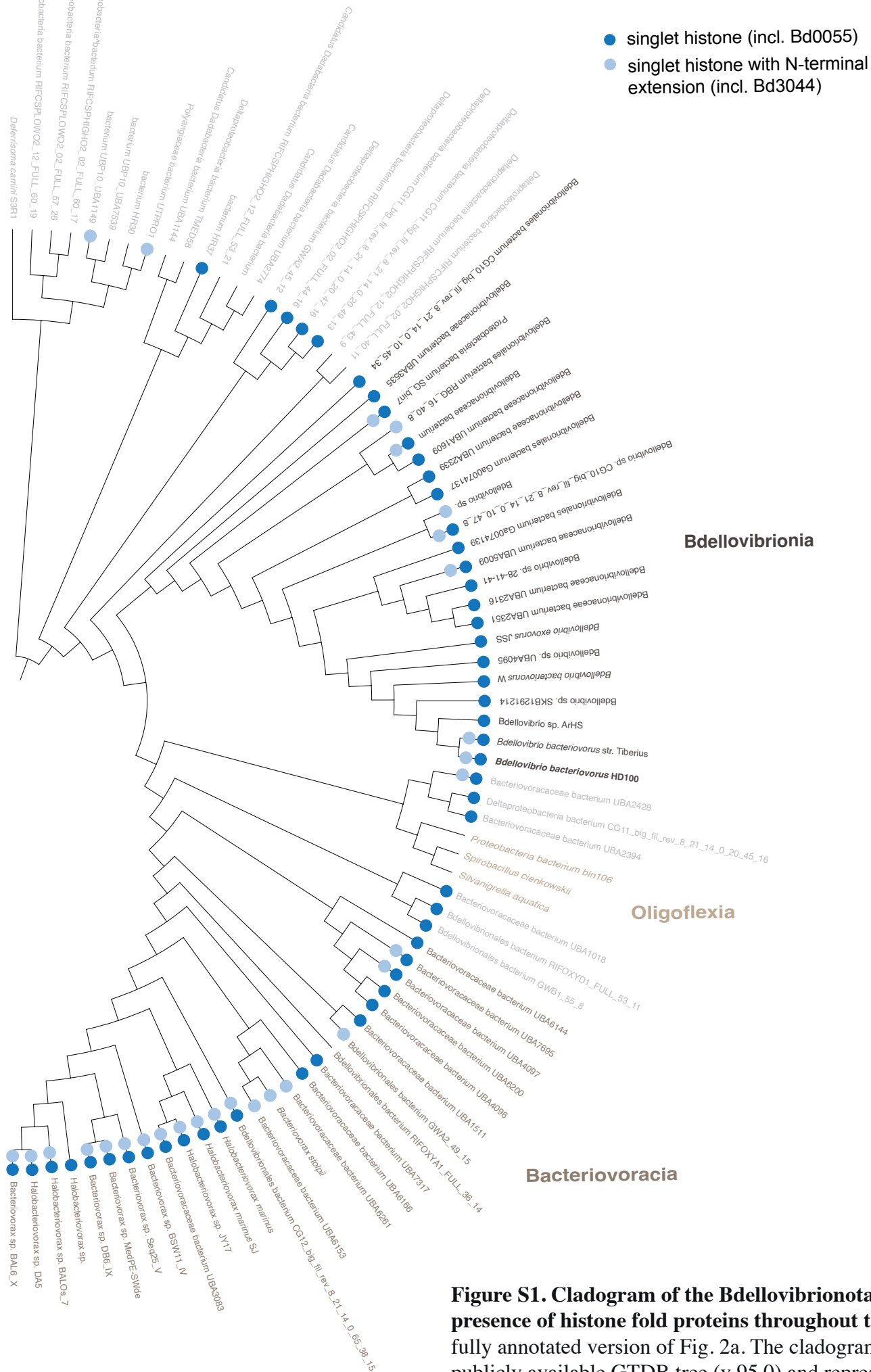
**b**



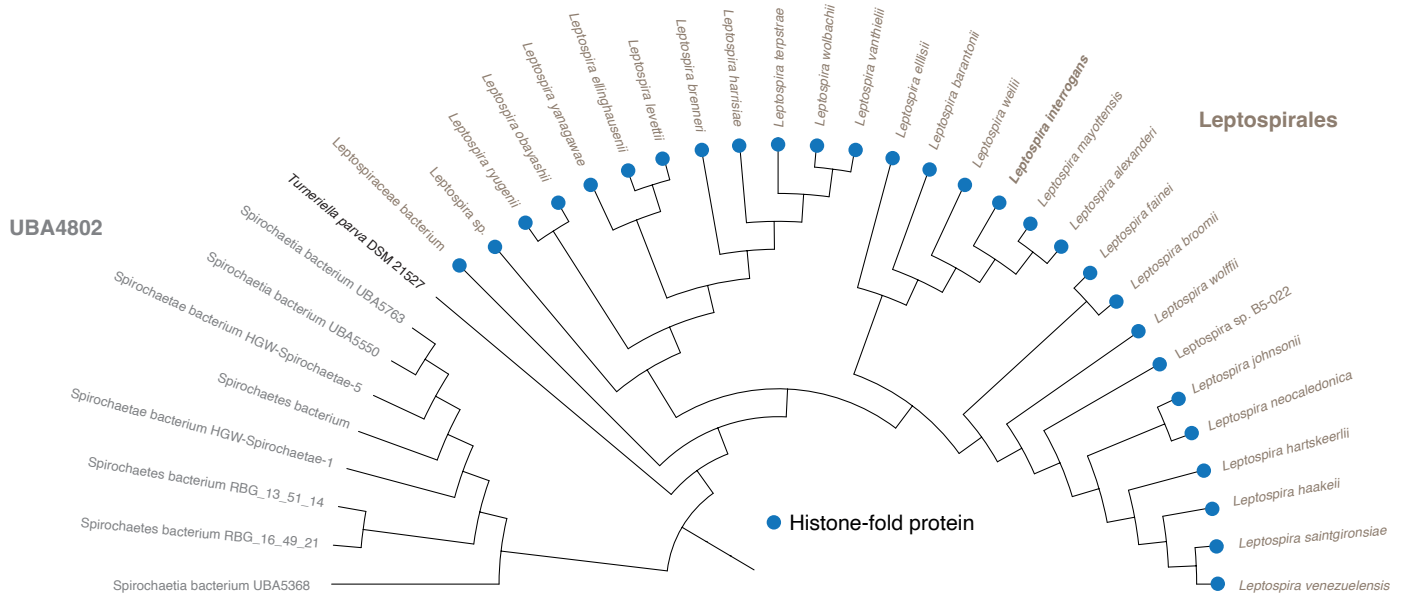
**c**



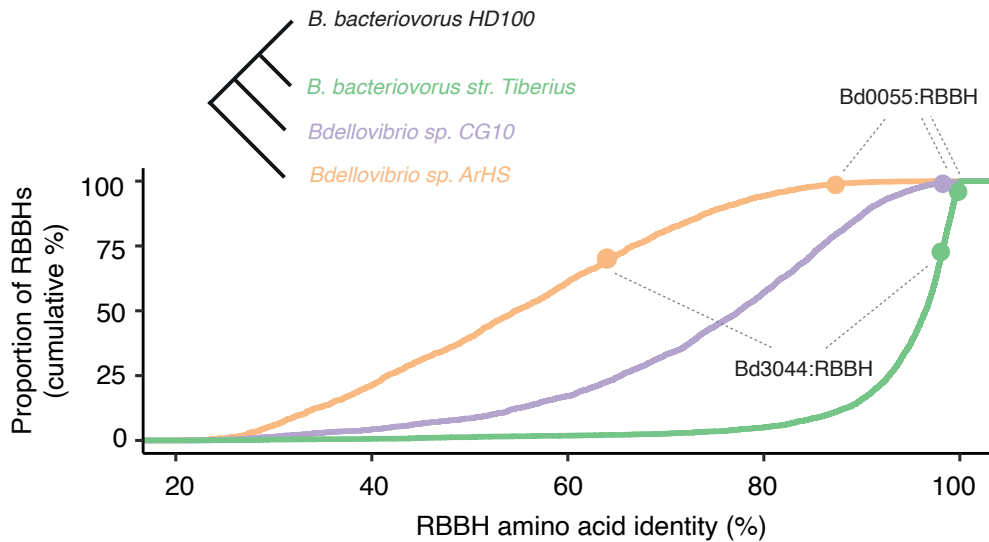
**Figure 6. Histones in other bacteria.** (a) High expression of histone fold proteins at the RNA and/or protein level is evident in bacteria from distant phylogenetic clades. Data sources: *L. interrogans* (Schmidt *et al.* 2011); *Waddlia chondrophila* (Beder and Saluz 2018); *Thermus thermophilus* (Swarts *et al.* 2015). (b) Phyletic distribution of singlet histones across the Leptospirales. (c) High abundance and prominent nucleoid enrichment of the *L. interrogans* histone fold protein (Uniprot ID: A0A2H1XGH2, gene ID: LA\_2458).



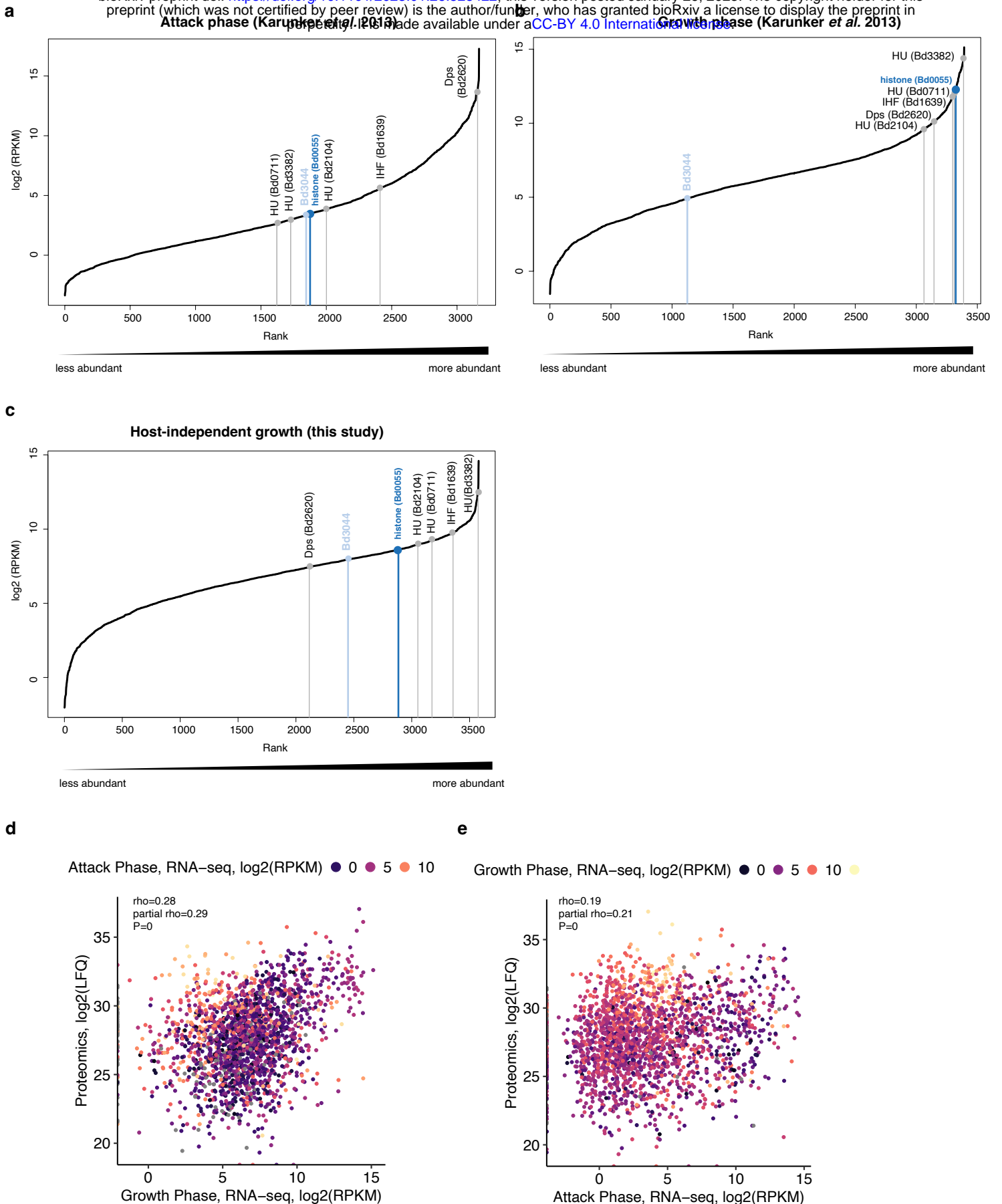
**Figure S1. Cladogram of the Bdellovibrionota, indicating the presence of histone fold proteins throughout the clade.** This is a fully annotated version of Fig. 2a. The cladogram is based on the publicly available GTDB tree (v 95.0) and represents the subset of genome assemblies that were used in this study and are also present in GTDB.



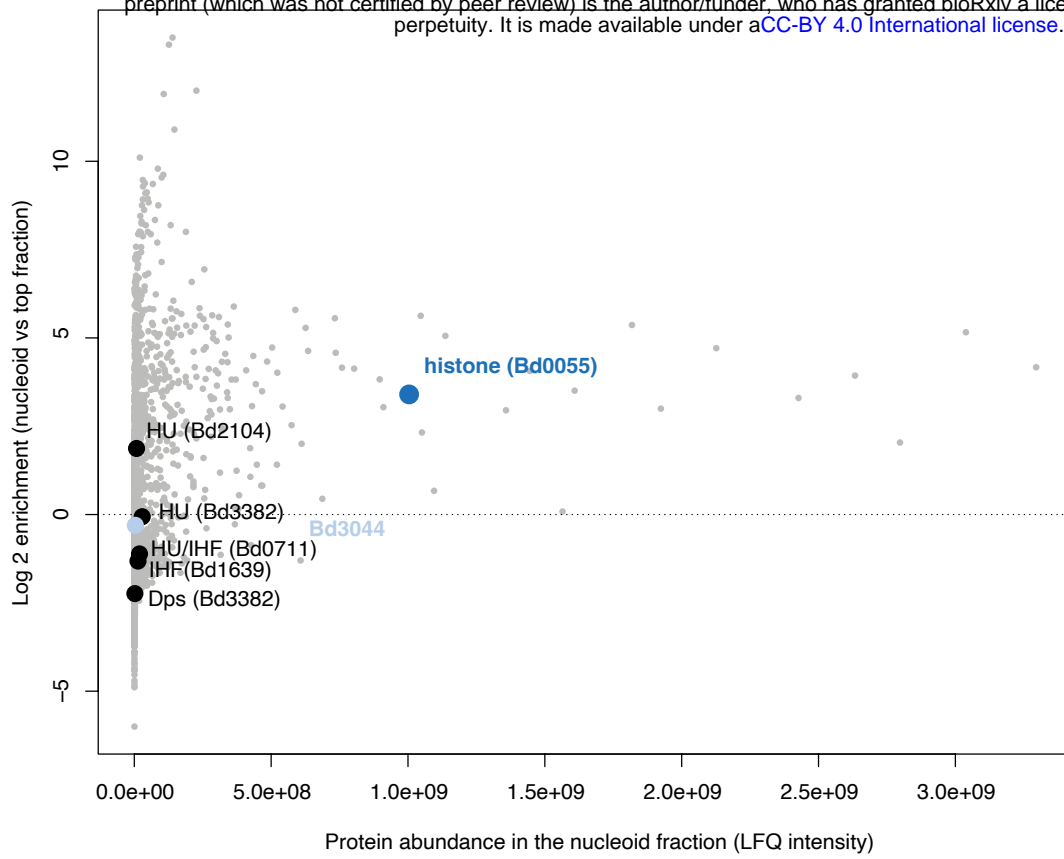
**Figure S2. Cladogram of the class Leptospirales and its UBA4802 sister clade, indicating the presence of histone fold proteins throughout the clade.** The cladogram is based on the publicly available GTDB tree (v 95.0) and represents the subset of genome assemblies that were used in this study and are also present in GTDB.



**Figure S3. Conservation of Bd0055 at the amino acid level in the context of other proteins encoded by *B. bacteriovorus*.** Homologs of proteins in *B. bacteriovorus* HD100 were identified in three other *Bdellovibrio* genomes of different divergence levels (*B. bacteriovorus* strain Tiberius, *Bdellovibrio* sp. CG10, and *Bdellovibrio* sp. ArHS) using a reciprocal best blast hit (RBBH) approach (see Methods). For each pairwise comparison, proteins were aligned and then ranked by amino acid identity. Bd0055 ranks amongst the most highly conserved proteins at different levels of proteome divergence. Bd3044 is comparatively less well conserved and absent from *Bdellovibrio* sp. CG10 (assembly ID: GCA\_002773975.1).

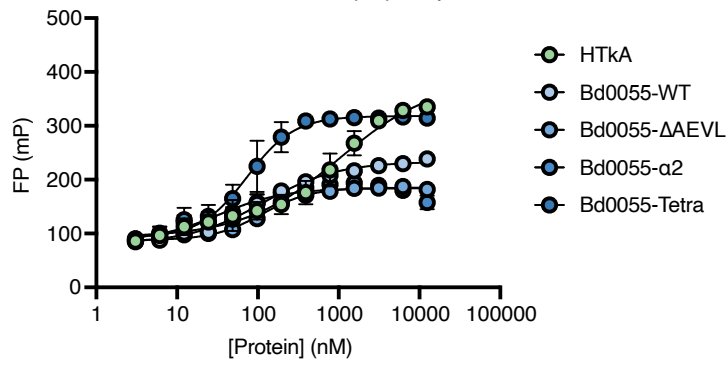


**Figure S4.** Ranked abundance of *B. bacteriovorus* transcripts during (a) attack phase, (b) growth inside the *E. coli* host, and (c) host-independent growth (strain HID13), highlighting transcript levels of Bd0055 relative to homologs of classic bacterial NAPs encoded in the **B. bacteriovorus** genome. (d, e) Correlation between protein abundance in attack phase cells (this study, whole cell extract) and RNA abundance during attack phase and growth inside the host (from Karunker *et al.* 2013).



**Figure S5.** Protein abundance and nucleoid enrichment of Bd0055 in *B. bacteriovorus* in the context of “classic” bacterial NAPs, for which homologs were detected in the *B. bacteriovorus* HD100 genome.

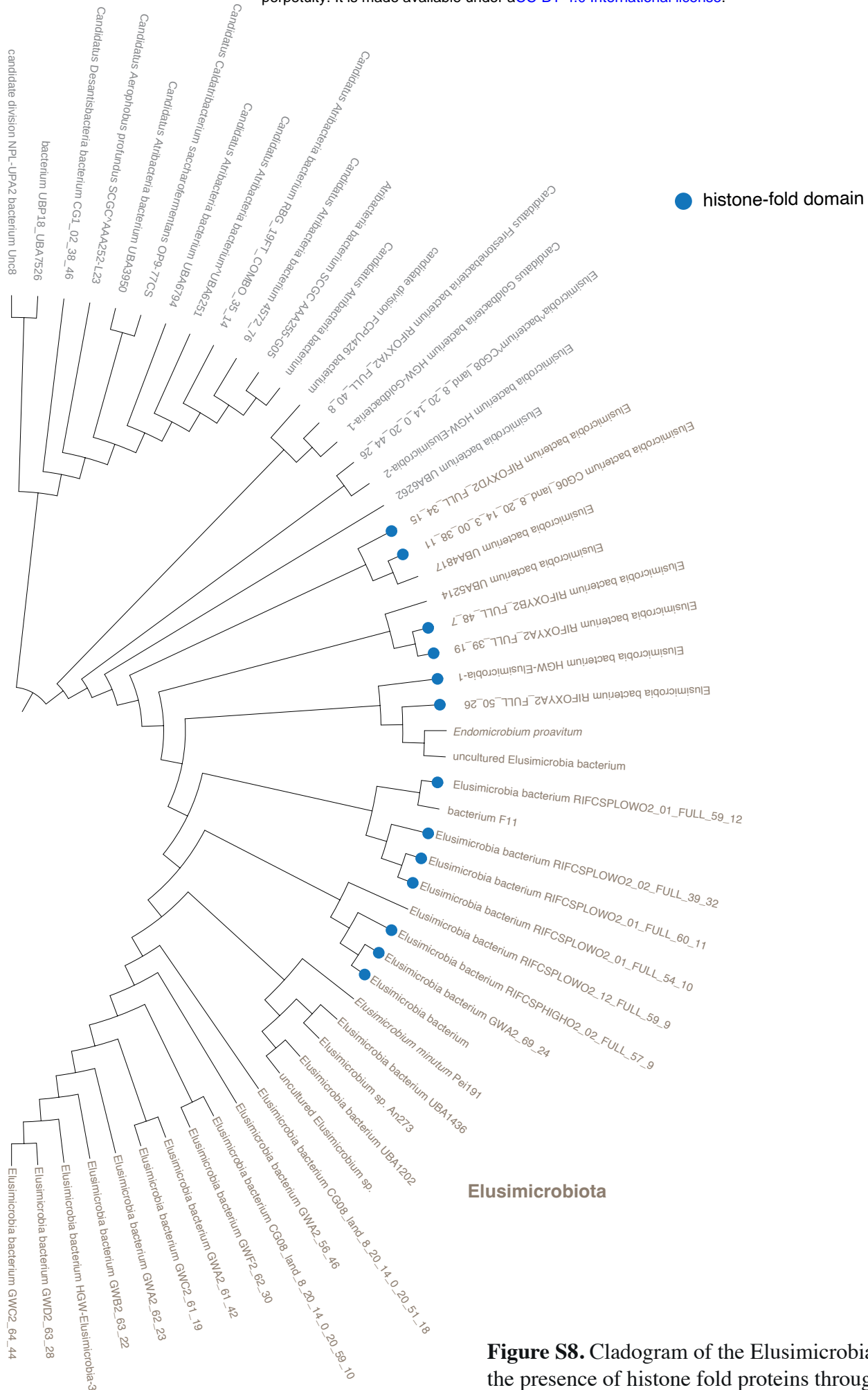




	HTkA	Bd0055-WT	Bd0055-ΔAEVL	Bd0055-α2	Bd0055-Tetra
IC50	1011	117.1	113.0	32.76	73.87
HillSlope	0.7363	0.9448	1.425	1.084	1.490

**Figure S6.** Fluorescence polarization (FP) of Bd0055 mutants binding to the same DNA as in Fig. 5d. FP is monitored using the Alexa488 flourophore (n=3). Error bars represent standard deviation of the mean.

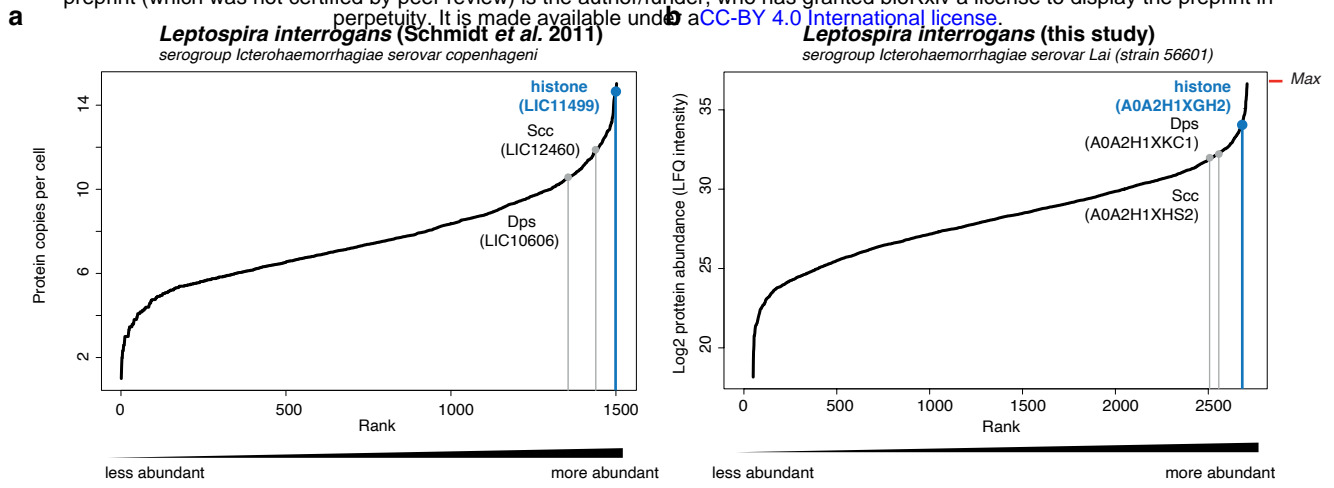




**Figure S8.** Cladogram of the Elusimicrobia, indicating the presence of histone fold proteins throughout the clade. The cladogram is based on the publicly available GTDB tree (v 95.0) and represents the subset of genome assemblies that were used in this study and are also present in GTDB.

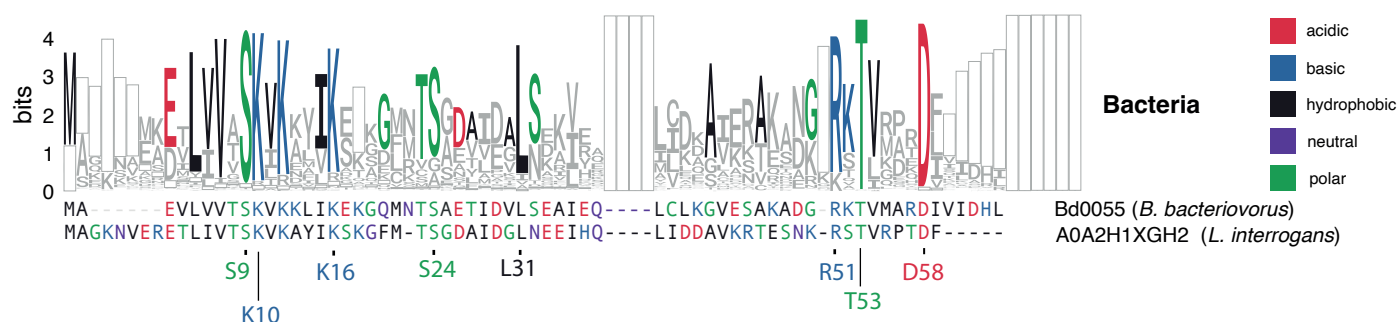




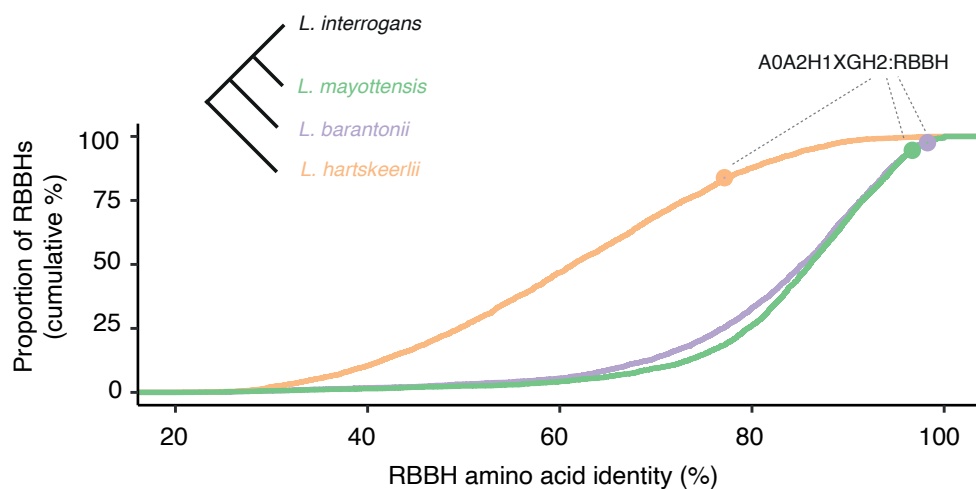


**Figure S11.** Ranked abundance of *L. interrogans* proteins from (a) Schmidt *et al.* (2011) and (b) this study, highlighting the abundance of the *L. interrogans* histone in the context of other known NAPs in this species. All quantified proteins are plotted.

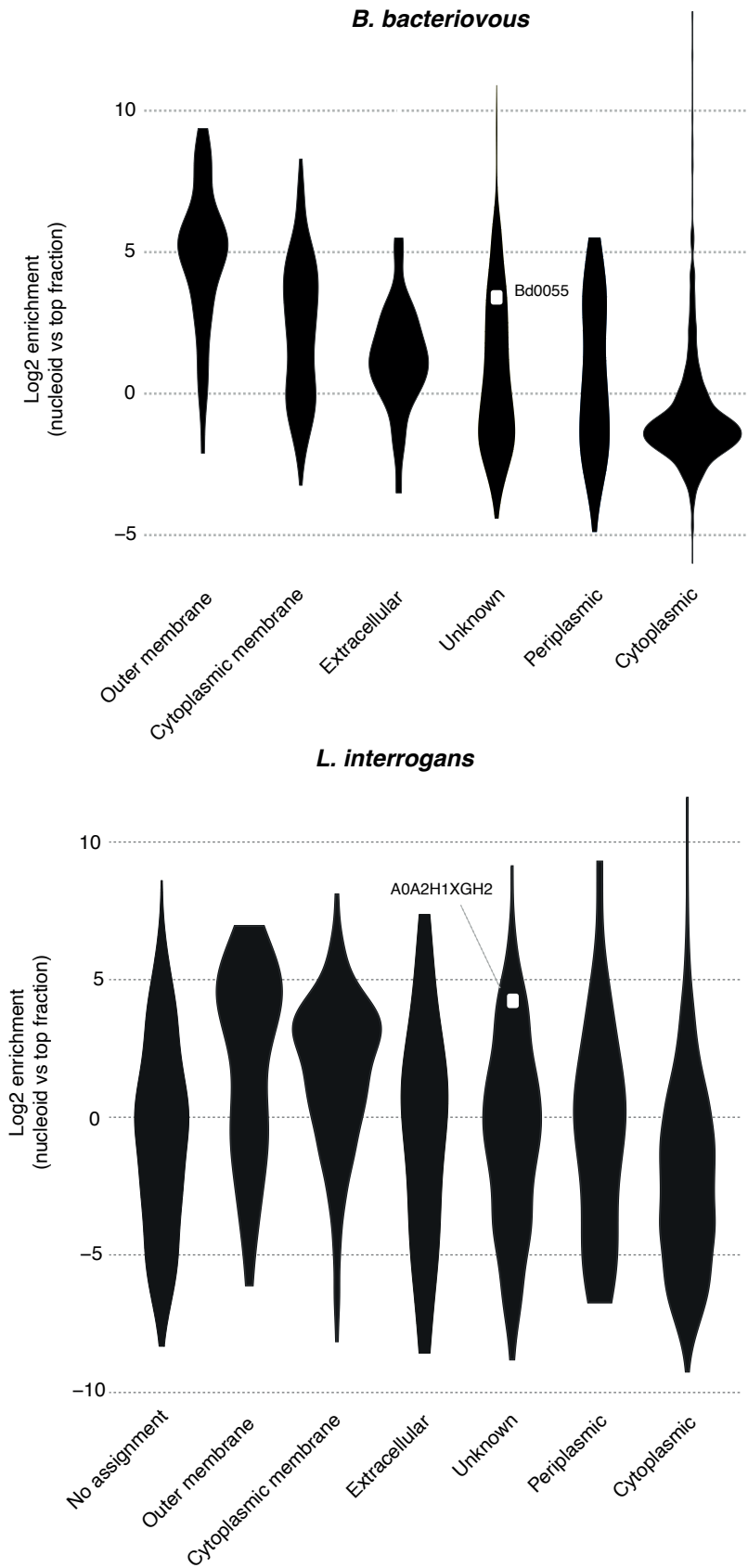
**a**



**b**



**Figure S12. (a)** Alignment of Bd0055 (*B. bacteriovorus*) and A0A2H1XGH2 (*L. interrogans*) in the context of overall amino acid usage across bacterial singlets (see Fig. 1c). **(b)** Conservation of the *L. interrogans* singlet histone (A0A2H1XGH2) at the amino acid level in the context of other proteins encoded by *L. interrogans*. Homologs of proteins in *L. interrogans* were identified in three other *Leptospira* genomes of different divergence levels (*L. mayottensis*, *L. barantonii*, *L. hartskeerlii*) using a reciprocal best blast hit approach (see Methods). For each pairwise comparison, proteins were aligned and then ranked by amino acid identity. A0A2H1XGH2 ranks amongst the most highly conserved proteins at different levels of proteome divergence.



**Figure S13.** Nucleoid enrichment of proteins in *B. bacteriovorus* and *L. interrogans* as a function of their cellular localization as predicted by PSORTdb (<https://db.psort.org/>).



**Table S2. X-ray diffraction data collection and refinement statistics (MR).**

	Bd0055 apo	Bd0055 bound to 35 bp dsDNA
<b>Data collection</b>		
Space group	P 2 21 21	I 21 21 21
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	26.7722 34.4118 56.484	33.081 103.143 111.145
$\alpha$ , $\beta$ , $\gamma$ (°)	90 90 90	90 90 90
Resolution (Å)	29.39 - 1.8 (1.865 - 1.8)	27.79 - 2.0 (2.072 - 2.0)
<i>R</i> <sub>merge</sub>	0.05163 (0.1029)	0.06045 (0.5857)
<i>I</i> / $\sigma$ <i>I</i>	26.79 (10.42)	16.49 (3.86)
Completeness (%)	99.42 (95.25)	98.37 (97.89)
Redundancy	8.4 (5.2)	7.1 (7.4)
<b>Refinement</b>		
No. reflections	5,156 (461)	13,136 (1,302)
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.1889 (0.2095) / 0.2331 (0.2670)	0.2218 (0.2320) / 0.2536 (0.2508)
No. atoms		
Macromolecules	476	1051
Ligand	0	0
Solvent	43	52
Protein residues	63	126
<i>B</i> -factors		
Macromolecules	11.88	46.07
Solvent	20.65	42.45
R.m.s. deviations		
Bond lengths (Å)	0.007	0.008
Bond angles (°)	0.82	0.92

Values in parentheses are for highest-resolution shell. One crystal was used for each structure.