# Hybrid analysis reveals how DNA sequence governs genomic location and DNA contacts of bacterial chromatin H-NS filaments

Beth A. Shen,[1] Christine M. Hustmyer[1], Daniel Roston[1,2], Michael B. Wolfe[1], Eric D. Jessen[1], and Robert Landick[1,3]

[1]Department of Biochemistry, University of Wisconsin–Madison, Madison, WI 53706
[2]Current address: Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA 92093
[3]Department of Bacteriology, University of Wisconsin–Madison, Madison, WI 53706

*To whom correspondence should be addressed; Email: landick@bact.wisc.edu

**Keywords:**

bacterial chromatin ; H-NS ; ChIP-seq ; FeBABE ; MD simulation ; filament ; gene regulation

**Abbreviations:**

DBD – DNA binding domain ; DBP – DNA binding protein ; FeBABE - Fe(III) (S)-1-(p-Bromoacetamidobenzyl)ethylenediaminetetraacetate ; Fe·EDTA–H-NS – H-NS derivatized with a the $Fe^{2+}$·EDTA nuclease ; MD – molecular dynamic ; ·OH – hydroxyl radical ; RNAP – RNA polymerase

## ABSTRACT

Gene silencing in bacteria is mediated by chromatin proteins, of which *Escherichia coli* H-NS is a paradigmatic example. H-NS forms multimeric nucleoprotein filaments with either one or two DNA duplexes, but the structures of these linear or bridged filaments remains unknown. Neither the determinants of H-NS genomic distribution nor the arrangements of H-NS DNA binding domains (DBDs) in linear or bridged filaments are currently defined. To address these questions, which are central to understanding gene silencing in bacteria, we developed a hybrid approach that combines a new method for *in vitro* chromatin reconstitution and mapping (Nitro-seq), ChIP-seq, nuclease-tagged DBDs to map DNA contacts, ·OH footprinting, molecular dynamics, and robust bioinformatic analyses. We find that DNA sequence principally governs genomic location of all filament types with indistinguishable sequence specificity, that DBD–DNA contacts vary in orientation and precise position with ~10-bp spacing, and that unchanging contacts in linear and bridged filaments support a hemi-sequestration model of linear-to-bridged filament switching.

## INTRODUCTION

Extensive study of nucleosome-based chromatin structure has provided broad insight into mechanisms of chromatin-mediated gene regulation in eukaryotes (Cook and Marenduzzo, 2018; Luger et al., 1997; Talbert et al., 2019). For bacteria, however, an analogous structural understanding remains elusive in part due to the diverse and complex DNA-binding patterns of the multiple DNA-binding proteins (DBPs) that generate bacterial chromatin (Dame et al., 2020; Grainger, 2016; Shen and Landick, 2019). The γ-proteobacterial histone-like nucleoid structuring protein (H-NS; 15 kDa monomer) provides a paradigmatic example of this complexity. H-NS, which contains two oligomerization sites in its N-terminal domain (NTD) and a C-terminal DNA binding domain (DBD), forms gene-silencing, oligomeric filaments on AT-rich DNA, notably on pathogenicity operons, surface-antigen operons, and horizontally acquired genes (Grainger et al., 2006; Kahramanoglou et al., 2011; Lucchini et al., 2006; Navarre et al., 2006).

H-NS oligomerizes on DNA through head-head and tail-tail NTD contacts (Arold et al., 2010; Gordon et al., 2011; Ulissi et al., 2014) into filaments that adopt either a bridged or linear conformation, in which H-NS interacts with two or one segments of duplex DNA, respectively (Kotlajich et al., 2015; Liu et al., 2010). The DBD forms an AT-hook that binds the narrow minor groove of AT-rich DNA (Gordon et al., 2011). A flexible linker connecting the DBD and NTD enables DBD-DNA contacts (Gao et al., 2017; Gulvady et al., 2018; van der Valk et al., 2017). Both bridged and linear filaments inhibit transcription initiation by RNA polymerase (RNAP) by occluding promoters, whereas bridged but not linear filaments can inhibit transcript elongation by RNAP through a topological trapping mechanism (Boudreau et al., 2018; Kotlajich et al., 2015; Lim et al., 2012). However, neither the overall structure nor the rules governing DBD locations for either filament type are known.

The extent to which DNA sequence alone *versus* competing effects of transcription and other DPBs control the location of the ~0.2-20 kb H-NS filaments that cover ~15% of the *E. coli* K-12 genome is unclear (Kahramanoglou et al., 2011; Noom et al., 2007). Various DNA sequence features, including a 10-bp high-affinity motif (TCGATAAATT), A-tracts, AT-content, and TA dinucleotide steps, have been proposed to be features of H-NS binding sites

(Bouffartigues et al., 2007; Gordon et al., 2011; Gulvady et al., 2018; Japaridze et al., 2017; Lang et al., 2007; Owen-Hughes et al., 1992; Riccardi et al., 2019; Yamada et al., 1991). However, both active transcription and other DBPs are known to impact H-NS binding patterns. Although H-NS can block RNAP (Dole et al., 2004; Kotlajich et al., 2015; Lim et al., 2012; Singh et al., 2014), RNAP also can displace H-NS at some loci (Chandraprakash and Seshasayee, 2014; Rangarajan and Schnetz, 2018). Additionally, DBPs can alter H-NS filament locations *in vivo* both by displacing and by promoting H-NS binding, for instance in anti-silencing mechanisms (Will et al., 2015). A more complete understanding of H-NS-mediated gene regulation requires knowing the extent to which the genome-wide patterns of H-NS binding are driven by DNA sequence alone *versus* by active transcription and by other DPBs.

It is unclear how DPBs are arranged within H-NS filaments and when filaments switch from linear to bridged conformations. DNase and ·OH footprinting experiments reveal nucleation at a high-affinity sites and H-NS oligomerization with an apparent periodicity of DBD contacts every ~10 bp (Bouffartigues et al., 2007; Ulissi et al., 2014; Will et al., 2018). However, 20- or 30-bp periodicities of DBD footprints also are reported, and a prevalent model for H-NS filament organization based on an NTD crystal structure suggests a 30-bp spacing of DBDs (Arold et al., 2010; Dame et al., 2006). Given the flexible linker connecting the DBD, a key question is whether DBDs necessarily follow a regular arrangement within H-NS filaments. New approaches are needed to determine how DNA sequence governs DBD location and the arrangement of DBDs in an H-NS filament.

Illustrating the complexity of bacterial chromatin, the properties of H-NS nucleoprotein filaments in *E. coli* K-12 are modulated by the H-NS paralog, StpA, by small Hha/YmoA-family regulators (*e.g.*, Hha), and by solutes (Dame et al., 2020). StpA forms heterodimers with H-NS and Hha binds to the NTD of H-NS, but not to DNA (Ali et al., 2013; Leonard et al., 2009); both proteins stimulate bridging by H-NS (Boudreau et al., 2018; van der Valk et al., 2017; Wang et al., 2014). Both Hha and StpA enhance silencing at a subset of H-NS regulated genes, but it is unclear whether or how these mixed filaments differ from H-NS in their DNA-binding properties (Banos et al., 2009; Lucchini et al., 2009; Muller et al., 2006; Nieto et al., 2000; Srinivasan et al.,

2013; Uyar et al., 2009). ChIP-seq results suggest that H-NS, StpA, or Hha bind similar locations in the genome (Srinivasan et al., 2013; Uyar et al., 2009), but their relative DNA-sequence specificities and why StpA and Hha affect H-NS-mediated gene regulation differentially remain to be explained.

To gain insight into the determinants of H-NS filament formation and DBD locations in filaments, we have developed and applied a hybrid approach that combines a new method to reconstitute and map using HT-sequencing bacterial chromatin *in vitro* (Nitro-seq), a novel DBD-mapping method using a tethered chemical nuclease, hydroxyl radical (·OH) footprinting, ChIP-seq, computational molecular dynamics, and rigorous statistical analysis and modeling. Our results reveal that DNA sequence, rather than transcription or other DBPs, principally guide the pattern of gene silencing *in vivo* and that DBDs contact DNA irregularly in an H-NS filament in patterns unaffected by Hha or by bridging *versus* linear conformations.

## RESULTS

### H-NS filament locations are not principally determined by transcription

Although transcribing RNAP can influence H-NS filaments by either displacing H-NS bound at a promoter (Rangarajan and Schnetz, 2018) or more extensively when ρ is inhibited genome-wide (Chandraprakash and Seshasayee, 2014), the extent to which these effects govern H-NS filament locations across the genome is unclear. To address this question, we first used ChIP-seq to determine RNAP and H-NS locations in the presence or absence of the transcription inhibitor rifampicin (Rif; **Figure 1**). Rif prevents extension of initial RNA transcripts past 3 nt, causing RNAP to relocate *in vivo* to promoters or promoter-like sequences (Campbell et al., 2001; Herring et al., 2005).We hypothesized that if transcribing RNAPs strongly contribute to H-NS filament patterns in cells, then trapping all RNAP at promoters with Rif would alter H-NS locations detected by ChIP-seq relative untreated cells (**Figure 1A**). We treated wild-type *E. coli* K-12 cells (strain RL3000; Ghosh and Landick, 2016) with Rif during early exponential phase growth in minimal glucose medium and then generated ChIP-seq samples from treated and untreated cells using antibodies specific for RNAP or H-NS. Comparison of the RNAP ChIP patterns in the presence or absence of Rif revealed the expected Rif-induced redistribution of RNAP to discrete binding sites consistent with RNAP trapping at known promoters throughout the genome (**Figure 1B**; **Figure 1 – figure supplement 1**; Herring et al., 2005; Kim et al., 2012; Thomason et al., 2015). The majority (72%) of identifiable RNAP peaks in Rif-treated cells are within 30 bp of a known promoter based on a genome-scale transcription start site (TSS) analysis (Thomason et al., 2015), with a mean distance of ~10 bp vs. ~140 bp for a randomized control ($p < 10^{-4}$; **Figure 1 – figure supplement 2**). In contrast, the H-NS binding pattern exhibited only minimal changes upon Rif treatment.

Visual inspection suggested that Rif inhibition of transcription had little effect on H-NS filaments. To detect less obvious changes, we used MOSAiCS to define filaments and compare those in Rif-treated *versus* untreated cells (Kuan et al., 2011). MOSAiCS identifies enriched regions in the immunoprecipitated (IP) sample relative to the DNA sample prior to IP. We chose MOSAiCS because it can identify broad enriched regions typical of filament-forming chromatin

proteins like H-NS or active transcription units in addition to the symmetrical Gaussian-like peaks typical for single-site DNA-binding proteins. From duplicate samples, MOSAiCS revealed 301 filaments in the untreated cells and 289 filaments in Rif-treated cells (**Figure 1C**). For both conditions, filament lengths averaged ~2 kb and ranged from 0.2-16 kb (**Figure 1C**, Table 1). The H-NS binding patterns in our untreated cells are similar to those reported using ChIP-chip in strain MG1655 growing in the same medium (Myers et al., 2013) and to the ChIP-seq patterns reported for flag-tagged H-NS in MG1655 growing in rich medium (LB 0.5% NaCl; **Figure 1 – figure supplement 1B**; Kahramanoglou et al., 2011).

To determine more precisely whether halting transcription changes H-NS filament length, we determined the extent of overlap of filaments present at the same genomic location in untreated and Rif-treated cells. Only ~10% of the filament regions were present in one condition and not the other (**Figure 1C**; see Methods). To classify similarities and differences, we defined four types of filament changes upon Rif treatment: (*i*) unchanged; (*ii*) a filament with boundaries that receded (2b) or expanded (2a) by more than 200 bp (~10% of the ~2 kb average filament length); (*iii*) a new filament; or (*iv*) dissociation of an existing filament. For any region differentially bound by H-NS between the two datasets, differences in the average ChIP density of that region were used to confirm the change in H-NS binding (see Methods). Class 1 (filaments unchanged by Rif treatment) included ~80% of all filaments present in untreated cells (265 of 301) (**Figure 1C**). Even when transcription was directly adjacent to an H-NS filament, such in *cysHIJ* adjacent to the H-NS-bound *cas* operon, H-NS filaments were not dramatically altered upon Rif treatment (**Figure 1B**). However, Rif treatment did affect the lengths or existence of 43 of the H-NS filaments (**Figure 1C**). Rif treatment caused changes at one or both ends of some filaments (class 2; $n$=27), appearance of a few filaments (class 3; $n$=3), and disappearance of other filaments (class 4; $n$=12) (**Supplemental Table S5**).

We next investigated whether changes in RNAP binding explained the changes in H-NS filaments by identifying H-NS filaments present only with or without Rif treatment. We hypothesized that upon Rif treatment, relocation of RNAP to promoters could either disrupt an existing H-NS filament or expose AT-rich sequences in transcribed regions to H-NS binding. Of

12 filaments that disappeared in Rif-treated cells (class 4), 9 overlapped a Rif-induced RNAP-binding site (**Supplemental Table S5**). For example, displacement of H-NS at the *dps* promoter was associated with binding of RNAP in Rif-treated cells (**Figure 1D**), consistent with known H-NS regulation of the *dps* promoter (Grainger et al., 2008). Similarly, of the 3 filaments that appeared with Rif treatment, 2 overlapped a location vacated by RNAP (**Supplemental Table S5)**.

To test for more subtle changes, we examined the boundaries of filaments that grew or receded upon Rif treatment. Of 18 filaments that receded from one boundary (class 2b), 16 were associated with Rif-induced binding of RNAP at the filament boundary (**Supplemental Table S5**). Of 11 filaments that grew when transcription was blocked (class 2a), six were associated with a region vacated by RNAP after Rif treatment. Together, these results indicate that active transcription plays a minimal role in determining locations of H-NS binding in *E. coli*, although small effects can occur at a limited number of sites where effects on gene expression could be significant.

**H-NS nucleoprotein filaments can be reconstituted using genomic DNA *in vitro***

Even though active transcription does not principally determine H-NS filament locations, it remained possible that other DBPs could play dominant roles in guiding H-NS filament formation *in vivo*, either by exclusion or via cooperative interactions. To test this possibility, we sought to reconstitute bacterial chromatin *in vitro* using purified proteins and chromosomal DNA to compare to genomic locations *in vivo*. This type of reconstitution has been achieved with eukaryotic chromatin and has proven informative (Krietenstein et al., 2016).

As a first step toward *in vitro* reconstitution of bacterial chromatin, we compared H-NS filaments formed on purified, deproteinized genomic DNA (gDNA) to *in vivo* H-NS binding patterns determined by ChIP-seq. We predicted that H-NS filaments would nucleate at AT-rich, high-affinity sites (Bouffartigues et al., 2007; Japaridze et al., 2017; Lang et al., 2007) but that strong contributions of other DBPs to H-NS filament locations would make the *in vitro* and *in vivo* patterns different. To define locations of H-NS filament formation on sheared gDNA, we captured the reconstituted filaments by binding to nitrocellulose filters, and then determined

filament locations by HT-sequencing of DNA recovered from the filters. We call this method Nitro-seq (**Figure 2A**).

We optimized conditions for Nitro-seq by identifying an H-NS–DNA ratio that preferentially captured AT-rich DNA. In *E. coli* cells, H-NS binds ~15% of the chromosomal DNA (**Figure 2 – figure supplement 1**; (Grainger et al., 2006; Kahramanoglou et al., 2011; Myers et al., 2013; Navarre et al., 2006). Thus, we varied H-NS–DNA ratios using $^{32}$P-labeled PCR fragments that included AT-rich fragments from the *bgl* operon and GC-rich fragments from pBR322 in the presence of sheared gDNA (avg. size, 880 bp; ~3.7 μM bp·mL$^{-1}$ equivalent to ~500 μg·mL$^{-1}$; **Supplemental Table S3**). In our binding conditions, 100 nM and 200 nM H-NS (31 and 64 H-NS/kb, respectively), retained AT-rich fragments, but not GC-rich fragments (**Figure 2 – figure supplement 1**). These H-NS–DNA ratios are similar to those used to form bridged H-NS filaments (Kotlajich et al., 2015). By including different length PCR fragments in the test, we verified that short (<1 kb) fragments suitable for direct HT-sequence library preparation supported H-NS filament formation. We found that longer fragments were retained at slightly lower H-NS concentrations than shorter filaments (*e.g.,* 50 nM H-NS; 15 H-NS/kb; **Figure 2 – figure supplement 1**). We choose 100 and 200 nM H-NS (31 and 64 H-NS/kb DNA) for subsequent experiments because these conditions gave selective retention of AT-rich fragments short enough for direct HT-sequencing as well as ~15% of gDNA (**Figure 2 – figure supplement 1**).

To interpret the Nitro-seq results, we compared retention of deproteinized input DNA without H-NS to the protein-bound samples. The sheared input gDNA exhibited a uniform distribution of sequence reads prior to nitrocellulose filter-binding and far lower amounts of retained DNA compared to protein-bound samples (~1.5% DNA retained without protein; **Figure 2 – figure supplement 3**). Addition of H-NS to the sheared gDNA yielded nitrocellulose retention of 4-12% of input DNA (**Figure 2 – figure supplement 1B**) with reproducible and distinct patterns of enrichment (peaks) across the genome (**Figure 2B; Figure 2 – figure supplement 4**). MOSAiCS analysis identified 407 H-NS binding locations at 100 nM H-NS (average of ~1400 bp long) and similar results for 200 nM H–NS (**Table 1**). Importantly, the H-

NS binding pattern was only ~20% similar to the DNA binding patterns generated by Fis (Jaccard statistic; see Methods), which binds AT-rich DNA but exhibits a different binding pattern than H-NS on chromosomal DNA *in vivo* (Cho et al., 2008; Kahramanoglou et al., 2011; **Figure 2 – figure supplement 4**; **Figure 3 – figure supplement 1**). These results verify that Nitro-seq identifies specific patterns of protein-DNA interactions and is suitable for tracking *in vitro* reconstitution of bacterial chromatin.

**The *in vivo* pattern of H-NS filaments is primarily determined by DNA sequence alone.** To ask if DNA sequence-specific binding by H-NS alone explains the distribution of H-NS *in vivo*, we compared the binding pattern of H-NS in our *in vitro* Nitro-seq assay to *in vivo* ChIP-seq results (**Figure 1**). The *in vivo* ChIP-seq and *in vitro* Nitro-seq patterns exhibited ~60% similarity (**Figure 2B, Figure 3B**), although 112 more H-NS filaments were identified *in vitro* than *in vivo* (**Figure 3B**, pink dots; **Table 1**). Of the 297 filaments that overlapped, most (255) overlapped for ≥90% of the length of the Nitro-seq filament. Overall, we conclude the *in vitro* and *in vivo* H-NS filaments are highly similar (**Figure 3B**).

We hypothesized the 112 *in vitro*-only filaments could be explained by competition for H-NS binding *in vivo* from other DBPs. To test this idea, we compared the 112 Nitro-seq-specific regions to known transcription factor (TF) binding sites from RegulonDB (Santos-Zavaleta et al., 2019), extended protein occupancy domains (EPODs) indicative of gene silencing by chromatin proteins like H-NS (tsEPODs), or highly expressed EPODs (heEPODs) (Vora et al., 2009) (**Figure 3 – figure supplement 1D**). If the *in vitro*-only (Nitro-seq-specific) regions were bound by other TFs *in vivo*, we expected to see statistically significant similarity between *in vitro*-only H-NS filaments and TF-binding sites, but not between the overlapping filaments and TF sites. We found the 112 Nitro-seq-specific H-NS filaments were more similar to RegulonDB TF-binding sites than the 255 overlapping regions were, but the Nitro-seq-specific regions were not more similar to heEPODs than the overlapping regions (difference in jaccard similarity, $p<0.001$; **Figure 3 – figure supplement 1D**). Consistent with the role of H-NS in gene silencing, this comparison revealed a high similarity between tsEPODs in the 255 overlapping regions ($p<0.001$). From this analysis, we conclude that the 112 Nitro-seq specific region likely reflect

exclusion of H-NS binding by TF, but not RNAP, *in vivo*, consistent with the idea that some TFs act by counter-silencing H-NS and that transcription largely does not affect H-NS filaments (Will et al., 2015; **Figure 1**). Overall, however, our results indicate that the most H-NS filament locations in *E. coli* cells are determined by DNA sequence alone with a minority of potential filament locations blocked by other DBPs.

**H-NS and StpA exhibit indistinguishable DNA binding specificities**

The H-NS paralog StpA can replace H-NS in H-NS nucleoprotein filaments and promotes bridging (Boudreau et al., 2018; Uyar et al., 2009), but whether its DNA sequence specificity is identical to H-NS or differs is unknown. The H-NS modulator Hha binds the N-terminal oligomerization domain of H-NS filaments, also promotes bridging, and may contact DNA, but its effect on the DNA sequence specificity of filament formation is unknown. Both StpA and Hha appear to be required for silencing at subsets of H-NS locations *in vivo* (Srinivasan et al., 2013; Uyar et al., 2009), but their ChIP patterns generally resemble those of H-NS and do not explain their roles in silencing (Banos et al., 2009; Fukui et al., 2016; Nieto et al., 2000; Uyar et al., 2009).

To determine if their sequence specificities differ from H-NS-only filaments, we applied the Nitro-seq assay to StpA and Hha–H-NS filaments. Using optimized binding conditions and MOSAiCS analysis (see Methods), we found that H-NS, StpA, and Hha–H-NS produced virtually indistinguishable binding patterns on gDNA (**Figure 2B; Figure 2 – figure supplement 1; Figure 3; Figure 3 – figure supplement 1A; Table 1**). Comparison of both the average density of Nitro-seq reads and the similarity of MOSAiCS regions for each set revealed a high degree of similarity between 100 nM H-NS and 50 nM StpA (Spearman correlation = 0.984 with 73% similarity between MOSAiCS regions; **Figure 3C; Figure 3 – figure supplement 1A**). Although there were more filaments in 50 nM StpA, all of the regions bound by H-NS were also bound by StpA. Similarly, we observed a high degree of overlap between the 100 nM H-NS and Hha–H-NS data sets (Spearman coefficient = 0.993 with 86% similarity between MOSAiCS regions; **Figure 3D**; **Figure 3 – figure supplement 1A**). Because their Nitro-seq binding patterns were nearly equivalent, we conclude that StpA, Hha–H-NS and H-NS

exhibit indistinguishable sequence specificity. However, it is possible that changing conditions, such as temperature or solutes could reveal differences relevant to their effects on gene expression (see Discussion).

**An H-NS-tethered artificial nuclease can probe H-NS DBD locations in H-NS filaments**

Having established that sequence not transcription nor other DBPs primarily dictates H-NS, StpA, and Hha–H-NS filament locations *in vivo*, we next sought to determine how sequence dictated the locations of DNA-binding domains (DBDs) within an individual filament. Previous footprinting experiments report discrete H-NS binding sites on DNA (Bouffartigues et al., 2007; Ulissi et al., 2014; Will et al., 2018), but it remains unclear how H-NS is organized in filaments to generate the footprint patterns and if filament length or H-NS modulating factors, like Hha and $Mg^{2+}$, affect DBD arrangement. To investigate the organization of H-NS in filaments, we developed a tethered artificial nuclease approach in which the nuclease ($Fe^{2+}$·EDTA) is attached to the H-NS DBD at a specifically positioned Cys residue by reaction with FeBABE (Fe(III) (*S*)-1-(*p*-Bromoacetamidobenzyl)ethylenediaminetetraacetate) (**Figure 4A-B**). The tethered $Fe^{2+}$·EDTA generates hydroxyl radicals that cleave the DNA backbone in a restricted radius (~12 Å) around the chelated $Fe^{2+}$ (Cheal et al., 2009). The resulting cleavage pattern identifies nucleotides near the tethered nuclease and thus would define the location and orientation of the H-NS DBD in the nucleoprotein complex. Here, we refer to tethered nuclease-induced cleavage as "tethered cleavage" to distinguish it from cleavage caused by free $Fe^{2+}$·EDTA in solution in the traditional ·OH footprinting assay.

To tether $Fe^{2+}$·EDTA to the H-NS DNA-binding domain, we constructed and purified H-NS C21S E102C, which positioned a single Cys residue in H-NS near, but not interfering with, the DNA binding interface of H-NS (**Figure 4A**; Gordon et al., 2011). FeBABE was then reacted with the free thiol on H-NS with an overall ~50% conjugation yield (see Methods). As a control for any potential off-target change in H-NS due to the conjugation process, we treated Cys-free H-NS C21S with FeBABE alongside H-NS C21S E102C. The mock-treated and conjugated H-NS were tested using previously described gel-shift and *in vitro* transcription elongation assays using a segment of the H-NS silenced *bgl* operon (Boudreau et al., 2018; Kotlajich et al., 2015).

Both H-NS C21S and H-NS C21S E102C exhibited decreased binding at low [H-NS] after reaction with FeBABE (**Figure 4 – figure supplement 1**). Decreased binding of the treated C21S protein suggests the loss of binding did not result from conjugation of $Fe^{2+}$·EDTA *per se*, but rather that H-NS, which is known to lose activity during some types of purification, may be less than 100% active after reaction. Nonetheless, at ~2 μM both H-NS C21S and H-NS C21S E102C treated with FeBABE formed complexes with EMSA behaviors similar to filaments formed with untreated or wild-type proteins. We conclude that the conjugation reaction reduces the active fraction of H-NS, but that the activity of H-NS is not dramatically altered by the presence of $Fe^{2+}$·EDTA compared to non-conjugated protein.

We next assessed the ability of $Fe^{2+}$·EDTA-modified H-NS (Fe·EDTA–H-NS) to stimulate pausing by RNAP in bridged, but not linear, filaments as observed for wild-type H-NS (Boudreau et al., 2018; Kotlajich et al., 2015) (**Figure 4 – figure supplement 2**). Halted ECs were formed on a *bgl*-operon template before addition of H-NS and pausing was then assayed by addition of all 4 NTPs. Pseudo-densitometry plots (Boudreau et al., 2018) of a 16-min time point revealed similar effects of WT, C21S, or C21S E102C H-NS. Lower H-NS–DNA ratios (66-100 H-NS/kb), which form bridged filaments, stimulated pausing, whereas higher ratios (>200 H-NS/kb), which form linear filaments, did not (**Figure 4 – figure supplement 2A-C**). These effects were consistent with previous observations (Kotlajich et al., 2015). Filaments formed with either the mock-treated H-NS C21S or Fe·EDTA–H-NS C21S E102C at 200 or 250 H-NS/kb also stimulated pausing by RNAP (**Figure 4 – figure supplement 2D**). The shift in behavior of 200 C21S H-NS/kb to greater pause stimulation after mock treatment suggests a fraction of conjugated H-NS was inactivated by treatment such that a lower concentration of active H-NS yielded bridged filaments (Kotlajich et al., 2015). However, we were unable to raise H-NS concentrations high enough in transcription assays using 1.5 kb DNA templates to determine if Fe·EDTA–H-NS could be switched back to a linear conformation. Despite differences in total concentration required to form filaments, the effects of the Fe·EDTA–H-NS filaments on transcript elongation suggest that neither the presence of $Fe^{2+}$·EDTA nor the conjugation process interfere with the formation of functional bridged H-NS filaments.

We next used tethered cleavage to characterize the locations of H-NS DBDs in H-NS–DNA filaments formed on the downstream regulatory element (DRE) segment of the *bgl* operon. The 146-bp *bgl* DRE fragment contained the sequence 610-756 bp downstream of the *bgl* operon promoter ($P_{bgl}$) with known high-affinity H-NS binding sites (Dole et al., 2004), could accommodate ~10-15 H-NS monomers (assuming a binding site is 10-15 bp; Amit et al., 2003; Ulissi et al., 2014), and enabled resolution of multiple tethered-DBD cleavage events on either the template (T) or non-template (NT) $^{32}$P-end-labeled DNA strand by PAGE (**Figure 4C, Figure 4 – figure supplement 3**). $H_2O_2$ and ascorbic acid were used to generate free radicals by the Fenton reaction, which leads to phosphodiester backbone cleavage *via* hydrogen abstraction (Balasubramanian et al., 1998). The cleavage products were resolved by denaturing PAGE to identify locations of the H-NS DBDs (**Figure 4D**). At Fe·EDTA–H-NS above 2 μM (640 H-NS/kb), we observed robust cleavage of both the T and NT strands in multiple locations, indicating formation of a stable filament on the short *bgl* DRE fragment. Importantly, DNA cleavage was dependent on the presence of ascorbic acid, $H_2O_2$, and Fe·EDTA tethered to H-NS, indicating that the cleavage events report H-NS binding and not DNA degradation or free radical formation from untethered Fe·EDTA (**Figure 4D**, top panel).

To quantify Fe·EDTA-induced cleavage, peak areas for each position were determined (Das et al., 2005; see Methods). This analysis revealed regions of robust cleavage at specific positions spaced irregularly at an average interval of 11-12 nt along both strands of DNA (**Figure 4E and F;** NT, red arrows; T, blue arrows). The location of the tethered cleavage peaks generally remained unchanged up to 5000 H-NS/kb, although the intensity of cleavage increased at a subset of positions. We quantified these H-NS-dependent changes by calculating the change in normalized peak area as a function of H-NS concentration (**Figure 4 – figure supplement 4C-D**). We found sites with large, positive slopes indicating robust H-NS-dependent cleavage ~10-15 bp apart, which is consistent with previous H-NS footprint analyses using untethered probes on different DNA fragments in which regions of protected DNA occurred with ~10–20 bp periodicity (Bouffartigues et al., 2007; Dame et al., 2006; Ulissi et al., 2014; Will et al., 2018).

Finally, we confirmed the pattern of Fe·EDTA-induced cleavage resulted from H-NS binding to specific sequences and was not dictated by phasing from H-NS interactions at or near the DNA ends. We tested this possibility by comparing Fe·EDTA–H-NS cleavage patterns on DNA fragment with different DNA ends, the 146-bp *bgl*-DRE fragment described above and a 151-bp fragment of identical internal sequence to which 5 bp were added to one end (**Figure 4 – figure supplement 5**). The identical cleavage patterns as a function of sequence not end location established that the patterns resulted from H-NS sequence-specific binding, as expected (Bouffartigues et al., 2007), rather than being set by the position of the DNA end.

**The H-NS nuclease cleavage pattern differs from the hydroxyl radical footprint pattern**

We next confirmed the cleavage pattern was caused by the location of the Fe·EDTA–H-NS nuclease and not by the intrinsic susceptibility of different DNA sequences to ·OH cleavage, since some sequence bias in ·OH cleavage patterns on DNA is well-documented (Burkhoff and Tullis, 1988). To define the intrinsic ·OH cleavage pattern of the DRE, we performed conventional ·OH footprinting with untethered Fe·EDTA and unmodified WT H-NS to characterize the *bgl*-DRE fragment (**Figure 5A; Figure 5 – figure supplement 1**). Using the control condition lacking H-NS, we compared the untethered Fe·EDTA-induced cleavage patterns to the tethered-cleavage patterns and found that untethered cleavage pattern exhibited cleavage biases much different than the cleavage patterns obtained with Fe·EDTA–H-NS (**Figure 5 – figure supplement 2**).

We also used the footprinting data to test whether the tethered cleavage pattern differed from a conventional ·OH footprinting pattern of H-NS (*i.e.,* protection against cleavage by H-NS). WT H-NS filaments protect nearly all positions in the *bgl*-DRE DNA to varying degrees (**Figure 5 – figure supplement 3A**). To identify the bases most robustly protected by H-NS, we calculated the change in the peak area as a function of H-NS concentration for each base and assumed that a steeper decrease in peak area indicated the most protected bases (**Figure 5 – figure supplement 3B-C**). We identified multiple regions of protection along the DNA on both strands. The locations of the protected regions differed unambiguously from the pattern of weak and strong cleavage produced by Fe·EDTA tethered to H-NS (**Figures 5B and C**; compare

positions of strong tethered cleavage, yellow stars on black traces, to the conventional H-NS footprint patterns, blue T-strand and red NT-strand traces). Together, these initial results confirmed that Fe·EDTA–H-NS is a viable reporter to determine the location of H-NS DBDs within filaments.

**Cleavage patterns modeled by MD simulation position H-NS DBDs in an H-NS filament**

Although the Fe·EDTA–H-NS artificial nuclease gave a unique pattern of DNA cleavage on the *bgl* DRE, the relationship between the cleavage pattern and the precise location and orientation of the H-NS DBDs was not immediately obvious. Because Fe·EDTA is tethered by a flexible linker and because ·OH can diffuse ~10 Å before being quenched by $H_2O$ (Cheal et al., 2009), we sought to predict the expected cleavage patterns when the DBD is bound to a site on DNA. To generate this prediction, we used molecular dynamics (MD) simulations to predict the position of the conjugated Fe·EDTA — and thus the position of the catalytic $Fe^{2+}$— with respect to DNA (**Figure 6**). The simulations generated a probability distribution of distances between the $Fe^{2+}$ and each nucleotide (nt) in the two duplex DNA strands (here arbitrarily referred to as strand A and strand B for modeling) to which H-NS is bound. We assumed that the cleavage propensity at each nt is directly proportional to the local concentration of hydroxyl radicals at that nt, which we estimated by a diffusion model for ·OH centered at the $Fe^{2+}$ (Cheal et al., 2009). Because some nt are shielded from the solvent by the bound protein, we scaled the concentration of radicals according to the solvent-accessible surface area of the putative cleavage site (**Figure 6 – figure supplement 2**; Balasubramanian et al., 1998). Modeling based on these assumptions yielded an estimate of the relative cleavage propensity at each nt in the duplex DNA bound by the DBD **(Figure 6A**). When displayed on B-form DNA, hot spots for cleavage of each DNA strand offset from the location of the DBD in the minor groove were evident (**Figure 6B**). The maximum peak of cleavage was predicted to occur at nt +10 of strand A and was flanked by two smaller predicted cleavage peaks in strand B. Adjacent to these sites of high predicted cleavage were sites of particularly low predicted cleavage, which resulted from shielding by the bound H-NS.

To identify specific H-NS binding sites in the *bgl*-DRE filament, we developed a statistically rigorous comparison the predicted cleavage pattern (**Figure 6**) and the experimentally measured tethered cleavage pattern over a rolling 15 bp window (**Figure 4 – figure supplement 4C**). Because H-NS DBDs could adopt forward or reverse orientations, we compared the alignment of the predicted cleavage pattern in the forward or reverse direction to the experimental data; here, forward and reverse indicate the direction of the Fe·EDTA tether with respect to the location of the promoter of the gene (**Figure 6B**). For each alignment, the Manhattan distance between the simulated pattern and the experimentally-determined cleavage pattern was calculated to determine how well the simulation matched the experiment, where a lower score is a closer match. This analysis revealed seven unique binding sites with match scores lower than 95% of all match scores from 1000 different sets of shuffled data, which we interpreted to be statistically significant (**Figure 6C**; see Methods). Due to the noise in both the predictions and experimental data, we interpreted binding sites that were separated by ≤3 bp as being the same binding site (*e.g.,* n663F and n664R are one binding site; n indicates assignment is from tethered nuclease data; F or R indicate forward or reverse). To visualize this arrangement of H-NS in 3D, we placed the center of the H-NS DBD (black triangle in **Figure 6B**) on B-form DNA at the locations with a statistically significant match score in the appropriate orientation (**Figure 6 – figure supplement 1**). 3D visualization of the H-NS DBD confirmed that occupation of these seven binding sites on one duplex DNA is sterically feasible. Overall, this analysis pipeline predicted the tethered nuclease cleavage data resulted from seven unique H-NS DNA binding sites where four were in the forward orientation (n642F, n676F, n695F, and n711F), two were in the reverse orientation (n685R and n706F) and one could be from an H-NS DBD bound in either orientation within a 2-bp window (n663F/n664R), thus with indeterminate (or alternating) orientation.

To validate the placement of H-NS DBD based on the tethered cleavage and MD simulations, we utilized a completely independent model to identify potential H-NS binding sites based on the predicted solvent accessibility of DNA bound by unmodified H-NS (**Figure 6 – figure supplement 2A**) and ·OH footprinting patterns (**Figure 5 – figure supplement 3B**). Because the solvent accessibility pattern was asymmetrical, we matched this pattern in either

orientation using the same directions defined for the tethered-cleavage comparison. The Manhattan distance was calculated for potential matches where sites with scores lower than 95% of scores calculated for 1000 different sets of shuffled data were considered statistically significant (**Figure 6 – figure supplement 2B**). This comparison identified four unique binding sites. Again, we visualized these alignments at the predicted locations on B-form DNA (**Figure 6 – figure supplement 2C**). In total, this comparison predicted that two of the H-NS DBDs were in the forward orientation (p694F and p711F; p for footprinting) and two had an indeterminate (or alternating) orientation (p665F/p668R and p685F/p688R).

Overall, there was good agreement between the two independent predictions of H-NS DBD placement within a filament, which allowed us to assign seven unique H-NS binding sites within 70 bp of duplex *bgl* DRE DNA with high confidence (**Figure 6D**). Three features stood out about the arrangement of the DBD within a filament. First, we observed 5-20 bp intervals between DBD, suggesting the DBD arrangement is not limited to one precise spacing between neighboring binding sites although most were separated by ~10 ± 1 bp (see Discussion). Second, both the forward and reverse orientations were observed for the DBD, suggesting that H-NS is not locked in one orientation within a filament. Finally, some of the binding sites had indeterminate or alternate orientations. In particular, we could not assign one particular location or orientation for DBDs binding between positions 663 and 669, suggesting a combination of binding sites are occupied in the filaments formed in the bulk assay. Taken together, these results suggest that H-NS forms a heterogenous structure within a single filament with variable spacing and orientation of the DBD and nominal spacing of 10 ± 1 bp in most cases.

**H-NS DBD locations are the same regardless of filament length**

The flexibility in DBD positioning within H-NS filaments highlights the question of whether DBDs change locations in different situations. To address this question, we first asked if the H-NS binding pattern changed in filaments formed on DNA longer than 146 bp as would occur *in vivo* (**Figure 1**) by comparing the tethered cleavage pattern on a 549-bp *bgl* DRE-containing fragment to the 146-bp fragment. The longer DRE fragment was extended toward $P_{bgl}$ but retained the same downstream end, so cleavage patterns of the 5′-$^{32}$P-labeled T strands could be

compared directly (**Figure 7A**). Due of the cooperative nature of H-NS binding (Bouffartigues et al., 2007), it took less H-NS to form a filament on the long DNA than on the shorter DNA. Even so, the tethered cleavage pattern was similar overall for both DNA fragments (**Figure 7A; Figure 7 – figure supplement 1**), except for a new peak at position ~645 on the long fragment. The new peak likely resulted from stabilization of the filament by H-NS binding to the long DNA extended in the upstream direction. We conclude that the locations of DBDs flexibly bound within the DRE did not change when the H-NS filament was extended in the upstream direction, indicating that sequence-specific binding dictates flexible DBD location in filaments regardless of length (see Discussion).

**H-NS DBD locations are unchanged in obligately bridged filaments**

We next asked if these DBD locations remained the same when H-NS is constrained to form only bridged filaments. We formed obligately bridged filaments by using a previously developed bridging pull-down assay (van der Valk et al., 2017; **Figure 7B; Figure 7 – figure supplement 2A**) in which H-NS–DNA filaments are captured from solution by DNA-coated magnetic beads, which requires that H-NS bridge between the bead-bound and unbound DNAs. To form bridged filaments, first, a 685-bp biotinated DNA (68% AT) was coupled to streptavidin beads and then incubated with a mixture of Fe·EDTA-H-NS and WT H-NS and the 549-bp DRE fragment that lacked biotin and was 5′-$^{32}$P-end-labeled on the T strand (**Figure 7B**). Bridged filaments attached to beads were magnetically separated from unattached filaments, resuspended in buffer, and treated with ascorbic acid and $H_2O_2$. To verify the mixture of WT H-NS and Fe·EDTA–H-NS bridges like WT H-NS in this assay, we checked that both captured equivalent amounts of DNA (**Figure 7 – figure supplement 2B**). Quantification of the cleavage patterns revealed that obligately bridged Fe·EDTA–H-NS cleaved the 549-bp *bgl* DRE DNA in a pattern indistinguishable from that generated by unconstrained Fe·EDTA–H-NS binding (compare **Figure 7A to 7B**; **Figure 7 – figure supplement 1**). We conclude that flexible DBD locations determined for the *bgl* DRE H-NS filament (**Figure 6D**), which includes footprint locations for WT H-NS at concentrations expected to form linear filaments (**Figure 5 – figure supplement 1**), remain unchanged in obligately bridged H-NS filaments.

**H-NS modulators, Hha and $Mg^{2+}$, do not alter H-NS DBD position**

As an additional test of DBD locations in different types of H-NS filaments, we determined whether H-NS modulators change the locations or orientations of DBDs in filaments. Hha binds to the NTD of H-NS and stimulates bridging by H-NS, but the location of H-NS DBDs in an Hha–H-NS filament are unknown. We confirmed that Hha stimulates formation of bridged filaments and stimulates H-NS-induced pausing by RNAP, although its effect on Fe·EDTA–H-NS-induced pausing was less than seen previously and confirmed here for wild-type H-NS (**Figure 7 – figure supplement 3;** Boudreau et al., 2018). We observed essentially no difference in the tethered cleavage pattern on the 146 bp DNA in the presence and absence of equimolar Hha (**Figure 7C**), except for a small increase in cleavage at position ~650 in the presence of Hha. These results suggest that Hha can bind to a bridged H-NS nucleoprotein filament without displacing H-NS DBDs from their locations.

Small cations, such as $Mg^{2+}$, and H-NS/DNA ratio also can modulate the conformation of H-NS filaments (Kotlajich et al., 2015; Liu et al., 2010; van der Valk et al., 2017; Will et al., 2018). Thus, we also tested whether H-NS DBD locations change in filaments formed at high (8 mM) and low (~0.3 mM) $Mg^{2+}$ or high (5 μM) vs. low (1 μM) total H-NS concentration. We found that DBD locations were unaffected by either $Mg^{2+}$ or H-NS concentration, although minor differences in the extent but not location of tethered nuclease cleavage were observed between 1 and 5 μM H-NS (**Figure 7 – figure supplement 4**). We conclude that $Mg^{2+}$ concentration has at best a minor influence on DBD locations, consistent with the conclusions of (Will et al., 2018).

**H-NS DBD positioning is associated with sequence-determined DNA structure**

Having established DBD location and orientation appears to be set by features of DNA sequence within a filament rather than being precisely fixed by filament geometry and that DNA sequence is the principal determinant of the genome-scale distribution of H-NS, we next asked if we could identify sequence parameters that best explain the genome-wide distribution. Previous work has demonstrated that H-NS exhibits variable preference for different 4mer sequences with an A-tract (*i.e.*, AAAA, AAAT, AATT, ATTT, TTTT) or TA-step locations (*i.e.*, ATAA, ATAT,

ATTA, AATA, TAAA, TAAT, TATA, TATT, TTAA, TTAT, TTTA) depending on the surrounding A/T content, with TA steps being most favored in regions of high A/T content (Gordon et al., 2011). These sequences were postulated to be relevant to H-NS binding as they represent regions of DNA that can form a narrow minor groove width (MGW), which H-NS has been shown to bind. Algorithms that enable high-throughput prediction of DNA structural parameters such as MGW, propeller twist (ProT), and electrostatic potential in the minor groove (EP) from sequence at the bp level have been developed (Chiu et al., 2017; Chiu et al., 2016; Zhou et al., 2013). Thus, we first attempted to use these algorithms to interrogate which sequence-defined DNA structure best explains H-NS DBD arrangement in the *bgl* DRE by aligning our identified binding locations with sequence and structural features determined for the *bgl* DRE (**Figure 8A**). Consistent with the previously determined preferences for individual H-NS DBDs, we found that our high-confidence DBD locations correspond to A-tracts with a narrow MGW, negative propeller twist, and highly negative electrostatic potential.

Using this view of the *bgl* DRE H-NS filament as a guide, we next sought to predict the genome-wide distribution of H-NS filaments from a minimal set of sequence and sequence-implied DNA structural features. We generated a dataset by considering the bound status of 1000 bp windows staggered every 250 bp across the *E. coli* genome and various sequence and structure features associated with those windows (see Methods). We used least absolute shrinkage and selection (lasso) logistic regression trained on H-NS ChIP-seq data to predict the binding pattern of H-NS, StpA, and Hha–H-NS. Lasso regression yielded a minimal set of DNA features that best explained the distribution of H-NS (**Figure 8B**). We found that a simple model with only TA-step abundance, A/T content, and A-tract abundance was able to predict H-NS ChIP-seq regions with high precision and recall (**Figure 8C**). Furthermore, we found that this model trained on H-NS ChIP-seq data also predicts the binding locations of H-NS *in vivo* in the presence of Rif, H-NS *in vitro*, StpA *in-vitro*, and Hha–H-NS complexes *in vitro* consistent with the model that sequence primarily drives H-NS binding (**Figure 8C**). As expected, the model does a poor job of predicting Fis binding sites *in vitro*, confirming its discriminatory power to differentiate H-NS peaks from Fis peaks despite the similarity in A/T preference of these two proteins. Somewhat surprisingly, only DNA sequence and not structural features were selected

by lasso for the final model. Possibly, the lack of DNA structure features in the final model may indicate that all structure information relevant to genome-wide distribution of H-NS is already fully captured by the sequence features. Models trained on H-NS Nitro-seq data performed similarly, selecting the same three variables to predict H-NS-bound regions with high precision and recall. Furthermore, models trained on the Fis Nitro-seq data identified the RegulonDB Fis PWM as the most important feature for predicting Fis binding sites, suggesting that this method is capable of identifying features important for protein binding unique to each protein.

## DISCUSSION

The biological function of the prototypical bacterial chromatin protein H-NS depends on its complex structural properties, which include alternating head-to-head–tail-to-tail oligomerization, flexibly connected DBDs, paralogous subunits, and small-protein modulators that combine to form as-yet unresolved filament structures with DNA. This complexity enables formation of multiple filament types, notably: (*i*) bridged nucleoprotein filaments that inhibit both transcription initiation and transcription elongation, and (*ii*) linear filaments that principally affect only initiation. Our hybrid approach combining new genome-scale reconstitution and sequencing methods (Nitro-seq) with direct detection of DBD locations and with molecular dynamic and bioinformatic modeling yielded three principal insights into this complex regulator of gene expression: (*i*) the genome-wide locations of H-NS binding *in vivo* are determined principally by DNA sequence, which is unchanged by the paralog StpA or modulator Hha; (*ii*) the DBDs in an H-NS filament do not contact DNA with precisely fixed spacing but rather adjust both contact point and orientation along adjacent minor groove positions accommodated by its flexible linker region; and (*iii*) H-NS DBD locations do not change in the switch from linear to bridged filaments, supporting a model of DBD hemi-sequestration in linear filaments. We discuss here each of these points in the context of an overall model of H-NS binding in which DNA sequence dictates both the location of H-NS filaments genome-wide and the arrangement of DBDs within the filaments.

### H-NS genomic distribution is primarily determined by DNA sequence

Multiple contributions to H-NS genomic distribution have been described, including AT-rich-sequence-specific binding (Grainger et al., 2006; Kahramanoglou et al., 2011; Lang et al., 2007; Lucchini et al., 2006; Navarre et al., 2006), competition with transcription (Chandraprakash and Seshasayee, 2014; Rangarajan and Schnetz, 2018), and competition with other DBPs (*e.g.*, in anti-silencing gene regulation; Will et al., 2015), but the extent to which each of these factors contributes to H-NS genomic distribution is currently uncertain. We predicted that if transcription or DBPs played a dominant role in determining genome-scale H-NS filament patterns, *in vitro* reconstitution of H-NS filaments on genomic DNA in the absence of

transcription of other DBDs would yield different H-NS binding patterns than observed *in vivo* by ChIP-seq. In fact, we found only minor differences attributable to limited cases of anti-silencing regulation or transcription (**Figures 2 and 3**). The limited effect of transcription on H-NS binding *in vivo* was further confirmed by minimal effects on H-NS of blocking all transcription with Rif (**Figure 1**). Further, a bioinformatic model based on only three DNA sequence features (A/T richness, TA steps, and A tracts) could robustly predict H-NS binding genome-wide (**Figure 8**). These findings do not contradict limited roles of active transcription or of anti-silencing by DBPs in H-NS binding, but together strongly support a view that H-NS genomic distribution is primarily determined by DNA sequence (**Figure 8E**).

Interestingly, both StpA and Hha confer greater regulation on a subset of genes bound by H-NS (Banos et al., 2009; Lucchini et al., 2009; Muller et al, 2006), raising the question of whether these modulators may alter the sequence specificity of H-NS filament formation. Our *in vitro* reconstitution and Nitro-seq results demonstrate that StpA and Hha do not alter the sequence preferences of H-NS filaments but rather bind the same DNA sequences as H-NS-only filaments (**Figure 2B; Figure 2 – figure supplement 1; Figure 3; Figure 3 – figure supplement 1A; Table 1**). These results agree with previous ChIP results showing extensive overlap between H-NS, StpA, and Hha throughout the genome (Ueda et al., 2013; Uyar et al., 2009), but raise an obvious question. If StpA and Hha do not alter the sequence-specificity of H-NS binding, then how do they accomplish preferential regulation of subsets of H-NS bound genes? One possibility is that different sequences give rise to different affinities for H-NS or H-NS modulator binding. If StpA or Hha increase binding affinity of H-NS filaments for DNA, as is consistent both with our *in vitro* reconstitution data (**Figure 2**) and with other reports (van der Valk et al., 2017; Srinivasan et al., 2013), then expression of genes whose sequences exhibit high affinity for H-NS alone may not be changed by StpA or Hha. However, expression of genes whose sequences confer weaker affinity for H-NS could be affected by StpA or Hha because these modulators increase the affinity for these genes without altering sequence specificity, especially at higher temperatures where H-NS affinity may be less (Shahul Hameed et al., 2019). It is also possible that differential effects on gene expression could be grounded in differential effects of sequences

on formation of bridged filaments (see below), since both StpA and Hha confer constitutive bridging on H-NS filaments (Boudreau et al., 2018).

Our results may also prove useful in prediction of H-NS binding locations. Previous attempts at predicting H-NS binding locations relied on similarity to the high-affinity consensus sequence TCGATAAATT (Lang et al., 2007). Our analysis of predictive strength of sequence or sequence-based structural features found that A/T richness, TA steps, and A tracts had maximal predictive power and could account for the genome-scale distribution of filaments *in vivo* and of H-NS-only, H-NS–StpA, and H-NS–Hha filaments *in vitro* but that the PWM based on the high-affinity consensus sequence added negligible predictive value (**Figure 8B**). Application of our predictive methods to other enterobacterial species that rely on H-NS may prove useful.

**DBD position and orientation in H-NS filaments is flexible and set by DNA sequence features**

Our finding that both H-NS DBD position and orientation are flexible is consistent with H-NS–DNA binding patterns reported for diverse DNA fragments. Footprinting experiments from other groups revealed semi-regular spacing of ~10 bp between most H-NS protected regions, but also variable spacing of ~5-30 bp in some locations (Badaut et al., 2002; Bouffartigues et al., 2007; Tupper et al., 1994; Ulissi et al., 2014; Will et al., 2018). Further, single molecule analysis of H-NS arrangement in a bridged filament revealed up to ~30 bp of spacing between adjacent strong DBD contacts to DNA (Dame et al., 2006). Our observed correlation between DBD locations and DNA sequence suggests that DNA sequence may dictate the lowest energy arrangement of DBDs within the constraints imposed by its flexible tether to the H-NS head-to-head–tail-to-tail oligomer. Specifically, DBDs were located on the DRE fragment at sequence and shape features that confer high-affinity for H-NS (*e.g.,* TA steps, A-tracts leading to narrow minor groove width and negative EP), but were not observed at lower-affinity sites lacking these features (Gordon et al., 2011; **Figure 8A**). Thus, the semi-regular spacing of DBDs of 10 ± 1 bp is likely dictated by the minor groove complement in shape and charge to the DBD within reach of the linker, and the larger spacings of 20 bp of more in contacts may reflect the absence of suitable binding features near a given tail-tail junction in the H-NS oligomer.

What can this DBD arrangement tell us about the structure of H-NS filaments? Overall, our results do not allow us to postulate a specific conformation of the H-NS head-to-head–tail-to-tail oligomer. However, the DBD locations we find are inconsistent with an H-NS filament structure observed in an H-NS NTD crystal structure that places the tail-to-tail junction to which the DBDs are tethered at ~30 bp intervals relative to B-form DNA (Arold et al., 2010). If DBDs contact B-form DNA every ~10 bp in a bridged filament in which one of the two DBDs at each junction binds one DNA duplex and the other DBD binds the other duplex, then the H-NS NTD must adopt a conformation in which the tail-tail junctions are ~10 bp or 34 Å apart. The crystal structure the NTD oligomer places these junctions ~95 Å apart, nearly three times the 10-bp distance on B-form DNA. Consistent with this idea, other conformations of the NTD have been observed by MD simulation (Shahul Hameed et al., 2019; van der Valk et al., 2017), which might enable closer, yet variable spacing between neighboring DBDs. Future studies, possibly by cryo-EM, of a complete H-NS–DNA filament are needed to resolve how H-NS monomers assemble to enable the semi-regular but flexible orientations we observed for the H-NS DBDs.

**H-NS DBD contacts to DNA do not change in the switch from bridged to linear conformations**

One of the most intriguing features of H-NS-based gene regulation is the capacity of H-NS filaments to switch from binding one DNA duplex as a linear filament to binding two DNA duplexes in a bridged conformation. The extent of DNA bridging *in vivo* and whether it varies among genes or conditions remains unknown. However, the fact that H-NS can block transcript elongation only in the bridged conformation where it creates topological traps to elongating RNAP (Kotlajich et al., 2015; Boudreau et al., 2018) and the clear evidence that H-NS modulates transcription elongation *in vivo* by stimulating ρ-dependent termination (Peters et al., 2012; Chandraprakash and Seshasayee, 2014) seemingly provides strong evidence that bridging occurs to some extent *in vivo*. A structural basis for switching of H-NS between linear and bridged filaments remains to be established.

The Dame and Arold groups have proposed a hemi-sequestration model of H-NS DBD location that can explain linear-to-bridged switching (van der Valk et al., 2017; Shahul Hameed

et al., 2018). In the hemi-sequestration model, one of the two DBDs present at each tail-tail dimerization interface in an H-NS filament binds to DNA and the second DBD is sequestered through interactions with the N-terminal dimerization domain possibly aided by buckling of the α-helix connecting the two dimerization interfaces. Thus, the hemi-sequestered or closed conformation enables only linear filament formation whereas the open conformation, which is favored by binding of $Mg^{2+}$ or Hha and possibly constitutive in StpA, releases the sequestered set of DBDs so they can bridge to a second DNA duplex (**Figure 8D**).

Our results are consistent with the hemi-sequestration model because we observe no changes in the locations of DBDs (**Figure 6D**) between filaments that are obligately bridged using a bridging pull-down assay (**Figure 7B**) and filaments formed with wild-type H-NS at very high H-NS/DNA ratios known to favor linear filament formation (**Figure 6 – figure supplement 2**, **Figure 5 – figure supplement 1**; Kotlajich et al., 2015). Even though the precise contacts and orientation of the DBDs to the *bgl* DRE appear to be set by local sequence features and in some cases are weak or absent at the ~10 bp intervals generally seen, these patterns show no evident change in conditions favoring linear conformations where twice as many DBDs could be available for contact. This same observation of no change in footprints in different H-NS/DNA ratios or changing $Mg^{2+}$ concentration thought to govern a linear *versus* bridged preference has been reported by several other groups (Badaut et al., 2002; Bouffartigues et al., 2007; Tupper et al., 1994; Ulissi et al., 2014; Will et al., 2018). Taken together, these data strongly favor the hemi-sequestration model of H-NS linear-to-bridged filament switching.

## Conclusion

Our hybrid approach to understanding the genome-scale basis of H-NS filament formation and the rules governing H-NS DBD contact to DNA have yielded a greatly improved understanding of how linear and bridged H-NS filaments are organized on DNA. The challenges now are to translate this understanding to molecular resolution structures of these filament conformations, possibly using cryo-EM, and to develop approaches that can distinguish locations of linear *versus* bridged filaments in living cells.

## MATERIALS AND METHODS

### Plasmid preparation

Sequences of plasmids (**Supplemental Table S1**) described here were all confirmed with Sanger sequencing. An expression plasmid for WT H-NS (pWTHNS) without a tag was constructed by Gibson assembly (NEB). The *hns* coding sequence was amplified from MG1655 genomic DNA using primers #11779 and #11780 (**Supplemental Table S2**) and the pET21d backbone was amplified from pHNS using primers #11781 and #11782 before assembly. To introduce point mutations in pWTHNS or pHisHNS, site-directed mutagenesis (SDM) was performed with Q5 DNA polymerase (NEB) using the primers indicated to introduce mutations (**Supplemental Table S2**). Primer #8964 mutated the Cys codon (TGT) to a Ser codon (AGT) in pHNSC21S. Primers #8964 and #8963, which mutated the Gln codon (GAA) to the Cys codon (TGT), were used to create pHNSE102C.

### Protein Purification

Native StpA and Hha were purified as previously described in (Boudreau et al., 2018). C-term $His_{6x}$-tagged H-NS was purified as in (Kotlajich et al., 2015).

*H-NS C21S purification*

H-NS C21S was essentially purified as described for non-tagged H-NS in (Boudreau et al., 2018) with the following modifications. Briefly, expression of H-NS C21S was induced in BL21 λDE3 cells with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 2 hours at 37 °C. Pelleted cells were lysed by sonication. H-NS C21S was precipitated with addition of 0.6% (w/v) PEI and eluted with PEI elution buffer containing 0.8 M NaCl followed by ammonium sulfate precipitation. Precipitated H-NS was dissolved in nickel binding buffer and applied to a HisTrap $Ni^{2+}$ column (GE Healthcare). Fractions containing H-NS were mixed with TEV protease to remove the N-terminus $His_{6x}$-tag. H-NS C21S was reapplied to the nickel column as before. Non-tagged protein was collected in the flow through and dialyzed against heparin binding buffer (20 mM Tris-HCl pH 7.5, 130 mM NaCl, 100 μM EDTA, 5% glycerol, 1 mM DTT). H-NS C21S was then applied to a HiTrap Hep column (GE Healthcare) and eluted with heparin binding buffer containing 1 M NaCl. Fractions containing H-NS C21S were pooled and dialyzed

against Resource Q binding buffer (20 mM Tris-HCl pH 7.2, 130 mM NaCl, 10% glycerol, 1 mM DTT). To concentrate H-NS, the protein was applied to a Resource Q column (GE Healthcare). Any remaining impurities were removed by a wash with 130-200 mM NaCl gradient over 7 column volumes. H-NS C21S was eluted in about 3 mL by block elution from 200-500 mM NaCl. Fractions containing H-NS were dialyzed against H-NS storage buffer (20 mM Tris-HCl pH 7.5, 0.3 M potassium acetate, 10% glycerol) before storing at -80 °C. Protein purity (~90%) and amount (~5 µg) were determined by SDS-PAGE and Qubit, respectively.

*Native WT and C21S E102C H-NS purification*

BL21 λDE3 cells were transformed with pWTHNS or pHNSE102C before deleting *hns* by P1 transduction with lysate from Δ*hns*::*kan* Keio collection strain (CGSC #9111; Ref. (Baba et al., 2006). After confirming *hns* deletion, 2 L of cells were grown in LB supplemented with 100 µg ampicillin/mL at 37 °C to $OD_{600}$ 0.5 before inducing H-NS expression with 1 mM IPTG and 250 µg carbenicillin/mL for 16 hours at 25 °C. Cells were pelleted at 3, 000 × *g* for 15 min at 4 °C and resuspended in 40 mL of lysis buffer (20 mM Tris-HCl, pH 7.2, 130 mM NaCl, 10% glycerol, 1 mM DTT, 1 mM β-mercaptoethanol supplemented with phenylmethylsulfonyl fluoride (PMSF) and protease inhibitor cocktail (PIC; 31.2 mg benzamide/mL, 0.5 mg chymostatin/mL, 0.5 m leupeptin/mL, 0.1 mg pepstatin/mL, 1 mg aprotonin/mL, 1 mg antipain/mL). Cells were lysed by sonication and the lysate was cleared by centrifugation at 27, 000 × *g* for 45 min. H-NS containing cell lysate was directly applied to a HiTrap Heparin HP column (GE Healthcare) equilibrated with binding buffer (20 mM Tris-HCl, pH 7.2, 130 mM NaCl, 10% glycerol, 1 mM DTT). H-NS was eluted off the column at about 600 mM NaCl with a gradient elution from 0% to 100% heparin binding buffer + 1 M NaCl over 20 column volumes. Fractions containing H-NS were pooled and dialyzed against binding buffer. H-NS was applied to a 1 mL Resource Q column equilibrated with binding buffer. Remaining impurities were eluted over a gradient of 130-200 mM NaCl in 10 mL. H-NS was eluted and concentrated in a peak over 3 mL by block elution with 500 mM NaCl. Fractions containing H-NS were dialyzed against H-NS storage buffer (20 mM Tris-HCl, pH 7.2, 300 mM potassium acetate, 10% glycerol, 5 mM TCEP). Purity (~90%) and amount of protein (~4 µg) was determined by SDS-PAGE and Qubit analysis, respectively. For tethered cleavage or ·OH cleavage assays, wt

H-NS was exchanged into 3X H-NS BABE buffer (120 mM HEPES-KOH, pH 8.0, 300 mM potassium glutamate, and 15% glycerol).

**ChIP-seq sample preparation**

RL3000 cells (MG1655 F-, λ-, *rph-1*) (Ghosh and Landick, 2016) were grown in MOPs defined minimal medium (with 0.2% glucose) at 37 °C with vigorous shaking to an OD ~0.4 before cultures were treated with either 100 µg rifampicin/mL or methanol (solvent) for 30 min. Cell preparation, crosslinking, and immunoprecipitation was completed largely as described (Mooney et al., 2009; Myers et al., 2013) with the following modification. After formaldehyde crosslinking, cells were resuspended in IP buffer (100 mM Tris, pH 8.0, 300 mM NaCl, 1% TritonX-100, 1 mM PMSF), sonicated for 1 hour at 60% power, 10 sec on and 10 sec off, and treated with 0.1 mg RNaseA/mL. To generate the input DNA control, a 1/10th of the resulting lysate was removed and incubated with a 50-50 slurry of Sepharose Protein A and Protein G overnight. The remaining lysate was pre-cleared with Sepharose beads at 4°C for 3 hours. Crosslinked protein and DNA were precipitated by adding either 10 µL of polyclonal H-NS antibody (Harlan Bioproducts) or 2 µL Beta antibody (Neoclone #WP0022)(Peters et al., 2012) . The next day, DNA/protein/antibody complexes were purified with the Sepharose bead slurry and washed extensively before incubating at 65 °C for 30 min to the crosslinks. Eluted DNA and input DNA were purified with Qiagen QIAquick PCR purification kit and eluted in 58 µL of EB. DNA was quantified by Qubit.

ChIP-seq libraries were prepared in house. DNA samples were first end-repaired by incubating with 0.4 mM dNTPs (Fisher Scientific), T4 DNA polymerase, Klenow Enzyme, and T4 PNK in 1x T4 DNA ligase buffer (NEB) for 30 min at 20 °C. DNA was purified with Qiagen QIAquick PCR purification kit. An A-tail was added to the DNA by incubating purified DNA with 1 mM dATP and Klenow Exo 3'-5' minus (NEB) in Klenow Buffer for 30 min at 37 °C and cleaned up with the Qiagen MinElute PCR purification kit. Adapter and index oligo mixtures were made by incubating 30 µM universal adapter oligo and 30 µM index primer for 5 min at 100 °C and slowly cooling back to room temp. Annealed oligos were ligated onto DNA fragments with DNA Ligase (NEB) in 1x DNA ligase buffer supplemented with 100 mM ATP

for 15 min at 20 °C. DNA was purified with MinElute Kit and resolved on a 2% E-gel (ThermoFisher) to select for fragments ~270 bp long. DNA fragments were PCR amplified with primers 1, 2, and the respective index primer by Phusion DNA polymerase. Final PCR fragments were purified with QIAquick PCR purification kit. Quality and quantity of the finished libraries were assessed using an Agilent DNA1000 chip and Qubit® dsDNA HS Assay Kit, respectively. Libraries were standardized to 2 nM. Cluster generation was performed using the Illumina cBot. Paired-end, 50bp sequencing was performed, using standard SBS chemistry on an Illumina HiSeq2500 sequencer. Images were analyzed using the standard Illumina Pipeline, version 1.8.2.

**Nitrocellulose-seq assay**

Genomic DNA (gDNA) was isolated from stationary phase *E. coli* MG1655 cells using a previously described CTAB isolation protocol (Wilson, 2001). Briefly, 1 OD of cells were treated with 2.6 mg lysozyme/mL at room temperature for 5 min. Protein and polysaccharides were removed with treatment with 0.5% SDS and 100 µg proteinase K/mL for 1 hour at 37 °C and then treatment with 630 mM NaCl and 37% CTAB. DNA was cleaned with three phenol:chloroform extractions and precipitated with isopropanol. Precipitated DNA was resuspended in TE with 20 µg RNaseA/mL. Purity and concentration were determined with Nanodrop. gDNA was sheared to fragments ~200-700 bp using the Misonix sonicator for 30 min at 60 watts and 50% duty cycle.

To determine the concentration of protein that bound only desired fragments, the nitrocellulose filter binding experiment was performed in the presence of low amounts of radiolabeled DNA fragments with different GC content or known Fis binding sites (**Supplemental Table S3**). Radiolabeled DNA was prepared with OneTaq PCR (NEB) using indicated primers, either genomic DNA, pMK110, or pBR322 as a template and the addition of 2.5 µCi of α[$P^{32}$]dCTP per 50 µL reaction. Excess label was removed with G50 spin columns (GE Healthcare). In a 250 µL reaction, gDNA (500 ng) was mixed with 1 µL of each indicated PCR product and a titration of either C-term $His_{6x}$-tagged H-NS, StpA, or Fis. Complexes were formed for 20 min at 20 °C in binding buffer (40 mM HEPES•KOH, pH 8.0; 100 mM potassium glutamate; 8 mM magnesium aspartate; 5% glycerol) before passing over a 0.45 µm

nitrocellulose filters (Millipore) at room temperature. DNA was eluted, ethanol precipitated, separated by 6% urea-PAGE, and visualized with the Typhoon (GE Healthcare) (**Figure 2 – figure supplements 1 and 2**). Counts retained on filter before and after elution were measured for H-NS and we found DNA eluted with ~70% efficiency (**Figure 2 – figure supplement 1A**). H-NS and StpA were expected to bind only AT rich fragments from the *bgl* operon, but not GC rich fragments from pBR322. For StpA, 50 or 100 nM protein retained the AT-rich but not GC-rich PCR fragments with selectivity similar to 100 nM H-NS as described in the results. Because Hha–H-NS filaments form at 1:1 Hha:H-NS ratio (Ali et al., 2013), we performed the assays with 100 nM Hha plus 100 nM H-NS. We anticipated that *in vitro* Fis would to bind to fragments that Fis bound to *in vivo* (Kahramanoglou et al., 2011). We found that 5 and 12 nM Fis retained Fis-specific PCR fragments on the nitrocellulose filter (**Figure 2 – figure supplement 2**).

For the Nitro-seq experiment, either H-NS, StpA, or Fis were mixed with 500 ng sheared gDNA at concentrations that selected AT-rich DNA fragments, but not GC rich fragments (100 and 200 nM H-NS; 50 and 100 nM StpA; 5 and 12 nM Fis; **Figure 2A**; **Figure 2 – figure supplements 1 and 2**) in binding buffer (40 mM HEPES•KOH, pH 8.0; 100 mM potassium glutamate; 8 mM magnesium aspartate; 5% glycerol). Reactions were incubated for 20 min at 20 °C. Protein-DNA complexes were isolated by applying samples to nitrocellulose filters and were washed twice with binding buffer (Papoulas, 1996). Bound DNA was eluted with elution buffer (20 mM Tris-HCl, pH 7.8, 0.2% SDS, 0.3 M sodium acetate), ethanol precipitated, and resuspended in water. Amount of DNA retained was quantified with the Qubit® dsDNA HS Assay Kit. Libraries were prepared for each sample using NEBNext Ultra II DNA library kit (NEB, Catalog#E7103). Quality and quantity of the finished libraries were assessed using an Agilent DNA1000 chip and Qubit® dsDNA HS Assay Kit, respectively. Libraries were standardized to 2nM. Paired-end, 50bp sequencing was performed, using standard SBS chemistry on an Illumina HiSeq2500 sequencer. Images were analyzed using the standard Illumina Pipeline, version 1.8.2.

**ChIP-seq and Nitrocellulose-seq sequence processing and alignment**

Sequencing adapter reads and reads shorter than 30 nt were removed with Trimmomatic (version 0.38) and quality of remaining read were confirmed by FastQC. Remaining paired end reads were aligned to the MG1655 genome (Accession: NC_000913.3) using Bowtie2 (version 2.3.4.2) with the parameters (-q --end-to-end -I 100 -X 800 --fr -t). Proper pair reads were retained, duplicate reads were removed with Samtools (version 1.9), and bedgraph files for visualization were created with BEDtools (version 2.8) (Quinlan and Hall, 2010). The vast majority of reads mapped for ChIP-seq (**Supplemental Table S4**) and Nitro-seq experiments (**Supplementary Table S7**) were of high quality and mapped properly.

**ChIP-seq and Nitro-seq data analysis**

ChIP- or Nitro-seq patterns, read coverage in each condition was scaled by reads per million. To normalize reads ChIP-seq to input sequences, the $\log_2$ of the IP and corresponding input scaled read coverage was calculated and the input was subtracted from the IP giving normalized read coverage. The RNAP baseline was set to the ChIP signal level in regions known to be transcriptionally silenced (Mooney et al., 2009) and the H-NS baseline was set at the lowest IP/input values. A similar analysis was done for Nitro-seq samples by normalizing the nitrocellulose retained reads to the input reads. All $\log_2$(ChIP/input) or $\log_2$(Nitro/input) samples to be compared (*i.e.,* all 4 H-NS ChIP samples) were quantile normalized in R. To compare binding pattern of Nitro-seq samples, the genome was divided into 200 bp bins and the average $\log_2$(Nitro/input) was calculated for each bin using a custom script. Density of read coverage in each bin was compared and the Spearman coefficient was used to calculate the correlation between datasets (**Figure 3 – figure supplement 1**).

To determine the enriched areas more rigorously, we used the MOSAiCS (Model-based one and two Sample Analysis and Inference for ChIP-Seq Data) two-sample analysis pipeline (Kuan et al., 2011), which also included MOSAiCS with the hidden Markov Model (MOSAiCS-HMM) architecture (Chung et al., 2014). This allowed for determination of enriched regions in an IP or nitrocellulose retained sample compared to a matched input sample. For H-NS and WT Beta ChIP, 200 bp bins were constructed from .sam files and duplicate reads were capped at 5 before

fitting a robust method of moment background model of the data with the MOSAiCS-HMM fitting parameters. Using the viterbi algorithm for the decoding parameter, peaks were called with a false discovery rate (FDR) ≤ 0.05. Peak boundaries were adjusted, and false positive peaks were filtered out with default settings. Enriched regions present in both H-NS ChIP replicates were determined using BEDtools intersect (default parameters) and were used as the enriched regions for WT or rif-treated cells. For the Beta ChIP (RNAP) in +Rif treated cells, settings similar to WT Beta ChIP were used find peaks , but the non-HMM MOSAiCS fit for background and the following mosaicsPeak parameters (FDR=0.05, maxgap=200, minsize=50, thres=10) were used instead. For Nitro-seq, MOSAiCS was performed as for H-NS ChIP-seq except all reads were used for creating bins in the MOSAiCS package and peaks were identified using the posterior decoding algorithm. Features of MOSAiCS regions for each condition are summarized in Table 1.

To compare MOSAiCS-identified regions between conditions, we used a series of BEDtools (Quinlan and Hall, 2010). First, we identified regions that overlapped between conditions using for BEDtools intersect with options to report the original regions (-wa -wb) and then the lengths of these regions were plotted on a scatter plot in Excel. Regions only in one condition were determined using BEDtools subtract (default parameters). The similarity between the two conditions was determined with the BEDtools jaccard, which calculates the percentage of the length of filaments that overlap out of the total length of filaments found in both samples that do not overlap (Favorov et al., 2012). Thus, 0% is completely dissimilar, non-overlapping data sets and 100% indicates two data sets that completely overlap.

For each H-NS differentially bound region, the average chip signal from both replicates was calculated using a custom script. A ratio of WT H-NS density/Rif-treatment H-NS density ≥1.2 indicated H-NS depletion in rif-treatment whereas a ratio ≤1.2 indicated enrichment of H-NS in rif-treatment (**Supplemental Table S5**). Similar calculations were made for Beta ChIP to determine differences in RNAP occupancy in H-NS differentially bound regions. A correlation between a H-NS differentially bound region and a change in RNAP binding (Beta ChIP) was confirmed if either MOSAiCS or average ChIP signal showed a change in RNAP binding.

To compare H-NS filaments either only found *in vitro* (Nitro-seq specific) or found both *in vitro* and *in vivo* (overlapping) to known TF binding sites, we calculated the BEDtools jaccard similarity between the known TF binding sites and the H-NS filament data set. Known binding sites were from either RegulonDB (Santos-Zavaleta et al., 2019), tsEPODs, or heEPODs (Vora et al., 2009). To calculate the p-value, chance overlaps between H-NS filaments and known TF binding sites were calculated using a two-sided permutation test (n = 1000) of shuffled H-NS filament class labels (Nitro-seq-specific or overlapping) with the difference in Jaccard similarity between the classes as the test statistic.

**FeBABE Conjugation**

H-NS C21S and H-NS C21S E102C were exchanged into conjugation buffer (10 mM MOPS, pH 8.0, 0.5 mM EDTA, 200 mM NaCl, 5% glycerol) using PD SpinTrap G-25 columns (GE Healthcare). FeBABE was made by incubating BABE (Santa Cruz Biotechnology) with $FeCl_3 \bullet 6H_2O$ as previously described (Greiner et al., 1997). H-NS variants were incubated with 10-fold excess FeBABE for 2 hours at room temperature to conjugate FeBABE to H-NS. After conjugation, aliquots were combined with mPEG for an additional 2 hours at room temperature to determine the percent conjugated (~50% for H-NS C21S E102C). Excess FeBABE was removed by applying conjugated H-NS to a clean PD SpinTrap G-25 column equilibrated with 3X H-NS BABE Buffer (120 mM HEPES•KOH, pH 8.0, 300 mM potassium glutamate, 15% glycerol). Conjugation was also confirmed by the appearance of a second, slower migrating band on 20% PhastGel (GE Healthcare). Fe·EDTA–H-NS was stored in aliquots at -80°C for 1 year. Activity of the conjugated protein was confirmed by electrophoretic mobility shift assay and in vitro transcription assays done as previously described (Boudreau et al., 2018; Kotlajich et al., 2015).

**Fe·EDTA Tethered Cleavage Assay**

The 146 bp DRE DNA template for the tethered cleavage assay was made by PCR amplifying the *bgl* downstream regulatory element (DRE) from a plasmid containing the *bgl* operon (pMK110) using primers 7486 and 12638. Single primers were 5′ end-labeled with [γ-$P^{32}$]ATP before adding to the PCR mixture. The resulting PCR product was gel purified and

cleaned up with the Qiagen PCR clean-up kit. Increasing concentrations of Fe·EDTA–H-NS were incubated with DRE DNA for 20 min at 20 °C in 7.5 µL reactions in cleavage buffer (40 mM HEPES-KOH, pH 8.0, 100 mM potassium glutamate, 8 mM magnesium aspartate, 5% glycerol). Hydroxyl radical cleavage was initiated by the rapid addition of 1 µL 2 mM $H_2O_2$ and 1 µL 2 mM ascorbic acid. Cleavage proceeded for 2 min at 20 °C before stopping the reaction by adding to one-to-one mixture of TE-saturated phenol, pH 7.9 and cleavage stop solution (20 mM EDTA, 20 µg glycogen). After phenol extraction, cleavage products were concentrated by a subsequent ethanol precipitation and resuspended in formamide stop dye (95% formamide, 20 mM EDTA, 0.02% xylene cyanol, 0.02% bromophenol blue). Samples were then loaded onto a 7% urea poly-acrylamide gel. Samples were run alongside a Maxam-Gilbert A+G ladder for alignment (Maxam and Gilbert, 1977) and end-labeled MspI-digested pBR322 (NEB) for size reference. Cleaved DNA was quantified with the semi-automated footprinting analysis (SAFA) software (Das et al., 2005). Resulting peak areas were first normalized to total signal in each sample determined from ImageQuant analysis of a short, not-overexposed scan of the gel and then normalized to the average peak area in each sample. Modifications to the tethered cleavage assay were performed as follows. For tethered cleavage with Hha and H-NS, H-NS: Hha filaments were formed at a 1:1 ratio on the 146-bp DRE fragment and then cleavage and analysis were completed as described (**Figure 7**). Finally, for the offset DNA experiment, the cleavage assay was performed as above except using a151-bp DRE DNA template, which was amplified from pMK110 using primer "add 5-bp" 5′ end-labeled with [γ-$P^{32}$]ATP and cold primer 12386. For tethered cleavage on the longer, 549 bp fragment, the DNA template was amplified from pMK110 using primer 7486 5′ end-labeled with [γ-$P^{32}$]ATP and cold primer 7491. DNA was purified as above. A titration of a 40-60 mixture of Fe·EDTA–H-NS and wt H-NS was added to ~18 nM DNA in a 20 µL reaction in cleavage buffer supplemented with 0.02% NP-40 and 100 µg acetylated BSA/mL to compare to bridged filament cleavage described below. Cleavage was initiated by addition of 1 µL 50 mM ascorbic acid and 4 µL 12.5 mM $H_2O_2$. Cleavage continued for 2 min at RT before stopping the reaction with 25 µL 22.4 mM thiourea and 20 µL TE. DNA was phenol extracted, ethanol precipitated, and resuspended in formamide stop dye. Cleavage products were visualized and quantified as described for cleavage on the 146 bp DNA template.

To test the effect of $Mg^{2+}$ on the H-NS DBD location and orientation, tethered cleavage assay was performed as described except the amount of magnesium aspartate was varied in the cleavage buffer (**Figure 7**).

**Hydroxyl radical footprinting**

H-NS filaments were formed on the same 146 bp DRE DNA template used in the tethered cleavage assay. A titration of WT H-NS (no-tag) from 50 nM to 4 μM (17 – 1360 H-NS/kb) was added to ~20 nM 5′ end-labeled DRE DNA in 50 μL reactions. Filaments were allowed to form at 20 °C for 20 min. •OH cleavage was initiated by rapid, sequential addition of 0.6 μL of 100 mM ascorbic acid, 4 μL of 0.15% $H_2O_2$, and 6 μL of fresh 50 mM $Fe^{2+}$-EDTA (50 mM $(NH_4)_2Fe(SO_4)_2$ and 100 mM EDTA) and incubated for 2 min at 20 °C. To determine DNA quality, 10.6 μL water was added to the no cleavage control. •OH were quenched by the addition of 80 μL 20 mM thiourea and 60 μL of a 0.3 M NaCl and 1 mM EDTA solution. DNA was phenol extracted, ethanol precipitated, and resuspended in formamide stop dye. Samples were then loaded onto a 7% urea poly-acrylamide gel alongside a Maxam-Gilbert A+G ladder for alignment (Maxam and Gilbert, 1977) and end-labeled MspI-digested pBR322 (NEB) for size reference. Cleaved DNA was quantified with the SAFA software (Das et al., 2005).

**Bridging cleavage assay**

Bridged filaments were formed for the bridging cleavage assay using the previously published bridging pull down assay with the following modifications (van der Valk et al., 2018; van der Valk et al., 2017). Instead of prey DNA that is the same sequence as the bait, prey DNA for cleavage was made by PCR amplification from pMK110 with primer 7486 5′ end-labeled with [$\gamma$-$P^{32}$]ATP and cold primer 7491 to make a 549 bp fragment from the *bgl* operon containing the DRE and part of *bglG* (55% AT). Prey DNA was gel purified and cleaned with Qiagen PCR clean up kit. To form bridged filaments, 5 nM biotinylated 685 bp DNA fragment (68% AT) coupled to M280 streptavidin dynabeads (ThermoFisher) was incubated with 5 nM prey DNA and variable H-NS concentrations in incubation buffer (40 mM HEPES-KOH, pH 8.0, 100 mM potassium glutamate, 8 mM magnesium aspartate, 5% glycerol, 0.02% NP-40, and 100 μg acetylated BSA/mL). H-NS was added at 5, 6, or 8 μM where 40% of the protein was

Fe·EDTA–H-NS and 60% was wt H-NS to have enough protein to form bridged filaments. In this assay, high levels of H-NS (>5 μM) are required to form the bridged filaments, presumably due to DNA sequestration on the positively charged streptavidin beads (**Figure 7 – figure supplement 2B**; (van der Valk et al., 2017). Bridged filaments were formed for 20 min at 20 °C shaking at 1000 rpm on an Eppendorf shaker before being purified from unbound protein and DNA using a magnetic rack. Beads containing bridged filaments were resuspended in 20 μL incubation buffer followed by rapid addition of 1 μL 50 mM ascorbic acid and 4 μL 12.5 mM $H_2O_2$. Cleavage continued for 2 min at RT before stopping the reaction with 25 μL 22.4 mM thiourea and 20 μL TE. DNA was phenol extracted, ethanol precipitated, and resuspended in formamide stop dye. Amount of radioactivity retained on beads by the conjugated-wt mixture was measured by scintillation counting alongside counts from bridging with wt H-NS and prey DNA only. Percent bridging was calculated as a fraction of counts retained on beads out of the total prey signal. Fraction bridged was likely decreased by any DNA lost during sample clean-up. Cleavage products were then visualized and quantified as described above for the tethered cleavage assay.

**MD simulation of FeBABE-H-NS on DNA movement**

All molecular dynamics simulations were run in the program OpenMM (Eastman et al., 2017) with the Charmm36 force field (Best et al., 2012; Hart et al., 2012). Initial setup and post-processing analyses were conducted using the program CHARMM (Brooks et al., 2009). We obtained an initial starting model of H-NS bound to a 12-bp strand of duplex DNA from Vreede and coworkers (Riccardi et al., 2019). We "mutated" the sidechain of Cys102 in that model to an FeBABE moiety and generated force field parameters for the organic component of FeBABE by analogy to existing functional groups in the Charmm36 force field. We modeled the $Fe^{2+}$ using a cationic dummy-atom model (Liao et al., 2017). We dissolved the H-NS/FeBABE/DNA system in a cubic box of water with 70 Å edges. We added $Na^+$ and $Cl^-$ ions to reach an ionic strength corresponding to 150 mM and to neutralize overall charge. The complete system contained 31,518 atoms. Following a short geometry optimization, the system heated from 48 K to 298 K during 1 ns and equilibrated at 298K and 1 atm for an additional 4 ns prior to production

simulations. The simulations employed the NPT ensemble using the Andersen thermostat and MC barostat. Non-bonded interactions were cutoff above a distance of 12 Å with a switching function from 10 Å to 12 Å and the integration time step was 1 fs. We placed a harmonic restraint with force constant of 1 kcal/mol•$A^2$ on DNA backbone atoms to prevent unraveling at the ends of the duplex, which we observed in initial tests. Under these conditions, we simulated the system for 500 ns, taking snapshots every 10 ps.

In each frame, we calculated the distance from the $Fe^{2+}$ to every C5', because FeBABE-induced cleavage is thought to occur by •OH radical reacting with C5'-H (Balasubramanian et al., 1998). We modeled the concentration of •OH radicals at C5' based on a diffusion model where the concentration of radicals around the $Fe^{2+}$ decays according to a Gaussian distribution (Cheal et al., 2009) To account for cleavage sites that are shielded from solvent by the bound protein, the average concentration of •OH at each site was scaled linearly by the average solvent accessible surface area of the corresponding C5' methylene group (**Figure 6 – figure supplement 2A**). The scaled concentration at each site yields the cleavage propensities displayed in **Figure 6A**. The simulations used a relatively short 12bp duplex in order to achieve reasonable computational efficiency and run long enough simulations to adequately sample the conformational flexibility of the FeBABE. To model cleavage propensities for bps at greater distances from the H-NS binding site we assumed that long-range interactions with DNA do not influence the overall behavior of FeBABE and thus estimated those cleavage propensities by adding additional DNA to the saved trajectories after the simulation was complete.

**Modeling of DBDs in H-NS filaments**

The •OH footprinting data for each H-NS concentration tested (0-4 μM titration), each strand, and each independent replicate was normalized by dividing the signal in each lane by the average signal within a lane. The resulting ratio was $\log_2$ transformed to center the data at zero. The $\log_2$-normalized deviance from the average signal from each strand was separately fit to a Bayesian multi-level model of the following form:

$$y_i \sim \text{Student's } t(\mu_i, \sigma, \nu)$$
$$\mu_i = \alpha_{\text{bp}[i]} + \beta_{\text{bp}[i]}[\text{HNS}]$$

$$\begin{bmatrix}\alpha_{\text{bp}}\\ \beta_{\text{bp}}\end{bmatrix} = \text{MVNormal}\left(\begin{bmatrix}\alpha\\ \beta\end{bmatrix}, \boldsymbol{S}\right)$$
$$\boldsymbol{S} = \begin{pmatrix}\sigma_\alpha & 0\\ 0 & \sigma_\beta\end{pmatrix}\boldsymbol{R}\begin{pmatrix}\sigma_\alpha & 0\\ 0 & \sigma_\beta\end{pmatrix}$$
$$\boldsymbol{R} = \begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}$$

With the following weakly regularizing priors:

$$\alpha \sim \text{Normal}(0, 10)$$
$$\beta \sim \text{Normal}(0, 10)$$
$$\sigma \sim \text{HalfCauchy}(0, 1)$$
$$\sigma_\alpha \sim \text{HalfCauchy}(0, 1)$$
$$\sigma_\beta \sim \text{HalfCauchy}(0, 1)$$
$$\boldsymbol{R} \sim \text{LKJcorr}(2)$$

The models were fit using the R package brms (version 2.10.0) (Bürkner, 2017) and default priors were used for any parameter for which a prior is not specified above.

Similarly, the tethered-cleavage data for each concentration (0-3 μM titration) was normalized by first dividing each lane by a constant factor resulting from quantification of a short scan of the entire lane. The data was then centered by subtracting the average signal within each lane. The centered and normalized signal for each strand separately was then fit to a Bayesian multilevel, of the form described above, with the following priors:

$$\alpha \sim \text{Normal}(0, 0.1)$$
$$\beta \sim \text{Normal}(0, 0.1)$$

Default priors were used for every other parameter not specified above. Notably, the 4 μM and 5 μM concentrations were not included in the tethered cleavage fits due to their substantial deviance from linearity. Point estimates for each slope at each bp in the models above were estimated from the average of the posterior distribution for that slope. These estimates can be interpreted as the H-NS concentration-dependent change in signal at each bp with data-driven shrinkage towards the grand mean of slopes for every bp.

To match the MD simulations with the *in vitro* data, the slope estimates for each strand in each data type were Robust-Z normalized using the following formula:

$$y_{\text{scaled}} = \frac{y - \text{median}(y)}{1.4826 \times \text{MAD}(y)}$$

Where, MAD is the median absolute deviation and 1.4826 is a scaling factor to convert the MAD to an estimate for the standard deviation (Hampel, 1974; Rousseeuw and Croux, 1993). Similarly, the expected 15 bp cleavage pattern of a single H-NS DBD (**Figure 6A**) was also scaled to a standard Normal using the formula above. Using a sliding window, the Manhattan distance between the scaled tethered cleavage estimates and the scaled 15 bp expected cleavage pattern was calculated for each base centered on the sixth bp of the cleavage pattern (triangle in **Figure 6B**). Due to the asymmetry of the expected cleavage pattern we performed this calculation in both directions along the DNA corresponding to two different possible orientations of the H-NS DNA binding domain. Since this score is a distance, lower values indicate a closer match between the expected cleavage pattern and the tethered cleavage data.

To assess the statistical significance of the MD match scores, we performed this same procedure on 1000 different random shuffles of the tethered cleavage data for each direction of the H-NS DBD. We considered MD match scores below the 5$^{th}$ percentile of the random MD match scores to signify a statistically significant match between the expected cleavage pattern and the tethered cleavage data.

To determine possible localizations of the H-NS DBD in the ·OH footprinting data, we performed the same procedure described above using the solvent accessibility of the 5′ carbon of the DNA duplex as the expected pattern of protection imparted by the H-NS DBD (**Figure 6 – figure supplement 2A**) and matched that pattern to the ·OH footprinting data (**Figure 5 – figure supplement 3**).

**Predicting H-NS DNA binding pattern genome-wide**

To generate a dataset for genome-wide prediction of the locations of H-NS filaments, the *E. coli* genome (accession U00096.3) was divided into a tiling array of 1000 bp windows staggered every 250 bp. A window was considered to be within a peak if over 50% of the window was covered by a called peak for that dataset. Any window that did not meet this criterion was considered as part of the unbound negative set. For each window the following sequence features

were generated: a count of the As and Ts (A/T content), a count of 4mers matching the criteria for an AT tract, a count of 4mers matching the criteria for a TA step, a count of motif matches to the Lang et al. H-NS PWM (Lang et al., 2007), and a count of motif matches to the RegulonDB (Santos-Zavaleta et al., 2019) Fis motif. Motif matches we generated by running FIMO (version 5.05) (Grant et al., 2011) over the entire *E. coli* genome with a third-order Markov background model, searching on both strands, and a p-value cutoff of 0.001 with no FDR correction. Each of these counts was then normalized by the length of the window. DNA structure features were generated using the DNAshapeR package (Chiu et al., 2016) run over the entire *E. coli* genome. The MGW, ProT, and EP bp level features were then Robust-Z normalized using data from the entire genome and the number of bases below a -2 cutoff on the Robust-Z scale for each feature was determined and normalized by the length of the window.

To predict genome-wide locations of H-NS filaments a LASSO logistic regression model was trained on the H-NS ChIP data using glmnet (Friedman et al., 2010)(version 2.0-16). All features were centered and scaled before training. The LASSO model reveals the minimal set of DNA features needed for the prediction of a given dataset by only including non-zero coefficients for a feature if that feature provides independent information from other non-zero coefficients already present in the model. Since all input data was centered and scaled before training the model, the relative magnitude of each coefficient can be interpreted as the importance for that feature in predicting the outcome, with positive values indicating an increase in the probability of a region being bound by a protein when the associated feature is a higher value. Models were trained using 10-fold cross-validation with stratified sampling and the penalty parameter, $\lambda$, was chosen as the value one standard error away from the minimum binomial deviance. $\lambda$ was determined individually for each subsampled dataset using 10-fold cross-validation internally to test for sensitivity to the exact data used to train the model. Predictive performance for each ChIP and Nitro-seq dataset was tested by calculating the area under the precision-recall curve (AUC-PRC) using the method of Davis and Goadrich (Davis and Goadrich, 2006) with the R package PRROC (version 1.3.1)(Grau et al., 2015). The average, minimum, and maximum AUC-PRC was determined using the fitted models for each subsampled dataset to predict the binding locations for each ChIP and Nitro-seq dataset.

## ACKNOWLEDGEMENTS

We thank Jessica Vera for guidance and scripts for analyzing ChIP-seq and Nitro-seq data, Jocelyne Vreede (University of Amsterdam) for sharing coordinates for the H-NS DNA binding domain and DNA structure and modeling the FeBABE moiety into the coordinates, Wilma Ross for kindly providing purified Fis, advice on footprinting experiments, and equipment for nitrocellulose binding experiments, and members of the Landick lab for helpful discussions and comments on this manuscript. We acknowledge the University of Wisconsin Biotechnology Center DNA Sequencing Facility for providing HiSeq2500 sequencing facilities and services. Computational resources were provided by an XSEDE allocation (TG-MCB180084). This work was supported by NIH grant R01 GM38660 to R.L.

## REFERENCES


Ali, S.S., Whitney, J.C., Stevenson, J., Robinson, H., Howell, P.L., and Navarre, W.W. (2013). Structural insights into the regulation of foreign genes in Salmonella by the Hha/H-NS complex. J Biol Chem *288*, 13356-13369. DOI: 10.1074/jbc.M113.455378.

Altuvia, S., Almiron, M., Huisman, G., Kolter, R., and Storz, G. (1994). The *dps* promoter is activated by OxyR during growth and by IHF and sigma-S in stationary phase. Mol. Microbiol. *13*, 265-272

Amit, R., Oppenheim, A.B., and Stavans, J. (2003). Increased bending rigidity of single DNA molecules by H-NS, a temperature and osmolarity sensor. Biophys. J. *84*, 2467-2473

Arold, S.T., Leonard, P.G., Parkinson, G.N., and Ladbury, J.E. (2010). H-NS forms a superhelical protein scaffold for DNA condensation. PNAS *107*, 15728-15732. DOI: 10.1073/pnas.1006966107/-/DCSupplemental.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol *2*, 2006 0008. DOI: 10.1038/msb4100050.

Badaut, C., Williams, R., Arluison, V., Bouffartigues, E., Robert, B., Buc, H., and Rimsky, S. (2002). The Degree of Oligomerization of the H-NS Nucleoid Structuring Protein Is Related to Specific Binding to DNA. J. Biol. Chem. *277*, 41657-41666. DOI: 10.1074/jbc.M206037200.

Balasubramanian, B., Pogozelski, W.K., and Tullis, T.D. (1998). DNA strand breaking by the hydroxyl radical is governed by the accessible surface area of the hydrogen atoms of the DNA backbone. Proc. Natl. Acad. Sci. USA *95*, 9738-9743

Banos, R.C., Vivero, A., Aznar, S., Garcia, J., Pons, M., Madrid, C., and Juarez, A. (2009). Differential regulation of horizontally acquired and core genome genes by the bacterial modulator H-NS. PLoS Genet. *5*, e1000513. DOI: 10.1371/journal.pgen.1000513.

Best, R.B., Zhu, X., Shim, J., Lopes, P.E.M., Mittal, J., Feig, M., and MacKerell, A.D. (2012). Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone phi, psi and Side-Chain chi(1) and chi(2) Dihedral Angles. Journal of Chemical Theory and Computation *8*, 3257-3273. DOI: 10.1021/ct300400x.

Boudreau, B.A., Hron, D.R., Qin, L., van der Valk, R.A., Kotlajich, M.V., Dame, R.T., and Landick, R. (2018). StpA and Hha stimulate pausing by RNA polymerase by promoting DNA-DNA bridging of H-NS filaments. Nucleic acids research *46*, 5525-5546. DOI: 10.1093/nar/gky265.

Bouffartigues, E., Buckle, M., Badaut, C., Travers, A., and Rimsky, S. (2007). H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. Nat. Struct. Mol. Biol. *14*, 441-448. DOI: 10.1038/nsmb1233.

Brooks, B.R., Brooks, C.L., Mackerell, A.D., Nilsson, L., Petrella, R.J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S.*, et al.* (2009). CHARMM: The Biomolecular Simulation Program. J. Comput. Chem. *30*, 1545-1614. DOI: 10.1002/Jcc.21287.

Burkhoff, A.M., and Tullis, T.D. (1988). Structural details of an adenine tract that does not cause DNA to bend. Nature *331*, 455-457

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software *80*. DOI: 10.18637/jss.v080.i01.

Campbell, E.A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., and Darst, S.A. (2001). Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. Cell *104*, 901-912

Chandraprakash, D., and Seshasayee, A.S.N. (2014). Inhibition of factor-dependent transcription termination in Escherichia coli might relieve xenogene silencing by abrogating H-NS-DNA interactions in vivo. J. Biosci. (Bangalore) *39*, 53-61. DOI: 10.1007/s12038-014-9413-4.

Cheal, S.M., Ng, M., Barrios, B., Miao, Z., Kalani, A.K., and Meares, C.F. (2009). Mapping protein-protein interactions by localized oxidation: consequences of the reach of hydroxyl radical. Biochemistry *48*, 4577-4586. DOI: 10.1021/bi900273j.

Chiu, T.-P., Rao, S., Mann, R.S., Honig, B., and Rohs, R. (2017). Genome-wide prediction of minor-groove electrostatic potential enables biophysical modeling of protein–DNA binding. Nucleic Acids Res. *45*, 12565-12576. DOI: 10.1093/nar/gkx915.

Chiu, T.P., Comoglio, F., Zhou, T., Yang, L., Paro, R., and Rohs, R. (2016). DNAshapeR: an R/Bioconductor package for DNA shape prediction and feature encoding. Bioinformatics *32*, 1211-1213. DOI: 10.1093/bioinformatics/btv735.

Cho, B.K., Knight, E.M., Barrett, C.L., and Palsson, B.O. (2008). Genome-wide analysis of Fis binding in Escherichia coli indicates a causative role for A-/AT-tracts. Genome Res. *18*, 900-910. DOI: 10.1101/gr.070276.107.

Chung, D., Zhang, Q., and Keles, S. (2014). MOSAiCS-HMM: A model-based approach for detecting regions of histone modifications from ChIP-seq data. In Statistical Analysis of Next Generation Sequencing Data, S. Data, and D. Nettleton, eds. (Springer).

Cook, P.R., and Marenduzzo, D. (2018). Transcription-driven genome organization: a model for chromosome structure and the regulation of gene expression tested through simulations. Nucleic Acids Res. *46*, 9895-9906. DOI: 10.1093/nar/gky763.

Dame, R.T., Noom, M.C., and Wuite, G.J. (2006). Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. Nature *444*, 387-390. DOI: 10.1038/nature05283.

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. DOI: 10.1038/s41576-019-0185-4.

Das, R., Laederach, A., Pearlman, S.M., Herschlag, D., and Altman, R.B. (2005). SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. RNA *11*, 344-354. DOI: 10.1261/rna.7214405.

Davis, J., and Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. In Proceedings of the 23rd international conference on Machine learning (Pittsburgh, Pennsylvania, USA, Association for Computing Machinery), pp. 233–240. DOI: 10.1145/1143844.1143874.

Dole, S., Nagarajavel, V., and Schnetz, K. (2004). The histone-like nucleoid structuring protein H-NS represses the *Escherichia coli bgl* operon downstream of the promoter. Mol Microbiol *52*, 589-600. DOI: 10.1111/j.1365-2958.2004.04001.x.

Eastman, P., Swails, J., Chodera, J.D., McGibbon, R.T., Zhao, Y.T., Beauchamp, K.A., Wang, L.P., Simmonett, A.C., Harrigan, M.P., Stern, C.D.*, et al.* (2017). OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. PLoS Comp. Biol. *13*, 17. DOI: 10.1371/journal.pcbi.1005659.

Favorov, A., Mularoni, L., Cope, L.M., Medvedeva, Y., Mironov, A.A., Makeev, V.J., and Wheelan, S.J. (2012). Exploring Massive, Genome Scale Datasets with the GenometriCorr Package. PLoS Comp. Biol. *8*, e1002529. DOI: 10.1371/journal.pcbi.1002529.

Friedman, J.H., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. 2010 *33*, 22 Journal of Statistical Software. DOI: 10.18637/jss.v033.i01.

Fukui, N., Oshima, T., Ueda, T., Ogasawara, N., and Tobe, T. (2016). Gene Activation through the Modulation of Nucleoid Structures by a Horizontally Transferred Regulator, Pch, in Enterohemorrhagic Escherichia coli. PLoS One *11*, e0149718. DOI: 10.1371/journal.pone.0149718.

Gao, Y., Foo, Y.H., Winardhi, R.S., Tang, Q., Yan, J., and Kenney, L.J. (2017). Charged residues in the H-NS linker drive DNA binding and gene silencing in single cells. PNAS *114*, 12560-12565. DOI: 10.1073/pnas.1716721114.

Ghosh, I.N., and Landick, R. (2016). OptSSeq: High-Throughput Sequencing Readout of Growth Enrichment Defines Optimal Gene Expression Elements for Homoethanologenesis. ACS Synth Biol *5*, 1519-1534. DOI: 10.1021/acssynbio.6b00121.

Gordon, B.R., Li, Y., Cote, A., Weirauch, M.T., Ding, P., Hughes, T.R., Navarre, W.W., Xia, B., and Liu, J. (2011). Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. PNAS *108*, 10690-10695. DOI: 10.1073/pnas.1102544108.

Grainger, D.C. (2016). Structure and function of bacterial H-NS protein. Biochem. Soc. Trans. *44*, 1561-1569. DOI: 10.1042/BST20160190.

Grainger, D.C., Goldberg, M.D., Lee, D.J., and Busby, S.J. (2008). Selective repression by Fis and H-NS at the Escherichia coli dps promoter. Mol Microbiol *68*, 1366-1377. DOI: 10.1111/j.1365-2958.2008.06253.x.

Grainger, D.C., Hurd, D., Goldberg, M.D., and Busby, S.J. (2006). Association of nucleoid proteins with coding and non-coding segments of the Escherichia coli genome. Nucleic acids research *34*, 4642-4652. DOI: 10.1093/nar/gkl542.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017-1018. DOI: 10.1093/bioinformatics/btr064.

Grau, J., Grosse, I., and Keilwagen, J. (2015). PRROC: computing and visualizing precision-recall and receiver operating characteristic curves in R. Bioinformatics *31*, 2595-2597. DOI: 10.1093/bioinformatics/btv153 %J Bioinformatics.

Greiner, D.P., Miyake, R., Moran, J.K., Jones, A.D., Negishi, T., Ishihama, A., and Meares, C.F. (1997). Synthesis of the protein cutting reagent Iron (*S*)-1-(*p*-

Bromoacetamidobenzyl)ethylenediaminetetraacetate and conjugation to cysteine side chains. Bioconj. Chem. *8*, 44-48

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). circlize implements and enhances circular visualization in R. Bioinformatics *30*, 2811-2812. DOI: 10.1093/bioinformatics/btu393 %J Bioinformatics.

Gulvady, R., Gao, Y., Kenney, L.J., and Yan, J. (2018). A single molecule analysis of H-NS uncouples DNA binding affinity from DNA specificity. Nucleic acids research *46*, 10216-10224. DOI: 10.1093/nar/gky826.

Hampel, F.R. (1974). The Influence Curve and its Role in Robust Estimation. Journal of the American Statistical Association *69*, 383-393. DOI: 10.1080/01621459.1974.10482962.

Hart, K., Foloppe, N., Baker, C.M., Denning, E.J., Nilsson, L., and MacKerell, A.D. (2012). Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium. Journal of Chemical Theory and Computation *8*, 348-362. DOI: 10.1021/ct200723y.

Henggee-Aronis, R. (2002). Starionary phase gene regulation: what makes an Escherichia coli promoter σS-selective? Curr. Opin. Microbiol. *5*, 591-595

Herring, C.D., Raffaelle, M., Allen, T.E., Kanin, E.I., Landick, R., Ansari, A.Z., and Palsson, B.O. (2005). Immobilization of Escherichia coli RNA polymerase and location of binding sites by use of chromatin immunoprecipitation and microarrays. J Bacteriol *187*, 6166-6174. DOI: 10.1128/JB.187.17.6166-6174.2005.

Japaridze, A., Renevey, S., Sobetzko, P., Stoliar, L., Nasser, W., Dietler, G., and Muskhelishvili, G. (2017). Spatial organization of DNA sequences directs the assembly of bacterial chromatin by a nucleoid-associated protein. J Biol Chem *292*, 7607-7618. DOI: 10.1074/jbc.M117.780239.

Kahramanoglou, C., Seshasayee, A.S., Prieto, A.I., Ibberson, D., Schmidt, S., Zimmermann, J., Benes, V., Fraser, G.M., and Luscombe, N.M. (2011). Direct and indirect effects of H-NS and Fis on global gene expression control in Escherichia coli. Nucleic acids research *39*, 2073-2091. DOI: 10.1093/nar/gkq934.

Kim, D., Hong, J.S., Qiu, Y., Nagarajan, H., Seo, J.H., Cho, B.K., Tsai, S.F., and Palsson, B.O. (2012). Comparative analysis of regulatory elements between Escherichia coli and Klebsiella pneumoniae by genome-wide transcription start site profiling. PLoS Genet. *8*, e1002867. DOI: 10.1371/journal.pgen.1002867.

Kotlajich, M.V., Hron, D.R., Boudreau, B.A., Sun, Z., Lyubchenko, Y.L., and Landick, R. (2015). Bridged filaments of histone-like nucleoid structuring protein pause RNA polymerase and aid termination in bacteria. Elife *4,* e04970. DOI: 10.7554/eLife.04970.

Krietenstein, N., Wal, M., Watanabe, S., Park, B., Peterson, C.L., Pugh, B.F., and Korber, P. (2016). Genomic Nucleosome Organization Reconstituted with Pure Proteins. Cell *167*, 709-721 e712. DOI: 10.1016/j.cell.2016.09.045.

Kuan, P.F., Chung, D., Pan, G., Thomson, J.A., Stewart, R., and Keles, S. (2011). A Statistical Framework for the Analysis of ChIP-Seq Data. J Am Stat Assoc *106*, 891-903. DOI: 10.1198/jasa.2011.ap09706.

Lang, B., Blot, N., Bouffartigues, E., Buckle, M., Geertz, M., Gualerzi, C.O., Mavathur, R., Muskhelishvili, G., Pon, C.L., Rimsky, S.*, et al.* (2007). High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. Nucleic acids research *35*, 6330-6337. DOI: 10.1093/nar/gkm712.

Leonard, P.G., Ono, S., Gor, J., Perkins, S.J., and Ladbury, J.E. (2009). Investigation of the self-association and hetero-association interactions of H-NS and StpA from Enterobacteria. Mol Microbiol *73*, 165-179. DOI: 10.1111/j.1365-2958.2009.06754.x.

Liao, Q.H., Pabis, A., Strodel, B., and Kamerlin, S.C.L. (2017). Extending the Nonbonded Cationic Dummy Model to Account for Ion-Induced Dipole Interactions. Journal of Physical Chemistry Letters *8*, 5408-5414. DOI: 10.1021/acs.jpclett.7b02358.

Lim, C.J., Lee, S.Y., Kenney, L.J., and Yan, J. (2012). Nucleoprotein filament formation is the structural basis for bacterial protein H-NS gene silencing. Sci Rep *2*, 509. DOI: 10.1038/srep00509.

Liu, Y., Chen, H., Kenney, L.J., and Yan, J. (2010). A divalent switch drives H-NS/DNA-binding conformations between stiffening and bridging modes. Genes & development *24*, 339-344. DOI: 10.1101/gad.1883510.

Lucchini, S., McDermott, P., Thompson, A., and Hinton, J.C. (2009). The H-NS-like protein StpA represses the RpoS (sigma 38) regulon during exponential growth of Salmonella Typhimurium. Mol Microbiol *74*, 1169-1186. DOI: 10.1111/j.1365-2958.2009.06929.x.

Lucchini, S., Rowley, G., Goldberg, M.D., Hurd, D., Harrison, M., and Hinton, J.C. (2006). H-NS mediates the silencing of laterally acquired genes in bacteria. PLoS Pathog *2*, e81. DOI: 10.1371/journal.ppat.0020081.

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 Å resolution. Nature *389*, 251-260

Maxam, A.M., and Gilbert, W. (1977). A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA *74*, 560-564

Mooney, R.A., Davis, S.E., Peters, J.M., Rowland, J.L., Ansari, A.Z., and Landick, R. (2009). Regulator trafficking on bacterial transcription units in vivo. Mol Cell *33*, 97-108. DOI: 10.1016/j.molcel.2008.12.021.

Muller, C.M., Dobrindt, U., Nagy, G., Emody, L., Uhlin, B.E., and Hacker, J. (2006). Role of histone-like proteins H-NS and StpA in expression of virulence determinants of uropathogenic Escherichia coli. J Bacteriol *188*, 5428-5438. DOI: 10.1128/JB.01956-05.

Myers, K.S., Yan, H., Ong, I.M., Chung, D., Liang, K., Tran, F., Keles, S., Landick, R., and Kiley, P.J. (2013). Genome-scale analysis of escherichia coli FNR reveals complex features of transcription factor binding. PLoS Genet. *9*, e1003565. DOI: 10.1371/journal.pgen.1003565.

Navarre, W.W., Porwollik, S., Wang, Y., McClelland, M., Rosen, H., Libby, S.J., and Fang, F.C. (2006). Selective silencing of foreign DNA with low GC content by the H-NS protein in Salmonella. Science *313*, 236-238. DOI: 10.1126/science.1128794.

Nieto, J.M., Madrid, C., Prenafeta, A., Miquelay, E., Balsalobre, C., Carrascal, M., and Juarez, A. (2000). Expression of the hemolysin operon in *Escherichia coli* is modulated by a nucleoid-protein complex that includes the proteins Hha and H-NS. Mol. Genet. Genomics *263*, 349-358

Noom, M.C., Navarre, W.W., Oshima, T., Wuite, G.J., and Dame, R.T. (2007). H-NS promotes looped domain formation in the bacterial chromosome. Current biology : CB *17*, R913-914. DOI: 10.1016/j.cub.2007.09.005.

Owen-Hughes, T., Pavitt, G., Santos, D., Sidebotham, J., Hulton, C., Hinton, J., and Higgins, C. (1992). The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. . Cell *71*, 255-265

Papoulas, O. (1996). Rapid Separation of Protein-Bound DNA from Free DNA Using Nitrocellulose Filters. Curr Protoc Mol Biol, 12. DOI: 10.1002/0471142727.mb1208s36.

Peters, J.M., Mooney, R.A., Grass, J.A., Jessen, E.D., Tran, F., and Landick, R. (2012). Rho and NusG suppress pervasive antisense transcription in Escherichia coli. Genes & development *26*, 2621-2633. DOI: 10.1101/gad.196741.112.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842. DOI: 10.1093/bioinformatics/btq033.

Rangarajan, A.A., and Schnetz, K. (2018). Interference of transcription across H-NS binding sites and repression by H-NS. Mol Microbiol *108*, 226-239. DOI: 10.1111/mmi.13926.

Renault, M., Garcia, J., Cordeiro, T.N., Baldus, M., and Pons, M. (2013). Protein oligomers studied by solid-state NMR--the case of the full-length nucleoid-associated protein histone-like nucleoid structuring protein. FEBS J. *280*, 2916-2928. DOI: 10.1111/febs.12297.

Riccardi, E., van Mastbergen, E.C., Navarre, W.W., and Vreede, J. (2019). Predicting the mechanism and rate of H-NS binding to AT-rich DNA. PLoS Comput Biol *15*, e1006845. DOI: 10.1371/journal.pcbi.1006845.

Rousseeuw, P.J., and Croux, C. (1993). Alternatives to the Median Absolute Deviation. Journal of the American Statistical Association *88*, 1273-1283. DOI: 10.1080/01621459.1993.10476408.

Santos-Zavaleta, A., Salgado, H., Gama-Castro, S., Sanchez-Perez, M., Gomez-Romero, L., Ledezma-Tejeida, D., Garcia-Sotelo, J.S., Alquicira-Hernandez, K., Muniz-Rascado, L.J., Pena-Loredo, P.*, et al.* (2019). RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in E. coli K-12. Nucleic acids research *47*, D212-D220. DOI: 10.1093/nar/gky1077.

Shahul Hameed, U.F., Liao, C., Radhakrishnan, A.K., Huser, F., Aljedani, S.S., Zhao, X., Momin, A.A., Melo, F.A., Guo, X., Brooks, C.*, et al.* (2019). H-NS uses an autoinhibitory conformational switch for environment-controlled gene silencing. Nucleic acids research *47*, 2666-2680. DOI: 10.1093/nar/gky1299.

Shen, B.A., and Landick, R. (2019). Transcription of Bacterial Chromatin. J. Mol. Biol. DOI: 10.1016/j.jmb.2019.05.041.

Singh, S.S., Singh, N., Bonocora, R.P., Fitzgerald, D.M., Wade, J.T., and Grainger, D.C. (2014). Widespread suppression of intragenic transcription initiation by H-NS. Genes & development *28*, 214-219. DOI: 10.1101/gad.234336.113.

Srinivasan, R., Chandraprakash, D., Krishnamurthi, R., Singh, P., Scolari, V.F., Krishna, S., and Seshasayee, A.S. (2013). Genomic analysis reveals epistatic silencing of "expensive" genes in Escherichia coli K-12. Mol Biosyst *9*, 2021-2033. DOI: 10.1039/c3mb70035f.

Talbert, P.B., Meers, M.P., and Henikoff, S. (2019). Old cogs, new tricks: the evolution of gene expression in a chromatin context. Nat. Rev. Genet. *20*, 283-297. DOI: 10.1038/s41576-019-0105-7.

Thomason, M.K., Bischler, T., Eisenbart, S.K., Forstner, K.U., Zhang, A., Herbig, A., Nieselt, K., Sharma, C.M., and Storz, G. (2015). Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in Escherichia coli. J Bacteriol *197*, 18-28. DOI: 10.1128/JB.02096-14.

Touchon, M., and Rocha, E.P. (2010). The small, slow and specialized CRISPR and anti-CRISPR of Escherichia and Salmonella. PLoS One *5*, e11126. DOI: 10.1371/journal.pone.0011126.

Tupper, A.E., Owen-Hughes, T.A., Ussery, D.W., Santos, D.S., Ferguson, D.J.P., Sidebotham, J.M., Hinton, J.C., and Higgins, C.F. (1994). The chromatin-associated protein H-NS alters DNA topology *in vitro*. EMBO J *13*, 258-268

Ueda, T., Takahashi, H., Uyar, E., Ishikawa, S., Ogasawara, N., and Oshima, T. (2013). Functions of the Hha and YdgT proteins in transcriptional silencing by the nucleoid proteins, H-NS and StpA, in Escherichia coli. DNA Res. *20*, 263-271. DOI: 10.1093/dnares/dst008.

Ulissi, U., Fabbretti, A., Sette, M., Giuliodori, A.M., and Spurio, R. (2014). Time-resolved assembly of a nucleoprotein complex between Shigella flexneri virF promoter and its

transcriptional repressor H-NS. Nucleic acids research *42*, 13039-13050. DOI: 10.1093/nar/gku1052.

Uyar, E., Kurokawa, K., Yoshimura, M., Ishikawa, S., Ogasawara, N., and Oshima, T. (2009). Differential binding profiles of StpA in wild-type and h-ns mutant cells: a comparative analysis of cooperative partners by chromatin immunoprecipitation-microarray analysis. J Bacteriol *191*, 2388-2391. DOI: 10.1128/JB.01594-08.

van der Valk, R.A., Qin, L., Moolenaar, G.F., and Dame, R.T. (2018). Quantitative Determination of DNA Bridging Efficiency of Chromatin Proteins. In Bacterial Chromatin: Methods in Molecular Biology, R.T. Dame, ed. (New York, NY: Humana Press).

van der Valk, R.A., Vreede, J., Qin, L., Moolenaar, G.F., Hofmann, A., Goosen, N., and Dame, R.T. (2017). Mechanism of environmentally driven conformational changes that modulate H-NS DNA-bridging activity. Elife *6*, e27369. DOI: 10.7554/eLife.27369.

Vora, T., Hottes, A.K., and Tavazoie, S. (2009). Protein occupancy landscape of a bacterial genome. Mol Cell *35*, 247-253. DOI: 10.1016/j.molcel.2009.06.035.

Wang, H., Yehoshua, S., Ali, S.S., Navarre, W.W., and Milstein, J.N. (2014). A biomechanical mechanism for initiating DNA packaging. Nucleic acids research *42*, 11921-11927. DOI: 10.1093/nar/gku896.

Will, W.R., Bale, D.H., Reid, P.J., Libby, S.J., and Fang, F.C. (2015). Evolutionary expansion of a regulatory network by counter-silencing. Nat Commun *5*, 5270. DOI: 10.1038/ncomms6270.

Will, W.R., Whitham, P.J., Reid, P.J., and Fang, F.C. (2018). Modulation of H-NS transcriptional silencing by magnesium. Nucleic acids research *46*, 5717-5725. DOI: 10.1093/nar/gky387.

Wilson, K. (2001). Preparation of Genomic DNA from Bacteria. Curr Protoc Mol Biol *56*, 2.4. DOI: 10.1002/0471142727.mb0204s56.

Yamada, H., Yoshida, T., Tanaka, K., Sasakawa, C., and Mizuno, T. (1991). Molecular analysis of the Escherichia coli has gene encoding a DNA-binding protein, which preferentially recognizes curved DNA sequences. . Mol. Gen. Genet. *230*, 332-336

Zhou, T., Yang, L., Lu, Y., Dror, I., Dantas Machado, A.C., Ghane, T., Di Felice, R., and Rohs, R. (2013). DNAshape: a method for the high-throughput prediction of DNA structural features on a genomic scale. Nucleic acids research *41*, W56-62. DOI: 10.1093/nar/gkt437.

## Figures

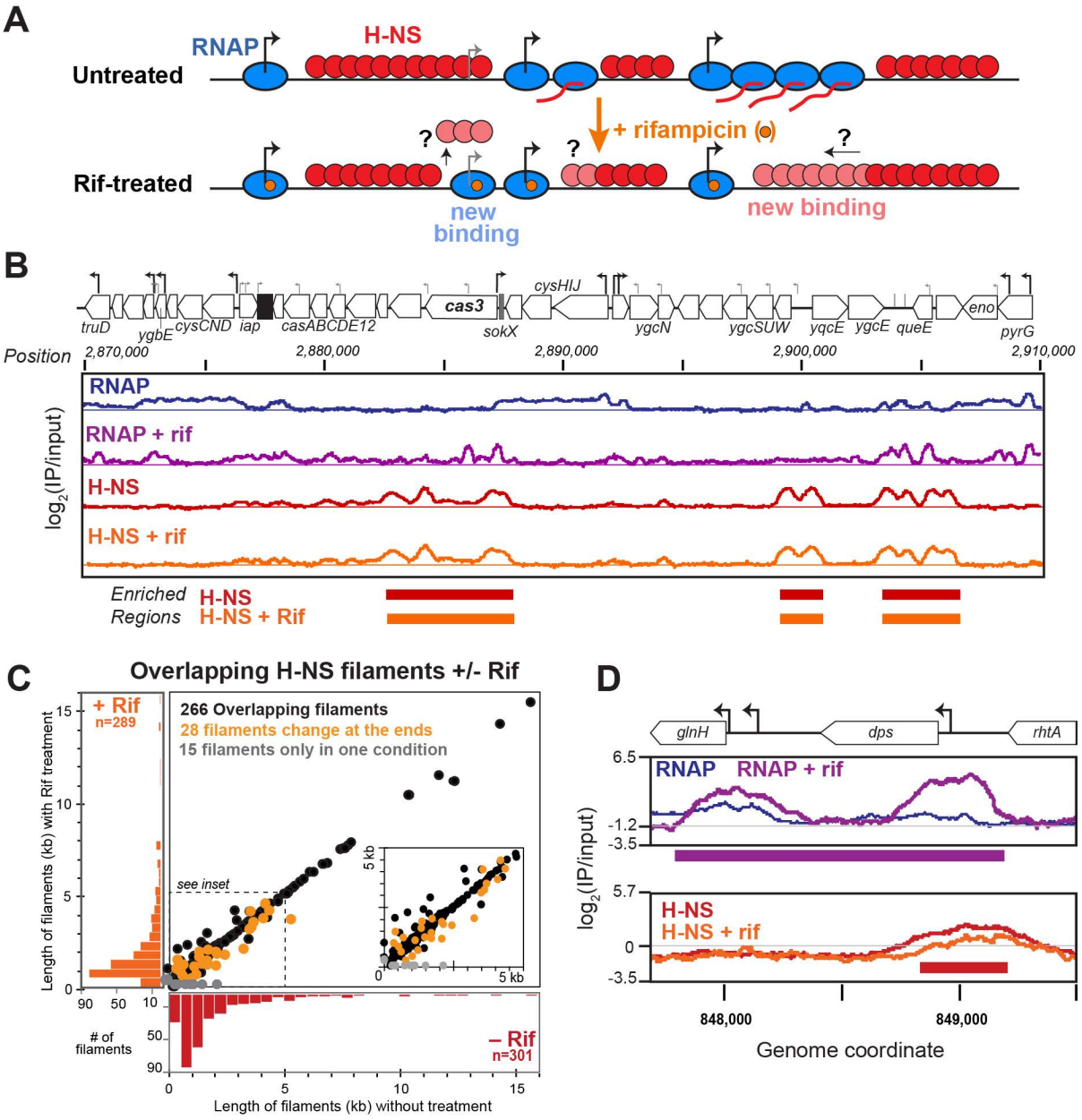


**Figure 1. Halting transcription does not dramatically change H-NS binding *in vivo*.** (A) If H-NS (red) filament boundaries are determined principally by RNAP (blue) actively transcribing genes or bound at promoters, then treatment with rifampicin (Rif), which traps RNAP at promoters, should change the H-NS binding pattern. (B) ChIP-seq locations on the *E. coli* genome of RNAP in the absence (blue) or presence (purple) of Rif and of H-NS in the absence (red) or presence (orange) of Rif. Coverage was normalized to reads per million bp. The RNAP baseline was set to the ChIP signal level in regions known to be transcriptionally silenced

(Mooney et al., 2009) and the H-NS baseline was set at the lowest IP/input values. Genes (white boxes) are shown above the coverage tracks. Strong TSSs are indicated by black bent arrows and weaker TSS are indicated by gray bent arrows (Thomason et al., 2015). Traces are representative of two replicates. H-NS enriched regions identified by MOSAiCS are shown as boxes below ChIP traces. (C) Distribution of the length of H-NS enriched regions (*i.e.,* filaments) identified by MOSAiCS in Rif-treated (red) or untreated (orange) cells is shown on each axis. The scatter plot shows the comparison of length of H-NS filaments overlapping between treatment and no treatment. Filaments that changed in H-NS density at the ends are indicated by golden dots. Filaments only present in one condition are indicated by gray dots. Number of filaments (n) shown in each graph are indicated. (D) H-NS filament at the *dps* promoter in untreated cells (red line) is lost after Rif-treatment (orange line), likely due to the appearance of RNAP trapped at the *dps* promoter (purple line). Purple and red bars indicate enriched regions identified by MOSAiCS. For A and C, values for $\log_2$(IP/input) were quantile normalized between H-NS IP samples or between Beta IP samples and are on different scales.

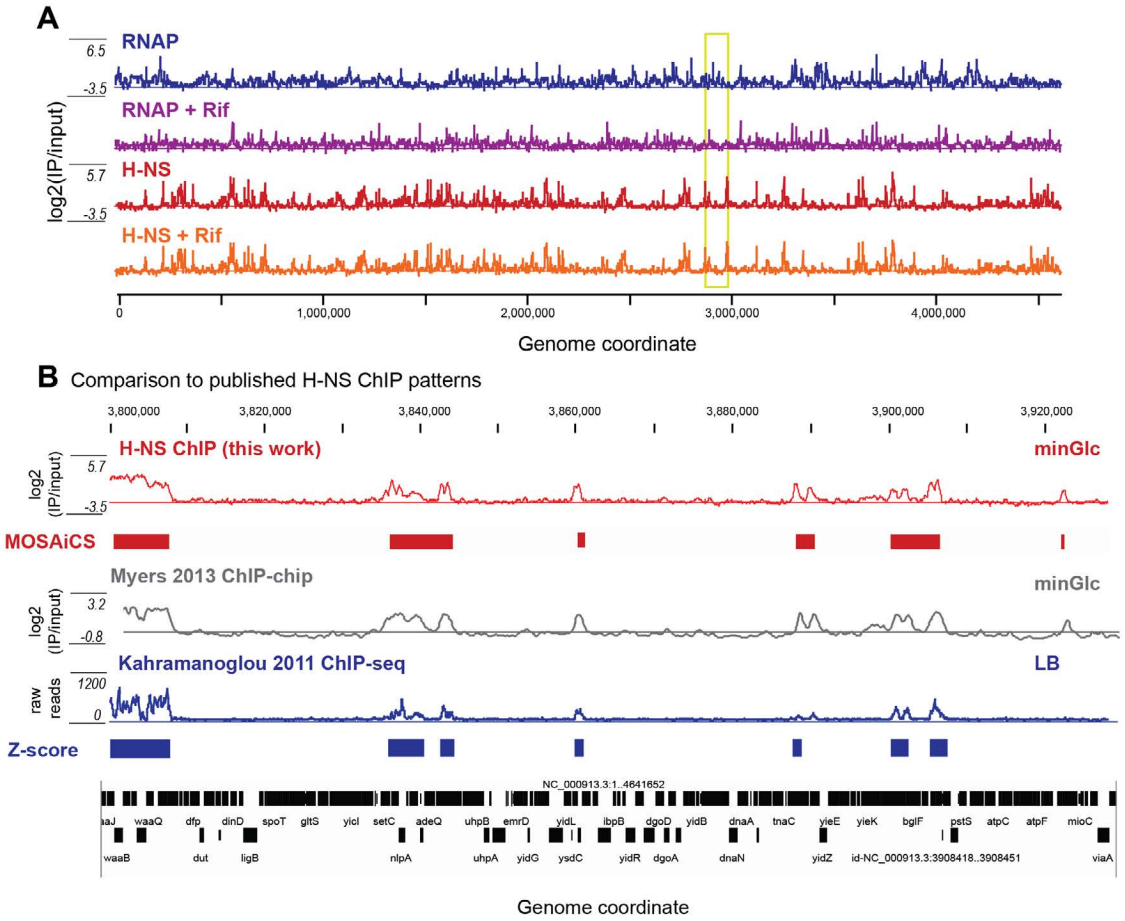


**Figure 1 - figure supplement 1. H-NS and RNAP ChIP-seq in untreated and Rif-treated cells.**

(A) Normalized read coverage (log2(IP/input)) of the entire genome for H-NS and RNAP (Beta) IP in the presence (+Rif) or absence of rifampicin. Beta and H-NS ChIP were normalized separately, and the scales are indicated. Yellow box indicates the region shown in Figure 1B. (B) Comparison of H-NS filaments as identified in this work by MOSAiCS (Figure 1) to both H-NS binding pattern identified by ChIP-chip (Myers et al., 2013) (gray) and H-NS filaments identified in (Kahramanoglou et al., 2011) with a Z-score analysis (blue). The culture medium was either Luria-Broth (LB) or minimal medium with glucose (minGlc). H-NS binding pattern is similar in each condition. We only observed minor differences between our results and previously published data sets in the number of binding regions and average lengths of filaments likely due to the different experimental and analysis methods used.

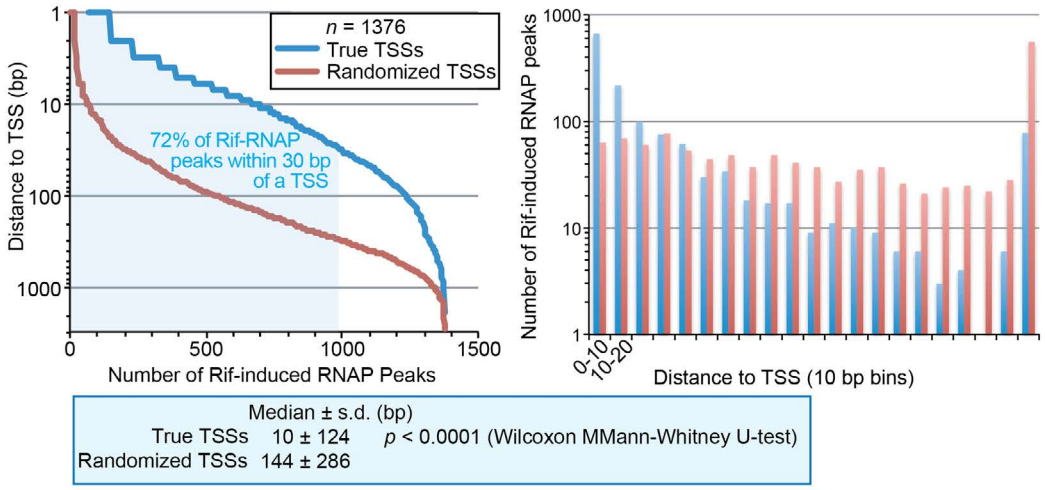


**Figure 1 - figure supplement 2. RNAP relocates to known TSS upon Rif-treatment**

Comparison of 1376 Rif-induced RNAP peaks to 11944 TSSs reported by Thomason et al., 2015 for M9Glc exp phase. Distances from Rif-induced RNAP peaks to the nearest TSS were compared to the same calculation for randomized TSSs (E. coli genome rotated 90°).

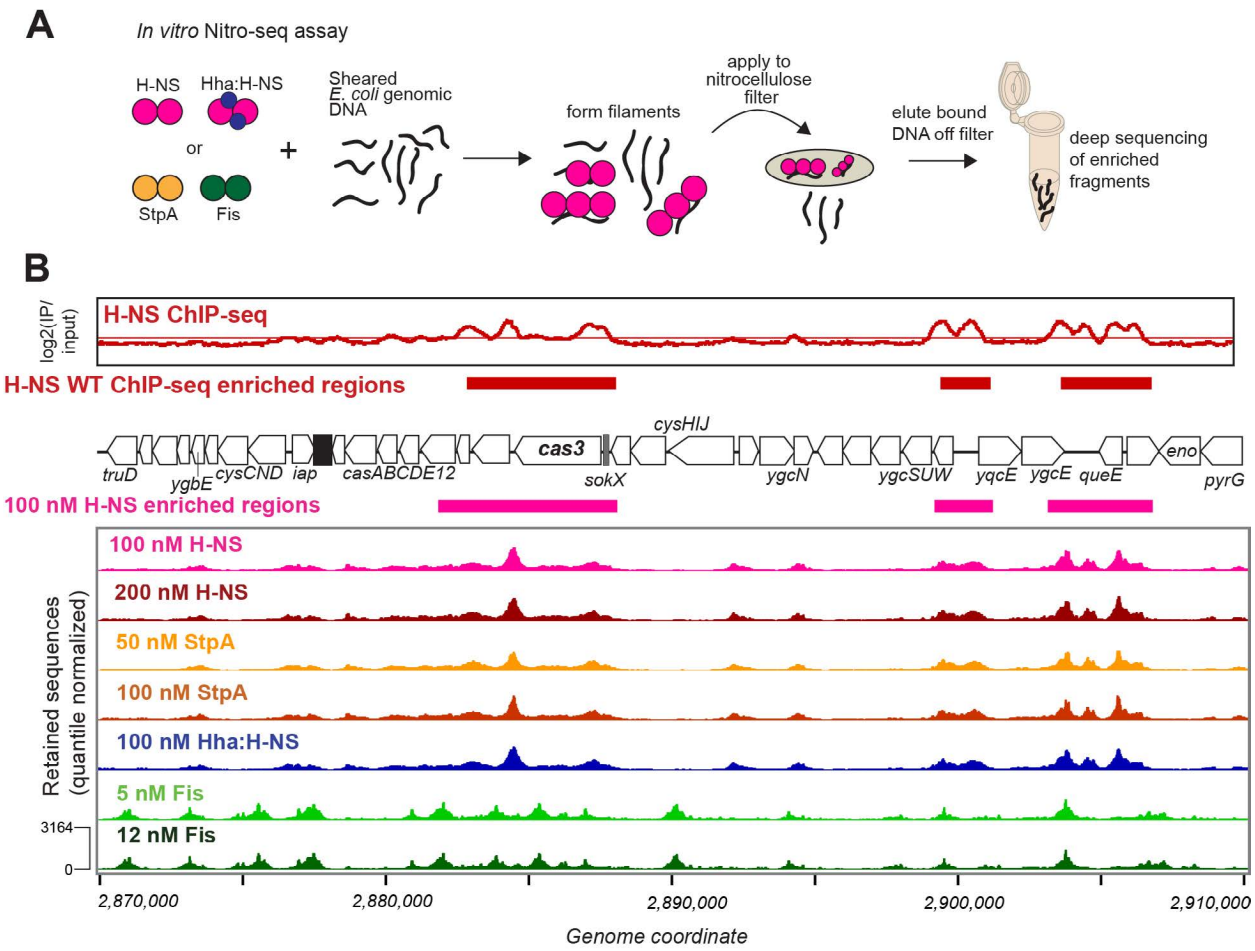


**Figure 2. H-NS filaments can be reconstituted *in vitro***

(A) Schematic of the Nitro-seq assay to determine sequence binding preference for H-NS, StpA, and Hha:H-NS. Proteins were mixed with sheared genomic DNA (500 ng) under conditions known to select AT rich fragments (**Figure 2 – figure supplements 1 and 2**). Filaments formed at 20 °C were applied to the nitrocellulose filter. Retained fragments were eluted and cleaned up before library prep and deep sequencing (see Methods). (B) Coverage of a 40 kb region of genome by H-NS (pink and red), StpA (yellow and orange), Hha:H-NS (blue) and Fis (light and dark green). Reads were quantile normalized to aid comparison between samples. Traces are representative coverage of three replicates. Normalized H-NS ChIP-seq coverage (**Figure 1**) shown for comparison at the top (bright red). Relevant genes in this region are indicated by white boxes. CRISPR repeats are indicated by a black rectangle (Touchon and Rocha, 2010). H-NS enriched regions (*i.e.,* filaments) identified in Nitro-seq (100 nM H-NS) and ChIP-seq (WT cells)

as determined from MOSAiCS are indicated by filled rectangles. Nitro-seq enriched regions were present in at least 2 of the three replicates and ChIP-seq enriched regions were present in both replicates.

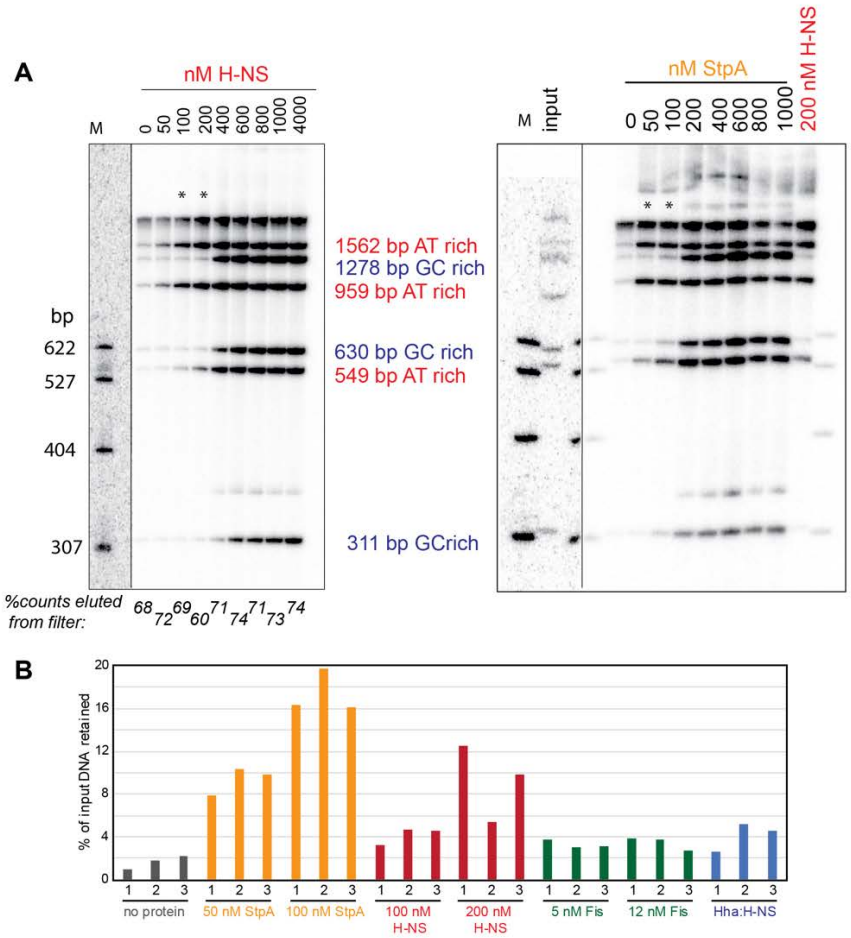


**Figure 2 – figure supplement 1. Determining selectivity of H-NS and StpA in nitrocellulose binding assay**

(A) Labeled fragments were mixed with sheared gDNA and either H-NS or StpA at indicated concentrations and incubated at 20 ˚C for 20 min. Complexes were applied to nitrocellulose filters, washed, eluted, ethanol precipitated and then separated by 6% Urea-PAGE. Fragment sizes are indicated. Colors indicate category of fragment (Supplemental Table S2). We expected that H-NS and StpA would bind to AT-rich, but not GC-rich fragments (Bouffartigues et al., 2007; Dole et al., 2004). Asterisks indicate concentrations selected for the final experiment. M indicates pBR322 MspI digest. The far-left side of the gel (with marker & input) had low signal compared to samples eluted from nitrocellulose filters, so the signal intensity was increased for clarity. Experiment was completed one time. Counts before and after elution were measured by liquid scintillation counting. (B) Fraction of sheared gDNA retained in each replicate of the nitrocellulose binding experiment.

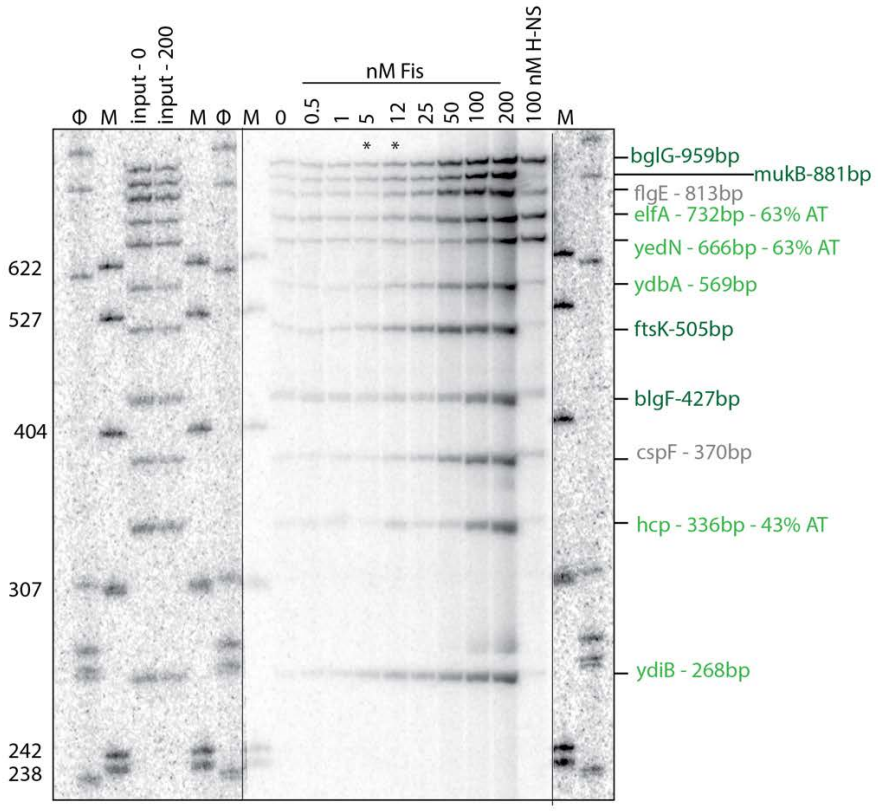


**Figure 2 – figure supplement 2. Determining selectivity of Fis in nitrocellulose binding assay**

To optimize conditions for Fis binding in the Nitro-seq assay, labeled PCR fragments were mixed with sheared, deproteinated gDNA and Fis at indicated concentrations and incubated at 20 ˚C for 20 min. Complexes were applied to nitrocellulose filters, washed, eluted, ethanol precipitated and then separated by 6% Urea-PAGE. Fragment sizes are indicated. Colors indicate category of Fis predicted binding (Supplemental Table S3): No binding (gray), some Fis binding precited (light green), and more Fis binding precited (dark green). As expected, the *flgE* and *cspF* fragments, which are not predicted to contain Fis sites, were not retained on the filter whereas other fragments like *ftsK*, which contains predicted Fis sites, were retained. This result suggests that these conditions capture specific Fis binding events. 5 and 12 nM Fis (asterisks) were selected for the final experiment. M indicates pBR322 MspI digest; Φ indicates HaeIII digest of ΦX174 DNA. The signal intensity of the far left and right sides of the gel had increased compared to samples eluted from nitrocellulose filters for clarity. Experiment was performed one time.

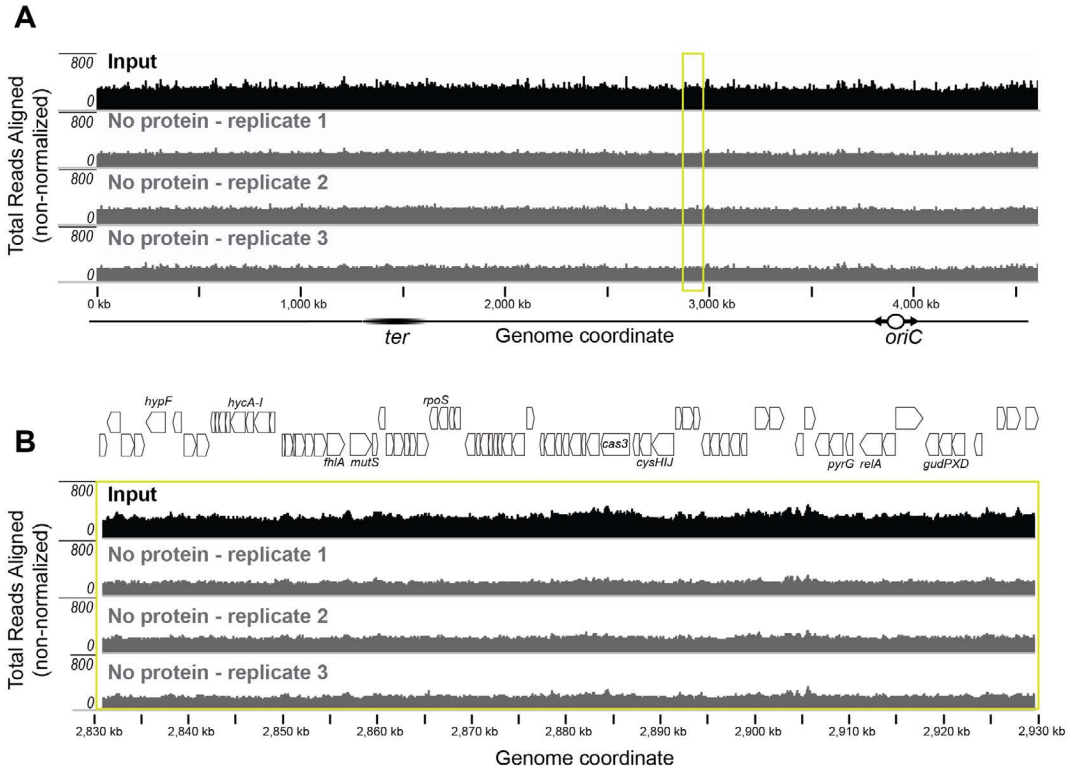


**Figure 2 – figure supplement 3. Nitro-seq gDNA was a diverse pool of sequences and free of protein contamination**

(A) View of reads aligned (reads were not normalized) across the genome in the input (black) and no protein samples (replicates 1-3; gray) generated in Integrated Genome Viewer from begraph files. (B) Zoom in of coverage in the yellow box in (A) showing even coverage in each sample. Genomic features shown in white boxes with limited gene names indicated. No protein samples were gDNA diluted in buffer and applied to the nitrocellulose filter. If protein contaminates were in the sheared gDNA, peaks of enrichment would have been present in the no protein samples.

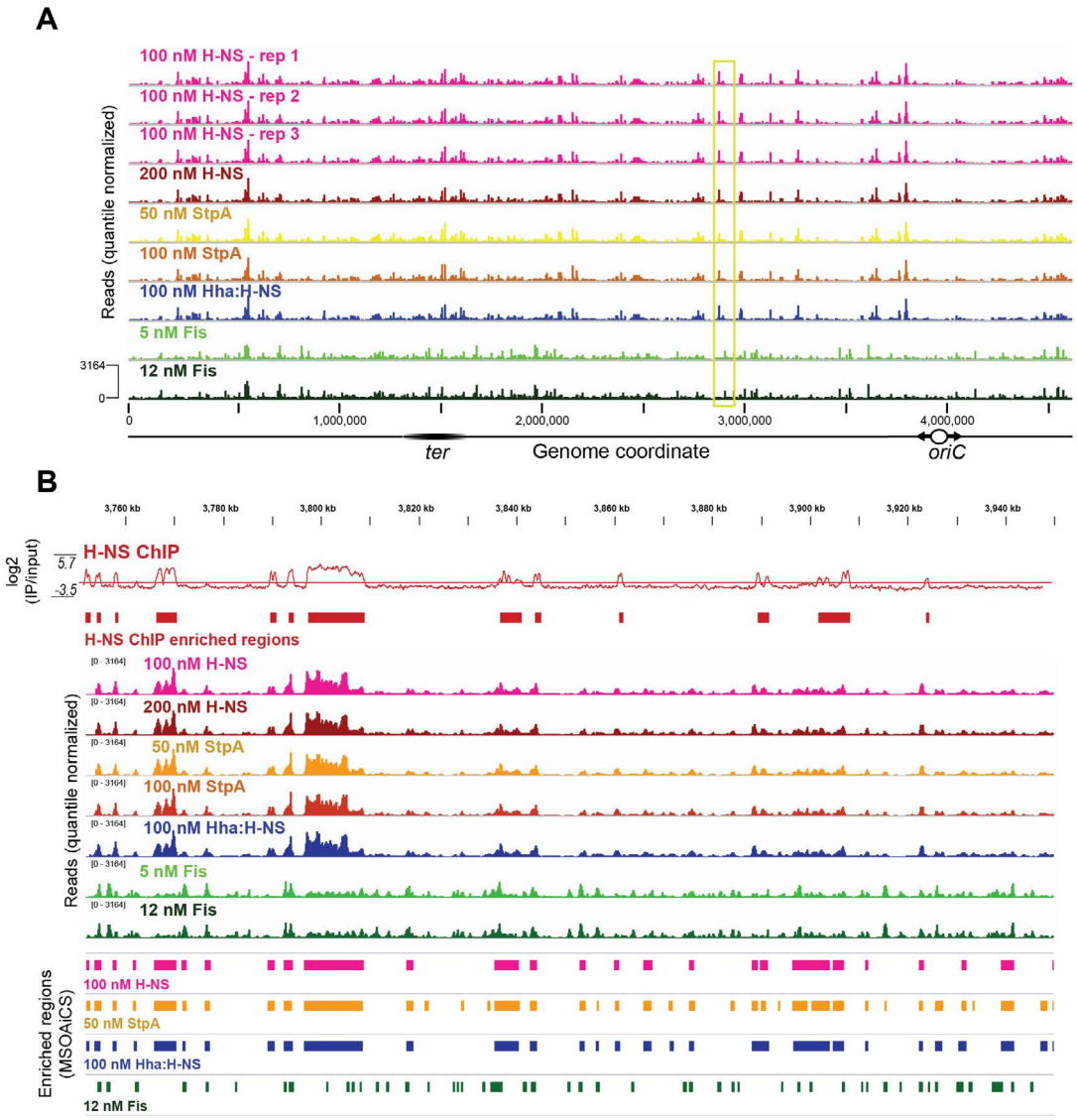


**Figure 2 - figure supplement 4. Binding of H-NS, StpA, Hha:H-NS, and Fis on genomic DNA *in vitro*.**

(A) Genome view of Nitro-seq results. Reads were quantile normalized to account for sequencing depth differences. Yellow box indicates region highlighted in Figure 2B. (B) View of Nitro-seq and H-NS ChIP-seq results at a 197 kb region of the genome. Traces are representative of triplicate samples. Reads for Nitro-seq were quantile normalized to each other, but not to ChIP-seq coverage. MOSAiCS enriched regions found in at least 2 of 3 replicates for each condition are shown by colored boxes below coverage plots.

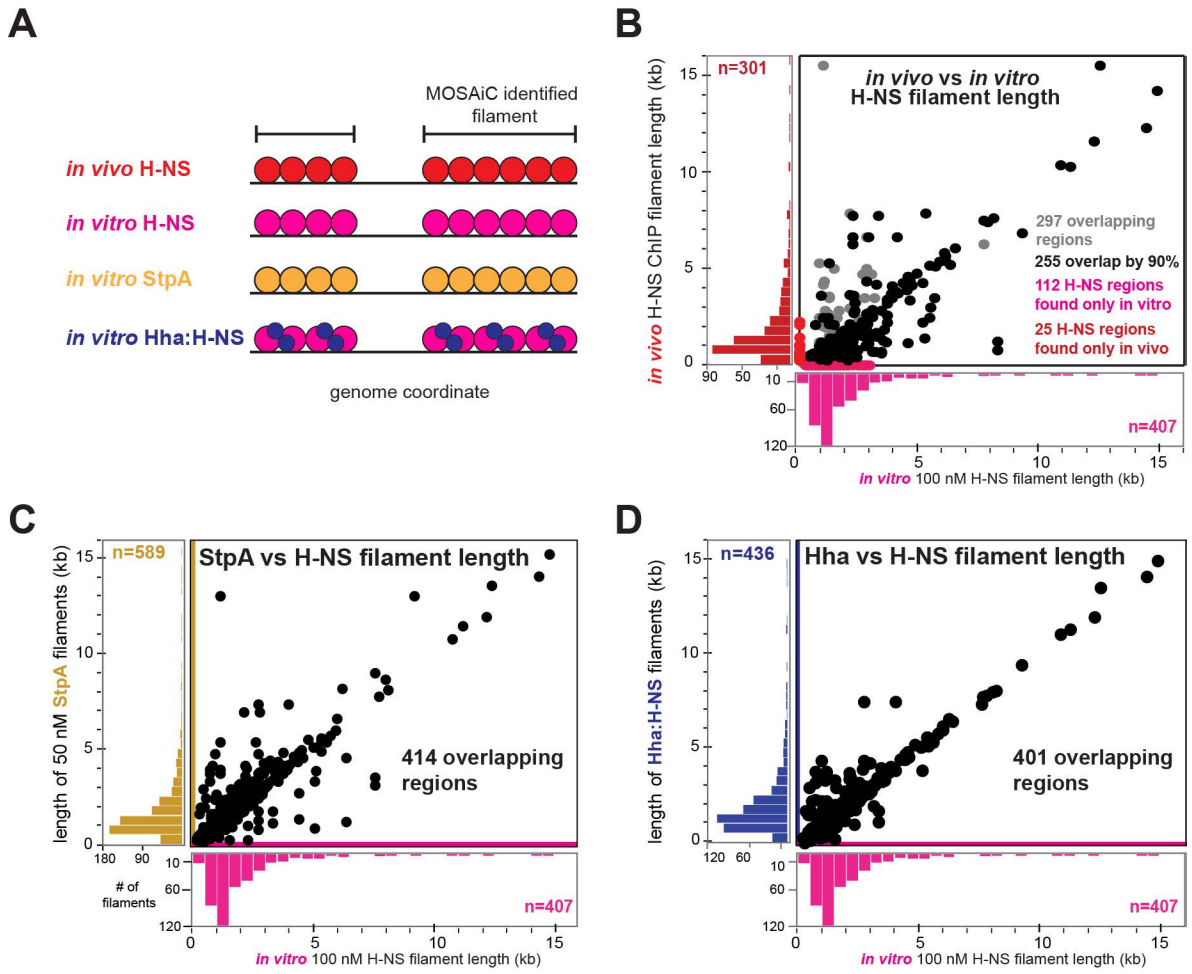


**Figure 3. Nitro-seq revealed similar sequence specificity between H-NS *in vivo*, H-NS *in vitro*, StpA, and Hha:H-NS.**

(A) Schematic showing the comparisons made between length of filaments of composed of different proteins at each genome coordinate. (B-D) Histograms showing lengths of different filaments as determined by MOSAiCS are shown on each axis. Length of *in vitro* 100 nM H-NS filaments is on each x-axis and the y-axis shows (B) *in vivo* H-NS, (C) 50 nM StpA, and (D) 100 nM Hha:H-NS filaments. Number of filaments (n) indicated for each sample. Scatter plots show number of *in vitro* H-NS filaments overlapping by at least 1 bp with each data set. (B) 297 filaments overlapped between *in vitro* and *in vivo* H-NS, but only 255 overlapped by 90% (*i.e.*, the percent of the length of the Nitro-seq filament that overlaps with the ChIP-seq filament). Filaments only found *in vitro* (112, pink) or *in vivo* (25, red) are shown. (B) Comparison of

lengths of overlapping filaments between 100 nM H-NS and 50 nM StpA. (C) Comparison of lengths of overlapping filaments between 100 nM H-NS and 100 nM Hha:H-NS.

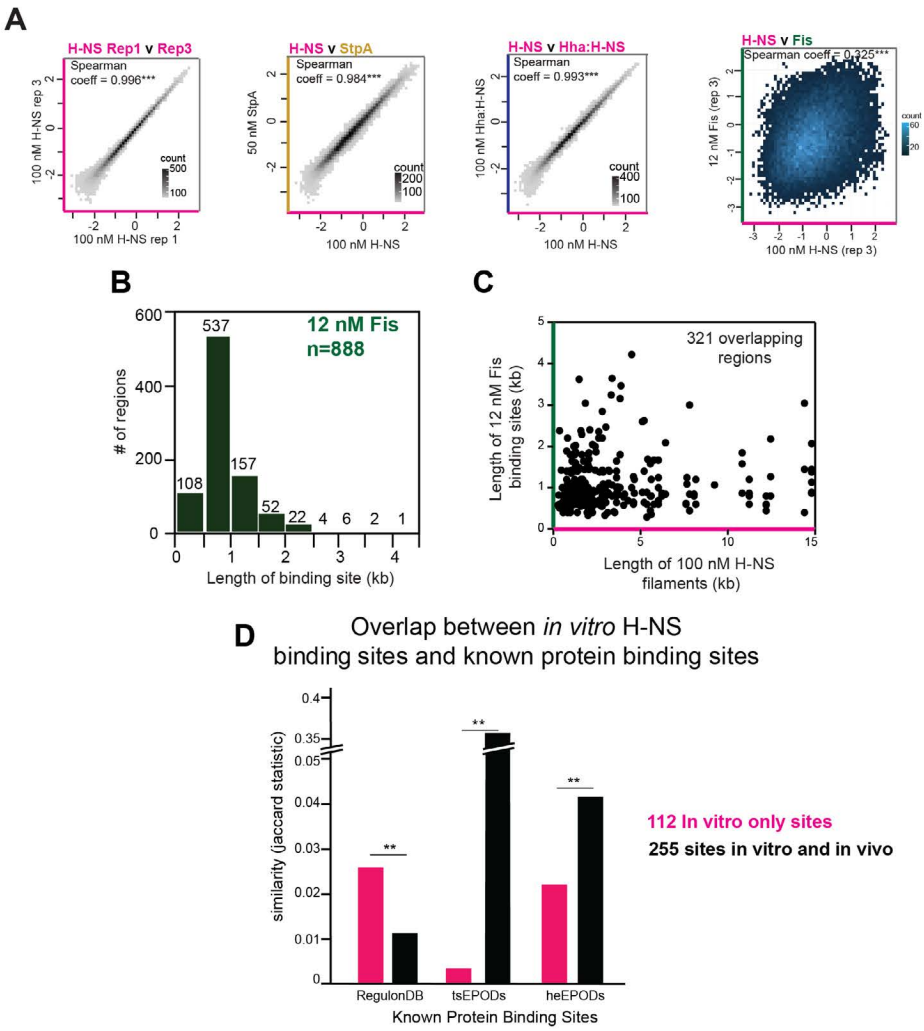


**Figure 3 - figure supplement 1. DNA binding pattern of H-NS in reconstituted filaments is distinct from Fis, but similar to H-NS *in vivo*, StpA, and Hha:H-NS.**

(A) Binned scatter plot comparing average log2(Nitro/input) in 200 bp bins across the genome between 100 nM H-NS replicate 1 and replicate 2 (pink axis), 100 nM H-NS and 50 nM StpA (gold axis), 100 nM H-NS and 100 nM Hha:H-NS (blue axis), or 100 nM H-NS to 12 nM Fis (green axis). *** indicates p-value $<2.2\times10^{16}$ for the Spearman coefficient. Unless specified, comparison of these replicates is representative of the all replicates in each sample. (B) Distribution of length of binding regions present in two of three 12 nM Fis replicates. Number of regions in each bin is labeled. (C) Comparison of lengths of 100 nM H-NS filaments (x-axis) overlapping with 12 nM Fis binding sites (y-axis). Only 321 regions are being compared and shows evidence of multiple Fis binding sites overlapping with a single H-NS filament. (D)

Similarity between Nitro-seq-specific (112 *in vitro* only) and overlapping (255 both *in vitro* and *in vivo*) H-NS binding regions and either RegulonDB TF binding sites(Santos-Zavaleta et al., 2019), transcriptionally silent extended protein occupancy domains (tsEPODS), or highly expressed EPODS (heEPODS) (Vora et al., 2009) as calculated by the jaccard statistic (Favorov et al., 2012). ** p-value <0.001.

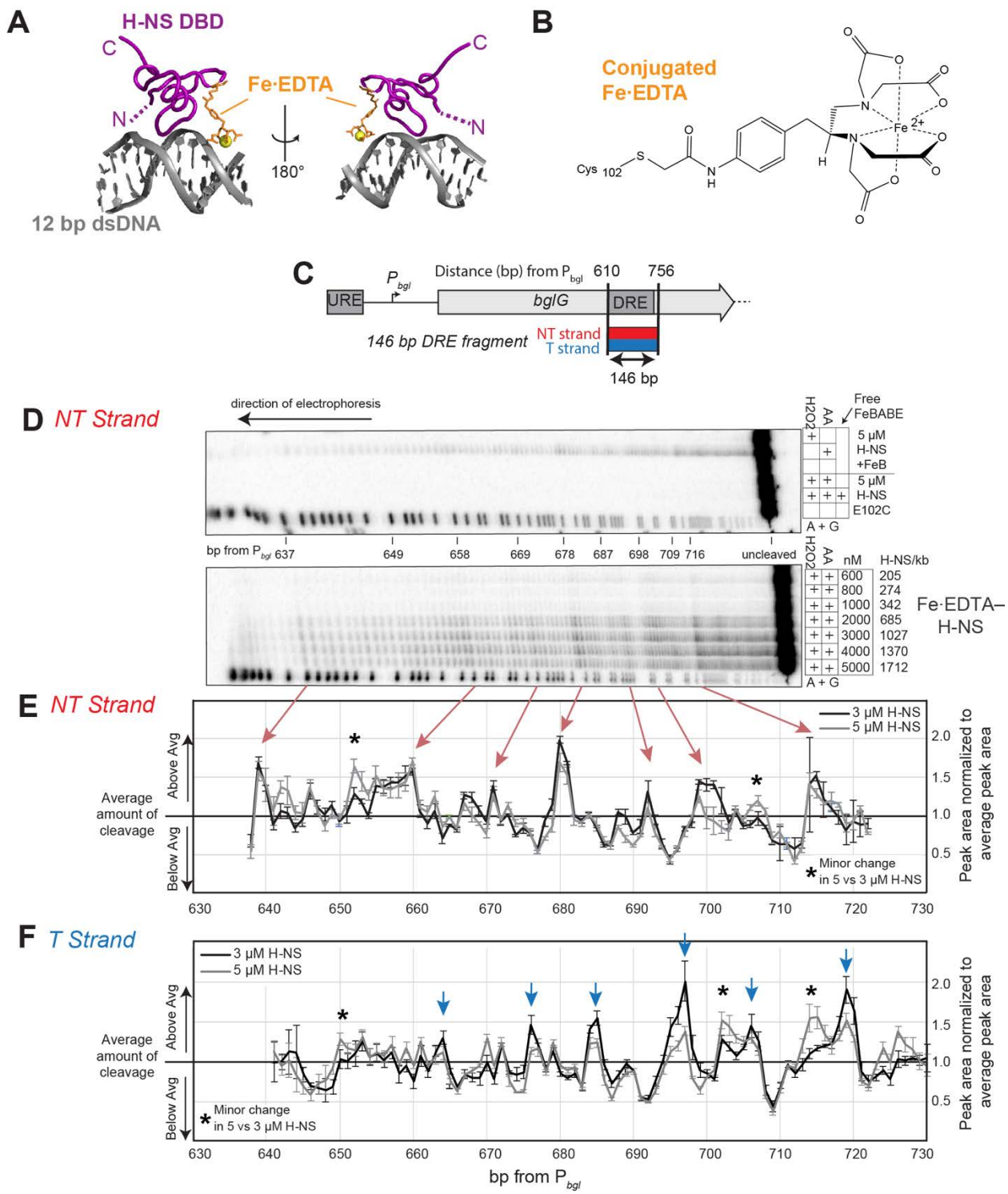


**Figure 4. Tethered cleavage of the DRE DNA from the *bgl* operon by Fe·EDTA–H-NS**

(A) Model of H-NS DNA binding domain (purple; PDB 2L93) bound to 12 bp duplex DNA (gray; Riccardi et al., 2019) with Fe·EDTA (orange) added at position 102. The linker and N-

terminal oligomerization domain would be connected the dashed line (N). (B) Chemical structure of Fe·EDTA conjugated at Cys102. (C) 146-bp DRE DNA fragment used in tethered cleavage assay is from the region of the *bgl* operon 610 to 756 bp downstream of the promoter ($P_{bgl}$), which encodes a known high-affinity binding region (DRE; downstream regulatory element; Dole et al., 2004). The DNA fragment was generated by PCR using 5′ end-labeled primers to label either the non-template (NT) strand (red) or the template (T) strand (blue). (D) 7% PAG showing Fe·EDTA–H-NS cleavage pattern of 146-bp DRE DNA labeled on the NT strand. Increasing concentrations of Fe·EDTA–H-NS were added to ~20 nM DNA alongside H-NS C21S E102 that does not contain Fe·EDTA. Cleavage was initiated with 2 mM $H_2O_2$ and 2 mM ascorbic acid. Controls with only 2 mM $H_2O_2$ or 2 mM ascorbic acid or water in the presence of high concentration of Fe·EDTA–H-NS. Addition of free FeBABE did not result in cleavage of the DNA. A+G indicates Maxam Gilbert ladder. (E & F) Quantification of peak areas using SAFA analysis at 3 and 5 μM Fe·EDTA–H-NS on the NT (E) or T (F) strands. Peak areas normalized to amount loaded in each lane and then the average peak area in each lane. Peaks areas above 1 indicate cleavage hot spots (arrows) and peak areas less than 1 indicate areas of little cleavage. Asterisks indicate areas that showed a slight increase in cleavage in 5 μM Fe·EDTA–H-NS compared to 5 μM Fe·EDTA–H-NS. Error bars show standard deviation of three replicates.

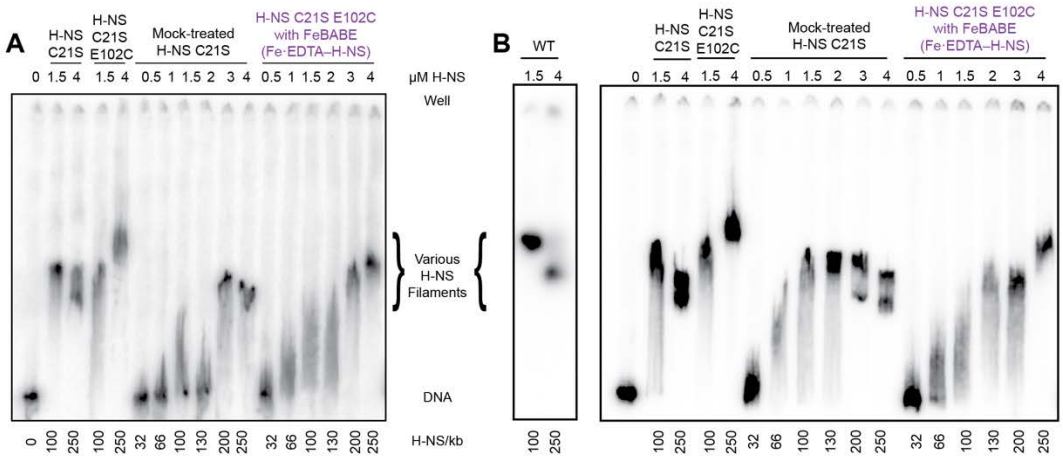


**Figure 4 - figure supplement 1. Fe·EDTA–H-NS can form filaments on DNA similar to non-conjugated protein.**

Electrophoretic mobility shift assay showing filament formation on 10 nM 1.5 kb DNA from the *bgl* operon (Boudreau et al., 2018; Kotlajich et al., 2015) by H-NS C21S or H-NS C21S E102C with or without FeBABE treatment. H-NS C21S was mock treated with FeBABE to account for loss of activity during conjugation process, whereas Fe·EDTA–H-NS has ~50% FeBABE conjugated to the free thiol in H-NS C21S E102C. Panels A and B represent filament formation by two independent preparations of H-NS-FeBABE. Filaments formed are similar between preps, but the activity of H-NS varied between preparations. Panel (A) shows filament formation with the first conjugation of H-NS and panel (B) shows the protein from a second independent conjugation. All filaments were run on the same gel in (B), but rearranged for clarity.

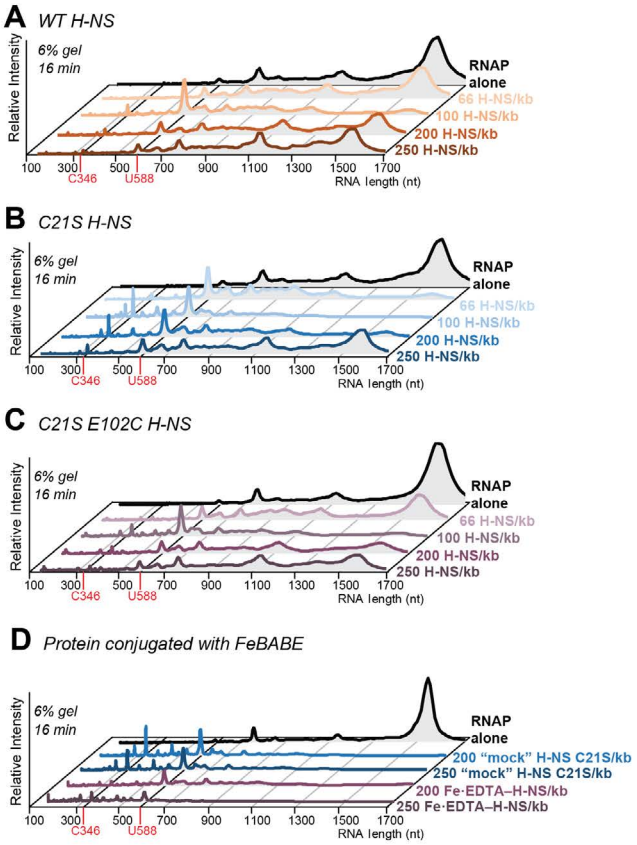


**Figure 4 - figure supplement 2. Effect of H-NS mutations and FeBABE conjugated protein on elongating RNAP.**

Pseudo-densitometry traces showing transcripts from *bgl* DNA present after 16 min of transcription with 30 µM NTPs through various H-NS filaments: (A) WT H-NS filaments, (B) C21S H-NS filaments, (C) C21S E102C H-NS filaments, (D) either C21S (blue) or C21S E102C (purple; Fe·EDTA–H-NS) H-NS that were treated with FeBABE. Concentrations of protein that formed filaments were used for transcription assays (see Figure 4 – figure supplement 1). Transcription was done at least one time for conditions shown.

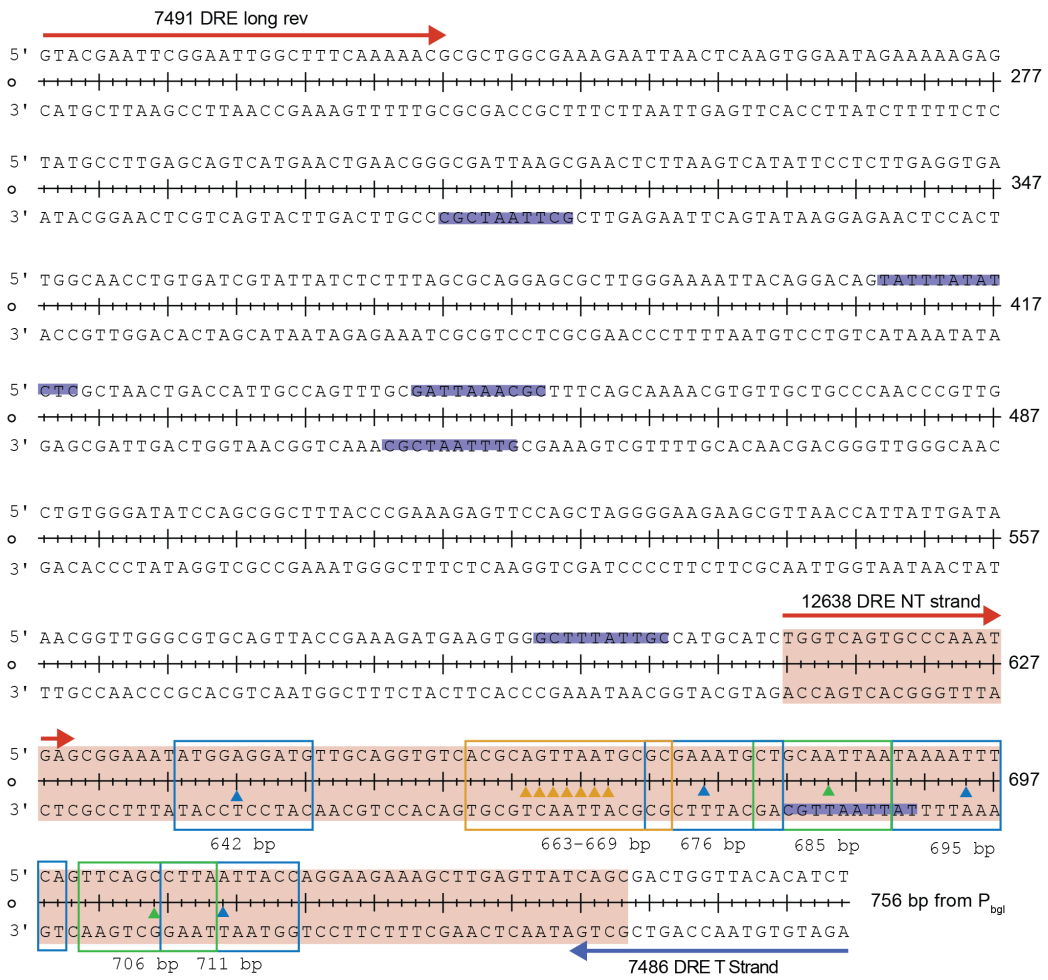


**Figure 4 - figure supplement 3. Sequence of DRE DNA fragments.**

146 bp DNA from the *bgl* operon was amplified from pMK110 using primers 7486 and12638. Longer fragments (549 bp) were amplified with primer 7486 and primer 7491. The location of the DRE as identified by (Dole et al., 2004) highlighted in light pink. High-affinity H-NS binding sites identified by Virtual Footprint (http://www.prodoric.de/vfp/vfp_regulon.php) are indicated by blue filled boxes. Distance of sequence from $P_{bgl}$ are shown on right. Light blue, green, and gold open boxes highlight locations of H-NS DNA binding domains in Figure 6D. Triangles indicate position where DBD was aligned on the DNA.

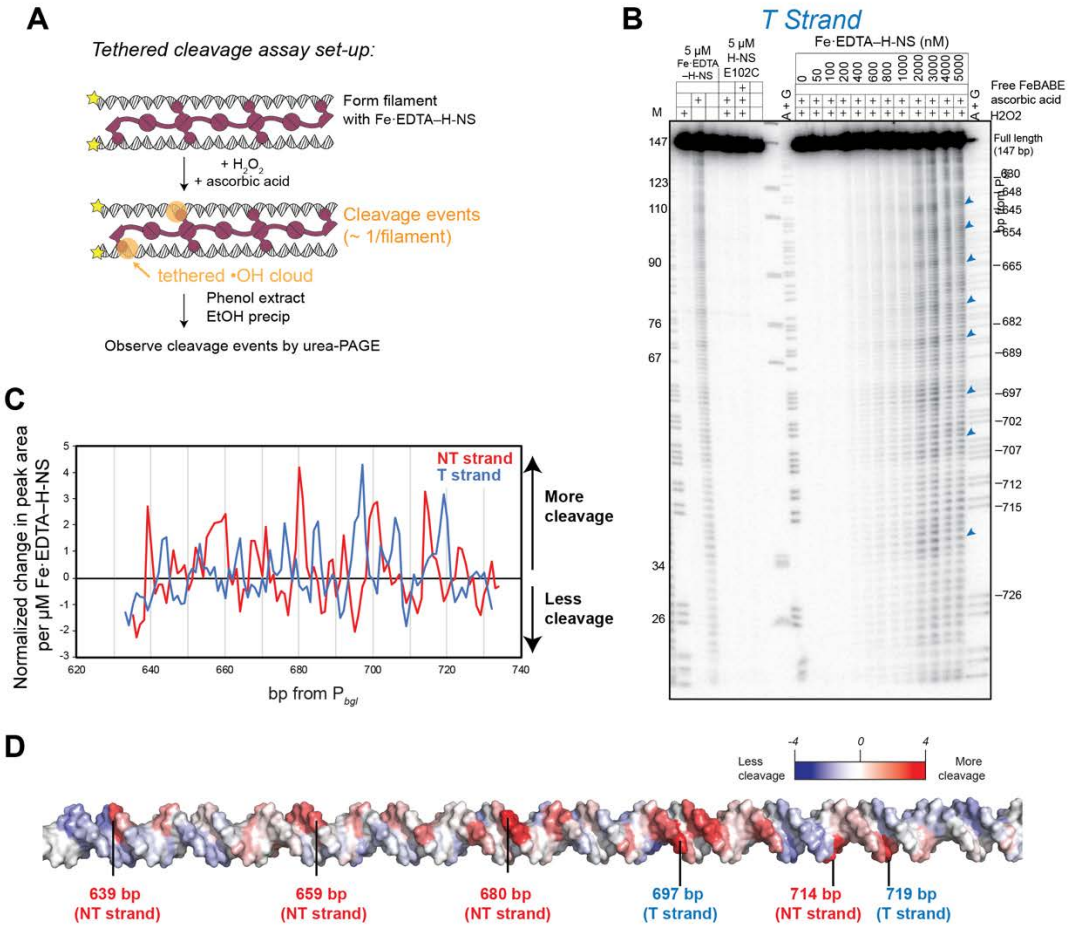


**Figure 4 - supplemental figure 4. Fe·EDTA–H-NS cleavage of short DRE template.**

(A) Schematic of tethered cleavage assay. Filaments are formed with Fe·EDTA–H-NS on single-end labeled DNA by incubating protein and DNA for 20 min at 20 °C. To initiate cleavage, $H_2O_2$ and ascorbic acid are added in rapid succession at concentrations resulting in ~1 cleavage event per filament. Cleavage events are limited to ~12 Å around the tether based on the diffusion of ·OH radicals (orange circle indicating ·OH cloud). Cleavage products are cleaned and concentrated with phenol extraction and ethanol precipitation, respectively, before resolving by urea-PAGE. (B) 7% PAG showing Fe·EDTA–H-NS cleavage pattern of 146 bp DRE DNA labeled on the T strand. Increasing concentrations of Fe·EDTA–H-NS were added to ~20 nM DNA alongside H-NS C21S E102 that does not contain Fe·EDTA. Cleavage was initiated with 2 mM $H_2O_2$ and 2 mM ascorbic acid. Controls with only 2 mM $H_2O_2$ or 2 mM ascorbic acid or water in the presence of high [Fe·EDTA–H-NS]. Addition of free FeBABE did not result in

cleavage of the DNA. A+G indicates Maxam Gilbert ladder. Blue arrows indicate hot spots of cleavage induced by Fe·EDTA–H-NS. Gel is quantified in Figure 4F. (C) The relationship between H-NS concentration and changes in mean-centered normalized peak area at each bp was determined using a Bayesian multilevel model. For each strand, a model was fit using data from experiments performed with H-NS concentrations from 0 to 3 μM and slope point-estimates were Robust-Z normalized within each strand for direct comparison between the strands (see Methods). More positive slopes indicate more H-NS dependent cleavage at that bp and likely coincide with robust H-NS binding. (D) Data in (C) displayed on B-form DNA where red indicates the most cut bases and blue are the least cut bases.

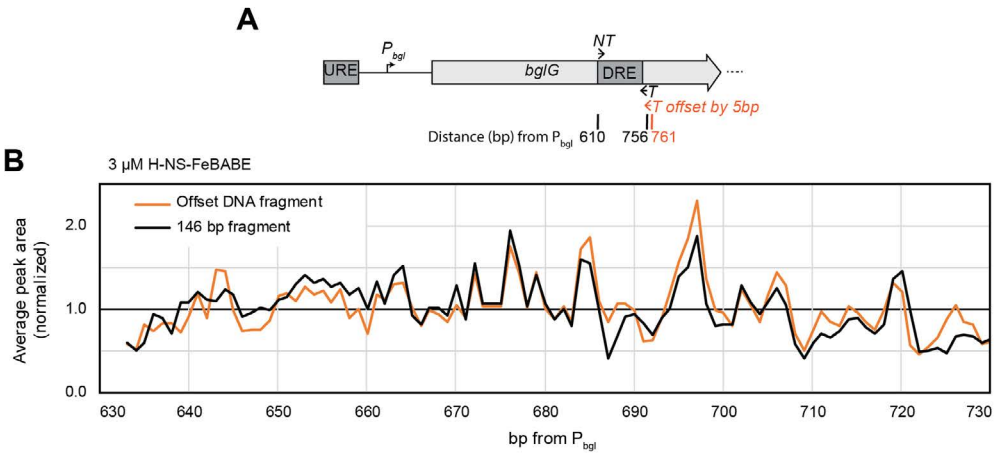


**Figure 4 - figure supplement 5. Fe·EDTA–H-NS cleavage of offset T strand.**

(A) Schematic of the DNA templates used to test if H-NS binding is sequence specific or initiates from the end of the DNA. The 146 bp DNA fragment (black arrow) was used in experiments shown in Figure 4. Labeling 5 bp upstream (orange arrow) would show a binding pattern offset by 5 bp compared to 146 bp DNA if H-NS bound from the end of the DNA and not at specific sequences. (B) Peak areas showing Fe·EDTA–H-NS (3 μM) cleavage pattern of 146 bp DRE DNA labeled on the T strand at either 146 bp DNA (black) or 151 bp offset DNA fragment (orange). Cleavage and quantification were done as described in Methods. Peak areas are an average of two experiments.

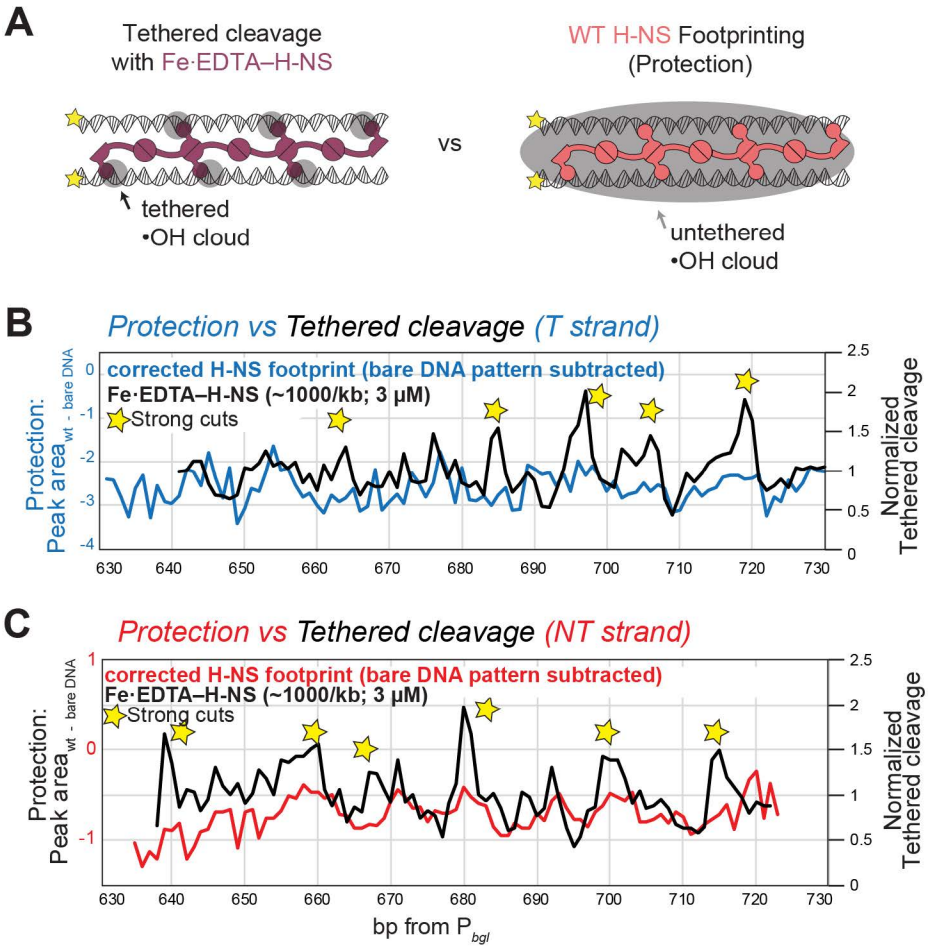


**Figure 5. Hydroxyl radical footprints of H-NS shows unique pattern of cleavage induced by Fe·EDTA–H-NS.**

(A) Schematic showing the difference between the tethered cleavage assay (Figure 4) and traditional ·OH footprinting. Tethered cleavage assay uses Fe·EDTA–H-NS (purple) to create a cloud of free-radicals that can only diffuse ~12Å (gray circle) from the tethered $Fe^{2+}$ (left). In traditional footprinting, excess $Fe^{2+}$·EDTA is added in solution creating a pool of free radicals that diffuse freely in solution. Any DNA not bound by protein (wt H-NS, pink) will be cleaved. (B) Blue line shows the corrected untethered ·OH footprints of H-NS on the T strand by subtracting the bare DNA cleavage pattern from the pattern of ·OH cleavage of the T strand in the presence of WT H-NS (1.5 μM). Valleys are protected regions. Black line shows cleavage by Fe·EDTA–H-NS (3 μM; 1027 H-NS/kb) demonstrating its distinct pattern, including cuts within H-NS footprints. The most notable cuts attributable to tethered Fe·EDTA are indicated by yellow

stars. (C) Same as (B) but for the NT strand. Lines represent an average of 3 experiments, but error bars are removed for clarity.

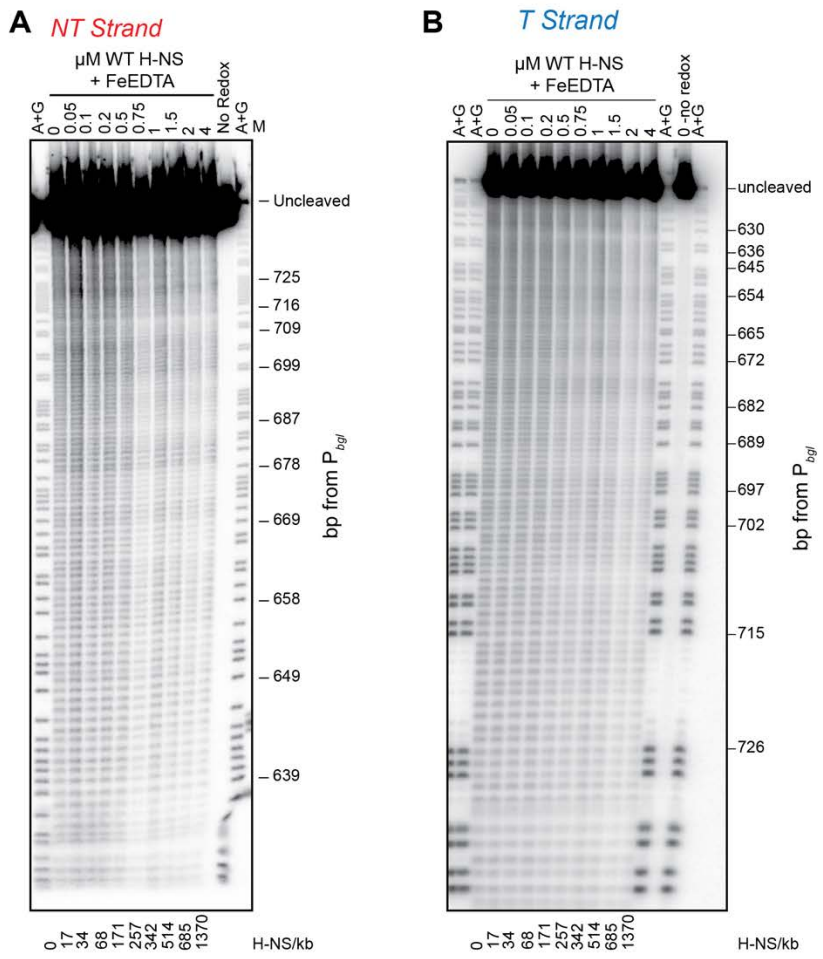


**Figure 5 - figure supplement 1. ·OH footprinting of H-NS filaments reveals phased binding of H-NS along the DRE.**

7% PAG showing hydroxyl radical cleavage pattern of 146 bp DRE DNA labeled on the non-template (A) or template (B) strand (Figure 4C) in the presence or absence of WT H-NS filaments formed at 0.05, 0.1, 0.2, 0.5, 0.75, 1, 1.5, 2, or 4 μM H-NS. The ratio of H-NS/kb is indicated below the gel. WT H-NS + untethered FeEDTA lanes were cleaved by sequential addition of $H_2O_2$, ascorbic acid, and FeEDTA. Gels are representative of at least 3 experiments.

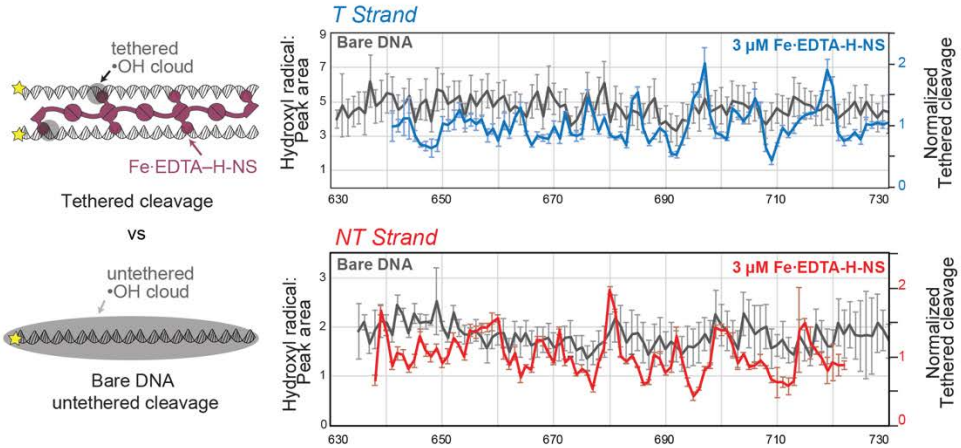


**Figure 5 – figure supplement 2. Fe·EDTA–H-NS tethered cleavage has different bias than intrinsic cleavage of bare DNA.**

To determine if Fe·EDTA–H-NS cleavage could reveal H-NS binding sites, the tethered cleavage pattern needed to be distinct from cleavage of bare DNA. So, the tethered cleavage pattern (top cartoon) was compared to the untethered cleavage pattern of bare DRE DNA (bottom cartoon). Transparent ovals represent the possible reach of ·OH cleavage in the assay. Peak area of ·OH cleavage of 146 bp DRE DNA T (top) or NT (bottom) strand in the absence of H-NS (gray) shows intrinsic cleavage propensity of the sequence used to form H-NS filaments. Intrinsic cleavage pattern (gray) is different from the tethered cleavage pattern introduced by 3 μM Fe·EDTA–H-NS on the template (blue) or non-template (red) strand, which indicates specific cleavage by Fe·EDTA–H-NS. Error bars indicate standard deviation of three replicates.

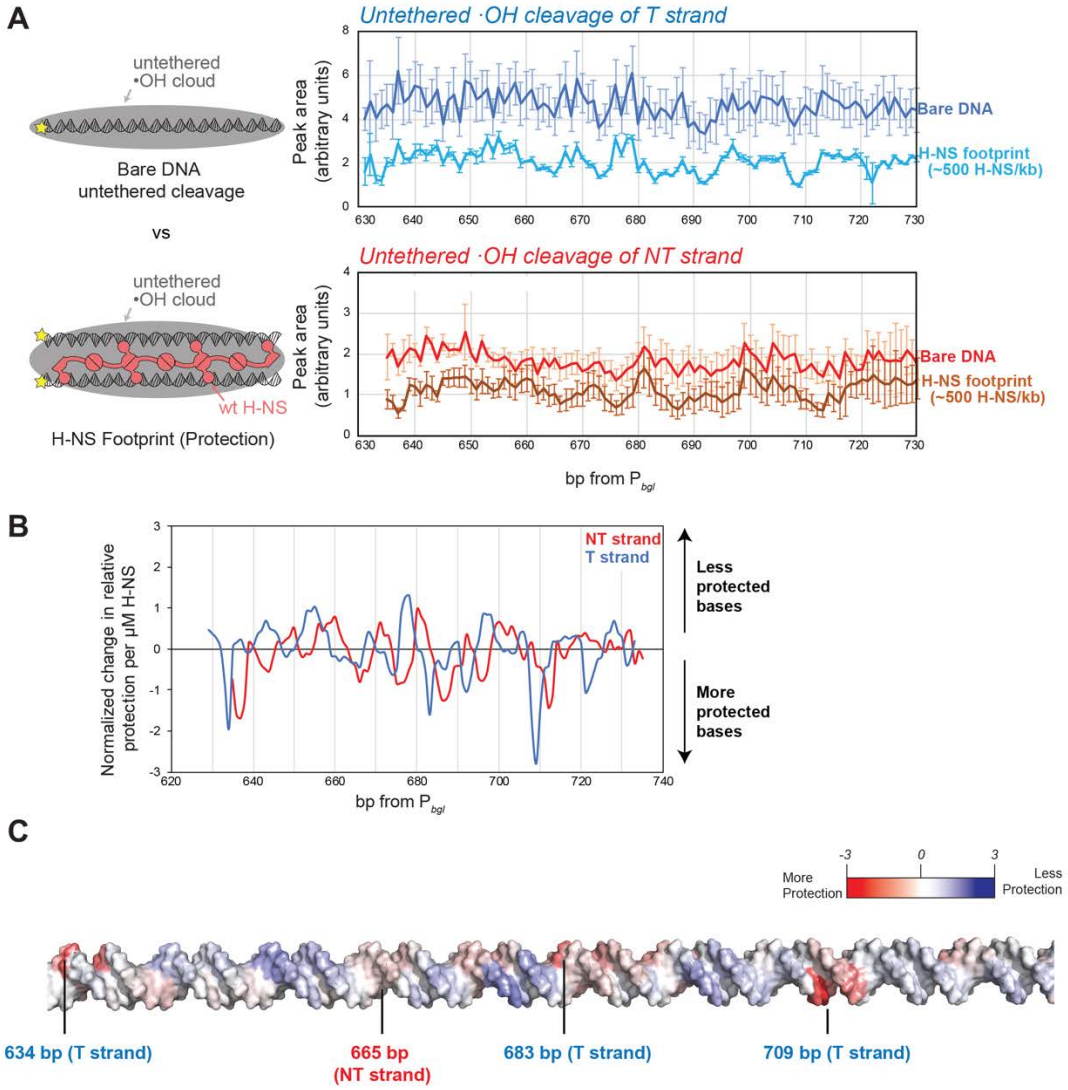


**Figure 5 – figure supplement 3. ·OH footprinting reveals regions protected by H-NS.**

(A) Comparison of the untethered cleavage of bare DNA versus an H-NS filament shows protection by H-NS. Cleavage of the T (top) or NT (bottom) strand by ·OH in the absence (bare DNA, blue) or presence of ~500 wild-type H-NS/kb (cyan; 1.5 µM H-NS). Peaks areas calculated from SAFA and reported as arbitrary units. Error bars represent standard deviation of at least 3 experiments. (B) Protected regions were identified by determining the relationship between H-NS concentration and the log2 transformed change from average protection at each bp using a Bayesian multilevel model. Slope point-estimates were Robust-Z normalized within each strand for direct comparison between the strands (see Methods for details). More negative slopes indicate more H-NS dependent protection at that bp and likely coincide with robust H-NS

binding. (C) Data in (B) displayed on B-form DNA where red indicates bases most protected by H-NS and blue indicates least protected bases. Select bases and strand highlighted.

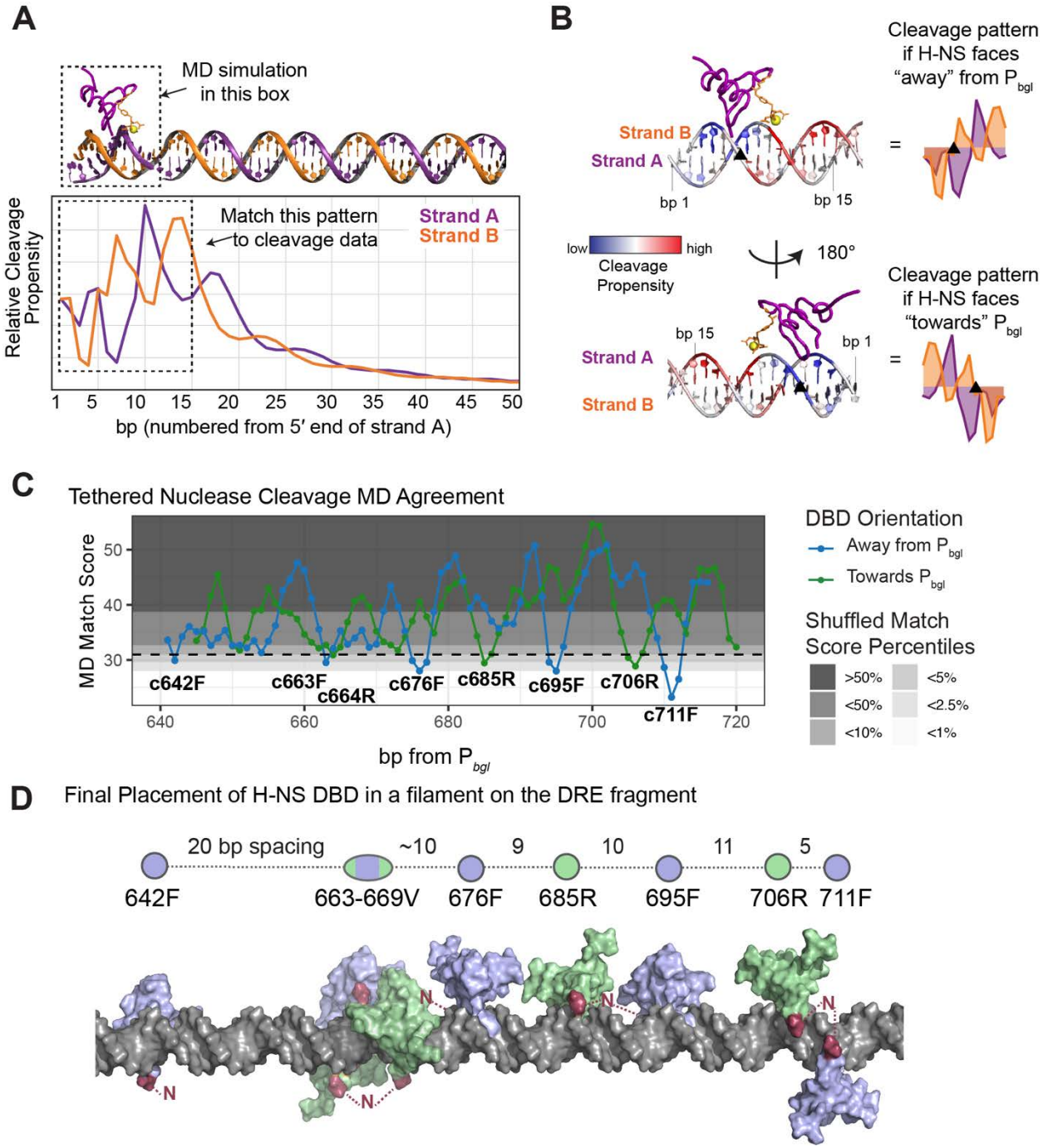


**Figure 6. MD simulations of conjugated Fe·EDTA conformational dynamics reveals arrangement of H-NS DBD**

(A) MD simulation of Fe·EDTA (orange) conjugated to the H-NS DBD (purple) bound to a 12 bp DNA duplex (strand A, purple; strand B, orange). The distance between the $Fe^{2+}$ and each base pair during the simulation was used to estimate the cleavage propensity at each bp.

Additional DNA (bp 13-50) was included after the MD simulations to estimate cleavage propensities at bp further from the binding site. (B) Overlay of the cleavage propensity on B-form DNA with H-NS bound (PDB 2L93; Riccardi et al., 2019). Cleavage propensity is on a blue-white-red spectrum where blue is low cleavage propensity and red is high cleavage propensity. Cleavage pattern of H-NS bound in the opposite orientation is also shown. (C) Concordance between expected and observed cleavage patterns as determined by the Manhattan distance (MD Match Score) between the pattern of predicted cleavage of the first 15 bp from the MD simulation and the tethered cleavage data of both strands (see Methods). Lower scores indicate greater concordance between the two patterns. The asymmetry of the MD simulation cleavage pattern enabled the determination of DBD orientation with respect to the $P_{bgl}$ promoter as described in B. Statistical significance was determined by recalculating the Manhattan distance across 1000 shuffled tethered cleavage datasets for the MD cleavage pattern in each orientation. Shaded regions indicate values below the indicated percentiles of MD match scores across all shuffled datasets. The dotted line indicates a MD match score at <5% of all match scores determined in the shuffled datasets. The location (bp) and orientation (F or R for forward or reverse) of 8 H-NS binding sites predicted by tethered nuclease cleavage are indicated for each DBD. (D) A final model placing 7 H-NS DBDs on the DRE DNA was generated from nuclease cleavage analysis and ·OH footprinting analysis (**Figure 6 – figure supplement 2**; see Methods). DBDs were either in the forward (slate), reverse (green), or variable orientation (pale orange; V). The spacing between each DBD is indicated on the 1D line graph above the 3D model. The NTD is not placed in our model but would be attached to the DBD at the raspberry residues (N).

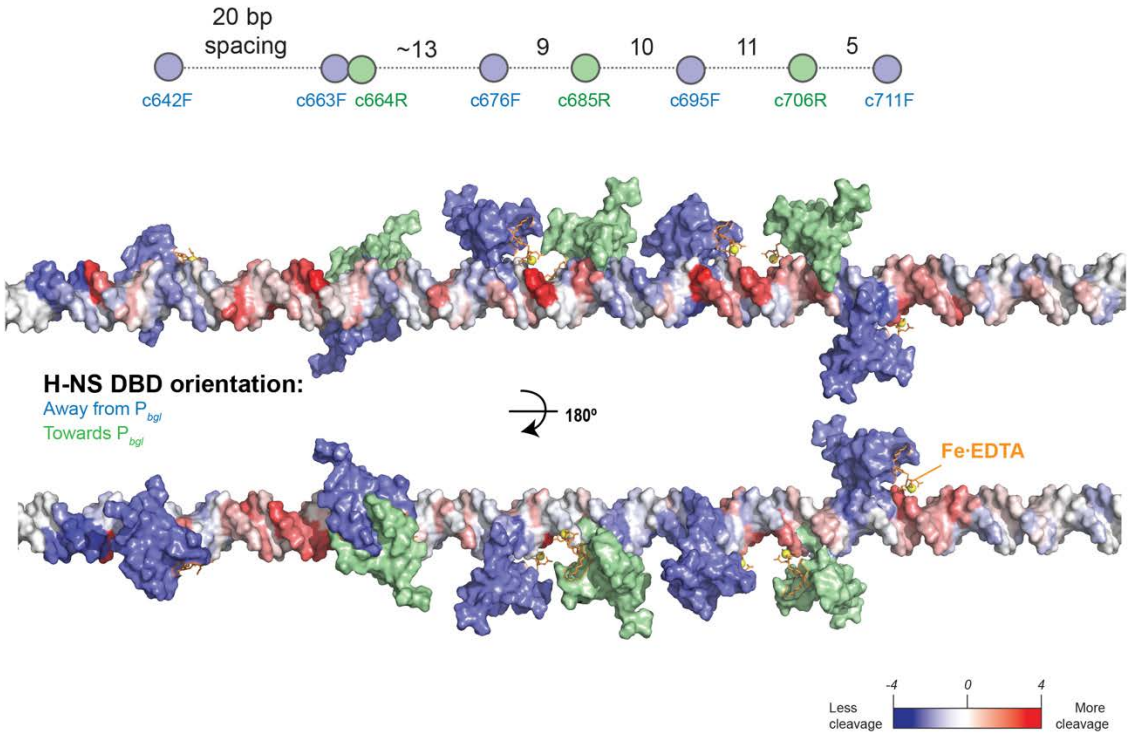


**Figure 6 – figure supplement 1. Modeling of H-NS DBD location based on match score between tethered cleavage pattern and MD simulation.**

The center of the H-NS DBD (purple) with the Fe·EDTA tether (orange sticks and yellow sphere, $Fe^{2+}$) was placed on B-form DNA at the position with the best match score in (Figure 6C). The prediction method (t for tethered cleavage), position on the DNA (bp from $P_{bgl}$), and the orientation (F or R) are indicated for each DBD. DNA is colored as in Figure 4 - figure supplement 4D.

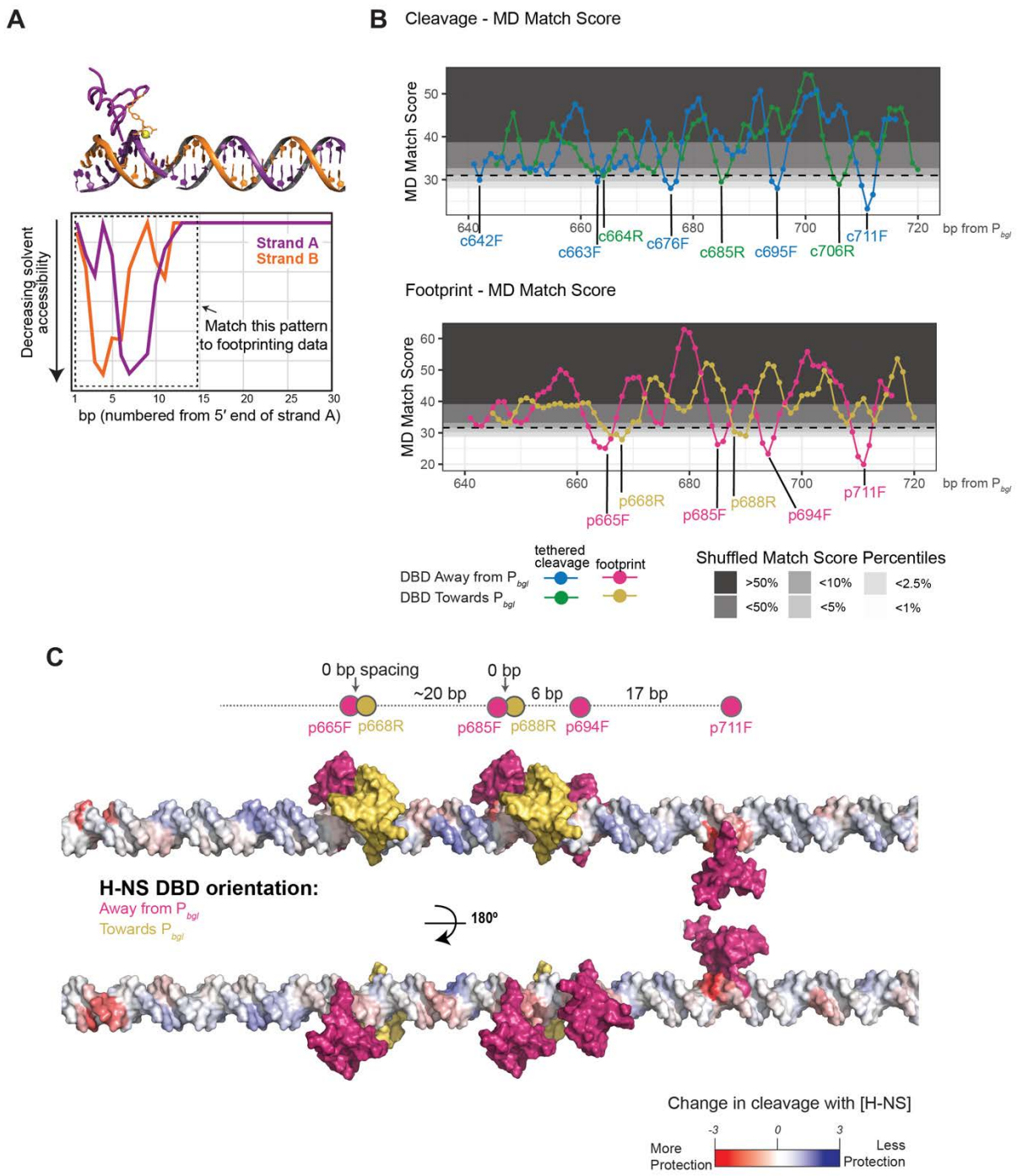


**Figure 6 – figure supplement 2. Modeling of H-NS DBD location based on ·OH footprinting and solvent accessibility patterns.**

(A) Solvent accessibility of DNA bound by one H-ND DBD calculated from structure shown at the top. 15 bp of this pattern was compared to the ·OH footprinting pattern (dashed box). (B) *Top:* repeat of data in Figure 6C so match scores can be compared. *Bottom:* Concordance

between expected and observed protection patterns as determined by the Manhattan distance (MD Match Score) between the pattern of predicted protection of the first 15 bp from the MD simulation (A) and the ·OH radical footprinting data of both strands (Figure 5; see Methods). Lower scores indicate greater concordance between the two patterns. The asymmetry of the MD simulation cleavage pattern enabled the determination of DBD orientation with respect to the $P_{bgl}$ promoter with the forward binding cleavage pattern indicating "Away from $P_{bgl}$" (blue trace) and the reverse binding cleavage indicating "Towards $P_{bgl}$" (green trace). Statistical significance was determined by recalculating the Manhattan distance across 1000 shuffled ·OH footprinting datasets for the MD protection pattern in each orientation. Shaded regions indicate values below the indicated percentiles of MD match scores across all shuffled datasets. The dotted line indicates a MD match score at <5% of all match scores determined in the shuffled datasets. (C) The center of the H-NS DBD (purple) was placed on B-form DNA at the position with the best match score in (B). The prediction method (p for footprint), position on the DNA (bp from $P_{bgl}$), and the orientation (F or R) are indicated for each DBD. Note that there is overlap between the two DBDs at 665 and 668 and the two DBDs at 685 and 688 suggesting that only one H-NS is bound in either of those 3 bp regions. DNA is colored as in Figure 5 - figure supplement 3.

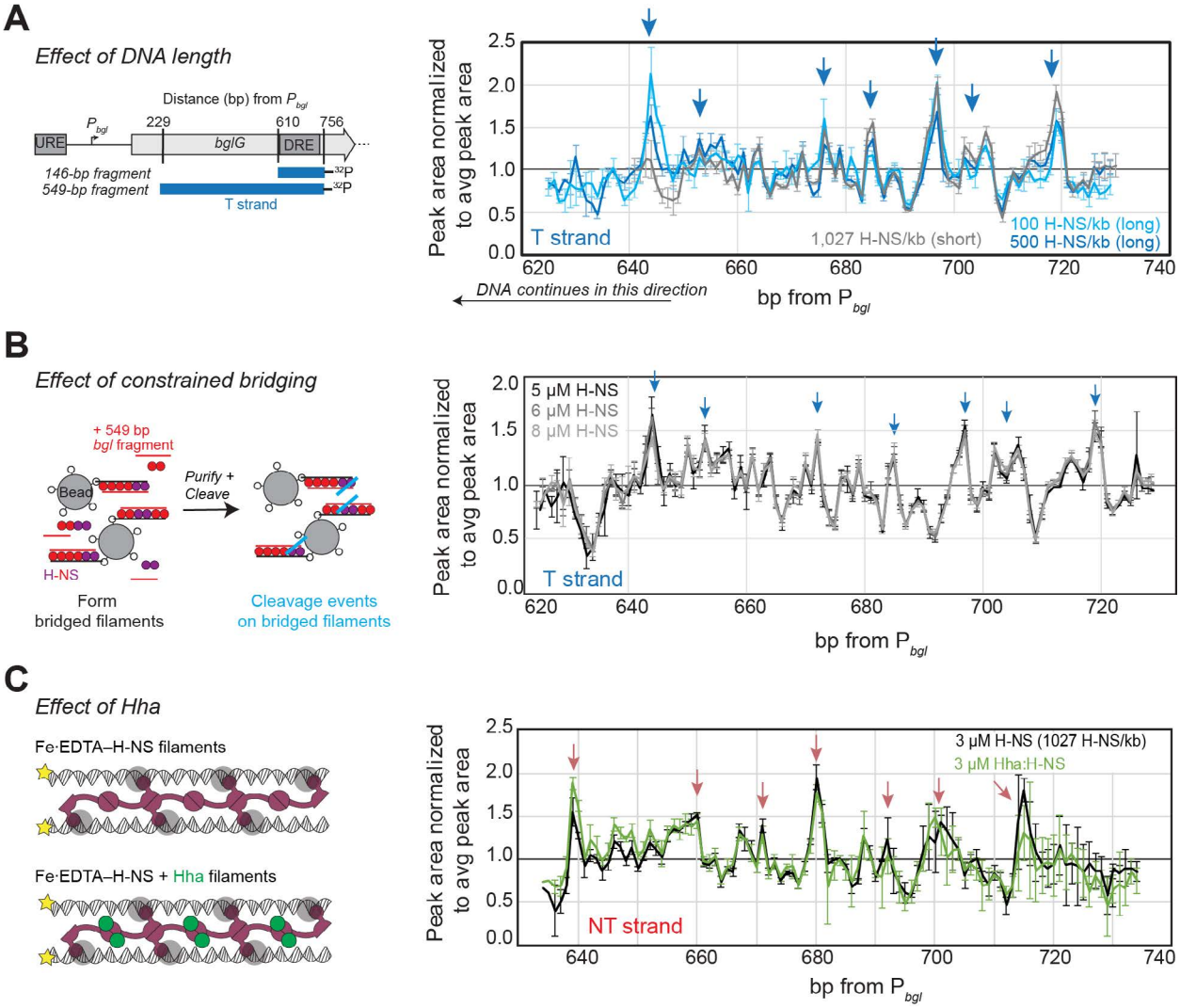


**Figure 7. H-NS binding pattern does not change dramatically under different filament forming conditions.**

(A) Cartoon of short (146 bp) and long (549 bp) DNA fragments generated from the *bgl* operon to test effects of DNA length on arrangement of H-NS DBD within a filament. Both fragments were labeled on the T strand. Fe·EDTA–H-NS cleavage pattern of the template strand was measured on both short (gray) or long (blue) DNA. Gray line matches data in **Figure 4F**. To form filaments on long DNA, H-NS was a mixture of 60% WT H-NS and 40% Fe·EDTA–H-NS since the starting concentration of Fe·EDTA–H-NS was too low to form pure Fe·EDTA–H-NS filaments. The direction of the remaining, unquantified, DNA sequence is indicated and shown in **Figure 7 – figure supplement 1**. (B) Bridged filaments were formed by mixing magnetic beads

with 685 bp bait DNA, 549 bp *bgl* DNA labeled on the T strand, and H-NS (60% WT H-NS, red + 40% Fe·EDTA–H-NS, purple) at desired concentration. Bridged filaments were purified using the magnetic beads and resuspended in cleavage buffer to initiate cleavage of DNA in bridged filaments only (**Figure 7 – figure supplement 2**). Cleavage pattern of a portion of the T strand in the presence of bridged filaments is shown. (C) Cartoon of either pure H-NS filament or an H-NS filament mixed with equimolar Hha (green circle). Formation of Hha:H-NS filament was confirmed by EMSA and *in vitro* transcription assay (**Figure 7 – figure supplement 3**). Cleavage pattern of NT strand on the 146 bp DRE fragment generated by Fe·EDTA–H-NS in the presence (green) or absence (black) of equimolar Hha. Red arrows indicate hotspots of cleavage, which are at the same locations for both H-NS and Hha:H-NS filaments. In all data panels, hotspots of cleavage are indicated by arrows and error bars represent average of at least three replicates.

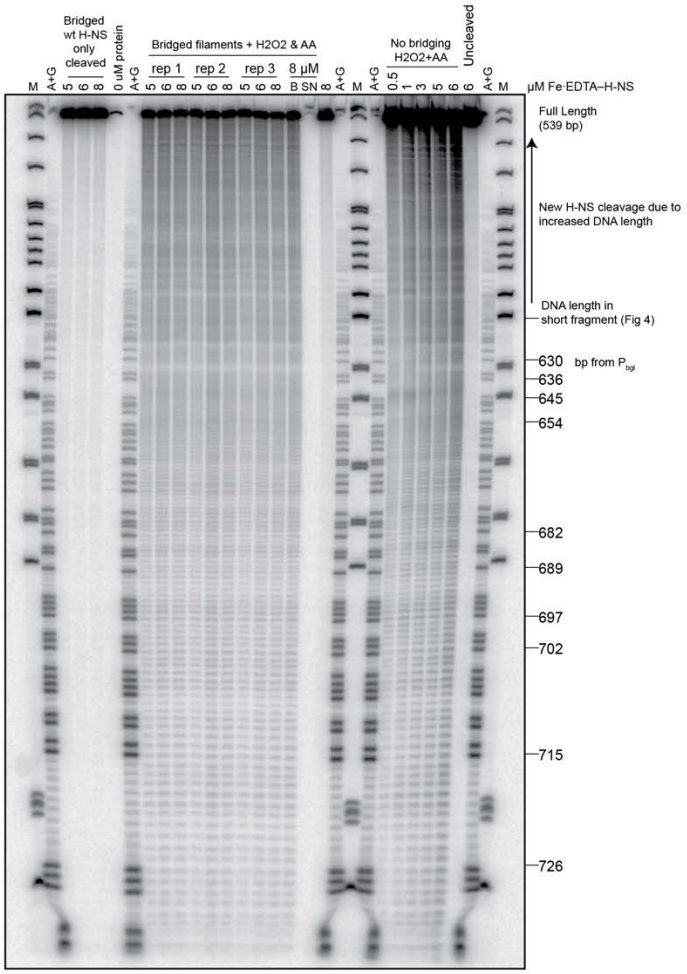


**Figure 7 – figure supplement 1. Fe·EDTA–H-NS cleavage of long DNA fragment in the presence or absence of bridging.**

7% PAG showing cleavage products generated by Fe·EDTA–H-NS in filaments formed on beads (bridged filaments) or in solution (no bridging). H-NS added to form filaments was 40% Fe·EDTA–H-NS and 60% WT H-NS at the indicated concentrations. 8 μM H-NS lanes marked "B" and "SN" indicate bridged filaments formed with 8 μM H-NS that were cleaved and then place back on the magnetic rack and the supernatant (SN) was removed which would contain any DNA that dissociated from the beads ("B") during the cleavage reaction. No bridging filaments were formed on 545 bp DNA without beads or bait DNA present at indicated concentrations of H-NS. A+G Maxam Gilbert A and G ladder and MspI digest pBR322 (M) were run alongside as size markers.

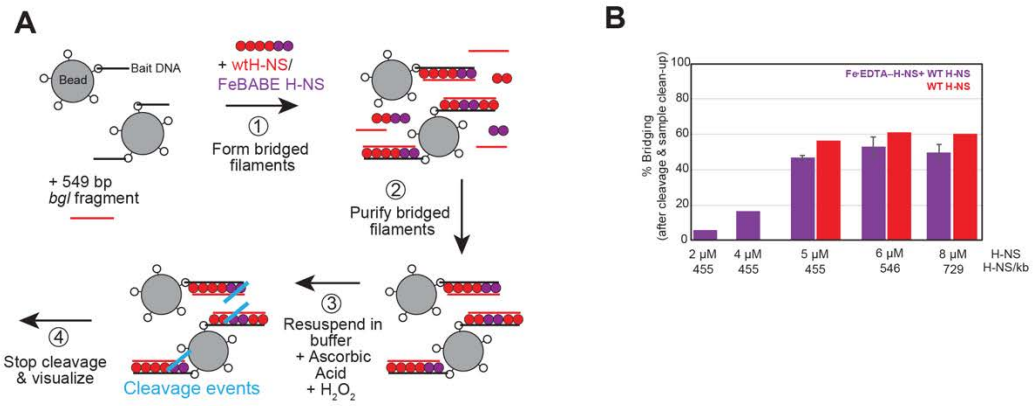


**Figure 7 – figure supplement 2. Fe·EDTA–H-NS cleavage of bridged filaments formed on beads.**

(A) Set-up of bridged cleavage assay. Filaments were formed with a 60-40 mix of WT H-NS (red) and Fe·EDTA–H-NS (purple) over a range of concentrations (5, 6 and 8 µM H-NS or 455, 546, and 729 H-NS/kb). Prey DNA was labeled on the T strand, but longer than fragments used in Figure 4 (549 bp vs 146 bp) since longer DNA was shown to work for the bridging pull-down assay. Purified bridged filaments (attached to beads) were gently resuspended in buffer, ascorbic acid, and $H_2O_2$ to induce cleavage. Reactions were stopped with thiourea and cleaned by phenol extraction and ethanol precipitation before quantification of bridging and separation of cleavage products by 7% PAGE. (B) Percent bridging of prey DNA in the presence of either WT H-NS (red) or a mixture of WT H-NS and Fe·EDTA–H-NS (purple).

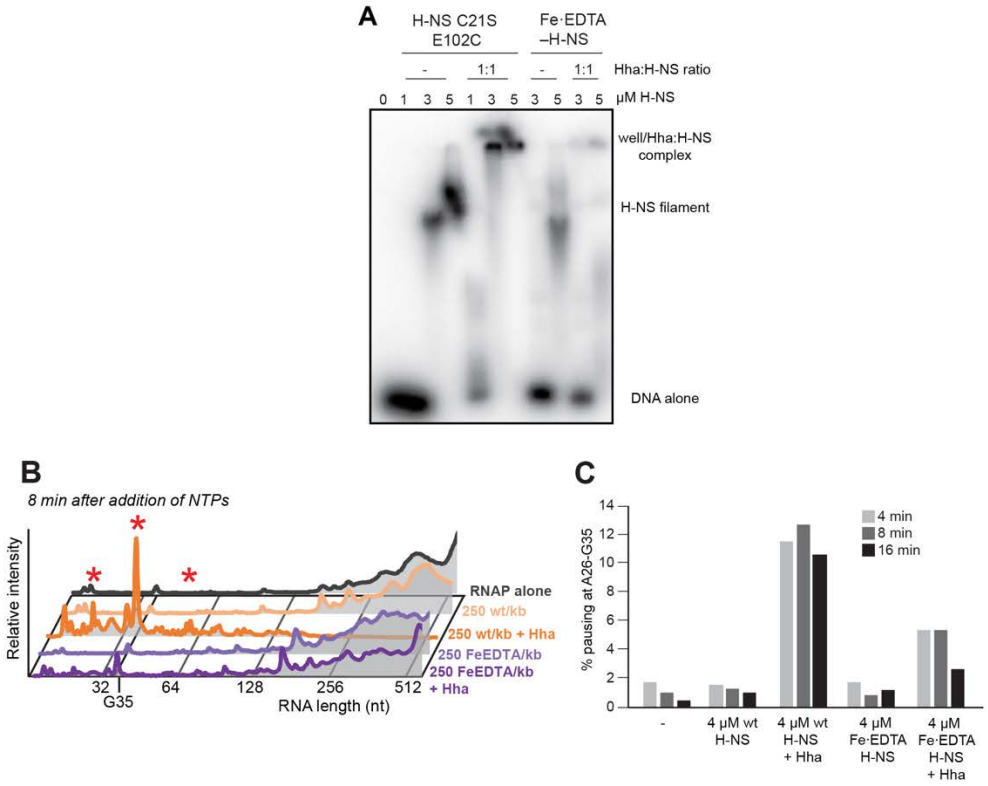


**Figure 7 – figure supplement 3. Formation of Hha:H-NS filaments with Fe·EDTA–H-NS.**

(A) 3% Native PAG showing H-NS or Hha:H-NS filament formation on 146 bp DRE DNA. Amounts of either H-NS C21S E102C or Fe·EDTA–H-NS indicated. Hha:H-NS filaments are retained in the well as previously observed (Boudreau et al., 2018). (B) Pseudo-densitometry plot showing transcripts made 8 min after addition of 30 μM NTPs in the presence of H-NS or Hha:H-NS filaments formed with WT or Fe·EDTA–H-NS in an *in vitro* transcription assay. Pauses stimulated by Hha:H-NS indicated with red asterisks. (C) Amount of pausing stimulated by H-NS and Hha:H-NS filaments. Hha added to both WT and Fe·EDTA–H-NS stimulated pausing at positions between A26 and G35 (Boudreau et al., 2018).

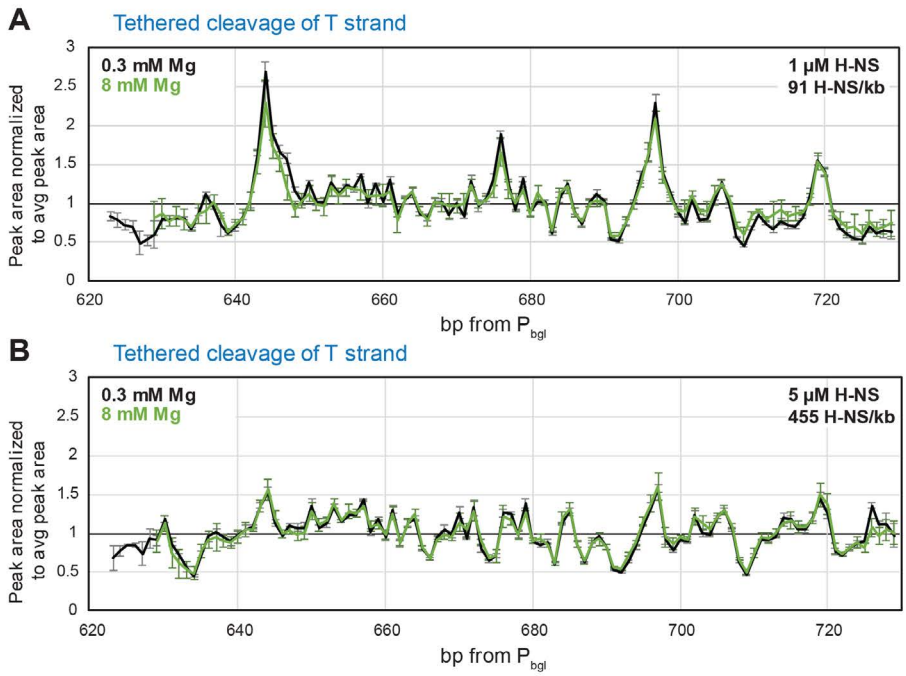


**Figure 7 – figure supplement 4. H-NS binds DNA in a similar pattern in the presence of low and high $Mg^{2+}$.**

Cleavage pattern of T strand generated by Fe·EDTA–H-NS at (A) 1 μM (91 H-NS/kb) or (B) 5 μM (455 H-NS/kb) in the presence 8 mM $Mg^{2+}$ (black) or 0.3 mM $Mg^{2+}$ (green). Filaments were formed on 549 bp fragment from the *bgl* operon (Figure 7A) labeled on the T strand. H-NS added was a 60-40 mixture of WT and Fe·EDTA–H-NS to match experiments done on beads (Figure 7B). Error bars represent the standard error of at least three replicates.

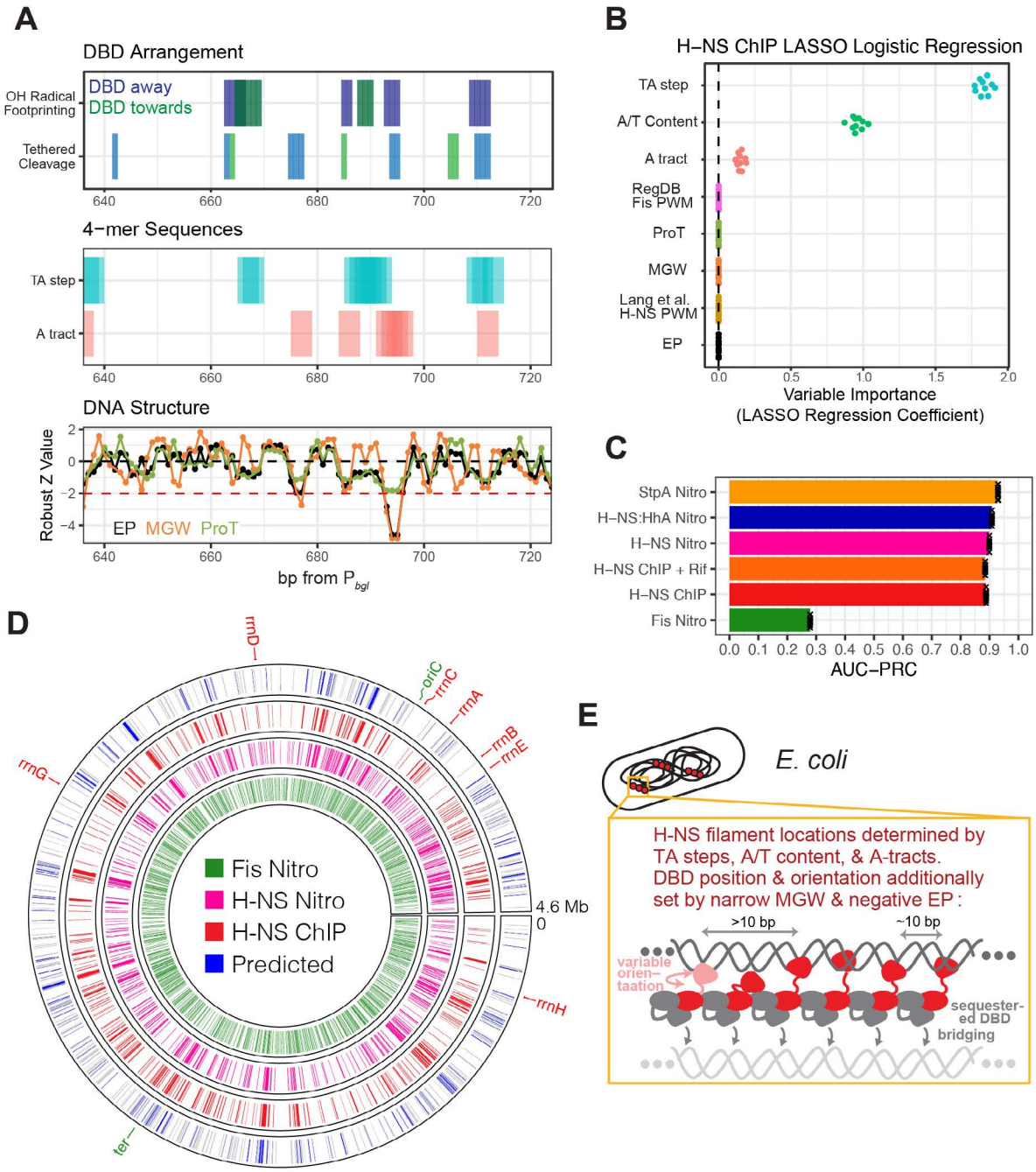


Figure 8.

**Figure 8. DNA sequence dictates H-NS DBD arrangement within filaments and H-NS filament location genome-wide**

(A) Summary of *in vitro* analysis of a single H-NS filament and association with DNA sequence and sequence-based structural features. (Top) Regions of high confidence DBD binding and

orientation determined from concordance between predicted and observed tethered cleavage patterns (see **Figure 6**) as well as predicted and observed DNA protection under ·OH radical footprinting (see **Figure 6 – supplement 2**). Each bar indicates regions where match scores between expected and observed patterns are below 95% of match scores obtained with 1000 shuffled datasets with each orientation of the DBD (see Methods). Blue bars indicate a DBD orientation away from $P_{bgl}$ whereas green bars indicate a DBD orientation towards $P_{bgl}$. (Middle) Location of 4mers consistent with a TA step or an A-tract (see Methods for definitions) within the DRE sequence. (Bottom) Overlay of predicted minor groove width (MGW, orange), propeller twist (ProT, green), and electrostatic potential in the minor groove (EP, black) across the DRE sequence. The values were Robust-Z scaled and centered using the genome-wide median and MAD (see Methods). The black dotted line indicates the genome-wide median whereas the red dotted line indicates the -2 cutoff used to determine regions with substantially lower values for that feature. (B) Determination of a minimal set of features to predict regions of H-NS binding across the genome. Final coefficients from a lasso logistic regression model trained on the H-NS ChIP data using the bound status of 1000 bp windows tiled every 250 bp across the genome are displayed with each point representing values determined from one of ten trained models using 10-fold cross validation with stratified sampling (see Methods). (C) Performance of models trained on predicting bound regions for each genome-wide dataset (from panel B). Each bar displays the average area under a precision recall curve (AUC-PRC) from the 10 replicates of models trained on stratified subsampled data (x indicates each replicate). (D) Circular plot (Gu et al., 2014) displaying the predicted genome-wide H-NS binding sites (black) as determined by a lasso logistic regression model trained on the full H-NS ChIP-seq data set and only considering windows with a predicted probability > 0.95 of being bound by H-NS as bound regions. For comparison, MOSAiCS determined peaks are shown for H-NS ChIP (red), H-NS Nitro-seq (pink), and Fis Nitro-seq (green). (E) DNA sequence and structure determine global and local H-NS filament arrangement. In *E. coli*, H-NS filaments form genome-wide at sequences with TA steps, A-tracts, and are AT-rich. Within each filament (yellow zoom in), most DBDs adopt an ordered arrangement with ~10 bp spacing between DBDs. Some DBDs are more flexible and adopt different orientations (pink arrow) or bind with >10 bp spacing along the

DNA (gray arrows). Importantly, the arrangement of H-NS coincides with DNA shape and sequence parameters. We propose our results show the same arrangement of the DBD in conditions forming either bridged or linear filaments. This observation could be explained by a filament structure where one DBD in an H-NS dimer is sequestered by interactions with the NTD in a linear filament but is able to bind DNA with ~10 bp spacing in a bridged filament (see Discussion).

**Table 1. Enriched ChIP-seq and Nitro-seq regions identified by MOSAiCS**

| SAMPLE | METHOD | TOTAL REGIONS | AVERAGE LENGTH | MEDIAN LENGTH | SMALLEST REGION | LARGEST REGION | FDR |
|---|---|---|---|---|---|---|---|
| **Untreated** | H-NS ChIP | 301 | 2005 | 1221 | 220 | 15629 | 0.02 |
| **Rif Treated** | H-NS ChIP | 289 | 2059 | 1279 | 211 | 15486 | 0.03 |
| **Untreated** | Beta ChIP | 294 | 1829 | 875 | 94 | 22190 | 0.06 |
| **Rif Treated** | Beta ChIP | 1352 | 601 | 400 | 200 | 3400 | 0.05 |
| **100 nM H-NS** | Nitro | 407 | 2066 | 1436 | 242 | 14836 | 0.05 |
| **50 nM StpA** | Nitro | 589 | 1817 | 1241 | 176 | 15281 | 0.05 |
| **Hha:H-NS** | Nitro | 436 | 2056 | 1490.5 | 122 | 14881 | 0.05 |
| **12 nM Fis** | Nitro | 888 | 917 | 818 | 209 | 4207 | 0.05 |