

Universiteit
Leiden
The Netherlands

The environmentally-regulated interplay between local three-dimensional chromatin architecture and gene expression

Rashid, F.Z.M.

Citation
Rashid, F. Z. M. (2021, June 22). *The environmentally-regulated interplay between local three-dimensional chromatin architecture and gene expression*. Retrieved from https://hdl.handle.net/1887/3192230

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3192230

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle https://hdl.handle.net/1887/3192230 holds various files of this Leiden University dissertation.

Author: Rashid, F.Z.M.
Title: The environmentally-regulated interplay between local three-dimensional chromatin architecture and gene expression
Issue Date: 2021-06-22

The environmentally-regulated interplay between local three-dimensional chromatin architecture and gene expression

Fatema Zahra Rashid

Table of contents

Chapter **Page**

Chapter 1:
Chromosome organisation in bacteria: mechanistic insights into genome structure and function 3

Chapter 2:
Hi-C in bacteria and archaea 46

Chapter 3:
Regulation of *proVWX* transcription by local chromatin re-modelling 68

Chapter 4:
HI-NESS: A family of genetically-encoded DNA labels based on a bacterial nucleoid associated protein 144

Chapter 5:
Outlook 174

Summary 182

Samenvatting 185

Chapter 1:
Chromosome organization in bacteria: mechanistic insights into genome structure and function

This chapter is based on:

Dame, R.T., Rashid, F.Z.M. and Grainger, D.C., 2020. Chromosome organization in bacteria: mechanistic insights into genome structure and function. *Nature Reviews Genetics*, pp.1-16.

Abstract

Bacterial chromosomes are folded to compact DNA and facilitate cellular processes. Studying model bacteria has revealed aspects of chromosome folding that are applicable to many species. Primarily controlled by nucleoid-associated proteins (NAPs), chromosome folding is hierarchical, from large-scale macrodomains to smaller-scale structures that influence DNA transactions including replication and transcription. Here, we review the environmentally regulated, architectural, and regulatory roles of NAPs and implications for bacterial cell biology. We also highlight similarities and differences in chromosome folding mechanisms of bacteria and eukaryotes.

Introduction

In all organisms, DNA is folded to fit inside the cell or its compartments. This is necessary because an organism's chromosome exceeds the cell's length by several orders of magnitude. Since the 1950's 'spreads' of liberated intracellular macromolecules, visualized by electron microscopy, have demonstrated the magnitude of this task. Genetic material readily spills out of lysed cells or nuclei to fill a volume many times larger than originally occupied (1). Precise mechanisms of DNA folding were first understood for eukaryotes. The basic structural units of eukaryotic folded DNA – nucleosomes (Figure 1.1a) – were identified as 'beads on a string' (2). The identification of higher-order structures (chromatin) was facilitated by the large size of eukaryotic cells that make them more amenable to light microscopy. Indeed, the basic dynamics of eukaryotic chromosomes during cell division were evident even before the genetic code was understood (3). It has taken much longer to understand chromosome organization in bacteria. Repeating structural units have never been identified, and early visualizations showed little more than a tangled mess (1). In retrospect, this is unsurprising. Bacteria lack most DNA folding factors present in eukaryotes so few cues can be taken. Furthermore, bacterial nucleoids undergo large changes in organization at different growth phases.

When we previously reviewed this topic in 2011 (4), evidence was emerging that bacterial chromosomes are not merely unstructured bodies of DNA, rather, the chromosomes fold into independent domains finely structured at the nanoscale. Advances in microscopy, structural biology, and genome-scale approaches (Box 1.1) have revealed many of the underlying molecular mechanisms. In this Review, we discuss these mechanisms and their impact on wider cell biology. Beginning at the level of individual DNA folding proteins we explain how DNA in bacteria is folded into myriad structures by looping, bending and twisting of the DNA. Subsequently, we explain how these DNA contortions influence not only nucleic acid compaction but also gene expression and DNA replication. On a whole-chromosome scale, we describe the characteristics of individual domains and discuss the possibility that the principles of chromosome folding are conserved across bacterial species. Throughout, we highlight similarities and differences in the DNA folding mechanisms used by bacteria and eukaryotes.

The principles of chromosome folding have been the subject of a long-standing research interest at Leiden University. The groups of Prof. dr. Pieter van de Putte and Prof. dr. Leendert Bosch advanced our understanding of the DNA binding and

structuring properties of factor for inversion stimulation (Fis) and Integration Host Factor (IHF), and highlighted their role in the regulation of genome transactions (5–11). The work presented in this thesis shows our continued interest in this field and carries on their heritage.

The nucleoid-associated proteins

Unlike eukaryotes, chromosomes of bacteria are not usually folded into regularly repeating structural units (Figure 1.1a). Instead, the chromosome is folded into a range of different conformations by nucleoid-associated proteins (NAPs) (Figure 1.1b-e). These are described further below.

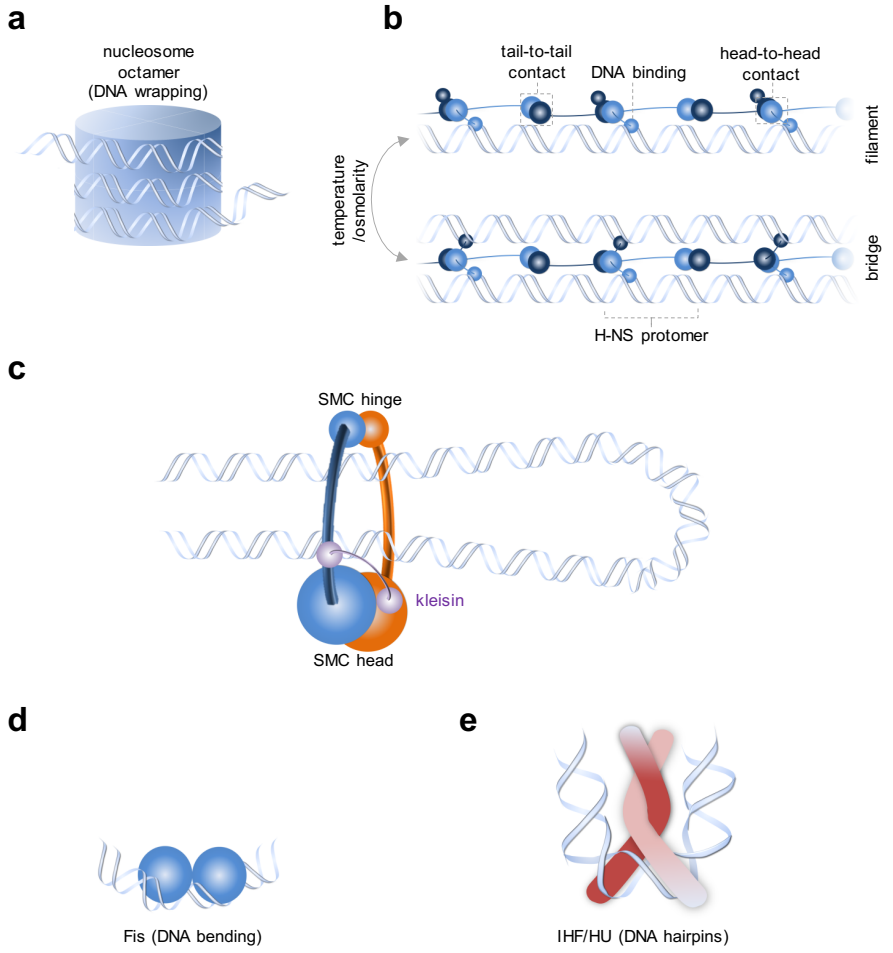


Figure 1.1: DNA is locally folded by NAPs in bacteria, histones in eukaryotes, and the evolutionarily conserved SMC complex. a: Eukaryotic chromosomes are folded into nucleosomes. Typically, a nucleosome is a regularly repeating structural unit that consists of 147 bp of DNA wrapped

around an octameric histone core; **b:** The 'daisy chaining' of H-NS along the DNA by head-to-head and tail-to-tail contacts forms H-NS:DNA filaments (top) and DNA:H-NS:DNA bridges (bottom), the latter of which results in the formation of DNA loops (12–14). *In vitro* studies indicate that the switch between the two modes of DNA binding by H-NS is mediated by changes in temperature and osmolarity (15–19); **c:** Structural maintenance of chromosomes (SMC) complexes are DNA looping proteins comprising a pair of SMC monomers, kleisin, and the 'kite' (kleisin interacting winged-helix tandem elements) or 'hawk' (HEAT repeat subunits containing proteins associated with kleisins) accessory/regulatory proteins. Each SMC monomer consists of a 'hinge' dimerization domain, an ATPase 'head' domain, and an anti-parallel coiled-coil 'arm' extending between the hinge and head domains. SMC complexes form DNA loops either by embracing a pair of DNA segments in a single ring, or by the dimerization of two rings that each trap a DNA segment (20–25) (also see Figure 1.2); **d:** Fis binds its target sequences as a dimer and induces a 50o–90o bend in the DNA (26, 27); **e:** IHF and HU also function as DNA bending proteins. IHF generates sharp 160o hairpin bends in the DNA, whereas HU functions as a flexible hinge (28) — it bends DNA less sharply, but over a range of different angles (29, 30).

Loop and filament formation by H-NS. The histone-like nucleoid structuring protein (H-NS) is a small (137 amino acids in *Escherichia coli*) polypeptide that binds the DNA minor groove via a C-terminal arginine hook motif (31). This is favoured for DNA with an elevated AT-content containing a TpA dinucleotide or 'step' (32). Hence, H-NS-bound genomic segments are AT-rich and have often been acquired by horizontal gene transfer (33–37). The N-terminal domain of H-NS contains two sites that facilitate 'daisy chaining' of the protein via head-to-head and tail-to-tail contacts (Figure 1.1b) (12). This drives the formation of lateral nucleoprotein filaments (Figure 1.1b, top) or loops between DNA segments bridged by H-NS (Figure 1.1b, bottom) (13, 14).

Proteins functionally similar to H-NS are found in diverse bacteria. Often, these have arisen via convergent evolution, that is, the independent evolution of the same function. For example, in *Burkholderia* spp. and *Mycobacterium tuberculosis*, the functional equivalents of *E. coli* H-NS are Bv3f and Lsr2, respectively. The proteins share structural similarity only in the arginine hook motif responsible for DNA binding (32, 38). *Bacillus subtilis* Rok shares no structural similarity with H-NS, Bv3f or Lsr2 yet fulfils the same physiological role (39, 40), by binding AT-rich DNA and having a strong preference for sequences containing a TpA step (41). The interaction, however, is not mediated by an arginine hook. Instead, lysine side chains in a winged helix make contacts with the DNA backbone (41). The MvaT protein of the *Pseudomonas* species also uses lysine residues in an AT-pincer motif to make similar contacts (42).

DNA looping by SMC proteins. Structural maintenance of chromosomes (SMC) complexes are tri-partite rings comprised of a pair of SMC monomers, kleisin, and the accessory/regulatory proteins 'kite' (kleisin interacting winged-helix tandem

elements) or 'hawk' (HEAT repeat subunits containing proteins associated with kleisins) (20–23). Each SMC monomer consists of a 'hinge' dimerization domain involved in the formation of a V-shaped SMC dimer, an ATPase 'head' domain, and an anti-parallel coiled-coil 'arm' extending between the hinge and head domains. The SMC dimer is bound to a kleisin complex to form a ring that captures DNA (20, 22, 24, 25), and, by encompassing two DNA segments, form a loop (Figure 1.1c; Figure 1.2, top) (43). Such a loop may also form by the dimerization of SMC dimers that each embrace a DNA segment (Figure 1.2, bottom) (44). Kleisins also recruit the regulatory kite and hawk proteins. Bacterial and archaeal SMC–kleisin complexes, and the eukaryotic Smc5–Smc6 complex, recruit the kite proteins. The condensin and cohesin SMC complexes of eukaryotes recruit the hawk proteins (21, 23).

In bacteria, three classes of SMC family proteins have been identified. SMC-ScpAB in *B. subtilis* and *Caulobacter crescentus*, SMC-like MukBEF in *E. coli* and other γ- and δ-proteobacteria, and the MukBEF-like MksBEF that has been detected in a wider range of bacterial species. These SMC family proteins are involved in segregation of newly replicated sister chromosomes (45–54). SMC–kleisin complexes are loaded onto the chromosome at the centromere-like *parS* sequences, positioned close to the origin of replication, by the *parS*-binding protein, ParB (55–57). Loading factors for SMC-like proteins, MukBEF and MksBEF, are currently unknown. Once associated with DNA, SMC complexes generate and maintain DNA loops, and are mechanistically characterized as loop extruding factors (Figure 1.2) (58, 59). By contrast, there is currently no evidence to suggest that SMC-like MukBEF and MksBEF play the same role. First proposed in 2001 (60), and formalized theoretically in 2012 (61), loop extrusion involves the clamping the protein complex around contiguous DNA sequences (Figure 1.2a) (60). The factor then 'pulls' the DNA through the clamp to produce a growing, unknotted loop of DNA (Figure 1.2b) (58). SMC proteins bi-directionally extrude DNA and progressively move along the chromosome (61) towards the terminus (Figure 1.2bc). The SMC extrusion complex may consist of a pair of DNA molecules pulled through one SMC ring referred to as the 'one ring, two DNA' model (Figure 1.2, top) (43), or one DNA molecule pulled through each ring of a 'handcuffed' pair (Figure 1.2, bottom) (44). The rate of loop extrusion by SMC is affected by transcription. Oppositely oriented highly-expressed genes (HEGs) attenuate the progression of SMC (Figure 1.2d) (56, 62). For instance, in *B. subtilis*, SMC progression can be slowed by >80% due to an oppositely-oriented HEG (62). As with other DNA looping proteins, SMC proteins have been suggested

to function by static loop formation — stably anchoring a pair of DNA loci to form a loop (63). However, as support that SMC proteins largely function by active loop extrusion in bacteria, chromosome arms progressively align from origin to terminus upon replenishment of the SMC loading factor ParB in *B. subtilis* (62), and they aberrantly align upon repositioning *parS* sites in *C. crescentus* (56).

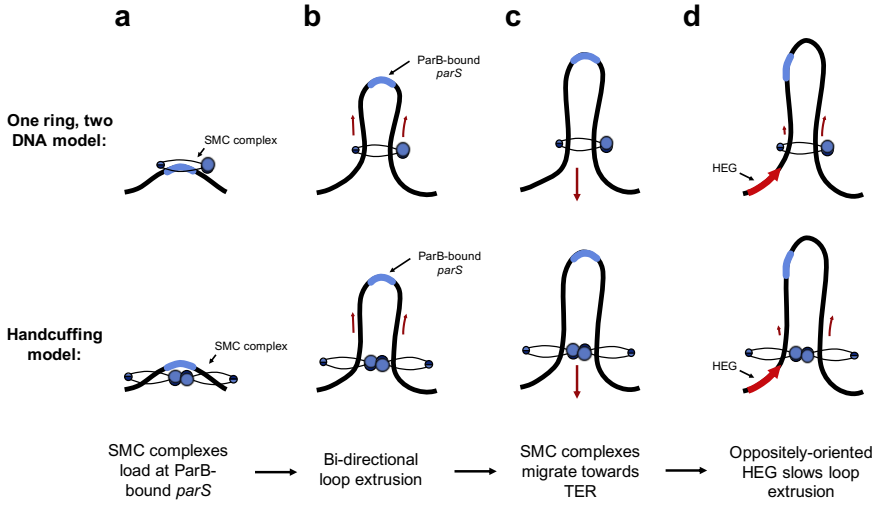


Figure 1.2: SMC proteins function as loop extruding factors. a: The structural maintenance of chromosomes (SMC) complex is loaded onto the DNA at ParB-bound parS sites in bacteria (55–57); **b:** The SMC complex then bi-directionally 'pulls' the DNA through its ring to extrude a growing, unknotted loop of DNA (58, 60, 61). The loop may be formed by an SMC complex entrapping two DNA strands within a single ring **(one ring, two DNA model (top))**, or by a dimer of two SMC rings that each trap one DNA segment **(handcuffing model (bottom))**; **c:** Loop extrusion allows bacterial SMC complexes to progressively move from the ParB-bound parS sites that are positioned close to the origin of replication towards the terminus (Ter) region (56, 62); **d:** The progression of SMC along the DNA is slowed by convergent transcription (56, 62), and may be slowed by up to 80% upon encountering an oppositely-oriented highly expressed gene (HEG) (62).

DNA bending by IHF, HU and Fis. Whereas H-NS and SMC proteins manage DNA loops (64), other NAPs primarily bend the DNA. For example, the factor for inversion stimulation (Fis) binds DNA as a dimer (Figure 1.1d) by virtue of a helix-turn-helix motif (26). Fis recognizes a 15 bp degenerate DNA palindrome characterized by a G at position 1 and a C at position 15 (5′ GNNVRWWWWWYVNNC-3′). Target recognition is driven by the shape of the minor groove resulting from the binding site sequence rather than the sequence itself (26). The degree of DNA bending induced by Fis binding can vary between 50° and 90° (Figure 1.1d) depending on the flanking DNA sequence (27). Fis is

often found at points where DNA duplexes cross (65, 66). This may stabilize plectonemes in supercoiled DNA.

DNA bending by integration host factor (IHF) is more severe, generating 160° bends (Figure 1.1e) (28). IHF binds its consensus sequence (5′-WATCAANNNNTTR-3′) (67) as a heterodimer, composed of α and β subunits. The minor groove is contacted by a β ribbon arm that protrudes from each subunit of the heterodimer (Figure 1.1e). This interaction is favoured by A-tracts. The insertion of a proline residue at the tip of each β-arm into the DNA base stack induces a hairpin bend in the DNA by kinking the DNA on either side of the hairpin apex (28). High intracellular concentrations of IHF permit non-specific interactions with many non-specific DNA targets, probably in a manner similar to HU (see below). The IHF protein has only been identified in gram-negative bacteria.

Heat-stable protein from *E. coli* strain U93 (HU) shares 40% sequence identity to IHF subunits (68). Unlike IHF, HU is found widely distributed among bacteria (69). In *E. coli*, HU forms heterodimers of α and β subunits. However, HU homodimers predominate across other bacteria in which, often, a single gene encoding HU is present. HU has no sequence specificity, but its mode of target recognition is similar to that of IHF. DNA is bent to a lesser extent by HU than IHF and over a range of different angles — akin to a flexible hinge (29, 30). Binding of HU also occurs preferentially at naturally bent or distorted DNA (70). The bends induced by each HU β-arm force the DNA out of a single angular plane (29, 71). Consequently, sequential binding of HU dimers induces coiling of the DNA around the bound proteins to form filaments. This means that HU can restrain negative supercoils in DNA and, alongside topoisomerase I, introduce negative supercoils in circular DNA (72–74). Generally, DNA is negatively supercoiled in bacteria to facilitate DNA transactions that require DNA melting (75).

Cross-talk between NAPs. The cross-talk between NAPs regulates the structural conformation adopted by the resulting nucleoprotein complex (76). *In vitro* atomic force microscopy studies indicate that the combinatorial effect of NAPs is influenced by DNA topology and the orientation of high-affinity NAP binding sites in the underlying DNA sequence (77, 78). In 42.6 kb DNA molecules, the differential structural effects of NAPs were only observed in negatively-supercoiled, hyperplectonemic structures. The binding of both H-NS and Fis unravelled the hyperplectoneme into individual plectonemes constrained by H-

NS-mediated bridges. DNA cross-over points in the complex were bound by Fis (77). In a pair of relaxed plasmids and corresponding linear DNA constructs that differed only in the order of high affinity Fis and H-NS binding regions — the upstream activating sequence of the tyrosine tRNA, and the negative regulatory element of the *proVWX* operon, respectively — the presence of both Fis and H-NS resulted in the folding of the DNA duplexes as a consequence of H-NS-mediated bridging. Fis occupied the cross-over points. Interestingly, the construct with high-affinity H-NS and Fis binding regions oriented in a Fis:H-NS:H-NS:Fis manner, but not the Fis:H-NS:Fis:H-NS construct, organised into a stem-loop formed by an H-NS-mediated bridge book-ended by Fis binding. The NAPs also 'phase separated' in the structure: the binding of Fis (a NAP conducive to transcription) excluded H-NS (a repressive NAP) from the looped DNA, a region that H-NS occupies in the absence of Fis (78). Collectively, these limited studies point to a model of the bacterial chromosome structured into chromatin islands of Fis-bound, transcriptionally-active, 'open' chromatin and H-NS-bound, transcriptionally-silent, 'closed' chromatin, the organisation of which is encoded as NAP binding sites in the genome and modulated by NAP availability and supercoiling density (77, 78).

Modulation of NAP function by other proteins. The architectural properties of NAPs may be regulated by paralogues and NAP modulators. For instance, DNA binding by H-NS is regulated by its paralogue StpA, and by Hha, a NAP modulator that belongs to the YmoA family of proteins (15, 79). StpA shares 58% sequence identity with H-NS and forms homodimers, and heterodimers with H-NS *in vitro*. Heterodimers are likely to predominate *in vivo* since the StpA homodimer is susceptible to proteolysis (80, 81). Therefore, H-NS-bound regions of the chromosome also contain StpA (82). StpA stimulates DNA bridging by H-NS and stabilizes the structure against changes in temperature and Mg^{2+} or K^{+} concentration (79). Hha is an 8 kDa protein involved in the regulation of H-NS-like proteins. Factors like Hha lack a DNA-binding domain and interact with the N-terminal domain of H-NS to enhance DNA bridging (15, 79, 83, 84). Disruption of the H-NS–Hha interaction relieves the repression of H-NS–Hha co-regulated operons, such as *hilA*, with minimal disruption of the H-NS binding profile at the operon (84).

Comparison with eukaryotic and archaeal DNA-folding proteins. The nomenclature for NAPs, for instance, 'histone-like nucleoid structuring protein' for H-NS, can imply a relationship to eukaryotic histones. However, there are few

similarities at the protein level. Most notably, H-NS is histone-like in only one regard: it is an abundant DNA-binding protein. Even so, there are many examples of eukaryotic DNA-binding proteins that utilize arginine hooks to bind AT-rich DNA (85). These can have global DNA-folding properties. For example, the metazoan special AT-rich sequence-binding protein 1 (SATB1) has genome-wide roles in DNA folding and, like H-NS, might link higher-order nucleoprotein structures and gene regulation (86–88). The HU protein, like histones, is able to induce DNA supercoiling (72–74). Indeed, the ability of HU to wrap DNA in filaments hints that the protein has the capacity to form structures similar to hypernucleosomes in archaea (89, 90). However, although the protein–DNA co-crystal structures are comparable for archaeal histones and HU, solution studies do not support this model (30). Structurally and functionally, SMC complexes in bacteria and eukaryotes have similar functions in managing DNA loops (see below).

Analogous to eukaryotic histones, bacterial NAPs also undergo post-translational modifications (PTMs). To date, 29 PTMs have been identified for *E. coli* H-NS that may fine-tune its properties. Acetylation neutralizes charges of Lys83, Lys87, Lys96, Lys120 and Lys121 that are known to facilitate DNA binding. Succinylation of Lys96, Lys120, and Lys121 may also interfere with DNA binding due to steric hindrance (91–94). Some HU proteins have terminal extensions enriched for lysine, proline or alanine repeats, reminiscent of the (S/T)PKK motifs found in eukaryotic histone H1 that are subject to post-translational modification (69).

Local patterns of DNA folding

Chromosome interaction domains. At a scale of tens to hundreds of kilobases, the bacterial chromosome (Figure 1.3Aa) is partitioned into chromosome interaction domains (CIDs) (Figure 1.3Ab) (95, 96), analogous to the topologically associating domains (TADs) in eukaryotes (compare Figures 1.3Ab and 1.3Bb) (97–99). CIDs and TADs exhibit a high degree of self-interaction and are insulated from flanking regions.

Hi-C in *C. crescentus* indicates that the chromosome is organized into 23 CIDs during exponential growth in rich medium and 29 CIDs in starvation conditions, with the length of these domains varying between 30 and 420 kb (95, 100). The boundaries between CIDs correspond to positions of HEGs that are >2kb in length (95, 100) (Figure 1.3Ab). In *C. crescentus*, these include, for instance, genes within the ATP synthase and NADH–quinone oxidoreductase gene clusters during

exponential growth in rich medium, and starvation-induced genes such as CCNA03169 (Lrp-like asnC family transcription regulator), and CCNA03327 (histidine kinase involved in signal transduction) during periods of starvation. House-keeping genes such as those within the ribosomal protein gene cluster form CID boundaries during both conditions, albeit of different strength. The ribosomal protein gene cluster forms a sharp CID boundary in exponential phase. At this stage of growth, the genes in the cluster exhibit a high transcription rate. During starvation, in accordance with the decreased rate of gene expression within the cluster, the sharpness of the boundary diminishes (95). Recombination-based experiments indicate that while HEGs generally form plectoneme-free regions that act as supercoil diffusion barriers, only long HEGs can generate extended barriers that insulate CIDs by physically separating flanking chromatin (100). Indeed, the ectopic insertion of a long, highly-expressed gene is sufficient to establish a CID boundary in the chromosome (95).

The *B. subtilis* chromosome is organized into 20 CIDs 50–300 kb in length. While 60% of the CID boundaries coincide with HEGs, ~30% overlap with sections of the genome bound by the Rok protein (96). This observation implies that Rok (and by extrapolation other bacterial NAPs) could function as domain barriers. The *E. coli* chromosome appears to be organized into 31 CIDs between 40 to ~300 kb in size. 22 of the CID boundaries correspond to the positions of HEGs, and 9 boundaries coincide with positions of genes that code for proteins with an export signal sequence (101). This may be relevant in light of the hypothesis that chromosomes are membrane-appended by coupled transcription–translation–translocation (102).

While multiple systems contribute to the formation of CID boundaries in bacterial chromosomes, the hierarchical structural organization that they contribute to is conserved. Bacterial CIDs exhibit a nested domain organization with each domain composed of smaller sub-domains (Figure 1.3Ad–e) (95, 96, 101). The smallest units of this organization may correspond to individual operons (Figure 1.3Ac)

In eukaryotic chromosomes (Figure 1.3B), TADs are typically formed by loop extrusion (61, 103, 104). Several lines of evidence suggest that SMC complexes including cohesin, condensin and Smc5/6 function as loop extruding factors. Unlike their bacterial counterparts, eukaryotic SMC complexes do not appear to be loaded onto the chromosome/chromatin at a specific DNA sequence. Upon clamping around the DNA, eukaryotic SMC complexes continue to extrude loops

until the complexes either spontaneously dissociate from the DNA, collide with another factor, or encounter an appropriately oriented TAD boundary element (61, 103, 104). TAD boundaries are encoded in the genome as CCCTC-binding factor (CTCF) binding sites. These sites exhibit a directionality, such that a TAD only forms between a pair of inward-facing CTCF sites (Figure 1.3Bb) (105–107). Indeed, deletion or inversion of CTCF binding sites disrupts TAD boundaries *in vivo* (104). Curiously, in *Drosophila melanogaster*, SMC complexes and CTCF are not markedly enriched at TAD boundaries. In flies, this role is played by the insulator complexes BEAF-32–CP190 and BEAF-32–Chromator (108–110).

As in bacteria, the TADs of eukaryotic genomes are nested (105, 109, 111). The smallest organizational units of TADs correspond to individual genes in *Saccharomyces cerevisiae* (Figure 1.3Be) (111). The functional relevance of this organization and the molecular determinants of the boundaries are still unclear.

Archaea are an evolutionary link between the bacterial and eukaryotic branches of life, with the physiology of archaeal cells often manifesting as an amalgamation of the physiologies of bacterial and eukaryotic cells (112). A 3C-based study of the model crenarchaea *Sulfolobus acidocaldarius* and *Sulfolobus islandicus* indicates that crenarchaeal chromosomes are compartmentalised into a transcriptionally active A-compartment and a transcriptionally repressed B-compartment – a feature characteristic of eukaryotic genomes (113). The chromatin in both compartments organises into self-interacting CIDs, the boundaries between which are formed by highly-transcribed genes (114) reminiscent of bacterial chromosomes (95, 96, 100). The maintenance of CIDs in the B-compartment, the repression of transcription units within the structures, and its separation from the A-compartment is facilitated by an SMC-like protein termed Coalescin (113, 114). The chromosome contact maps of *Sulfolobus* show evidence of the formation of transcription hubs. Genes involved in ribosomal biogenesis form long-range loops that appear to be dependent on active transcription. Decreased expression of the ribosomal biogenesis genes as a result of entry into stationary phase or treatment with actinomycin D reduces the strength of the loops (114). On the other hand, the organisation of chromosomes in the euryarchaeaota, represented at present by *Haloferax volcanii*, *Halobacterium salinarum*, and *Thermococcus kodakarensis*, involves folding into CIDs and absence of A/B compartmentalization (115). In *H. volcanii*, CIDs are formed by archaeal SMC proteins (a eukaryotic feature) and transcription (a bacterial feature) as evidenced by the loss of looping and CID

boundaries in a Δ*smc* mutant and in wild-type cells treated with actinomycin D (115).

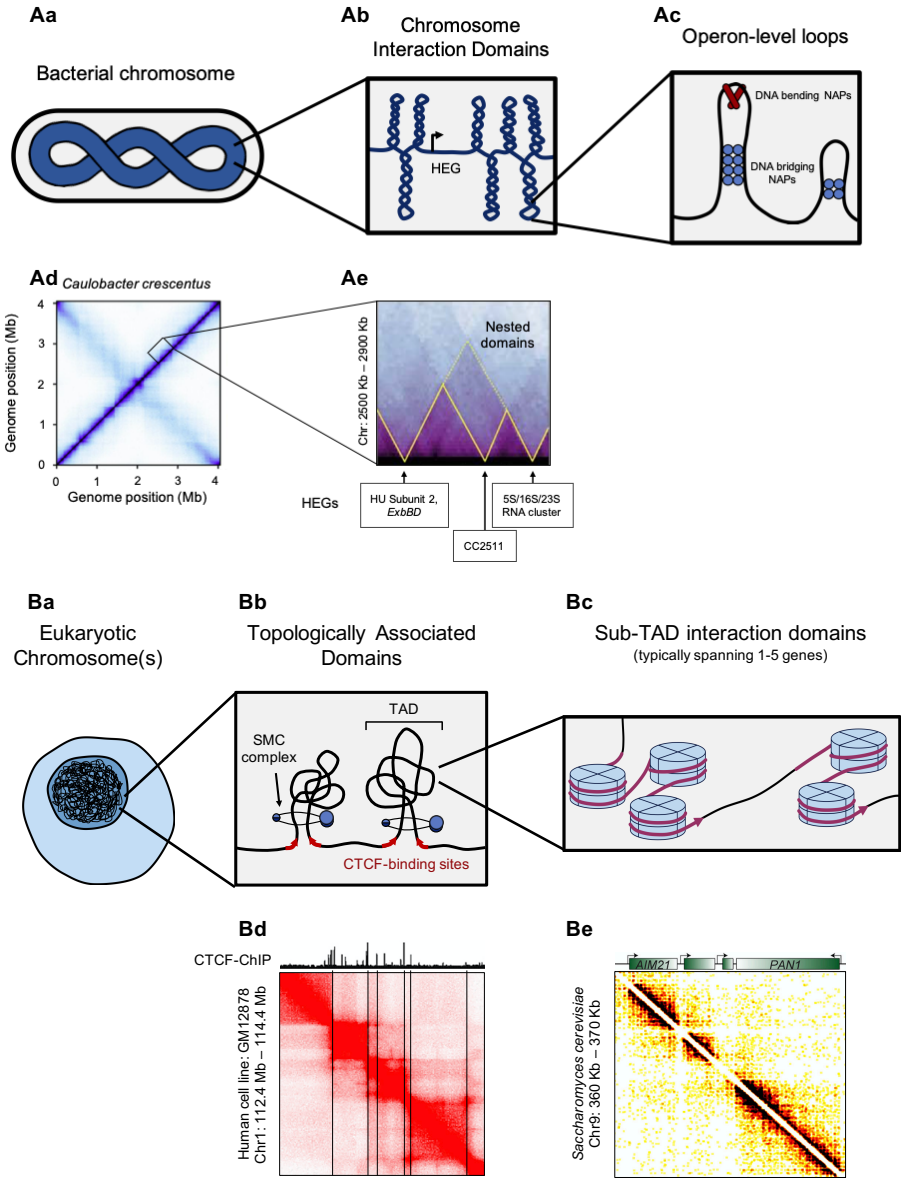


Figure 1.3: Chromosomes are hierarchically organized in bacteria and eukaryotes. A: Bacterial chromosome organization. At a global scale, the bacterial chromosome is spirally folded to fit within the bacterial cell **(Aa)**. Regions of the chromosome sequentially close to each other interact in three-dimensional space as evidenced by the presence of a primary diagonal of high interaction frequency in a Hi-C contact map **(Ad)**. Except for *Escherichia coli*, all reported bacterial chromosome interaction profiles also show a secondary diagonal of low interaction frequency that lies perpendicular to the primary diagonal **(Ad)**. This feature indicates interaction between the chromosomal arms that run

alongside each other in the spirally organized chromosome (inter-arm interaction) (95, 96, 101, 116). At the scale of tens to hundreds of kilobases, the chromosome is subdivided into chromosome interaction domains (CIDs) **(Ab)** (95, 96). CIDs exhibit self-interaction and are insulated from flanking chromatin. These structures are observed as squares along the primary diagonal of a Hi-C map **(Ad)** or as triangles when observing one half of the symmetric Hi-C map **(Ae)**. Bacterial CIDs are nested (95, 96, 101): larger domains (broken yellow line) are organized into smaller sub-domains (solid yellow line) **(Ae)** (95). The boundaries between the domains are typically formed by highly expressed genes (HEGs) >2kb in length that physically separate the flanking chromatin **(Ab)** (95, 100). The smallest structural unit of organization of the bacterial chromosome may correspond to loops formed at the level of individual operons by nucleoid-associated proteins (NAPs) **(Ac)**. **B: Eukaryotic chromosome organization.** The eukaryotic chromosome, localized inside the nucleus **(Ba)**, is organized into topologically associating domains (TADs) **(Bb)**, analogous to the bacterial CIDs. TADs are formed by loop extrusion (61, 103, 104). Eukaryotic structural maintenance of chromosomes (SMC) proteins load onto the chromosome and extrude DNA loops (Figure 1.2) until the complexes collide with inward-facing CTCF binding sites **(Bb)** (61, 103, 104). Indeed, TAD boundaries — identified as the region between two squares along the diagonal of a Hi-C matrix — occur at genomic regions enriched for CTCF **(Bd)** (105–107). Eukaryotic TADs are nested, with the smallest sub-TAD interaction domains typically comprising up to 5 genes **(Bc,Be)** (111). Hi-C contact maps in parts Ad and Ae are modified with permission from (95). GM12878 Hi-C map ad CTCF-ChIP profile in part Bd is modified with permission from (106). *Saccharomyces cerevisiae* Micro-C map in part Be is modified with permission from (111).

The impact of DNA supercoiling. Local patterns of DNA supercoiling influence DNA folding within topologically isolated regions of the *E. coli* chromosome (Figure 1.3Ab). However, tools to measure chromosome-wide patterns of DNA folding have only become accessible in recent years (117, 118). As noted above, DNA is in an average state of negative supercoiling. However, supercoiling density is unevenly distributed and varies across phases of growth. In particular, a gradient of increased negative supercoiling runs from the origin of replication to the terminus, along each arm of the chromosome, only in starved cells (117). This gradient requires the HU protein. The wrapping of DNA around HU, and the change in twist of the double-helix mediated by the protein are consistent with effects of HU on global DNA supercoiling (117). The expression levels of HU also vary strongly across different phases of growth (119), potentially explaining effects on DNA topology. Collectively, this may also explain why loss of HU has different effects on intra-chromosome interactions in different bacteria (95, 101) with different levels of DNA supercoiling (120).

Interactions between chromosomal arms. In bacteria, progression of SMC from the origin to the terminus mediates contacts between the right and left replichores, resulting in their parallel alignment. This manifests itself as a characteristic ‘secondary diagonal’, perpendicular to the main diagonal in Hi-C matrices of bacterial chromosomes (95, 96, 116) (Figure 1.3Ad). Curiously, this secondary diagonal is absent in the contact maps of the *E. coli* chromosome (101) despite the presence of the SMC-like MukBEF system. The MukBEF complex, in the absence

of ATP, consists of a V-shaped MukB dimer, the MukF kleisin that extends between the pair of MukB head domains, and four MukE kite proteins. Unlike other characterized kleisins, MukF is not monomeric (20, 21, 121). Instead, MukF forms a dimer via its N-terminal winged-helix domain and binds MukB via its C-terminal domain (122, 123). Upon binding ATP, the MukB ATPase heads dimerize and sterically displace one of the two MukF monomers, rendering the N-terminal of the MukB-bound MukF available for dimerizing with another ATP-bound MukBEF complex (123). The subsequently formed dimer is the minimal functional unit of MukBEF (124). ATP-bound MukBEF stably associates with the chromosome and is involved in its condensation while ATP hydrolysis results in MukBEF dissociation (124, 125). The hydrolysis of both ATP molecules contained within the dimerized head domains is required for a single MukBEF unit to release the DNA (123–125).

The MukBEF complex has been proposed to move along the chromosome as a 'rock climber'. In this model, ATP hydrolysis in a single unit of a MukBEF dimer releases it from the chromosome while the other unit remains bound. ATP binding then allows the released MukBEF unit to capture a different segment of the DNA and hence, move along the chromosome (124) (Figure 1.4A). The release–capture cycles of this model implies that the minimal functional unit of MukBEF in *E. coli* cannot promote and maintain inter-arm interactions as observed for loop-extruding SMCs such as SMC–ScpAB in *B. subtilis*. Other models which consider that a MukBEF dimer is not completely released from DNA upon ATP hydrolysis speculate that MukBEF may still carry a loop extrusion functionality (Figure 1.4B) (126).

Global chromosome organization

Over the past 10 years, advances in genome-scale approaches have improved our understanding of bacterial DNA folding at the micron scale (Box 1.1). Most notably, chromosomal patterns of NAP binding, and physical interaction frequencies have revealed independently organized macrodomains with distinct properties (4). Such structures are best defined for *E. coli* where the chromosome is divided into 4 macrodomains and two non-structured regions. All macrodomains exhibit reduced intracellular mobility compared to the non-structured chromosomal regions. Thus, macrodomains tend to interact with the non-structured regions but not with other macrodomains (127).

Constraint of Ori macrodomain mobility by MaoP. The Ori macrodomain contains the origin of chromosome replication *oriC* (128, 129). The constrained

mobility of Ori requires the *yifE* gene product, MaoP (macrodomain Ori protein) and a 17 bp motif in the upstream intergenic region (5′-CTAATACTCCGCGCCAT-3′) named *maoS* (macrodomain Ori sequence) (128). In otherwise wild-type cells, inactivation of *maoS*/MaoP specifically increased the mobility of Ori (128). It is not known how MaoP acts over long distances to constrain DNA mobility.

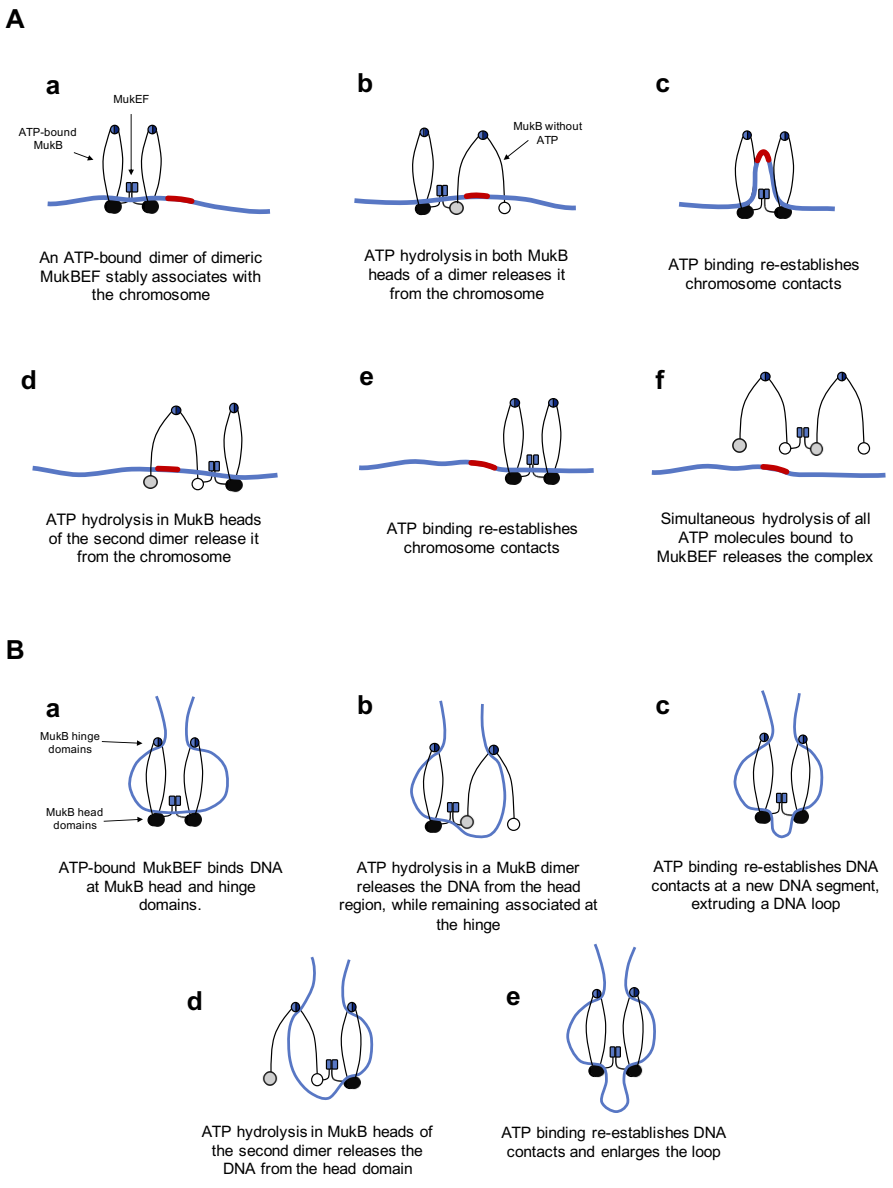


Figure 1.4: MukBEF moves along the chromosome as a 'rock climber'. A: MukBEF movement along a single DNA molecule (based on the rock climber model proposed in (124)**). Aa:** The minimal functional unit of the MukBEF complex corresponds to a dimer of dimers — $MukB_4E_4F_2$. When each of the MukB heads are bound to ATP, the MukBEF complex remains stably associated with the

chromosome. **Ab:** Hydrolysis of ATP in both MukB heads of the same dimer disengages the MukB heads, and releases the dimer from the chromosome. The MukBEF complex remains bound to the DNA via the ATP-bound dimer. **Ac:** ATP binding to MukB of the released dimer re-establishes chromosome contacts with a different chromosomal locus. **Ad–e:** A DNA segment release and capture cycle in the second MukBEF dimer allows the complex to move along the chromosome as a 'rock climer'. **Af:** Simultaneous hydrolysis of all four ATP molecules bound to a MukBEF complex releases it from the chromosome. This step may involve a MukBEF 'unloading' factor. **B: MukBEF as a loop extruding factor** (126)**. Ba:** ATP-bound MukBEF binds chromosomal DNA at the MukB head and hinge domains. **Bb:** ATP hydrolysis in a MukBEF dimer releases the chromosomal DNA segment bound at the MukB head domains. **Bc:** ATP binding re-establishes MukB head–DNA contacts at a new DNA segment, thus generating a DNA loop. **Bd–e:** A release–capture cycle in the second MukBEF dimer results in loop enlargement, hence, loop extrusion.

Condensation of Ter macrodomain structure by MatP. The Ter macrodomain is diametrically opposed to Ori (Figure 1.5Aa) and encompasses the replication terminus. A major breakthrough for understanding Ter was the identification of a sequence repeated 23 times in Ter but not elsewhere in the *E. coli* chromosome (5′ GTGACRNYGTCAC-3′) (130). The same sequence uniquely occurs in equivalent parts of many bacterial chromosomes (130). This DNA site, named macrodomain Ter sequence (*matS*) is the target of the macrodomain Ter protein (MatP). This interaction is highly specific, as shown by MatP exclusively binding these DNA targets in chromatin immunoprecipitation (ChIP) experiments (130). Loss of MatP activity leads to decondensation of the Ter macrodomain (130). MatP consists of three domains: an N-terminal 4 helix bundle, a central β-strand helix-helix and a C-terminal coiled-coil (131). Interaction of MatP with DNA is mediated by the β-strand helix-helix that resembles ribbon helix-helix structures found in other DNA-binding proteins (131). MatP binds DNA as a dimer mediated by interactions involving both the N-terminal and central domains. The C-terminal coiled-coil is required for tetramerization of MatP. Such tetramers generate bridges between distal matS sites on the chromosome, effectively condensing the Ter macrodomain (131, 132). This is evident in Hi-C experiments: deletion of matP specifically restructures Ter with reduced intradomain interactions being observed (101). Loss of MatP also prevents correct positioning of the DNA replication at mid-cell and this depends on an interaction between MatP and division-apparatus-associated protein ZapB (133).

Other proteins with macrodomain-specific DNA-binding properties. In *E. coli* at least two additional proteins, SeqA and SlmA, have macrodomain-specific DNA-binding properties (134, 135). However, unlike MaoP and MatP, there is no evidence that SeqA and SlmA contribute to the overall folding of these domains. Briefly, SeqA is involved in sequestration of the DNA replication origin after a new round of DNA replication has been initiated (136). This is permitted because

newly replicated DNA is hemimethylated at 5′ GATC-3′ motifs targeted by DNA adenine methylase (Dam) and SeqA (137). These 5′ GATC-3′ motifs are underrepresented in the Ter macrodomain and overrepresented elsewhere, particularly near to the origin of replication (135). Similarly, SlmA binds throughout the *E. coli* chromosome, except in the Ter macrodomain (134). SlmA recognizes the sequence 5′ GTGAGTACTCAC-3′ and is required for correct cell division (134). SlmA, SeqA, MatP and MaoP are co-conserved in bacteria encoding Dam methylase, suggesting that these bacteria use similar strategies to organize their chromosomes. Indeed, even in bacteria lacking Dam, proteins with similar patterns of chromosome-wide DNA binding have been identified. For instance, in *B. subtilis* the nucleoid occlusion (Noc) protein appears to be the functional equivalent of SlmA (138). Similarly, in *C. crescentus*, GapR targets a large region surrounding the origin of replication (139, 140).

Environmental regulation

The structure of the bacterial chromosome changes in response to the environment (Figure 1.5Ab). In part, this is because a small number of NAPs (most notably H-NS, and MvaT) can undergo conformational changes in response to specific ligands (15, 141). More commonly, the intracellular concentration of NAPs alters in response to environmental triggers (119). These two scenarios are discussed in more detail below.

Studies on environmentally-triggered conformational changes of NAPs have largely focused on H-NS and H-NS-like proteins (15, 141, 142). In this regard, helix α3 of *E. coli* H-NS plays an osmosensory role. The helix is unstable and frequently buckles. This folds one of the DNA-binding domains of the H-NS dimer onto the body of the protein (15). A similar conformation is also adopted by the H-NS family protein MvaT under low-osmolarity conditions; electrostatic interactions occur between a positively charged patch at the C-terminal DNA-binding domain and a negatively charged patch at the N-terminal domain (141). In the folded conformation, one of the two DNA-binding domains of the protein dimer is unavailable for DNA binding, thus favouring the formation of lateral filaments along DNA (Figure 1.1b, top) (15, 141). Magnesium ions stabilize helix α3 in H-NS to prevent buckling (15). Correspondingly, high-osmolarity conditions destabilize the electrostatic interaction between the N- and C-terminal domains of MvaT (141). Hence, both DNA-binding domains of the H-NS and MvaT dimers become available for DNA binding and bridged loops can form (Figure 1.1b,

bottom) (15, 141). The osmosensory role of helix α3 of *E. coli* H-NS was verified by stabilising the helix with E43A,E44A,S45A mutations (15).

H-NS is also temperature sensitive. High temperatures reduce the co-operativity of H-NS oligomerization and favour its dissociation from DNA (16–19, 142). Helix α4 of *Salmonella typhimurium* H-NS, located in the protein's dimer-dimer interaction domain, functions as a thermosensor. Heat-induced unfolding of this helix destabilises H-NS oligomerisation. The disruption also folds the negatively-charged N-terminal region of the protein onto its positively charged C-terminal, sequestering the DNA binding domain in an 'auto-inhibited' conformation (142). The structural change may relieve H-NS-mediated repression and trigger the expression of toxicity islands at human body temperatures. *In silico* Molecular Dynamics simulations and *in vitro* thermolysin assays indicate that in contrast to *E. coli* H-NS, the structure of helix α3 of *S. typhimurium* H-NS does not respond to changing osmolarity (or temperature), and therefore, might not function as an 'environmental sensor' in this protein (142).

A less subtle mechanism controlling chromosome dynamics is based on levels of NAP expression, which can change substantially (reviewed in (76, 143, 144)). This is most notable during stress and starvation when the nucleoid is reorganized into a condensed crystalline structure (Figure 1.5B) (145). Most NAPs are present at lower levels in starved cells, and Fis, which is among the most abundant DNA-binding proteins during periods of rapid cell division, is undetectable (146). Conversely, DNA-binding protein from starved cells (Dps) and curved DNA-binding protein A (CbpA) – NAPs that are undetectable during rapid growth – accumulate to 175,000 and 14,000 copies per cell in stationary phase (119, 147). Both bind the DNA highly co-operatively, and interactions between DNA-bound protein molecules lead to DNA compaction (66, 148). Electron micrographs of Dps–DNA complexes reveal that they are organized in a crystalline lattice *in vitro* (Figure 1.5B) (149). A similar structure is observed *in vivo* when Dps is expressed in exponentially growing Δ*fis* strains (149, 150). Complexes of Dps or CbpA with DNA are resistant to damage induced by chemical and biological nucleases (66, 151). Hence, expression of these proteins is thought to protect the integrity of the genetic material in harmful environments. However, how these highly condensed protein–DNA structures co-exist with other cellular processes has been a mystery. Recent work has shed light on the puzzle, revealing that super-condensed nucleoids of starved *E. coli* cells are phase-separated organelles (151). Phase-separation is an inherent physical property of macromolecules (such as proteins)

to self-organize into condensates or 'droplets' in a crowded environment such as the cell interior. Binding of Dps to DNA *in vitro* blocks access to nucleases and hydroxyl radicals, but the DNA remains fully permissive to transcription (147, 151). This is because the physical properties of some molecules allows them to move between separated phases whilst other molecules are trapped within a specific phase.

Interplay with genome transactions

As eluded to above, understanding how chromosome folding impacts other cellular processes has been a longstanding area of interest. Although Dps seems unable to impede transcription, the same is not true of other NAPs that can have specific effects on gene expression. Furthermore, additional roles have been identified for NAPs in chromosome replication and segregation, as well as cell cycle progression. Here, we describe the intricate interplay between NAPs, genome structure and diverse types of genome transactions.

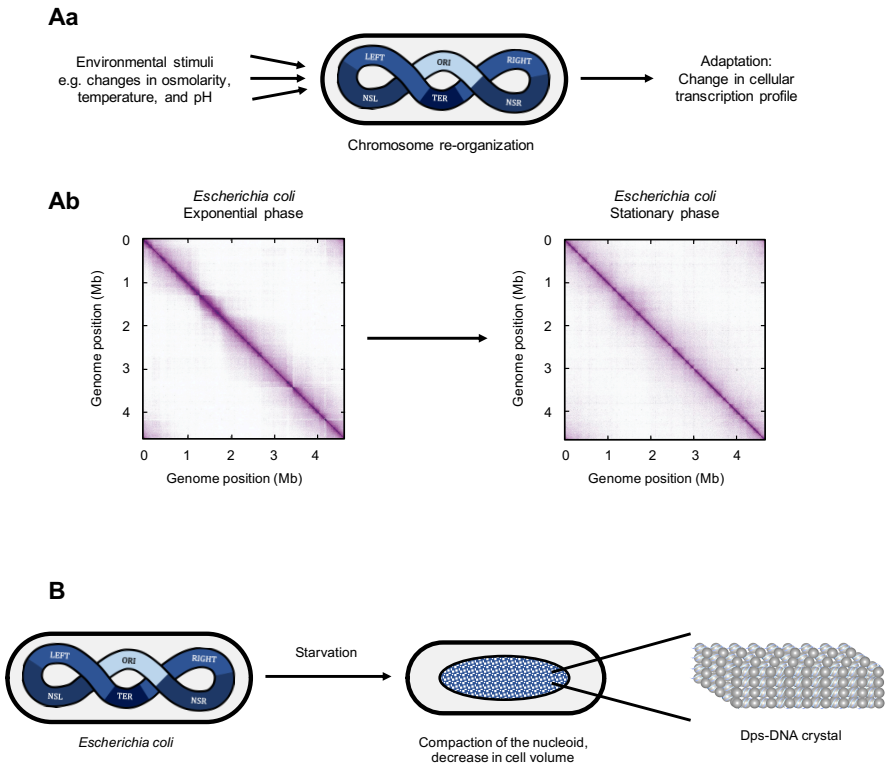


Figure 1.5: Environmental stimuli induce changes in chromosome organization. A: DNA re-organization in growing bacteria. Aa: Re-organization of the bacterial nucleoid is induced in response to environmental stimuli such as changes in osmolarity, temperature and pH. Consequently, activity of specific sets of genes required for environmental adaptation is altered. **Ab:** The transition between the exponential and stationary phases of growth of *Escherichia coli* is associated with a re-

organization of the bacterial chromosome. Specifically, the chromosome exhibits a weakening of compartmentalization into chromosome interaction domains (CIDs). This is observed as 'blurring' of the squares along the main diagonal (101). **B: Chromosome reorganization in starved bacteria.** Reorganization of the bacterial nucleoid can be induced by starvation or stress. Such changes are apparent by light microscopy and indicate compaction of the nucleoid, often accompanied by a reduction in the overall cell volume (145). The inset depicts the nucleoid structure in molecular detail as revealed by electron microscopy. Most notably, the Dps protein (pale blue spheres) drives the formation of an ordered crystal lattice that incorporates DNA (149). These structures are phase separated from other compartments of the cell and resistant to damage (66, 151). NSL, left non-structured region; NSR, right non-structured region; Ori, Ori macrodomain; Ter, Ter macrodomain. *E. coli* contact maps in part a are modified with permission from (101).

Silencing of horizontally acquired genes by H-NS. H-NS targets DNA sequences that have a high AT-content, often acquired by horizontal gene transfer. H-NS binding at these loci represses transcription (known as xenogeneic silencing) (35). Remarkably, the majority of transcription suppressed by H-NS at such loci is spurious in nature (152), arising due to the high probability of sequences that fortuitously resemble promoter elements for RNA polymerase in high AT-content DNA (153, 154). Left unchecked by H-NS, this transcription imposes a severe fitness defect due to titration of RNA polymerase and a global downshift in transcription of housekeeping genes (152).

Canonical gene regulation by H-NS. Although most promoters repressed by H-NS have spurious output, H-NS also plays a key role in regulating transcription of mRNAs. In these instances, the mechanisms by which H-NS influences promoter activity appear diverse. A common mechanism of repression by H-NS involves blocking the binding of RNA polymerase, or transcriptional activator proteins, completely (155, 156). Alternatively, at the *rrnB* P1 and *hdeAB* promoters, H-NS-induced DNA looping traps RNA polymerase, interfering with promoter escape (157, 158). Similarly, but not involving loop formation, a direct contact between RNA polymerase and H-NS can interfere with promoter clearance (159). Because H-NS-controlled looping is mediated by environmental factors, many H-NS-regulated genes are responsive to temperature and osmolarity. For instance, *proVWX* (*proU*) is an H-NS-regulated osmosensitive operon. Its regulation requires two elements, the upstream regulatory element (URE) positioned upstream of the transcription start site (TSS) and the downstream regulatory elements (DRE) in the coding region that extends across the TSS (160–162). The two elements operate synergistically in H-NS-mediated osmoregulation (163). Such synergy could imply lateral or bridge-mediated interactions between the elements. Although direct evidence is lacking, *in vitro* experiments showing that only H-NS-mediated bridging is sensitive to osmolarity lead us to hypothesize that H-NS

represses transcription of the *proU* operon by loop formation, and that relief of repression involves local restructuring of the chromosome (15, 19, 164). In pathogenic bacteria, H-NS can be utilized to control the expression of virulence factors during host colonization with contributions from additional proteins that alter the ability of H-NS to multimerize and/or bind DNA (165).

Regulation of transcription elongation by H-NS. As well as regulating the initiation of transcription, H-NS can control transcription elongation by impeding the progression of RNA polymerase. This depends on the type of H-NS–DNA complex. For example, RNA polymerase can transcribe through lateral H-NS–DNA filaments (Figure 1.6a, top panel), whereas H-NS–DNA bridges efficiently block transcript extension and are likely to trap RNA polymerase in the loops formed (Figure 1.6a, bottom panel) (19). In both cases, it is not known if RNA polymerase advancement removes H-NS from the DNA or if the nucleoprotein complex is transiently remodelled (166).

Roles of bacterial H-NS in regulating transcription versus eukaryotic nucleosomes. In summary regarding the transcriptional roles of H-NS, loci bound by H-NS are often not permissive to binding of RNA polymerase or regulatory proteins but can be remodelled for transcription to occur. By analogy, in eukaryotes, nucleosomes block transcription initiation, and so promoters are usually nucleosome free. Histone modifications lead to remodelling of chromatin that impacts transcription (167). For H-NS, transcription itself could lead to local remodelling of the nucleoprotein complex (166). Furthermore, 'anti-silencing' transcription factors can perturb repressive nucleoprotein filaments or interfere with their formation (31).

Activation and repression of specific promoters by Fis. In general, Fis activates the expression of genes encoding products that are important for rapid cell division (168). Conversely, Fis is often a repressor of genes that allow utilization of alternative carbon sources or terminal electron acceptors (169, 170). Interestingly, the DNA folding activity of Fis appears to be important for counteracting the super-condensation of chromosomes mediated by Dps (150). Taken together with the gene regulatory roles of Fis, this implies the protein is crucial to prepare cells for maximal rates of growth on exiting periods of starvation. This is consistent with observations that Fis is only present at detectable levels when cells are dividing rapidly (146). The ability of Fis to activate or repress transcription is dependent on the position of binding, and interactions with other regulators at a given

promoter. Hence, the mechanisms by which Fis activates and represses transcription are similar to those utilized by canonical transcription factors. For instance, to activate transcription of rRNA operons, Fis facilitates the recruitment of the transcriptional apparatus via a specific contact with the C-terminal domain of the RNA polymerase alpha subunit (alpha-CTD) (171). The same contact is made by many canonical transcriptional activators (172). Similarly, mechanisms of transcription repression by Fis are not unusual and involve occlusion of RNA polymerase or transcription factors (169, 170).

Stabilization of DNA repression loops by HU. Although it is unable to recognize specific DNA sequences, HU can regulate transcription from specific promoters. This results from the ability of HU to bind and stabilize certain deformations in DNA. This behaviour has been described for the *E. coli* galactose operon regulatory region. Two promoters at this locus are repressed by the activity of the repressor protein, GalR. Maximal repression by GalR is mediated by interactions between GalR molecules bound at distal sites to create a repression loop. HU binds at the apex of the DNA loop and stabilizes the complex, thus enhancing repression (173) (Figure 1.6b). As HU affects global patterns of DNA supercoiling, genes responsive to DNA topology are part of the HU regulon (174, 175).

IHF can regulate transcription by bending the DNA. Like Fis, IHF is able to activate and repress transcription by binding to specific sites near promoters. This can be due to the ability of IHF to sharply bend DNA. For instance, IHF binds upstream of many *E. coli* promoters dependent on the alternative σ factor, σ^{54}. σ factors are general transcription factors (functionally similar to those found in eukaryotes) that are used by bacterial RNA polymerases to bind selectively to specific promoters. By bending the DNA, IHF facilitates interactions between RNA polymerase and enhancer proteins bound upstream. This stimulates promoter opening to activate transcription (176) (Figure 1.6c). The binding and bending of DNA by IHF can also repress transcription. In one example, at the *E. coli nrf* promoter, DNA binding by IHF alters interactions with a bound activator to hinder transcription activation (177). The role of IHF as an activator or repressor depends on local nucleoprotein organization. Hence, there is no universal position upstream of a promoter from which IHF consistently exerts an activating versus repressive effect on transcription.

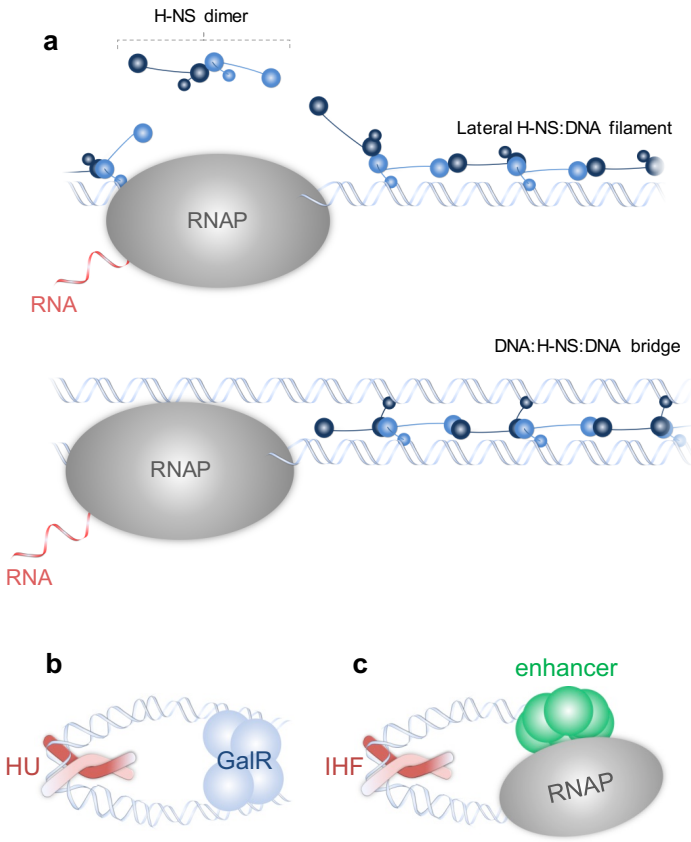


Figure 1.6: Modulation of transcription by nucleoid-associated proteins. a: Alternative types of H-NS:DNA complexes have different effects on transcription elongation. Top panel: lateral H-NS:DNA filaments can be invaded by RNA polymerase (RNAP). These H-NS:DNA complexes are unable to prevent transcription elongation and are either transiently displaced or remodelled as a result (19, 166). **Bottom panel:** bridged DNA:H-NS:DNA complexes are potent blocks to transcription and result in stalled elongation complexes (19); **b: Stabilization of a DNA bend by HU facilitates repression.** At the *Escherichia coli* gal operon the GalR repressor protein forms a repressosome that is stabilized by HU binding to the bent DNA (173); **c: Activation of transcription by DNA bending.** Promoters that are dependent on enhancer-binding proteins for transcription require IHF as a co-factor. The sharp DNA bend introduced by IHF brings the distally bound enhancer-binding protein into the proximity of RNA polymerase so transcription can be activated (176).

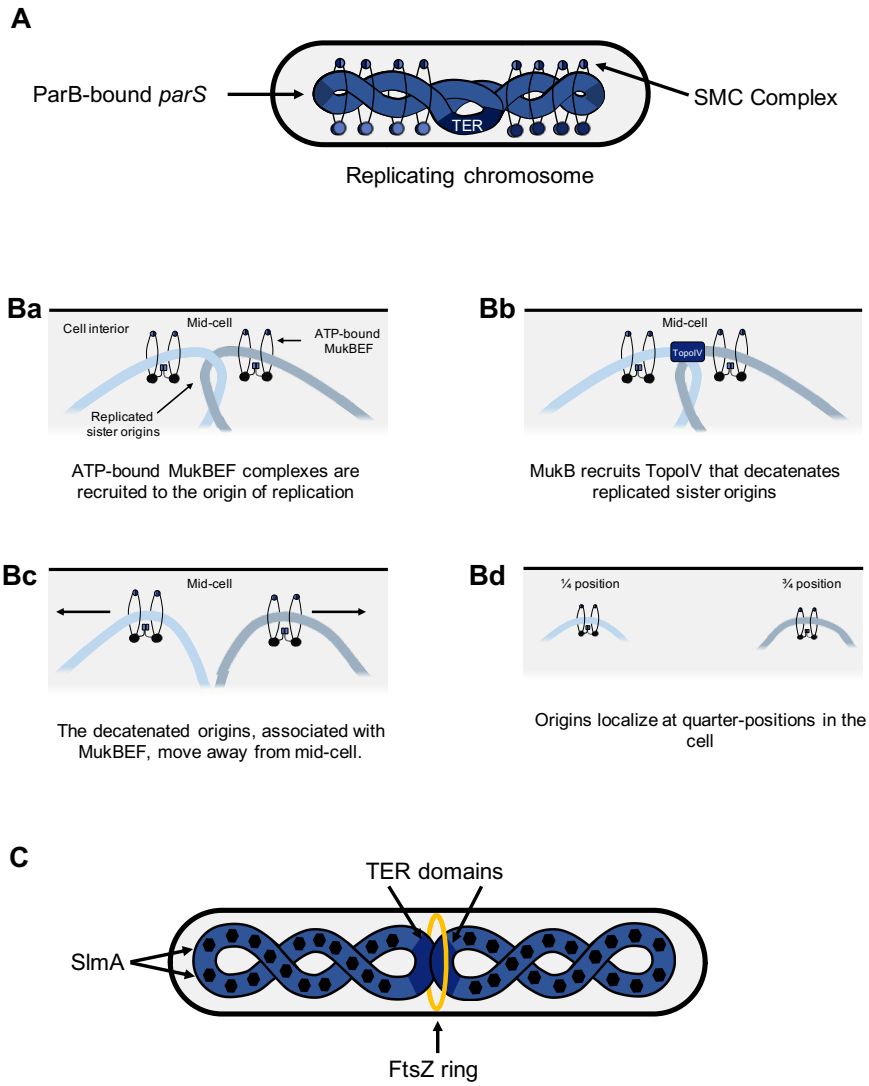


Figure 1.7: Chromosome organization has an impact on chromosome segregation and cell cycle progression. A: ParAB–*parS*-mediated chromosome segregation. Structural maintenance of chromosomes (SMC) complexes are loaded at ParB-bound *parS* sites in the origin domain (51, 55, 57). SMC complexes progressively move along the chromosome (not to scale) towards the terminus (Ter) domain, driving the alignment of chromosome replichores and promoting the segregation of sister chromosomes (56, 62, 96). For ease of representation, the handcuffing model (Figure 1.2, bottom) has not been shown in this figure. **B: MukBEF-mediated origin segregation. Ba:** ATP-bound MukBEF complexes are recruited to the replicated origins (52, 178, 179). **Bb:** MukB recruits TopoIV, a type II topisomerase, that decatenates the entangled sister origins (180, 181). **Bc–d:** Once decatenated, the sister origins, associated with MukBEF, move from the mid-cell position towards the quarter positions in the cell (52, 124). The segregation is proposed to be driven by a self-organizing gradient of MukBEF and the origin of replication (182). **C: Cell cycle progression.** Cell division requires FtsZ ring assembly. FtsZ polymerization occurs at mid-cell where the segregating terminus (Ter) domains are located and SlmA is occluded. DNA-bound SlmA promotes the depolymerization of FtsZ at non-Ter regions preventing 'guillotining' of the chromosome (134, 183–185).

Interplay with replication and chromosome segregation. Chromosome architecture and NAPs also influence chromosome replication and segregation. The bacterial equivalent of the mitotic apparatus, the ParAB–*parS* partitioning system, and SMC proteins that are proposed to regulate origin firing in *B. subtilis* (96), are evidently involved in the segregation of bacterial chromosomes into opposite cell halves concomitantly with replication (186–188) (Figure 1.7A). In *C. crescentus* and *B. subtilis*, the ParB partitioning protein binds to the *parS* sequences present close to OriC to form a nucleoprotein complex on both sister chromosomes (Figure 1.7A). The ParA ATPase is recruited to the complex and generates the free energy required for the resolution of the sister origins and their segregation (189). The loop extruding complex SMC-ScpAB is also recruited to the origin of replication by ParB (51, 55, 57, 190), from where it migrates along the chromosome to the terminus extruding disentangled DNA of a single chromosome. This structurally separates sister chromosomes and favours their segregation (Figure 1.7A) (56, 62, 96, 190).

The ParAB–*parS* partitioning system is absent in some species of the γ- and δ-proteobacteria. In these organisms, the SMC-like MukBEF complex participates in chromosome segregation. In *E. coli*, MukBEF complexes containing an ATP-bound MukB are recruited to the origin (52, 178, 179). MukB, in turn, recruits TopoIV, a type-II topoisomerase, that decatenates replicated sister origins (180, 181). Immediately after decatenation, the segregated origins, and the associated MukBEF clusters, move towards the quarter positions of the cell (Figure 1.7B) (52, 124). Computational modelling suggests that segregation is driven by a self-organizing gradient of MukBEF and ori (182). MukBEF is also recruited to *matS* sites in the Ter macrodomain. At these sites, MatP and ATP hydrolysis by MukB release MukBEF complexes and associated TopoIV enzymes. In Δ*matP* strains, and strains with an ATPase-defective MukB, MukBEF accumulates in the Ter macrodomain where it recruits TopoIV and promotes early resolution of the chromosome terminus (191).

Interplay with cell cycle progression. GapR, a conserved NAP of the α-proteobacteria is a master regulator of cell cycle progression. Its binding sites overlap with loci bound by other regulators of cell cycle progression including CtrA, MucR1, MucR2, and GcrA169. GapR-depleted and Δ*gapR* strains of *C. crescentus* are temperature sensitive and exhibit cell division defects, forming filamentous, undivided cells or anucleate cells (139, 140, 192). GapR binds to the

origin of replication of the *C. crescentus* chromosome where it is involved in the initiation of replication (140). GapR also binds DNA ahead of the replication fork (139) where it interacts with DNA gyrase or TopoIV to relieve positive superhelical stress (193). Indeed, GapR deletion is associated with a lengthened S-phase and stalling of the replication fork (139). Furthermore, GapR binds the *parS* locus at which it plays a role in the segregation of newly replicated sister origins (140). Sister chromosome segregation in *C. crescentus* constitutes a ParA-independent slow step that involves the separation of the pair of *parS*–ParB nucleoprotein complexes, and a ParA-dependent fast step that localizes one of the sister origins to the opposite pole. GapR regulates the initial slow step of segregation, as evidenced by the remerging of resolved *parS*–ParB complexes in Δ*gapR* cells (140).

The NAP SlmA also controls cell cycle progression (183). SlmA bound to SlmA-binding sites (SBSs) on the chromosome are involved in signalling the polymerization of the cytokinetic FtsZ ring (134). As a nucleoid occlusion factor, SlmA also ensures that the FtsZ ring is precisely positioned around the site of Ter decatenation to prevent the 'guillotining' of the chromosome (184). SlmA plays its role by regulating the dynamics of FtsZ polymerization within phase-separated FtsZ droplets. In membrane-bound phase-separated systems, FtsZ polymerizes within phase-separated droplets to form filaments at membrane boundaries. The presence of SBS-bound SlmA counteracts this polymerization (183). This suggests that in FtsZ phase-separated structures in the bacterial cytoplasm, SBS-bound SlmA antagonizes the assembly of the FtsZ ring (183). In *E. coli*, SlmA-binding sites occur throughout the chromosome except at the Ter region (134, 185). This way, an SlmA-free region is produced within the cell when replication reaches the chromosome terminus. The FtsZ ring assembles at this site to initiate bacterial cytokinesis (Figure 1.7C) (183). The precise positioning of the FtsZ ring is reinforced by MatP-mediated condensation of the Ter macrodomain (130). MatP also interacts with the ZapA and ZapB septal proteins to position Ter at mid-cell (133).

Conclusions and perspectives

The past 10 years have seen the establishment of broadly applicable models for the folding of bacterial chromosomes. DNA bending and bridging proteins play a key role in chromosome folding at the level of individual genes and in the formation of CIDs with sizes up to 300 kb. Higher-order chromosome folding leads to the formation of macrodomains. Although these principles have been best studied in *E. coli*, biased binding of proteins across the chromosomes of distantly

related bacteria suggests widespread relevance (138–140). The next challenges in the field of bacterial chromosome biology include better understanding local changes in DNA folding, and how these impact on other nucleic acid transactions within living cells. For example, biophysical techniques have defined the structures that H-NS can form with nucleic acids *in vitro* but it is still not clear whether and how such structures impact transcription *in vivo*. Furthermore, although we understand how individual NAPs organize DNA it is not obvious how the concerted efforts of all NAPs combine within cells.

This thesis addresses these challenges. Chapter 2 describes Hi-C, a proximity ligation-based technique that reveals the contact probability of genomic loci in three-dimensional space. In Chapter 3, RT-qPCR, 3C-qPCR – a modification of the Hi-C technique described in Chapter 2, single molecule Förster resonance energy transfer, and live-cell FRET are used to show that the local three-dimensional folding of the *proVWX* operon of *E. coli* responds to osmotic stress and that this response is associated with a change in the transcriptional profile of the operon. The RT-qPCR studies reported in Chapter 3 also demonstrate how a heteromeric complex of H-NS and StpA oligomerized over the *proV* open reading frame differs in functionality from a homomeric structure comprising of H-NS.

The dynamic organisation of the chromosome necessitates its visualisation in live cells. In Chapter 4, HI-NESS, a novel DNA label designed by the translational fusion of a fluorescent protein to the DNA binding domain of H-NS (H-NS-dbd), is described. The higher DNA dissociation constant of the H-NS-dbd (13) compared to full-length H-NS is exploited in HI-NESS to minimise the perturbation of chromosome structure and genomic transactions that are observed with the use of organic DNA labelling dyes and fluorescent protein fusions to full-length NAPs.

Whilst it is expected that general principles of DNA organization are conserved throughout the bacterial domain of life, and in fact all domains of life (194), we speculate that organisms occupying extreme environmental niches may have fine-tuned the molecular mechanisms to better cope with environmental challenges. In Chapter 5, preliminary chromosome contact profiles of *Haloquadratum walsbyi*, a hyperhalophilic archaeon, are discussed in context of chromosome organisation in model bacteria and archaea.

References:

1. Cairns,J. (1963) The Chromosome of Escherichia coli. *Cold Spring Harb. Symp. Quant. Biol.* , **28**, 43–46.
2. Olins,D.E. and Olins,A.L. (1972) Physical studies of isolated eucaryotic nuclei. *J. Cell Biol.*, 10.1083/jcb.53.3.715.
3. BAKER,J.R. (1955) The Cell-theory: a Restatement, History, and Critique. *Q. J. Microsc. Sci.*, **s3-96**, 449 LP – 481.
4. Dame,R.T., Kalmykowa,O.J. and Grainger,D.C. (2011) Chromosomal macrodomains and associated proteins: Implications for DNA organization and replication in gram negative bacteria. *PLoS Genet.*, 10.1371/journal.pgen.1002123.
5. Goosen,N. and van de Putte,P. (1984) Regulation of Mu transposition. I. Localization of the presumed recognition sites for HimD and Ner functions controlling bacteriophage Mu transcription. *Gene*, 10.1016/0378-1119(84)90103-3.
6. Goosen,N., van Heuvel,M., Moolenaar,G.F. and van de Putte,P. (1984) Regulation of Mu transposition II. The Escherichia coli HimD protein positively controls two repressor promoters and the early promoter of bacteriophage Mu. *Gene*, 10.1016/0378-1119(84)90017-9.
7. Zulianello,L., De Rosny,E.D.L.G., Van Ulsen,P., Van De Putte,P. and Goosen,N. (1994) The HimA and HimD subunits of integration host factor can specifically bind to DNA as homodimers. *EMBO J.*, 10.1002/j.1460-2075.1994.tb06415.x.
8. Nilsson,L., Vanet,A., Vijgenboom,E. and Bosch,L. (1990) The role of FIS in trans activation of stable RNA operons of E. coli. *EMBO J.*, 10.1002/j.1460-2075.1990.tb08166.x.
9. Verbeek,H., Nilsson,L. and Bosch,L. (1991) FIS-induced bending of a region upstream of the promoter activates transcription of the E coli thrU (tufB) operon. *Biochimie*, 10.1016/0300-9084(91)90051-2.
10. Nilsson,L., Verbeek,A., Vijgenboom,E., Van Drunen,C., Vanet,A. and Bosch,L. (1992) FIS-dependent trans activation of stable RNA operons of Escherichia coli under various growth conditions. *J. Bacteriol.*, 10.1128/jb.174.3.921-929.1992.
11. Verbeek,H., Nilsson,L. and Bosch,L. (1992) The mechanism of trans-activation of the escherichia coli operon thru(tufB) by the protein FIS. A model. *Nucleic Acids Res.*, 10.1093/nar/20.15.4077.
12. Arold,S.T., Leonard,P.G., Parkinson,G.N. and Ladbury,J.E. (2010) H-NS forms a superhelical protein scaffold for DNA condensation. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 15728–15732.
13. Dame,R.T., Noom,M.C. and Wuite,G.J.L. (2006) Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature*, 10.1038/nature05283.
14. Dame,R.T., Wyman,C. and Goosen,N. (2000) H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res.*, **28**, 3504–3510.
15. 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*, 10.7554/eLife.27369.
16. Ono,S., Goldberg,M.D., Olsson,T., Esposito,D., Hinton,J.C.D. and Ladbury,J.E. (2005) H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. *Biochem. J.*, **391**, 203–213.
17. Göransson,M., Sondén,B., Nilsson,P., Dagberg,B., Foreman,K., Emanuelsson,K. and Uhlin,B.E. (1990) Transcriptional silencing and thermoregulation of gene expression in Escherichia coli. *Nature*, 10.1038/344682a0.
18. 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.*, 10.1016/S0006-3495(03)75051-6.
19. 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*, 10.7554/eLife.04970.
20. Nolivos,S. and Sherratt,D. (2014) The bacterial chromosome: Architecture and action of bacterial SMC and SMC-like complexes. *FEMS Microbiol. Rev.*, 10.1111/1574-6976.12045.
21. Palecek,J.J. and Gruber,S. (2015) Kite Proteins: A Superfamily of SMC/Kleisin Partners Conserved Across Bacteria, Archaea, and Eukaryotes. *Structure*, 10.1016/j.str.2015.10.004.
22. Schleiffer,A., Kaitna,S., Maurer-Stroh,S., Glotzer,M., Nasmyth,K. and Eisenhaber,F. (2003) Kleisins: A superfamily of bacterial and eukaryotic SMC protein partners. *Mol. Cell*,

10.1016/S1097-2765(03)00108-4.
23. Wells,J.N., Gligoris,T.G., Nasmyth,K.A. and Marsh,J.A. (2017) Evolution of condensin and cohesin complexes driven by replacement of Kite by Hawk proteins. *Curr. Biol.*, **27**, R17–R18.
24. Melby,T.E., Ciampaglio,C.N., Briscoe,G. and Erickson,H.P. (1998) The symmetrical structure of structural maintenance of chromosomes (SMC) and MukB proteins: Long, antiparallel coiled coils, folded at a flexible hinge. *J. Cell Biol.*, **142**, 1595–1604.
25. Cuylen,S., Metz,J. and Haering,C.H. (2011) Condensin structures chromosomal DNA through topological links. *Nat. Struct. Mol. Biol.*, 10.1038/nsmb.2087.
26. Stella,S., Cascio,D. and Johnson,R.C. (2010) The shape of the DNA minor groove directs binding by the DNA-bending protein Fis. *Genes Dev.*, **24**, 814–826.
27. Hancock,S.P., Stella,S., Cascio,D. and Johnson,R.C. (2016) DNA sequence determinants controlling affinity, stability and shape of DNA complexes bound by the nucleoid protein Fis. *PLoS One*, **11**.
28. Rice,P.A., Yang,S.W., Mizuuchi,K. and Nash,H.A. (1996) Crystal structure of an IHF-DNA complex: A protein-induced DNA U-turn. *Cell*, 10.1016/S0092-8674(00)81824-3.
29. Swinger,K.K., Lemberg,K.M., Zhang,Y. and Rice,P.A. (2003) Flexible DNA bending in HU-DNA cocrystal structures. *EMBO J.*, 10.1093/emboj/cdg351.
30. Van Noort,J., Verbrugge,S., Goosen,N., Dekker,C. and Dame,R.T. (2004) Dual architectural roles of HU: Formation of flexible hinges and rigid filaments. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 6969–6974.
31. Grainger,D.C. (2016) Structure and function of bacterial H-NS protein. *Biochem. Soc. Trans.*, 10.1042/BST20160190.
32. Gordon,B.R.G., 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. *Proc. Natl. Acad. Sci. U. S. A.*, 10.1073/pnas.1102544108.
33. Grainger,D.C., Hurd,D., Goldberg,M.D. and Busby,S.J.W. (2006) Association of nucleoid proteins with coding and non-coding segments of the Escherichia coli genome. *Nucleic Acids Res.*, 10.1093/nar/gkl542.
34. Oshima,T., Ishikawa,S., Kurokawa,K., Aiba,H. and Ogasawara,N. (2006) Escherichia coli histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Res.*, **13**, 141–153.
35. 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 (80-.).*, 10.1126/science.1128794.
36. Lucchini,S., Rowley,G., Goldberg,M.D., Hurd,D., Harrison,M. and Hinton,J.C.D. (2006) H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog.*, 10.1371/journal.ppat.0020081.
37. Kahramanoglou,C., Seshasayee,A.S.N., 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 Res.*, **39**, 2073–2091.
38. Gordon,B.R.G., Li,Y., Wang,L., Sintsova,A., Van Bakel,H., Tian,S., Navarre,W.W., Xia,B. and Liu,J. (2010) Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in Mycobacterium tuberculosis. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 5154–5159.
39. Smits,W.K. and Grossman,A.D. (2010) The transcriptional regulator Rok binds A+T-rich DNA and is involved in repression of a mobile genetic element in Bacillus subtilis. *PLoS Genet.*, 10.1371/journal.pgen.1001207.
40. Qin,L., Erkelens,A.M., Ben Bdira,F. and Dame,R.T. (2019) The architects of bacterial DNA bridges: A structurally and functionally conserved family of proteins. *Open Biol.*, **9**.
41. Duan,B., Ding,P., Hughes,T.R., Navarre,W.W., Liu,J. and Xia,B. (2018) How bacterial xenogeneic silencer rok distinguishes foreign from self DNA in its resident genome. *Nucleic Acids Res.*, 10.1093/nar/gky836.
42. Ding,P., McFarland,K.A., Jin,S., Tong,G., Duan,B., Yang,A., Hughes,T.R., Liu,J., Dove,S.L., Navarre,W.W., *et al.* (2015) A Novel AT-Rich DNA Recognition Mechanism for Bacterial Xenogeneic Silencer MvaT. *PLoS Pathog.*, 10.1371/journal.ppat.1004967.
43. Gruber,S., Haering,C.H. and Nasmyth,K. (2003) Chromosomal cohesin forms a ring. *Cell*, **112**, 765–777.
44. Zhang,N., Kuznetsov,S.G., Sharan,S.K., Li,K., Rao,P.H. and Pati,D. (2008) A handcuff model for

the cohesin complex. *J. Cell Biol.*, **183**, 1019–1031.
45. Minnen,A., Attaiech,L., Thon,M., Gruber,S. and Veening,J.W. (2011) SMC is recruited to oriC by ParB and promotes chromosome segregation in Streptococcus pneumoniae. *Mol. Microbiol.*, **81**, 676–688.
46. Schwartz,M.A. and Shapiro,L. (2011) An SMC ATPase mutant disrupts chromosome segregation in Caulobacter. *Mol. Microbiol.*, 10.1111/j.1365-2958.2011.07836.x.
47. Hiraga,S., Niki,H., Ogura,T., Ichinose,C., Mori,H., Ezaki,B. and Jaffe,A. (1989) Chromosome partitioning in Escherichia coli: Novel mutants producing anucleate cells. *J. Bacteriol.*, **171**, 1496–1505.
48. Niki,H., Jaffe,A., Imamura,R., Ogura,T. and Hiraga,S. (1991) The new gene mukB codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of E. coli. *EMBO J.*, **10**, 183–193.
49. Jensen,R.B. and Shapiro,L. (1999) The Caulobacter crescentus smc gene is required for cell cycle progression and chromosome segregation. *Proc. Natl. Acad. Sci. U. S. A.*, 10.1073/pnas.96.19.10661.
50. Moriya,S., Tsujikawa,E., Hassan,A.K.M., Asai,K., Kodama,T. and Ogasawara,N. (1998) A Bacillus subtilis gene-encoding protein homologous to eukaryotic SMC motor protein is necessary for chromosome partition. *Mol. Microbiol.*, 10.1046/j.1365-2958.1998.00919.x.
51. Wang,X., Tang,O.W., Riley,E.P. and Rudner,D.Z. (2014) The SMC condensin complex is required for origin segregation in Bacillus subtilis. *Curr. Biol.*, 10.1016/j.cub.2013.11.050.
52. Danilova,O., Reyes-Lamothe,R., Pinskaya,M., Sherratt,D. and Possoz,C. (2007) MukB colocalizes with the oriC region and is required for organization of the two Escherichia coli chromosome arms into separate cell halves. *Mol. Microbiol.*, **65**, 1485–1492.
53. Petrushenko,Z.M., She,W. and Rybenkov,V. V. (2011) A new family of bacterial condensins. *Mol. Microbiol.*, **81**, 881–896.
54. Yu,W., Herbert,S., Graumann,P.L. and Götz,F. (2010) Contribution of SMC (Structural Maintenance of Chromosomes) and spoIIIE to chromosome segregation in staphylococci. *J. Bacteriol.*, **192**, 4067–4073.
55. Sullivan,N.L., Marquis,K.A. and Rudner,D.Z. (2009) Recruitment of SMC by ParB-parS Organizes the Origin Region and Promotes Efficient Chromosome Segregation. *Cell*, 10.1016/j.cell.2009.04.044.
56. Tran,N.T., Laub,M.T. and Le,T.B.K. (2017) SMC Progressively Aligns Chromosomal Arms in Caulobacter crescentus but Is Antagonized by Convergent Transcription. *Cell Rep.*, 10.1016/j.celrep.2017.08.026.
57. Gruber,S. and Errington,J. (2009) Recruitment of Condensin to Replication Origin Regions by ParB/SpoOJ Promotes Chromosome Segregation in B. subtilis. *Cell*, 10.1016/j.cell.2009.02.035.
58. Ganji,M., Shaltiel,I.A., Bisht,S., Kim,E., Kalichava,A., Haering,C.H. and Dekker,C. (2018) Real-time imaging of DNA loop extrusion by condensin. *Science (80-.).*, 10.1126/science.aar7831.
59. Kim,E., Kerssemakers,J., Shaltiel,I.A., Haering,C.H. and Dekker,C. (2020) DNA-loop extruding condensin complexes can traverse one another. *Nature*, 10.1038/s41586-020-2067-5.
60. Nasmyth,K. (2001) Disseminating the genome: Joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annu. Rev. Genet.*, **35**, 673–745.
61. Alipour,E. and Marko,J.F. (2012) Self-organization of domain structures by DNA-loop-extruding enzymes. *Nucleic Acids Res.*, **40**, 11202–11212.
62. Wang,X., Brandão,H.B., Le,T.B.K., Laub,M.T. and Rudner,D.Z. (2017) Bacillus subtilis SMC complexes juxtapose chromosome arms as they travel from origin to terminus. *Science (80-.).*, 10.1126/science.aai8982.
63. Hirano,T. (2002) The ABCs of SMC proteins: Two-armed ATPases for chromosome condensation, cohesion, and repair. *Genes Dev.*, 10.1101/gad.955102.
64. Cairns,J. (1963) The bacterial chromosome and its manner of replication as seen by autoradiography. *J. Mol. Biol.*, 10.1016/S0022-2836(63)80070-4.
65. Schneider,R., Lurz,R., Lüder,G., Tolksdorf,C., Travers,A. and Muskhelishvili,G. (2001) An architectural role of the Escherichia coli chromatin protein FIS in organising DNA. *Nucleic Acids Res.*, **29**, 5107–5114.
66. Cosgriff,S., Chintakayala,K., Chim,Y.T.A., Chen,X., Allen,S., Lovering,A.L. and Grainger,D.C. (2010) Dimerization and DNA-dependent aggregation of the Escherichia coli nucleoid protein

and chaperone CbpA. *Mol. Microbiol.*, **77**, 1289–1300.
67. Hales,L.M., Gumport,R.I. and Gardner,J.F. (1994) Determining the DNA sequence elements required for binding integration host factor to two different target sites. *J. Bacteriol.*, 10.1128/jb.176.10.2999-3006.1994.
68. Rouviere Yaniv,J. and Gros,F. (1975) Characterization of a novel, low molecular weight DNA binding protein from Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.*, **72**, 3428–3432.
69. Grove,A. (2011) Functional evolution of bacterial histone-like HU proteins. *Curr. Issues Mol. Biol.*, **13**, 1–12.
70. Swinger,K.K. and Rice,P.A. (2007) Structure-based Analysis of HU-DNA Binding. *J. Mol. Biol.*, **365**, 1005–1016.
71. Swinger,K.K. and Rice,P.A. (2004) IHF and HU: Flexible architects of bent DNA. *Curr. Opin. Struct. Biol.*, **14**, 28–35.
72. Bensaid,A., Almeida,A., Drlica,K. and Rouviere-Yaniv,J. (1996) Cross-talk between topoisomerase I and HU in Escherichia coli. *J. Mol. Biol.*, 10.1006/jmbi.1996.0086.
73. Ghosh,S., Mallick,B. and Nagaraja,V. (2014) Direct regulation of topoisomerase activity by a nucleoid-associated protein. *Nucleic Acids Res.*, **42**, 11156–11165.
74. Guo,F. and Adhya,S. (2007) Spiral structure of Escherichia coli HUαβ provides foundation for DNA supercoiling. *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 4309–4314.
75. Witz,G. and Stasiak,A. (2009) DNA supercoiling and its role in DNA decatenation and unknotting. *Nucleic Acids Res.*, **38**, 2119–2133.
76. Luijsterburg,M.S., Noom,M.C., Wuite,G.J.L. and Dame,R.T. (2006) The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: A molecular perspective. *J. Struct. Biol.*, 10.1016/j.jsb.2006.05.006.
77. Japaridze,A., Muskhelishvili,G., Benedetti,F., Gavriilidou,A.F.M., Zenobi,R., De Los Rios,P., Longo,G. and Dietler,G. (2017) Hyperplectonemes: A Higher Order Compact and Dynamic DNA Self-Organization. *Nano Lett.*, 10.1021/acs.nanolett.6b05294.
78. Japaridze,A., Yang,W., Dekker,C., Nasser,W. and Muskhelishvili,G. (2020) DNA sequence-directed cooperation between nucleoid-associated proteins. *bioRxiv*.
79. 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 Res.*, 10.1093/nar/gky265.
80. Johansson,J. and Uhlin,B.E. (1999) Differential protease-mediated turnover of H-NS and StpA revealed by a mutation altering protein stability and stationary-phase survival of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.*, 10.1073/pnas.96.19.10776.
81. Johansson,J., Eriksson,S., Sondén,B., Sun Nyunt Wai and Uhlin,B.E. (2001) Heteromeric interactions among nucleoid-associated bacterial proteins: Localization of StpA-stabilizing regions in H-NS of Escherichia coli. *J. Bacteriol.*, 10.1128/JB.183.7.2343-2347.2001.
82. Uyar,E., Kurokawa,K., Yoshimura,M., Ishikawa,S., Ogasawara,N. and Oshima,T. (2009) Differential binding profiles of StpA in wild-type and hns mutant cells: A comparative analysis of cooperative partners by chromatin immunoprecipitation- microarray analysis. *J. Bacteriol.*, 10.1128/JB.01594-08.
83. Madrid,C., Balsalobre,C., García,J. and Juárez,A. (2007) The novel Hha/YmoA family of nucleoid-associated proteins: Use of structural mimicry to modulate the activity of the H-NS family of proteins. *Mol. Microbiol.*, **63**, 7–14.
84. 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.*, 10.1074/jbc.M113.455378.
85. Aravind,L. and Landsman,D. (1998) AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Res.*, **26**, 4413–4421.
86. Galande,S., Purbey,P.K., Notani,D. and Kumar,P.P. (2007) The third dimension of gene regulation: organization of dynamic chromatin loopscape by SATB1. *Curr. Opin. Genet. Dev.*, 10.1016/j.gde.2007.08.003.
87. Naik,R. and Galande,S. (2019) SATB family chromatin organizers as master regulators of tumor progression. *Oncogene*, 10.1038/s41388-018-0541-4.
88. Yasui,D., Miyano,M., Cai,S., Varga-Weisz,P. and Kohwi-Shigematsu,T. (2002) SATB1 targets chromatin remodelling to regulate genes over long distances. *Nature*, **419**, 641–645.
89. Mattiroli,F., Bhattacharyya,S., Dyer,P.N., White,A.E., Sandman,K., Burkhart,B.W., Byrne,K.R.,

Lee,T., Ahn,N.G., Santangelo,T.J., *et al.* (2017) Structure of histone-based chromatin in Archaea. *Science (80-.).*, 10.1126/science.aaj1849.
90. Henneman,B., van Emmerik,C., van Ingen,H. and Dame,R.T. (2018) Structure and function of archaeal histones. *PLoS Genet.*, **14**.
91. Kuhn,M.L., Zemaitaitis,B., Hu,L.I., Sahu,A., Sorensen,D., Minasov,G., Lima,B.P., Scholle,M., Mrksich,M., Anderson,W.F., *et al.* (2014) Structural, kinetic and proteomic characterization of acetyl phosphate-dependent bacterial protein acetylation. *PLoS One*, **9**.
92. Schmidt,A., Kochanowski,K., Vedelaar,S., Ahrné,E., Volkmer,B., Callipo,L., Knoops,K., Bauer,M., Aebersold,R. and Heinemann,M. (2016) The quantitative and condition-dependent Escherichia coli proteome. *Nat. Biotechnol.*, 10.1038/nbt.3418.
93. Weinert,B.T., Iesmantavicius,V., Wagner,S.A., Schölz,C., Gummesson,B., Beli,P., Nyström,T. and Choudhary,C. (2013) Acetyl-Phosphate is a critical determinant of Lysine Acetylation in E.coli. *Mol. Cell*, 10.1016/j.molcel.2013.06.003.
94. Dilweg,I.W. and Dame,R.T. (2018) Post-translational modification of nucleoidassociated proteins: An extra layer of functional modulation in bacteria? *Biochem. Soc. Trans.*, **46**, 1381–1392.
95. Le,T.B.K., Imakaev,M. V., Mirny,L.A. and Laub,M.T. (2013) High-resolution mapping of the spatial organization of a bacterial chromosome. *Science (80-.).*, 10.1126/science.1242059.
96. Marbouty,M., Le Gall,A., Cattoni,D.I., Cournac,A., Koh,A., Fiche,J.B., Mozziconacci,J., Murray,H., Koszul,R. and Nollmann,M. (2015) Condensin- and Replication-Mediated Bacterial Chromosome Folding and Origin Condensation Revealed by Hi-C and Super-resolution Imaging. *Mol. Cell*, 10.1016/j.molcel.2015.07.020.
97. Sexton,T., Yaffe,E., Kenigsberg,E., Bantignies,F., Leblanc,B., Hoichman,M., Parrinello,H., Tanay,A. and Cavalli,G. (2012) Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell*, 10.1016/j.cell.2012.01.010.
98. Nora,E.P., Lajoie,B.R., Schulz,E.G., Giorgetti,L., Okamoto,I., Servant,N., Piolot,T., Van Berkum,N.L., Meisig,J., Sedat,J., *et al.* (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, 10.1038/nature11049.
99. Dixon,J.R., Selvaraj,S., Yue,F., Kim,A., Li,Y., Shen,Y., Hu,M., Liu,J.S. and Ren,B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 10.1038/nature11082.
100. Le,T.B. and Laub,M.T. (2016) Transcription rate and transcript length drive formation of chromosomal interaction domain boundaries. *EMBO J.*, 10.15252/embj.201593561.
101. Lioy,V.S., Cournac,A., Marbouty,M., Duigou,S., Mozziconacci,J., Espéli,O., Boccard,F. and Koszul,R. (2018) Multiscale Structuring of the E. coli Chromosome by Nucleoid-Associated and Condensin Proteins. *Cell*, 10.1016/j.cell.2017.12.027.
102. Woldringh,C.L. (2002) The role of co-transcriptional translation and protein translocation (transertion) in bacterial chromosome segregation. *Mol. Microbiol.*, 10.1046/j.1365-2958.2002.02993.x.
103. Fudenberg,G., Imakaev,M., Lu,C., Goloborodko,A., Abdennur,N. and Mirny,L.A. (2016) Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.*, **15**, 2038–2049.
104. Sanborn,A.L., Rao,S.S.P., Huang,S.C., Durand,N.C., Huntley,M.H., Jewett,A.I., Bochkov,I.D., Chinnappan,D., Cutkosky,A., Li,J., *et al.* (2015) Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. U. S. A.*, 10.1073/pnas.1518552112.
105. Rao,S.S.P., Huntley,M.H., Durand,N.C., Stamenova,E.K., Bochkov,I.D., Robinson,J.T., Sanborn,A.L., Machol,I., Omer,A.D., Lander,E.S., *et al.* (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, **159**, 1665–1680.
106. Rowley,M.J., Nichols,M.H., Lyu,X., Ando-Kuri,M., Rivera,I.S.M., Hermetz,K., Wang,P., Ruan,Y. and Corces,V.G. (2017) Evolutionarily Conserved Principles Predict 3D Chromatin Organization. *Mol. Cell*, **67**, 837-852.e7.
107. Vietri Rudan,M., Barrington,C., Henderson,S., Ernst,C., Odom,D.T., Tanay,A. and Hadjur,S. (2015) Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture. *Cell Rep.*, 10.1016/j.celrep.2015.02.004.
108. Cubenãs-Potts,C., Rowley,M.J., Lyu,X., Li,G., Lei,E.P. and Corces,V.G. (2017) Different enhancer classes in Drosophilabind distinct architectural proteins and mediate unique chromatin interactions and 3D architecture. *Nucleic Acids Res.*, **45**, 1714–1730.

109. Wang,Q., Sun,Q., Czajkowsky,D.M. and Shao,Z. (2018) Sub-kb Hi-C in D. melanogaster reveals conserved characteristics of TADs between insect and mammalian cells. *Nat. Commun.*, 10.1038/s41467-017-02526-9.
110. Ulianov,S. V., Khrameeva,E.E., Gavrilov,A.A., Flyamer,I.M., Kos,P., Mikhaleva,E.A., Penin,A.A., Logacheva,M.D., Imakaev,M. V., Chertovich,A., *et al.* (2016) Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. *Genome Res.*, **26**, 70–84.
111. Hsieh,T.H.S., Weiner,A., Lajoie,B., Dekker,J., Friedman,N. and Rando,O.J. (2015) Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. *Cell*, **162**, 108–119.
112. Eme,L., Spang,A., Lombard,J., Stairs,C.W. and Ettema,T.J.G. (2017) Archaea and the origin of eukaryotes. *Nat. Rev. Microbiol.*, 10.1038/nrmicro.2017.133.
113. Takemata,N., Samson,R.Y. and Bell,S.D. (2019) Physical and Functional Compartmentalization of Archaeal Chromosomes. *Cell*, 10.1016/j.cell.2019.08.036.
114. Takemata,N. and Bell,S.D. (2021) Multi-scale architecture of archaeal chromosomes. *Mol. Cell*, 10.1016/j.molcel.2020.12.001.
115. Cockram,C., Thierry,A., Gorlas,A., Lestini,R. and Koszul,R. (2021) Euryarchaeal genomes are folded into SMC-dependent loops and domains, but lack transcription-mediated compartmentalization. *Mol. Cell*, 10.1016/j.molcel.2020.12.013.
116. Umbarger,M.A., Toro,E., Wright,M.A., Porreca,G.J., Baù,D., Hong,S.H., Fero,M.J., Zhu,L.J., Marti-Renom,M.A., McAdams,H.H., *et al.* (2011) The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation. *Mol. Cell*, 10.1016/j.molcel.2011.09.010.
117. Lal,A., Dhar,A., Trostel,A., Kouzine,F., Seshasayee,A.S.N. and Adhya,S. (2016) Genome scale patterns of supercoiling in a bacterial chromosome. *Nat. Commun.*, 10.1038/ncomms11055.
118. Bermúdez,I., García-Martínez,J., Pérez-Ortín,J.E. and Roca,J. (2010) A method for genome-wide analysis of DNA helical tension by means of psoralen-DNA photobinding. *Nucleic Acids Res.*, 10.1093/nar/gkq687.
119. Azam,T.A., Iwata,A., Nishimura,A., Ueda,S. and Ishihama,A. (1999) Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *J. Bacteriol.*, **181**, 6361–6370.
120. Rovinskiy,N., Agbleke,A.A., Chesnokova,O., Pang,Z. and Higgins,N.P. (2012) Rates of Gyrase Supercoiling and Transcription Elongation Control Supercoil Density in a Bacterial Chromosome. *PLoS Genet.*, 10.1371/journal.pgen.1002845.
121. Uhlmann,F. (2016) SMC complexes: From DNA to chromosomes. *Nat. Rev. Mol. Cell Biol.*, 10.1038/nrm.2016.30.
122. Fennell-Fezzie,R., Gradia,S.D., Akey,D. and Berger,J.M. (2005) The MukF subunit of Escherichia coli condensin: Architecture and functional relationship to kleisins. *EMBO J.*, 10.1038/sj.emboj.7600680.
123. Woo,J.S., Lim,J.H., Shin,H.C., Suh,M.K., Ku,B., Lee,K.H., Joo,K., Robinson,H., Lee,J., Park,S.Y., *et al.* (2009) Structural Studies of a Bacterial Condensin Complex Reveal ATP-Dependent Disruption of Intersubunit Interactions. *Cell*, 10.1016/j.cell.2008.10.050.
124. Badrinarayanan,A., Reyes-Lamothe,R., Uphoff,S., Leake,M.C. and Sherratt,D.J. (2012) In vivo architecture and action of bacterial structural maintenance of chromosome proteins. *Science (80-.).*, 10.1126/science.1227126.
125. Chen,N., Zinchenko,A.A., Yoshikawa,Y., Araki,S., Adachi,S., Yamazoe,M., Hiraga,S. and Yoshikawa,K. (2008) ATP-induced shrinkage of DNA with MukB protein and the MukBEF complex of escherichia coli. *J. Bacteriol.*, **190**, 3731–3737.
126. Zawadzka,K., Zawadzki,P., Baker,R., Rajasekar,K. V., Wagner,F., Sherratt,D.J. and Arciszewska,L.K. (2018) MukB ATPases are regulated independently by the N-and C-terminal domains of MukF kleisin. *Elife*, 10.7554/eLife.31522.
127. Espeli,O., Mercier,R. and Boccard,F. (2008) DNA dynamics vary according to macrodomain topography in the E. coli chromosome. *Mol. Microbiol.*, **68**, 1418–1427.
128. Valens,M., Thiel,A. and Boccard,F. (2016) The MaoP/maoS Site-Specific System Organizes the Ori Region of the E. coli Chromosome into a Macrodomain. *PLoS Genet.*, **12**.
129. Duigou,S. and Boccard,F. (2017) Long range chromosome organization in Escherichia coli: The position of the replication origin defines the non-structured regions and the Right and Left macrodomains. *PLoS Genet.*, 10.1371/journal.pgen.1006758.

130. Mercier,R., Petit,M.A., Schbath,S., Robin,S., El Karoui,M., Boccard,F. and Espéli,O. (2008) The MatP/matS Site-Specific System Organizes the Terminus Region of the E. coli Chromosome into a Macrodomain. *Cell*, 10.1016/j.cell.2008.08.031.
131. Dupaigne,P., Tonthat,N.K., Espéli,O., Whitfill,T., Boccard,F. and Schumacher,M.A. (2012) Molecular Basis for a Protein-Mediated DNA-Bridging Mechanism that Functions in Condensation of the E. coli Chromosome. *Mol. Cell*, 10.1016/j.molcel.2012.09.009.
132. Thiel,A., Valens,M., Vallet-Gely,I., Espéli,O. and Boccard,F. (2012) Long-range chromosome organization in E. coli: A site-specific system isolates the ter macrodomain. *PLoS Genet.*, 10.1371/journal.pgen.1002672.
133. Espéli,O., Borne,R., Dupaigne,P., Thiel,A., Gigant,E., Mercier,R. and Boccard,F. (2012) A MatP-divisome interaction coordinates chromosome segregation with cell division in E. coli. *EMBO J.*, **31**, 3198–3211.
134. Cho,H., McManus,H.R., Dove,S.L. and Bernhardt,T.G. (2011) Nucleoid occlusion factor SlmA is a DNA-activated FtsZ polymerization antagonist. *Proc. Natl. Acad. Sci. U. S. A.*, 10.1073/pnas.1018674108.
135. Sánchez-Romero,M.A., Busby,S.J.W., Dyer,N.P., Ott,S., Millard,A.D. and Grainger,D.C. (2010) Dynamic distribution of SeqA protein across the chromosome of Escherichia coli K-12. *MBio*, **1**.
136. Lu,M., Campbell,J.L., Boye,E. and Kleckner,N. (1994) SeqA: A negative modulator of replication initiation in E. coli. *Cell*, **77**, 413–426.
137. Slater,S., Wold,S., Lu,M., Boye,E., Skarstad,K. and Kleckner,N. (1995) E. coli SeqA protein binds oriC in two different methyl-modulated reactions appropriate to its roles in DNA replication initiation and origin sequestration. *Cell*, 10.1016/0092-8674(95)90272-4.
138. Wu,L.J., Ishikawa,S., Kawai,Y., Oshima,T., Ogasawara,N. and Errington,J. (2009) Noc protein binds to specific DNA sequences to coordinate cell division with chromosome segregation. *EMBO J.*, **28**, 1940–1952.
139. Arias-Cartin,R., Dobihal,G.S., Campos,M., Surovtsev,I. V, Parry,B. and Jacobs-Wagner,C. (2017) Replication fork passage drives asymmetric dynamics of a critical nucleoid-associated protein in Caulobacter . *EMBO J.*, 10.15252/embj.201695513.
140. Taylor,J.A., Panis,G., Viollier,P.H. and Marczynski,G.T. (2017) A novel nucleoid-associated protein coordinates chromosome replication and chromosome partition. *Nucleic Acids Res.*, 10.1093/nar/gkx596.
141. Qin,L., Bdira,F. Ben, Sterckx,Y.G.J., Volkov,A.N., Vreede,J., Giachin,G., Van Schaik,P., Ubbink,M. and Dame,R.T. (2020) Structural basis for osmotic regulation of the DNA binding properties of H-NS proteins. *Nucleic Acids Res.*, **48**, 2156–2172.
142. Hameed,U.F.S., 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 Res.*, **47**, 2666–2680.
143. Hołówka,J. and Zakrzewska-Czerwińska,J. (2020) Nucleoid Associated Proteins: The Small Organizers That Help to Cope With Stress. *Front. Microbiol.*, 10.3389/fmicb.2020.00590.
144. Dame,R.T. (2005) The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol. Microbiol.*, 10.1111/j.1365-2958.2005.04598.x.
145. Meyer,A.S. and Grainger,D.C. (2013) The Escherichia coli Nucleoid in Stationary Phase. In *Advances in Applied Microbiology*.
146. Ball,C.A., Osuna,R., Ferguson,K.C. and Johnson,R.C. (1992) Dramatic changes in Fis levels upon nutrient upshift in Escherichia coli. *J. Bacteriol.*, **174**, 8043–8056.
147. Almiron,M., Link,A.J., Furlong,D. and Kolter,R. (1992) A novel DNA-binding protein with regulatory and protective roles in starved Escherichia coli. *Genes Dev.*, 10.1101/gad.6.12b.2646.
148. Karas,V.O., Westerlaken,I. and Meyer,A.S. (2015) The DNA-binding protein from starved cells (Dps) utilizes dual functions to defend cells against multiple stresses. *J. Bacteriol.*, **197**, 3206–3215.
149. Wolf,S.G., Frenkiel,D., Arad,T., Finkeil,S.E., Kolter,R. and Minsky,A. (1999) DNA protection by stress-induced biocrystallization. *Nature*, **400**, 83–85.
150. Ohniwa,R.L., Morikawa,K., Kim,J., Ohta,T., Ishihama,A., Wada,C. and Takeyasu,K. (2006) Dynamic state of DNA topology is essential for genome condensation in bacteria. *EMBO J.*, **25**, 5591–5602.

151. Janissen,R., Arens,M.M.A., Vtyurina,N.N., Rivai,Z., Sunday,N.D., Eslami-Mossallam,B., Gritsenko,A.A., Laan,L., de Ridder,D., Artsimovitch,I., *et al.* (2018) Global DNA Compaction in Stationary-Phase Bacteria Does Not Affect Transcription. *Cell*, **174**, 1188-1199.e14.
152. Lamberte,L.E., Baniulyte,G., Singh,S.S., Stringer,A.M., Bonocora,R.P., Stracy,M., Kapanidis,A.N., Wade,J.T. and Grainger,D.C. (2017) Horizontally acquired AT-rich genes in Escherichia coli cause toxicity by sequestering RNA polymerase. *Nat. Microbiol.*, 10.1038/nmicrobiol.2016.249.
153. Singh,S.S. and Grainger,D.C. (2013) H-NS Can Facilitate Specific DNA-binding by RNA Polymerase in AT-rich Gene Regulatory Regions. *PLoS Genet.*, **9**.
154. 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 Dev.*, 10.1101/gad.234336.113.
155. Myers,K.S., Yan,H., Ong,I.M., Chung,D., Liang,K., Tran,F., Keleş,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**.
156. Haycocks,J.R.J., Sharma,P., Stringer,A.M., Wade,J.T. and Grainger,D.C. (2015) The Molecular Basis for Control of ETEC Enterotoxin Expression in Response to Environment and Host. *PLoS Pathog.*, 10.1371/journal.ppat.1004605.
157. Shin,M., Song,M., Joon,H.R., Hong,Y., Kim,Y.J., Seok,Y.J., Ha,K.S., Jung,S.H. and Choy,H.E. (2005) DNA looping-mediated repression by histone-like protein H-NS: Specific requirement of Eσ70 as a cofactor for looping. *Genes Dev.*, 10.1101/gad.1316305.
158. Dame,R.T., Wyman,C., Wurm,R., Wagner,R. and Goosen,N. (2002) Structural basis for H-NS-mediated trapping of RNA polymerase in the open initiation complex at the rrnB P1. *J. Biol. Chem.*, **277**, 2146–2150.
159. Shin,M., Lagda,A.C., Lee,J.W., Bhat,A., Rhee,J.H., Kim,J.S., Takeyasu,K. and Choy,H.E. (2012) Gene silencing by H-NS from distal DNA site. *Mol. Microbiol.*, **86**, 707–719.
160. Dattananda,C.S., Rajkumari,K. and Gowrishankar,J. (1991) Multiple mechanisms contribute to osmotic inducibility of proU operon expression in Escherichia coli: Demonstration of two osmoresponsive promoters and of a negative regulatory element within the first structural gene. *J. Bacteriol.*, **173**, 7481–7490.
161. Gowrishankar,J. (1989) Nucleotide sequence of the osmoregulatory proU operon of Escherichia coli. *J. Bacteriol.*, 10.1128/jb.171.4.1923-1931.1989.
162. Lucht,J.M., Dersch,P., Kempf,B. and Bremer,E. (1994) Interactions of the nucleoid-associated DNA-binding protein H-NS with the regulatory region of the osmotically controlled proU operon of Escherichia coli. *J. Biol. Chem.*, **269**, 6578–6586.
163. Nagarajavel,V., Madhusudan,S., Dole,S., Rahmouni,A.R. and Schnetz,K. (2007) Repression by binding of H-NS within the transcription unit. *J. Biol. Chem.*, **282**, 23622–23630.
164. Rashid,F.Z.M. (2021) Chapter 3: Regulation of proVWX transcription by local chromatin re-modelling.
165. Madrid,C., Nieto,J.M. and Juárez,A. (2001) Role of the Hha/YmoA family of proteins in the thermoregulation of the expression of virulence factors. *Int. J. Med. Microbiol.*, **291**, 425–432.
166. Wade,J.T. and Grainger,D.C. (2018) Waking the neighbours: disruption of H-NS repression by overlapping transcription. *Mol. Microbiol.*, 10.1111/mmi.13939.
167. Mai,X., Chou,S. and Struhl,K. (2000) Preferential Accessibility of the Yeast his3 Promoter Is Determined by a General Property of the DNA Sequence, Not by Specific Elements. *Mol. Cell. Biol.*, 10.1128/mcb.20.18.6668-6676.2000.
168. Schneider,D.A., Ross,W. and Gourse,R.L. (2003) Control of rRNA expression in Escherichia coli. *Curr. Opin. Microbiol.*, 10.1016/S1369-5274(03)00038-9.
169. Browning,D.F., Cole,J.A. and Busby,S.J.W. (2000) Suppression of FNR-dependent transcription activation at the Escherichia coli nir promoter by Fis, IHF and H-NS: Modulation of transcription by a complex nucleo-protein assembly. *Mol. Microbiol.*, 10.1046/j.1365-2958.2000.02087.x.
170. Browning,D.F., Grainger,D.C., Beatty,C.M., Wolfe,A.J., Cole,J.A. and Busby,S.J.W. (2005) Integration of three signals at the Escherichia coli nrf promoter: A role for Fis protein in catabolite repression. *Mol. Microbiol.*, 10.1111/j.1365-2958.2005.04701.x.
171. Bokal,A.J., Ross,W., Gaal,T., Johnson,R.C. and Gourse,R.L. (1997) Molecular anatomy of a transcription activation patch: FIS-RNA polymerase interactions at the Escherichia coli rrnB P1

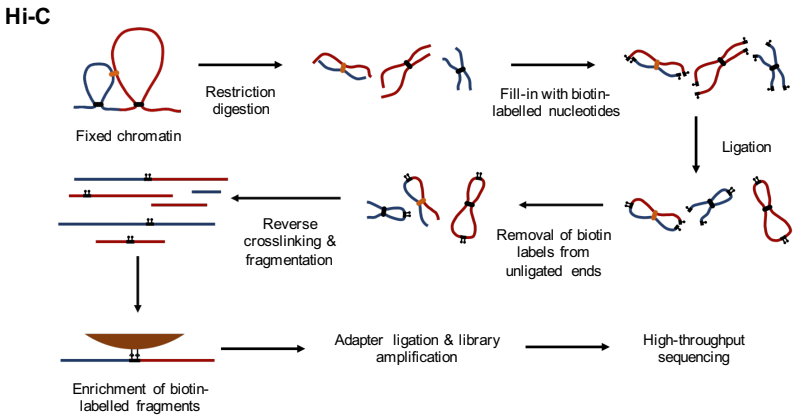
promoter. *EMBO J.*, 10.1093/emboj/16.1.154.
172. McLeod,S.M., Aiyar,S.E., Gourse,R.L. and Johnson,R.C. (2002) The C-terminal domains of the RNA polymerase α subunits: Contact site with Fis and localization during co-activation with CRP at the Escherichia coli proP P2 promoter. *J. Mol. Biol.*, 10.1006/jmbi.2001.5391.
173. Semsey,S., Tolstorukov,M.Y., Virnik,K., Zhurkin,V.B. and Adhya,S. (2004) DNA trajectory in the Gal repressosome. *Genes Dev.*, **18**, 1898–1907.
174. Oberto,J., Nabti,S., Jooste,V., Mignot,H. and Rouviere-Yaniv,J. (2009) The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. *PLoS One*, **4**.
175. Mangan,M.W., Lucchini,S., Cróinín,T.Ó., Fitzgerald,S., Hinton,J.C.D. and Dorman,C.J. (2011) Nucleoid-associated protein HU controls three regulons that coordinate virulence, response to stress and general physiology in Salmonella enterica serovar Typhimurium. *Microbiology*, **157**, 1075–1087.
176. Zhang,N., Darbari,V.C., Glyde,R., Zhang,X. and Buck,M. (2016) The bacterial enhancer-dependent RNA polymerase. *Biochem. J.*, 10.1042/BCJ20160741C.
177. Browning,D.F., Beatty,C.M., Wolfe,A.J., Cole,J.A. and Busby,S.J.W. (2002) Independent regulation of the divergent Escherichia coli nrfA and acsP1 promoters by a nucleoprotein assembly at a shared regulatory region. *Mol. Microbiol.*, 10.1046/j.1365-2958.2002.02776.x.
178. Adachi,S., Kohiyama,M., Onogi,T. and Hiraga,S. (2005) Localization of replication forks in wild-type and mukB mutant cells of Escherichia coli. *Mol. Genet. Genomics*, **274**, 264–271.
179. Badrinarayanan,A., Lesterlin,C., Reyes-Lamothe,R. and Sherratt,D. (2012) The escherichia coli SMC complex, MukBEF, shapes nucleoid organization independently of DNA replication. *J. Bacteriol.*, **194**, 4669–4676.
180. Nicolas,E., Upton,A.L., Uphoff,S., Henry,O., Badrinarayanan,A. and Sherratt,D. (2014) The SMC complex MukBEF recruits topoisomerase IV to the origin of replication region in live Escherichia coli. *MBio*, **5**.
181. Hayama,R. and Marians,K.J. (2010) Physical and functional interaction between the condensin MukB and the decatenase topoisomerase IV in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 18826–18831.
182. Hofmann,A., Mäkelä,J., Sherratt,D.J., Heermann,D. and Murray,S.M. (2019) Self-organised segregation of bacterial chromosomal origins. *Elife*, **8**.
183. Monterroso,B., Zorrilla,S., Sobrinos-Sanguino,M., Robles-Ramos,M.A., López-Álvarez,M., Margolin,W., Keating,C.D. and Rivas,G. (2019) Bacterial FtsZ protein forms phase-separated condensates with its nucleoid-associated inhibitor SlmA. *EMBO Rep.*, **20**.
184. Bernhardt,T.G. and De Boer,P.A.J. (2005) SlmA, a nucleoid-associated, FtsZ binding protein required for blocking septal ring assembly over chromosomes in E. coli. *Mol. Cell*, **18**, 555–564.
185. Tonthat,N.K., Arold,S.T., Pickering,B.F., Van Dyke,M.W., Liang,S., Lu,Y., Beuria,T.K., Margolin,W. and Schumacher,M.A. (2011) Molecular mechanism by which the nucleoid occlusion factor, SlmA, keeps cytokinesis in check. *EMBO J.*, **30**, 154–164.
186. Toro,E. and Shapiro,L. (2010) Bacterial chromosome organization and segregation. *Cold Spring Harb. Perspect. Biol.*, 10.1101/cshperspect.a000349.
187. Wang,X., Montero Llopis,P. and Rudner,D.Z. (2013) Organization and segregation of bacterial chromosomes. *Nat. Rev. Genet.*, **14**, 191–203.
188. Mierzejewska,J. and Jagura-Burdzy,G. (2012) Prokaryotic ParA-ParB-parS system links bacterial chromosome segregation with the cell cycle. *Plasmid*, 10.1016/j.plasmid.2011.08.003.
189. Toro,E., Hong,S.H., McAdams,H.H. and Shapiro,L. (2008) Caulobacter requires a dedicated mechanism to initiate chromosome segregation. *Proc. Natl. Acad. Sci. U. S. A.*, **105**, 15435–15440.
190. Böhm,K., Giacomelli,G., Schmidt,A., Imhof,A., Koszul,R., Marbouty,M. and Bramkamp,M. (2020) Chromosome organization by a conserved condensin-ParB system in the actinobacterium Corynebacterium glutamicum. *Nat. Commun.*, **11**.
191. Nolivos,S., Upton,A.L., Badrinarayanan,A., Müller,J., Zawadzka,K., Wiktor,J., Gill,A., Arciszewska,L., Nicolas,E. and Sherratt,D. (2016) MatP regulates the coordinated action of topoisomerase IV and MukBEF in chromosome segregation. *Nat. Commun.*, **7**.
192. Ricci,D.P., Melfi,M.D., Lasker,K., Dill,D.L., McAdams,H.H. and Shapiro,L. (2016) Cell cycle progression in Caulobacter requires a nucleoid-associated protein with high AT sequence

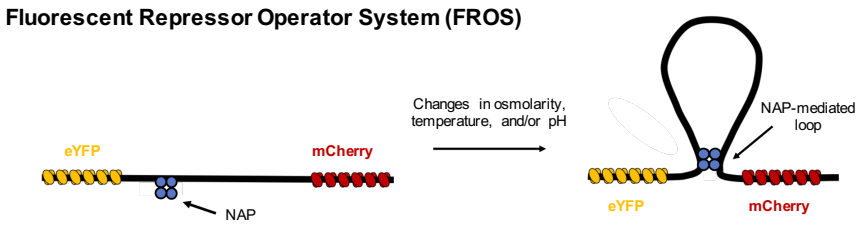
recognition. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, E5952–E5961.
193. Guo,M.S., Haakonsen,D.L., Zeng,W., Schumacher,M.A. and Laub,M.T. (2018) A Bacterial Chromosome Structuring Protein Binds Overtwisted DNA to Stimulate Type II Topoisomerases and Enable DNA Replication. *Cell*, **175**, 583-597.e23.
194. Luijsterburg,M.S., White,M.F., Van Driel,R. and Th. Dame,R. (2008) The major architects of chromatin: Architectural proteins in bacteria, archaea and eukaryotes. *Crit. Rev. Biochem. Mol. Biol.*, **43**, 393–418.
195. Dekker,J., Rippe,K., Dekker,M. and Kleckner,N. (2002) Capturing chromosome conformation. *Science (80-.).*, 10.1126/science.1067799.
196. Lieberman-Aiden,E., Van Berkum,N.L., Williams,L., Imakaev,M., Ragoczy,T., Telling,A., Amit,I., Lajoie,B.R., Sabo,P.J., Dorschner,M.O., *et al.* (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science (80-.).*, 10.1126/science.1181369.
197. van Berkum,N.L., Lieberman-Aiden,E., Williams,L., Imakaev,M., Gnirke,A., Mirny,L.A., Dekker,J. and Lander,E.S. (2010) Hi-C: A method to study the three-dimensional architecture of genomes. *J. Vis. Exp.*, 10.3791/1869.
198. Nagano,T., Lubling,Y., Stevens,T.J., Schoenfelder,S., Yaffe,E., Dean,W., Laue,E.D., Tanay,A. and Fraser,P. (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*, 10.1038/nature12593.
199. Nagano,T., Lubling,Y., Várnai,C., Dudley,C., Leung,W., Baran,Y., Mendelson Cohen,N., Wingett,S., Fraser,P. and Tanay,A. (2017) Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature*, 10.1038/nature23001.
200. Stevens,T.J., Lando,D., Basu,S., Atkinson,L.P., Cao,Y., Lee,S.F., Leeb,M., Wohlfahrt,K.J., Boucher,W., O'Shaughnessy-Kirwan,A., *et al.* (2017) 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature*, 10.1038/nature21429.

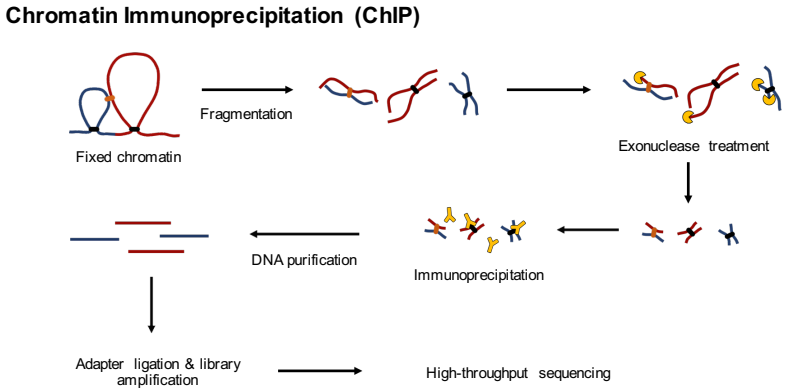
Acknowledgements
The authors thank the Netherlands Organization for Scientific Research [VICI 016.160.613] (R.T.D.), Wellcome Trust (212193/Z/18/Z) (D.C.G.), Leverhulme Trust (RPG-2018-198) (D.C.G.), BBSRC (BB/H010289/1) (D.C.G.) and the Human Frontier Science Program (HFSP) [RGP0014/2014] (R.T.D. and D.C.G.) for funding of current research in our labs.

Author contributions
All authors researched data for the article, substantially contributed to discussion of content, and wrote the article. R.T.D. and D.C.G. reviewed/edited the manuscript before submission.

Competing interests
The authors declare no competing interests.







Box 1.1: Techniques used to study chromosome structure and organization. Hi-C: Hi-C is a chromosome conformation capture (3C)-based method used to study the three-dimensional organization of the chromosome (195–197). It is a high-throughput technique that determines the probability of interaction between pairs of genomic loci in an unbiased manner at resolutions of up to 1kb (105, 196). Hi-C enables studies of chromosome organization in situ in the nucleus or nucleoid, changes therein in response to environmental stimuli, and — by alignment with genome-wide protein-occupancy profiles such as chromatin immunoprecipitation (ChIP) datasets — identification of proteins involved in chromosome structuring (56, 62, 95, 96, 100, 104–106, 196). Hi-C is generally an ensemble technique providing an averaged chromosome interaction profile. Single-cell Hi-C is gaining momentum in the field of chromosome biology, yet still has to be applied to bacterial organisms (198–200). The technique involves treating cells in culture with formaldehyde to chemically crosslink all

DNA–protein and protein–protein interactions within the cell, hence, fixing the structure of the chromosome (see the figure, part a). Of the remaining steps outlined in the figure part a, the key principle is that loci that are close to each other in three dimensional space are ligated into individual DNA ligation products regardless of their position along the primary genome sequence. Ligation products are read out using high-throughput sequencing to identify interacting pairs of genomic loci en masse. The sequencing data are represented as a heatmap of genome-wide interaction probabilities. **Flourescent repressor operator system (FROS):** FROS is a microscopy-based technique used to determine the position and track the dynamics of specific DNA loci in living cells. Loci of interest are marked with an array of 'operator' sequences that can be recognized and bound by ectopically expressed 'repressor' proteins that are translationally fused to flourescent protein labels (see the figure, part b). Different loci can be tagged and independently tracked in live cells to determine the spatial position and interrelationship of positions, establishing changes in structure in response to environmental stimuli. Changes in chromosome structure are evident as changes in distance between pairs of loci. FROS is a single-cell technique that reveals non-averaged chromosome dynamics. Despite limitations in its throughput, it is more powerful in establishing direct structure–function relations than Hi-C. **Chromatin immunoprecipitation (ChIP):** ChIP uses antibodies that target a DNA-binding protein of interest to isolate the factor and the chromosomal regions associated with it, following enzymatic, chemical or physical genome fragmentation. Often combined with high-throughput sequencing (ChIP-seq), the technique measures genome-wide patterns of DNA binding at single base-pair resolution. In this method, the chromosome of cells in culture is fixed with formaldehyde to crosslink all protein–protein and protein–DNA interactions (see the figure, part c). The chromosome is fragmented and antibodies are used to immunoprecipitate any DNA segments that are bound by the protein of interest. In ChIP-seq, the immunoprecipitated library is then purified and sequenced to determine the genome-wide DNA binding profile of the protein of interest.

Glossary

Chromosome
An essential molecule containing some or all of the genes required by an organism to survive and reproduce. Whereas chromosomes are made of DNA, not all DNA is chromosomal. Extra-chromosomal DNA molecules such as plasmids also encode genes, although these genes are not absolutely required for an organism's survival and reproduction.

Chromatin
A compact macromolecular complex of DNA and structuring proteins.

Nucleoid
A structure found in prokaryotic cells that contains chromosome(s), bound proteins and other associated molecules (e.g. RNAs). Nucleoids are functionally similar to the nuclei of eukaryotic cells but not enclosed within a membrane. Note that nucleoids can be found in eukaryotic organelles believed to be bacterial in origin.

Genome
The complete set of genes encoded by the DNA content of a given organism. The genome includes genes encoded by chromosomal and extra-chromosomal DNA, and intervening non-coding regions.

Nucleoid-associated proteins
(NAPs). A broad term to describe any proteins implicated in organizing bacterial chromosomes. Here, we consider structural maintenance of chromosomes (SMC) proteins as NAPs due to their association with the nucleoid and role in shaping nucleoid structure. SMC proteins — discovered later than other NAPs and initially studied primarily in the context of chromosome segregation — have historically (and in our view unjustly) not been classified as NAPs.

Nucleoprotein
A generic term, applicable to prokaryotes and eukaryotes, to describe DNA in complex with bound proteins.

Plectonemes

DNA loops in which the double-stranded DNA is wrapped around itself as a result of supercoiling.

Hyperplectoneme
Higher-order plectonemic structure formed by the winding of multiple plectonemes into a filament.

Supercoiled
Pertains to supercoiling, which is under- or over-winding of the double helix that causes the double-stranded DNA to fold into higher-order structures: plectonemes and toroids. To alter DNA supercoiling levels enzymatic breaking and rejoining of DNA strands is required.

Topoisomerase
An enzyme that alters DNA supercoiling by breaking and rejoining DNA strands. Mechanistically, topoisomerases are distinguished by whether they break and re-join either a single strand (type I) or both strands (type II).

Replichores
The sections of a chromosome between the origin and terminus of replication. Circular chromosomes are usually divided into a left and right replichore.

Chapter 2:
Hi-C in bacteria and archaea

Part of this chapter is based on: Crémazy, F.G., Rashid, F.M., Haycocks, J.R., Lamberte, L.E., Grainger, D.C., Dame, R.T., 2018. Determination of the 3D Genome Organization of Bacteria Using Hi-C. *Methods Mol Biol*, 1837:3-18. doi: 10.1007/978-1-4939-8675-0_1.

Abstract

The three-dimensional structure of the chromosome is encoded within its sequence and regulates activities such as replication and transcription. Hi-C combines proximity ligation and high-throughput sequencing to resolve spatial chromosome organization in relation to the underlying sequence, hence, revealing the principles of structural regulation of genome transactions. This chapter describes the preparation of Hi-C sequencing libraries for bacteria and archaea.

Introduction

In prokaryotes, the genome sequesters into a phase-separated nucleoid in the cytoplasm. Within this compartment, chromosome(s) organise into domains — physically insulated regions of the chromosome in which the chromatin interacts more frequently with itself than with flanking chromatin. Domain organisation is hierarchical, with larger domains formed from the preferential interaction of two or more smaller domains (1–3).

Three-dimensional chromosome organisation is determined by the physical properties of the DNA polymer and by the architectural properties of nucleoid associated proteins (NAPs) bound to the macromolecule. These proteins maintain chromosome structure and remodel it in response to environmental stimuli, directly affecting genome processes. For example, in *Escherichia coli*, the formation of a DNA:H-NS:DNA bridge between the downstream regulatory element (DRE) and the end of the *proU* operon silences its expression. Disruption of this bridge upon a hyperosmotic shock activates the operon (Chapter 3, this thesis). In *Bacillus subtilis*, a pair of hairpins formed by the directional movement of SMC-ScpAB complexes along the chromatin bring distantly placed DnaA boxes in close proximity to OriC to initiate chromosome replication (1). Conversely, genome processes affect chromosome structure. In *Caulobacter crescentus*, long, highly-transcribed genes form domain boundaries by generating extended supercoil diffusion barriers that physically separate flanking DNA (4). Furthermore, gene transcription in the Ter>Ori direction slows the movement of SMC complexes towards the chromosome terminus (5, 6).

The intricate interplay between chromosome architecture and its processes is encoded within the genome sequence as, for instance, binding sites of varying affinity for the repertoire of NAPs available within a cell, transcription initiation and termination sites of varying strengths, replication origins and termini, and loading and offloading sites for motor proteins that track along the DNA (Reviewed in (7)). The inherent sequence dependence of chromosome architecture necessitates describing chromosome structure in terms of its underlying sequence. To that end, chromosome conformation capture (3C)-based techniques detect the physical proximity of chromosome segments in three-dimensional space (8). Hi-C uses proximity ligation and next generation sequencing to probe physical interactions amongst all segments of the chromosome in an unbiased manner (9). The technique, therefore, describes the three-dimensional structure of the chromosome in relation to genome sequence.

As an ensemble technique, Hi-C determines the probability that two regions of the chromosome lie in physical proximity (9). This highlights common architectural features of an organism's chromosome such as its domains and domain boundaries, compartmentalization into hetero- and eu-chromatin, and, in some prokaryotes, the presence of inter-arm interactions (1–4, 10). As a single-cell technique, Hi-C identifies the mutual inclusivity or exclusivity of these features (11–13).

This chapter builds on our earlier work (14) and describes protocols for the preparation of ensemble Hi-C libraries compatible for sequencing on an Illumina NGS platform from *Escherichia coli*, *Saccharolobus solfataricus*, and *Haloquadratum walsbyi* cultures.

Materials

Cell fixation

1. 37% w/v formaldehyde solution, stored at room temperature away from direct sunlight.
2. 1X TE Buffer: 10 mM Tris-HCl, 1 mM EDTA·Na_2, pH 8.0. Store the buffer at room temperature and cool it to 4 °C prior to use.
3. Liquid nitrogen.
4. Refrigerated benchtop centrifuge with rotors for 50 mL, 15 mL, and 1.5 mL tubes.
5. Roller bench.
6. -80°C freezer.
7. 50 mL and 15 mL conical tubes.
8. Serological pipettes.
9. Pipette pump.
10. Elongated Pasteur pipette (made of soda lime silica glass): Heat a 1 cm space of the Pasteur pipette above the tip and stretch the heated region to reduce the diameter of the pipette bore. Use a pair of tweezers to break the pipette at the elongated region.
11. 2.5 µL, 20 µL, 200 µL, and 1000 µL micropipettes.
12. Low-retention micropipette tips.
13. Low-retention 1.5 mL microfuge tubes.

Additional materials for *Escherichia coli* cell fixation:

14. Methanol, ACS reagent ≥99.8%. Store the reagent at room temperature and cool it to 4 °C prior to use.

15. 1X PBS solution: 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.47 mM KH_2PO_4, pH 7.4. Store 1X PBS at room temperature and cool it to 4 ºC prior to use.
16. 2.5 M Glycine solution filter-sterilized using a sterile, non-pyrogenic 0.2 µm filter *(see note 1)*.

Additional materials for *Saccharolobus solfataricus* cell fixation:

14. Thermomixer (e.g. Eppendorf Thermomixer® C) with a thermoblock for 15 mL conical tubes (e.g. Eppendorf SmartBlock™ 15 mL).
15. Filter-sterilised 2.5 M glycine *(see note 1)*.

Additional materials for *Haloquadratum walsbyi* cell fixation:

14. DSMZ medium 1091 for *H. walsbyi*.
15. 10X *H. walsbyi* dialysis buffer: 60 mM L-glutamic acid, 270 mM KCl, 115 mM sodium citrate dihydrate, 800 mM $MgSO_4 \cdot 7H_2O$, 2.0 mM $FeCl_2 \cdot 4H_2O$, 20.0 µM $MnCl_2 \cdot 2H_2O$, pH 6.3-6.4. Store the buffer at room temperature (20-25 °C), preferably in the dark.
16. Dialysis tubing with a 3.5 kDa molecular weight cut-off.
17. Dialysis tubing clamps and rubber bands.
18. 2 L beakers.
19. Cling film or aluminium foil.
20. (Heated) magnetic stirring plate and stir bar.
21. Optional: 20% ethanol.

Hi-C library preparation

1. 1X TE Buffer: 10 mM Tris-HCl, 1 mM EDTA·Na_2, pH 8.0.
2. **For *E. coli*:** 40,000 U/µL Ready-Lyse Lysozyme (Epicentre).
3. **For *E. coli*:** 1.0 M NaCl.
4. **For *S. solfataricus* and *H. walsbyi*:** Glass beads, 0.5 mm in diameter (e.g. Z250465-1PAK from Sigma Aldrich).
5. Nuclease-free water.
6. 10% w/v Sodium dodecyl sulfate (SDS) solution: Dissolve 1.0 g of SDS (ACS reagent, ≥99.0%) in 10 mL of nuclease-free water.
7. 10% v/v Triton X-100 solution: Dilute 1.0 mL of Triton X-100 (molecular biology grade) with 9.0 mL of nuclease-free water. Store the solution at room temperature (20-25 °C) in the dark.
8. Restriction enzyme: 10 U/µL PsuI (ThermoFisher Scientific) *(see note 2)*.

9. 10X restriction enzyme buffer, as supplied with the selected restriction enzyme *(see note 2)*: 10X Buffer B (ThermoFisher Scientific).
10. 100 mM deoxynucleotides (ultrapure dNTPs as separate solutions). Repeated freeze-thaw cycles may result in the degradation of dNTPs. To minimize degradation of the 100 mM stock, dilute 20.0 µL of 100 mM dNTP with 180.0 µL of nuclease-free water. Store the 10 mM dNTP solution at -20 °C as 10.0 – 50.0 µL aliquots.
11. Separate solutions of 1 mM dNTP: Dilute one 10 mM dNTP aliquot tenfold with nuclease-free water. Store the dNTP solution at -20 °C and thaw on ice or at 4 °C when necessary.
12. 0.4 mM Biotin-14-dATP (*see note 3*).
13. 20 mg/mL Bovine serum albumin (BSA).
14. 5 U/µL DNA Polymerase I, Large (Klenow) fragment.
15. Qubit® dsDNA HS Assay Kit and a Qubit® fluorometer from ThermoFisher Scientific.
16. 10X T4 DNA Ligase buffer: 0.5 M Tris-HCl, 0.1 M $MgCl_2$, 10 mM ATP, 0.1 M DTT, pH 7.5.
17. 2000 U/µL T4 DNA Ligase.
18. 0.5 M EDTA, pH 8.0.
19. 10 mg/mL RNase A.
20. 20 mg/mL Proteinase K.
21. 5.0 M NaCl.
22. 25:24:1 Phenol:chloroform:isoamyl alcohol solution *(see note 4)*.
23. Chloroform *(see note 4)*.
24. 1.0 M Sodium acetate (NaOAc), pH 8.0.
25. 5.0 mg/mL Glycogen.
26. 100% and 70% (v/v) Ethanol. Keep the solutions cold at -20 °C.
27. 10 mM Tris, pH 8.0.
28. 10X T4 DNA Polymerase buffer: 0.5 M NaCl, 0.1 M Tris-HCl, 0.1 M $MgCl_2$, 1 mg/mL BSA, pH 7.9 (available as NEBuffer 2.1).
29. 3 U/µL T4 DNA Polymerase.
30. Thermomixer (e.g. Eppendorf Thermomixer® C) with a thermoblock for 1.5 mL microfuge tubes (e.g. Eppendorf SmartBlock™ 1.5 mL).
31. Benchtop centrifuge.
32. Refrigerated benchtop centrifuge.
33. Vortex mixer.
34. Agarose gel electrophoresis set-up.
35. -20 °C freezer.

36. 2.5 µL, 20 µL, 200 µL, and 1000 µL micropipettes.
37. Low-retention micropipette tips.
38. Low-retention 1.5 mL microfuge tubes.
39. 15 mL conical tubes

NGS library preparation

1. Solid Phase Reverse Immobilization (SPRI) beads (e.g. KAPA HyperPure Beads from KAPA Biosystems). Store the beads as 1.0 mL aliquots at 4 °C. Ensure that the beads are equilibrated to 25 °C and vortexed before use.
2. Nuclease-free water.
3. 10 mM Tris, pH 8.0.
4. 10 mM Tris, pH 8.0 + 0.1% Tween 20.
5. Qubit® dsDNA HS Assay Kit and a Qubit® fluorometer from ThermoFisher Scientific.
6. Illumina® library preparation kit (e.g. KAPA HyperPlus kit from KAPA Biosystems)
7. Illumina®-compatible paired-end sequencing adapters at a concentration of 15 µM (e.g. KAPA Single-Indexed Adapter Sets A and B from KAPA Biosystems). The sequencing adapters should be barcoded to allow multiplexing during NGS.
8. Streptavidin-coupled magnetic beads optimized for the enrichment of biotin-labelled nucleic acids (e.g. Dynabeads® MyOne™ Streptavidin T1 from ThermoFisher Scientific). Store the beads at 4 °C. Equilibrate the beads to 25 °C and vortex well before use.
9. 2X Beads Wash Buffer (2X BWB): 10 mM Tris pH 7.5, 1 mM EDTA, 2 mM NaCl, 0.2% Tween 20.
10. 1X Beads Wash Buffer (1X BWB): 5 mM Tris pH 7.5, 0.5 mM EDTA, 1 mM NaCl, 0.1% Tween 20.
11. Magnetic racks (e.g. DynaMag™-PCR Magnet and DynaMag™-2 Magnet from ThermoFisher Scientific).
12. Thermomixer (e.g. Eppendorf Thermomixer® C) with a thermoblock for 1.5 mL microfuge tubes (e.g. Eppendorf SmartBlock™ 1.5 mL).
13. Thermal cycler with a heated lid.
14. Vortex mixer.
15. PCR tubes.
16. Tapestation (Agilent Technologies), or Bioanalyser (Agilent Technologies), or an agarose gel electrophoresis set-up.
17. 2.5 µL, 20 µL, 200 µL, and 1000 µL micropipettes.

18. Low-retention micropipette tips.
19. Low-retention 1.5 mL microfuge tubes.

Methods

The protocols described here are optimized for HiC in *Escherichia coli*, *Saccharolobus solfataricus*, and *Haloquadratum walsbyi*. To process other prokaryotes using these protocols, optimize the cell lysis and solubilization steps by adjusting the treatment duration – and, if relevant, the concentrations – of Ready-Lyse Lysozyme, SDS, and physical shearing with glass beads.

Cell fixation

Escherichia coli

1. Transfer an aliquot of an *Escherichia coli* culture equivalent to ~$5.0x10^9$ cells into a chilled 50 mL conical tube.
2. Harvest the cells by centrifugation at 3000 xg for 5 minutes at 4 °C.
3. Decant the supernatant and carefully remove the rest with a micropipette. Resuspend the cell pellet in 4.0 mL of 1X PBS.
4. Add 16.0 mL of methanol to the cell suspension *(see note 5)*. Mix well by inversion and incubate the cells at 4 °C for 10 minutes on a roller bench at 30 rpm.
5. Collect the cells at 3000 xg for 5 minutes at 4 °C.
6. Discard the supernatant and resuspend the cells in 20.0 mL of 1X PBS to wash.
7. Centrifuge the cell suspension at 3000 xg for 5 minutes at 4 °C.
8. During the centrifugation, prepare 40 mL of 3% formaldehyde in 1X PBS by adding 3.2 mL of 37% formaldehyde solution to 36.8 mL of cold 1X PBS in a 50 mL conical tube. Mix well by inversion and place the solution on ice or at 4 °C.
9. Discard the supernatant of the centrifugation performed at step 7 of this section. Resuspend the cell pellet in 40.0 mL of 3% formaldehyde. Incubate the cells at 4 °C for 1 hour on a roller bench at 30 rpm.
10. Add 6.4 mL of 2.5 M glycine to the fixation reaction (to a final concentration of 0.375 M) to quench excess formaldehyde. Incubate the cells for 15 minutes at 4 °C on a roller bench at 30 rpm.
11. Collect the fixed cells by centrifugation at 3000 xg for 10 minutes at 4 °C.
12. Decant the supernatant and carefully remove the rest with a micropipette. Resuspend the pellet in 5.0 mL of 1X TE and divide the suspension into

five aliquots of 1.0 mL in sterile 1.5 mL microfuge tubes. Each aliquot contains approximately 10^9 cells.

13. Pellet the fixed cells by centrifugation at 10000 xg for 2 minutes at 4 °C.
14. Remove the supernatant with an elongated Pasteur pipette.
15. Proceed immediately with Hi-C library preparation or flash freeze the cells in liquid nitrogen. Frozen pellets can be stored at -80 °C.

Saccharolobus solfataricus

1. Heat up a 15 mL conical tube to 80 ºC in a Thermomixer.
2. Transfer an aliquot of a *Saccharolobus solfataricus* culture equivalent to ~5.0x10^9 cells to the conical tube at 80 ºC.
3. Add 37% formaldehyde solution to the culture to a final concentration of 1%. This corresponds to 28.8 µL of 37% formaldehyde per mL of culture. Mix the reaction by inversion and return the conical tube to the Thermomixer.
4. Change the set temperature of the Thermomixer to 75 ºC and allow the fixation reaction to cool to 75 ºC. This takes approximately 6 minutes with the Eppendorf Thermomixer® C.
5. Add 2.5 M glycine to the fixation reaction to a final concentration of 0.125 M. This corresponds to 54.1 µL of 2.5 M glycine per mL of culture. Mix well by inversion and incubate the reaction at 4 ºC on a roller bench at 30 rpm for 15 minutes.
6. Collect the cells by centrifugation at 3000 xg for 5 minutes at 4 ºC.
7. Discard the supernatant and resuspend the cell pellet in 10.0 mL of 1X TE to wash.
8. Centrifuge the cells at 3000 xg for 5 minutes at 4 °C.
9. Pour out the supernatant and carefully remove the rest with a micropipette. Resuspend the cell pellet in 5.0 mL of 1X TE and separate the suspension of fixed cells into five 1.0 mL aliquots in low-retention 1.5 mL microfuge tubes.
10. Pellet the fixed cells by centrifugation at 10000 xg for 2 minutes at 4 °C.
11. Remove the supernatant with an elongated Pasteur pipette.
12. Use the fixed cells immediately for Hi-C library preparation, or, flash freeze the pellets in liquid nitrogen and store the cells at -80 °C.

Haloquadratum walsbyi

1. Transfer 25 mL of a *Haloquadratum walsbyi* culture into a 50 mL conical tube.

2. Collect the cells by centrifugation at 2300 xg for 20 minutes at 25 ºC.
3. Decant the supernatant and gently resuspend the cell pellet in 1.0 mL of DSMZ medium 1091 for *H. walsbyi*. Vigorous pipetting or vortexing will rupture the cells.
4. Transfer the cell suspension into a 1.5 mL low retention microfuge tube containing 156.3 µL of 37% formaldehyde (final concentration: 5% formaldehyde). Mix by pipetting and incubate the fixation reaction in a Thermomixer at 25 ºC for 3 hours.
5. Dilute 300 mL of 10X *H. walsbyi* buffer in 2700 mL of MiliQ water (final concentration: 1X *H. walsbyi* buffer).
6. Pour out 1.5 L of 1X *H. walsbyi* dialysis buffer into each of two 2 L beakers. Cover one beaker with cling film or aluminium foil to be used at step 10 of this section.
7. Cut a ~5 cm strip of a dialysis tubing with a 3.5 kDa molecular weight cut-off and soak it in 1X *H. walsbyi* dialysis buffer for 20 seconds. Fold 1 cm of one end of the tubing onto itself and clamp the fold. Using a micropipette, fill the tubing with 1X dialysis buffer from the open end and verify that the clamped end does not leak. Leakages can be addressed by tightening the clamps using rubber bands.
8. Empty the dialysis tubing and pipette in the suspension of fixed *H. walsbyi* cells. Fold the open end of the tubing onto itself and clamp the fold. Use rubber bands if necessary. Place the tubing in the dialysis buffer and cover the beaker with cling film or aluminium foil.
9. Dialyse the fixed cells for 1 hour at 25 ºC with stirring. Check the set-up every 15 minutes for leakages. A well-clamped set-up will swell during dialysis.
10. Quickly transfer the dialysis tube to the second beaker of 1X *H. walsbyi* dialysis buffer set aside earlier (step 6 of this section). Cover the set-up with cling film or aluminium foil and let the dialysis continue overnight at 25 ºC with stirring.
11. Transfer the contents of the tubing into a 1.5 mL low-retention microfuge tube.
12. Pellet the fixed cells by centrifugation at 10000 xg for 2 minutes at 4 ºC.
13. Remove the supernatant with an elongated Pasteur pipette and proceed immediately with Hi-C library preparation.
14. Optional: Wash the dialysis tubing thoroughly with MiliQ water and store it in 20% ethanol. The tubing can be reused for the dialysis of fixed *H.*

walsbyi cells as described from steps 7 to 11 of this section following thorough washing with MiliQ water and 1X *H. walsbyi* buffer.

Hi-C library preparation

Lysis and solubilization: *Escherichia coli*

1. Dilute 0.5 µL of 40,000 U/µL Ready-Lyse Lysozyme (Epicentre) with 4.0 µL of 1X TE and 0.5 µL of 1.0 M NaCl. Place the dilution on ice.
2. Resuspend a pellet of fixed *Escherichia coli* cells in 50.0 µL of 1X TE and add 0.5 µL of the diluted lysozyme (2000 U) to the cell suspension. Mix by pipetting and incubate the tube at 37 ºC in a Thermomixer at 1000 rpm for 15 minutes.
3. Add 2.5 µL of 10% SDS to the lysis reaction and mix by pipetting. Incubate the tube for 15 minutes at 37 °C in a Thermomixer at 1000 rpm.
4. Add 45.0 µL of 1X TE to 5.0 µL of the lysed and solubilised cells. Store the preparation at -20 ºC as 'Chromatin'. This control is used to qualitatively examine the efficiency of cell lysis and solubilisation – (steps 2 and 3 of this section; Figure 2.1). Continue with processing Chromatin at step 20 of Hi-C library preparation.
5. Make up the volume of the lysed and solubilised cells to 53.0 µL with 5.0 µL of 1X TE and continue with Hi-C library preparation (step 6 of Hi-C library preparation).

Lysis and solubilization: *Saccharolobus solfataricus*, and *Haloquadratum walsbyi*

1. Resuspend a pellet of fixed cells in 60.0 µL of 1X TE. Add 25-30 glass beads to the cell suspension and vortex vigorously for 5-7 minutes to shear the cells. Ensure that the beads flow with the cell suspension on the sides of the tube during vortexing.
2. Centrifuge the tube for 5 seconds to spin down the disrupted cells and sediment the glass beads.
3. Add 2.8 µL of 10% SDS to the lysate and mix by briefly vortexing the tube. Incubate the tube for 15 minutes at 37 ºC in a Thermomixer at 1000 rpm.
4. Mix 5.0 µL of the lysed and solubilised cells with 45.0 µL of 1X TE in a 1.5 mL low-retention microfuge tube. Store the preparation at -20 ºC as the Chromatin control. Continue with processing this sample at step 20 of Hi-C library preparation.

5. Transfer 53.0 µL of the lysed and solubilised cells into a 1.5 mL low-retention microfuge tube to continue with Hi-C library preparation (step 6 of Hi-C library preparation).

Restriction digestion

6. Add 136.0 µL of nuclease-free water, 25.0 µL of 10% Triton X-100, and 25.0 µL of 10X Buffer B (ThermoFisher Scientific) to 53.0 µL of lysed and solubilised cells (step 5 of Hi-C library preparation).
7. Mix by inversion and incubate the sample at 37 °C for 10 minutes in a Thermomixer at 1000 rpm to sequester SDS in the cell lysate with Triton X-100.
8. Add 10.0 µL of 10 U/µL PsuI (ThermoFisher Scientific) *(see note 2)* to the digestion mix and incubate the reaction at 37 °C for 3 hours at 450 rpm.

Fill-in with biotin-labelled nucleotides

9. To 250 µL of the digestion reaction, add 3.0 µL of 1.0 mM dTTP, 3.0 µL of 1.0 mM dGTP, 3.0 µL of 1.0 mM dCTP, 7.5 µL of 0.4 mM biotin-14-dATP, 1.5 µL of 20 mg/mL BSA, 26.0 µL of nuclease-free water, and 6.0 µL of 5 U/µL DNA Polymerase I, Large (Klenow) fragment (*see note 3*). Mix well by pipetting.
10. Incubate the reaction for 45 minutes at 25 °C in a Thermomixer at 450 rpm.
11. Add 15.8 µL of 10% SDS to the fill-in reaction to a final concentration of 0.5%. Incubate the sample for 20 minutes at 25 °C in a Thermomixer at 1000 rpm to terminate the reaction. Avoid inactivation at higher temperatures to minimize premature reverse cross-linking of the chromatin fragments.

Fractionation:

12. Centrifuge the digested, biotin-labelled chromatin at 25000 xg for 1 hour at 4 °C.
13. Carefully remove the supernatant with an elongated Pasteur pipette and resuspend the gel-like pellet in 200.0 µL of nuclease-free water by vigorous pipetting and vortexing.
14. Use 1.0 µL of the sample to determine DNA concentration using the Qubit® dsDNA HS Assay Kit. Proceed with Hi-C library preparation if at least 1300 ng of digested chromatin is available.

15. Keep a 200-500 ng aliquot aside as the R+ control to visualize the efficiency of restriction digestion (Figure 2.1). Make up the volume of the control to 50.0 µL using 1X TE and store it at -20 °C. Continue with processing the R+ control at step 20 of Hi-C library preparation.

Ligation:

16. Adjust the volume of an aliquot of digested chromatin containing between 1.0-3.0 µg of DNA to 895 µL with nuclease-free water.
17. Add 100.0 µL of 10X T4 DNA Ligase Buffer, 5.0 µL of 20 mg/mL BSA, and 2.0 µL of 2000 U/µL T4 DNA Ligase to the sample and mix well by inversion. Multiple reactions can be prepared to use up all the fractionated chromatin.
18. Incubate the reaction(s) in a Thermomixer for 16 hours at 16 °C at 0 rpm, and for 1 hour at 25 °C at 450 rpm.
19. Terminate the ligation reaction(s) with 20.5 µL of 0.5 M EDTA.

Reverse cross-linking and DNA purification

20. Add 16.6 µL of 10 mg/mL RNase A to the ligation reaction(s). For Chromatin and R+ (steps 4 and 15 of Hi-C library preparation), use 0.8 µL of 10 mg/mL RNase A. Mix by pipetting and incubate the samples for 30 minutes at 37 °C in a Thermomixer at 450 rpm.
21. Add 12.5 µL of 20 mg/mL Proteinase K and 120.8 µL of 5.0 M NaCl to the ligation reaction(s). Use 0.64 µL of 20 mg/mL Proteinase K and 6.2 µL of 5.0 M NaCl for Chromatin and R+ (steps 4 and 15 of Hi-C library preparation). Incubate the samples for 16 hours at 65 °C in a Thermomixer at 450 rpm.
22. Transfer the ligation sample into a 15 mL conical tube. Pool corresponding samples if applicable (step 17 of Hi-C library preparation).
23. Add 1.0 volume of 25:24:1 phenol:chloroform:isoamyl alcohol to ligation sample and controls *(see note 4)*. Vortex vigorously for 10 seconds.
24. Centrifuge the samples for 10 minutes at 13000 xg at 4 °C.
25. Transfer the top aqueous layer into sterile 1.5 mL or 15 mL tubes, as necessary *(see note 4)*.
26. Repeat steps 23 to 25 of Hi-C library preparation.
27. Add 1.0 volume of chloroform to the ligation sample and controls and vortex vigorously for 10 seconds *(see note 4)*.
28. Centrifuge the tubes for 10 minutes at 13000 xg at 4 °C.

29. Transfer the top aqueous layer into sterile 1.5 mL or 15 mL tubes, as necessary *(see note 4)*.
30. Add 0.1 volumes of 1.0 M sodium acetate, pH 8.0, 0.025 volumes of 5.0 mg/mL glycogen, and 2.5 volumes of cold 100% ethanol to the ligation sample and controls. Vortex well and incubate the solutions overnight at -20 °C.
31. Centrifuge the solutions for 20 minutes at 25000 xg at 4 ºC.
32. Remove the supernatant with an elongated Pasteur pipette. A white pellet should be visible at the base of the tubes.
33. Add 500 µL of cold 70% ethanol to the tubes to wash the pellets.
34. Centrifuge the tubes at 25000 xg for 5 minutes at 4 °C.
35. Remove the supernatant with an elongated Pasteur pipette.
36. Repeat steps 33 to 35 of Hi-C library preparation.
37. Dry the pellets in a 50 °C incubator.
38. Dissolve each pellet in 31.0 µL of nuclease-free water in a Thermomixer at 60 ºC for 20 minutes at 1000 rpm.
39. Use 1.0 µL of the controls and the purified ligation sample, hereafter referred to as the Hi-C library, to determine DNA concentration using the Qubit® dsDNA HS Assay Kit.
40. Visualize between 100 and 500 ng of each sample on a 1.5% Agarose gel to qualitatively assess Hi-C library preparation. Typically, the Chromatin migrates as a heavy >10 kb band on a 1.5% agarose gel, R+ runs as a smear of digestion products, and a well-ligated Hi-C library shows a shift of the R+ profile to higher molecular weights (Figure 2.1). Ensure that at least 1.0 µg of the Hi-C library is available for the next steps of the protocol.

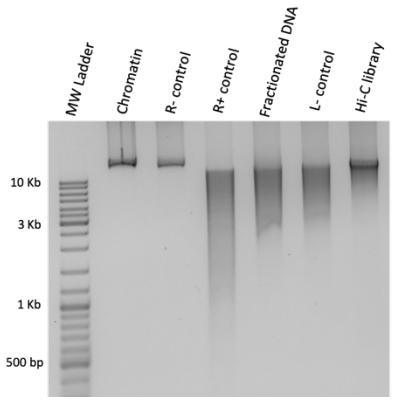


Figure 2.1: Typical results obtained for Hi-C library preparation. Chromatin: Chromatin extracted from fixed cells during lysis and solubilization runs as a high molecular weight band >10 kb. **R- control:** Chromatin that is incubated in the restriction digestion mix for three hours in the absence of

a restriction enzyme should appear as a high molecular weight band comparable to the Chromatin. A smear in the R- control indicates that the chromatin undergoes degradation during restriction digestion. **R+ control:** Chromatin digested with a 6-cutter should run as a smear with a bulk of the fragments having a size > 2 kb. **Fractionated DNA:** Fractionation pellets DNA that is cross-linked to protein (16). 'Free floating' DNA fragments that may contribute to random ligation and self-circularized products are eliminated. This step may be associated with a loss of shorter fragments. **L- control:** Fractionated chromatin in the ligation mix in the absence of T4 DNA Ligase should run on the agarose gel with a profile comparable to fractionated DNA. Degradation will be observed as a shift of the profile towards shorter DNA lengths. **Hi-C library:** A well-ligated Hi-C library shows a shift of the DNA profile towards high molecular weights.

Removal of biotin-labelled nucleotides from unligated restriction fragment ends

41. To 30.0 µL of the Hi-C library, add 8.75 µL of nuclease-free water, 5.0 µL of 10X NEBuffer 2.1, 5.0 µL of 1 mM dGTP, 0.25 µL of 20 mg/mL BSA, and 1.0 µL of 3 U/µL T4 DNA Polymerase. Mix by pipetting and incubate the reaction at 16 °C for 3 hours in a Thermomixer at 450 rpm (*see note 6*).

Hi-C library purification

42. Add 1.0 volume of 25:24:1 phenol:chloroform:isoamyl alcohol to Hi-C library and vigorously vortex the tube for 10 seconds *(see note 4)*.
43. Centrifuge the tube at 13000 xg for 10 minutes at 4 °C.
44. Transfer the top aqueous layer into a low-retention 1.5 mL microfuge tube *(see note 4)*.
45. Add 0.1 volumes of 3.0 M NaOAc pH 5.6, 0.025 volumes of 5.0 mg/mL glycogen, and 2.5 volumes of cold 100% ethanol. Mix well by vortexing and incubate the solution at -20 °C for at least 1 hour.
46. Centrifuge the solution for 20 minutes at 25000 xg at 4 °C.
47. Remove the supernatant with an elongated Pasteur pipette. A white pellet should be visible at the base of the tube.
48. Add 500 µL of cold 70% ethanol to the microfuge tube to wash the pellet.
49. Centrifuge the tube at 25000 xg for 5 minutes at 4 °C.
50. Remove the supernatant with an elongated Pasteur pipette.
51. Dry the pellet in a 50 °C incubator.
52. Dissolve the pellet in 21.0 µL of nuclease-free water in a Thermomixer at 60 °C for 20 minutes at 1000 rpm.
53. Use 1.0 µL of the library to determine DNA concentration using the Qubit® dsDNA HS Assay Kit. Ensure that at least 750 ng of the Hi-C library is available for NGS Library preparation.

NGS library preparation

1. Remove glycogen and excess salt from the Hi-C library by adding 60.0 µL of KAPA HyperPure beads with 20.0 µL of the Hi-C library in a low-retention 1.5 mL microfuge tube. Vortex to mix. Incubate the sample at 25 ºC in a Thermomixer for 5 minutes.
2. Wash the beads according to the manufacturer's instructions and elute the Hi-C library in 36.0 µL of 10 mM Tris, pH 8.0 (*see note 7*).
3. Use 1.0 µL of the eluted Hi-C library to determine DNA concentration using the Qubit® dsDNA HS Assay Kit. Proceed to the next step if more than 500 ng of the Hi-C library is available *(see note 8)*.
4. Add 5.0 µL of 10X KAPA Frag buffer and 10.0 µL of the KAPA Frag enzyme to the Hi-C library on ice. Vortex gently to mix and incubate the tube in a Thermomixer set at 37 ºC for 45 minutes.
5. Transfer the fragmentation reaction to ice and set the temperature of the Thermomixer to 65 ºC.
6. Mix 7.0 µL of the End Repair and A-Tailing Buffer and 3.0 µL of the End Repair and A-Tailing Enzyme Mix on ice (supplied with the KAPA HyperPlus kit). Transfer the mix (10.0 µL) to the fragmentation reaction and vortex gently.
7. Place the reaction in the Thermomixer at 65 ºC for 30 minutes.
8. Transfer the reaction to ice and set the temperature of the Thermomixer to 20 ºC.
9. Add the following to 60.0 µL of end repaired and A-tailed DNA: 5.0 µL of 15 µM Illumina® paired-end sequencing adapter (barcoded), 5.0 µL of nuclease-free water, 30.0 µL of ligation buffer (KAPA HyperPlus), and 10.0 µL of DNA Ligase (KAPA HyperPlus). To sequence several Hi-C libraries in a single Illumina® lane, ligate each library to a different sequencing adapter.
10. Vortex gently and incubate the reaction at 20 °C for 2 hours. Continue processing this sample at step 14 of NGS library preparation.
11. Transfer 20.0 µL of Dynabeads® MyOne™ Streptavidin T1 into a low-retention microfuge tube. Collect the beads with a magnetic rack and carefully pipette off the supernatant.
12. Re-suspend the beads in 50.0 µL of 1X BWB by vortexing. Place the beads back on the magnetic rack until the solution clears and pipette off the supernatant. Perform this step twice.
13. Re-suspend the beads in 50.0 µL of 2X BWB by vortexing.

14. Clean-up the adapter-ligated Hi-C library (steps 9 and 10 of NGS library preparation) with 220.0 µL of KAPA HyperPure beads. Wash the beads according to the manufacturer's instructions and elute the library in 50.0 µL of nuclease-free water.
15. Mix the eluted library with 50.0 µL of Dynabeads® MyOne™ Streptavidin T1 in 2X BWB (step 13 of NGS library preparation). Incubate the sample for 45 minutes at 37 °C in a ThermoMixer at 500 rpm.
16. Collect the beads (now coupled to the adapter-ligated Hi-C library) using a magnetic rack and pipette off the supernatant.
17. Re-suspend the beads in 100.0 µL of 10 mM Tris pH 8.0 + 0.1% Tween 20 by vortexing. Collect the beads with the magnetic rack and remove the supernatant. Perform this step three times.
18. Re-suspend the beads in 20.0 µL of 10 mM Tris pH 8.0 + 0.1% Tween 20 and transfer the suspension into a PCR tube.
19. To the adapter-ligated Hi-C library on beads, add: 25.0 µL of 2X KAPA HiFi HotStart ReadyMix, and 5.0 µL of 10X Library Amplification Primer Mix as supplied with the KAPA HyperPlus kit. Run the following program in a thermocycler to amplify the library:

Step	**Temperature**	**Time**	**Cycles**
Initial denaturation	98 °C	45 seconds	1
Denaturation	98 °C	15 seconds	6–8 (*see note 9*)
Annealing	60 °C	30 seconds	
Extension	72 °C	30 seconds	
Final extension	72 °C	1 minute	1
Hold	4 °C	∞	1

20. Collect the beads with a magnetic rack and transfer the supernatant (hereafter referred to as the NGS library) into a 1.5 mL low-retention microfuge tube. Place the tube on ice until processing at step 23 of NGS library preparation.
21. Re-suspend the beads in 100.0 µL of 10 mM Tris pH 8.0 + 0.1% Tween 20 by vortexing. Place the tube on a magnetic rack and pipette off the supernatant. Perform this step three times.
22. Re-suspend the beads in 20.0 µL of 10 mM Tris pH 8.0 + 0.1% Tween 20 and store the sample at 4 °C. This adapter-ligated Hi-C library on beads can be re-used on a later date to prepare the NGS library as outlined in step 19 of NGS library preparation.

23. Clean-up the NGS library (step 20 of NGS library preparation) with 100.0 µL of KAPA HyperPure beads and elute it in 20.0 µL of nuclease-free water (*see note 10*).
24. Add 10.0 µL of KAPA HyperPure beads to the eluted library. Vortex to mix and incubate the suspension at 25 °C in a Thermomixer for 5 minutes. *(see notes 7 and 10)*.
25. Place the tube on a magnetic rack and transfer the supernatant to a fresh low-retention microfuge tube. Discard the beads *(see note 10)*.
26. Purify the supernatant with 60.0 µL of KAPA HyperPure beads and elute the size-selected NGS library in 11.0 µL of nuclease-free water.
27. Use 1.0 µL of the NGS library to determine DNA concentration using the Qubit® dsDNA HS Assay Kit. The concentration of the library should be between 5 - 20 ng/µL.
28. Assess the quality of the NGS library by determining the size distribution of the DNA fragments using a Tapestation or Bioanalyzer (Agilent Technologies), or by agarose gel electrophoresis (*see note 11*) (Figure 2.2). 1.0-2.0 µL of the NGS library is sufficient for this purpose.
29. Libraries that are marked with distinct Illumina® paired-end sequencing adapters can be pooled for sequencing. The ratio at which the libraries are pooled determines the fraction of total reads available for each library.

Data analysis and generation of contact maps.

Reads from Illumina® paired-end sequencing of a Hi-C library are mapped to restriction fragments of the reference genome. Paired reads that correspond to self-circularized DNA molecules, unligated restriction fragments *(see note 6)*, and duplicates from PCR amplification of the adapter-ligated library on beads (step 19, NGS library preparation) are filtered out, and each valid junction is assigned to a pair of 'bins' by the midpoint coordinate of its corresponding restriction fragments. In this context, a 'bin' refers to a segment of a computationally partitioned genomic sequence that spans between one to several Kbs, for instance, the 4.64 Mb chromosome of *E. coli* can be partitioned into 464 bins spanning 10 kb. The interaction data is contained in symmetric square matrix where every value represents the interaction frequency between a pair of bins. The final, normalised contact map is obtained after balancing the matrix by equalising the sum of all its rows and columns. Detailed methods for processing Hi-C data are available in (15).

The resolution of a Hi-C contact map is determined by its bin size. Smaller bins increase the resolution of the map at the cost of a decreased signal-to-noise ratio. Thus, the selected bin size is a trade-off between the two factors and is heavily influenced by read depth. In our experience, 15 million valid junction reads are sufficient to generate a normalized contact map of the *E. coli* chromosome with a resolution of 5 kb.

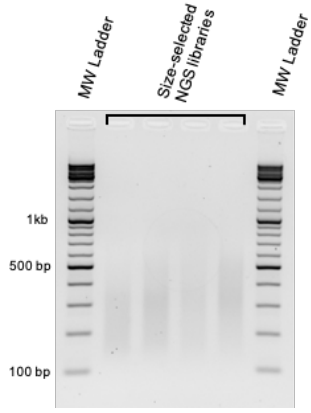


Figure 2.2: Typical run of a size-selected NGS library on a 2% agarose gel. A size-selected NGS library runs as a smear of DNA fragments between 150 – 600 bp.

Notes

1. 2.5 M Glycine should be stored at 20-25 °C. Glycine crystallises out of solution at lower temperatures.
2. Restriction enzymes define the resolution of the Hi-C library. The enzyme is selected based on the criterion that it produces a non-degenerate 5' overhang upon cleavage of its restriction site. This is necessary as 3' overhangs cannot be filled in, and degenerate sticky ends may interfere with biotin tagging. It is necessary that the enzyme functions optimally at 37 °C (or at an even lower temperature; for instance, the optimal temperature for CviAII activity is 25 °C) as restriction digestion at higher temperatures leads to premature reverse cross-linking of DNA-protein complexes. To ensure a good digestion efficiency during library preparation, it is important that the activity of the selected enzyme is not blocked by DNA methylation. The size distribution of the predicted restriction fragments generated by the enzyme should also be evaluated to ensure that the enzyme generates limited, if any, restriction fragments with sizes greater than the chosen Hi-C map resolution; long restriction fragments introduce biases in the contact matrix.

3. During the preparation of Hi-C libraries, restriction fragment overhangs are filled in with biotin-labelled nucleotides to tag ligation junctions generated during the subsequent proximity ligation step. Ensure that the fill-in of the 5' overhang generated by restriction digestion allows the incorporation of the selected biotin-labelled nucleotide. For instance, do not use biotin-14-dATP to fill-in overhangs generated by HpaII (C^CGG).
4. 25:24:1 phenol:chloroform:isoamyl alcohol and chloroform are neurotoxic organic solvents. Perform all steps requiring these solvents under a fume hood with necessary personal protective equipment.
5. Ensure that the *E. coli* cells are completely resuspended in 1X PBS prior to the addition of methanol. In an 80% methanol solution, even vigorous pipetting and vortexing only break a bacterial pellet into large clumps.
6. Some chromatin fragments may have been filled in with biotin-labelled nucleotides but may not have undergone ligation to give a chimeric product. Unligated 'dangling ends' that are bound by streptavidin-coated magnetic beads and thereafter sequenced reduce the number of useful reads per sequencing run. Biotin removal from such ends is performed with T4 DNA Polymerase under conditions that favor the 3' exonuclease activity of the enzyme in removing nucleotides past blunted ends. This condition is provided by maintaining the concentration of selected free nucleotides in the reaction mix below 100 µM. For instance, for the removal of biotin-14-dATP from the dangling ends of PsuI- or BglII-digested Hi-C libraries (RGA*TC), 100 µM of dGTP is added to the reaction mix, whereas dATP, dCTP, and dTTP are not (step 41 of Hi-C library preparation). This way, nucleotides up to but excluding G will be removed as a result of the activation of the 5' polymerase function of the enzyme at that site. The 3' exonuclease activity of T4 DNA Polymerase is also favored by increasing enzyme concentration, extending reaction times, and incubation at higher temperatures, for instance, 16 °C – as opposed to 12 °C – to favor the breathing of DNA ends.
7. Solid Phase Reverse Immobilization (SPRI) beads reversibly bind DNA fragments in solution in the presence of polyethylene glycol (PEG). The immobilization is size selective and depends on the concentration of PEG in the DNA-bead solution. Since PEG is present in the SPRI beads storage buffer (as supplied by the manufacturer), the size selective immobilization of DNA onto the SPRI beads is determined by the volumetric ratio rather than the concentration ratio of beads to DNA. Short DNA fragments can only be immobilized at high PEG concentrations. Therefore, to remove glycogen and excess salts from the HiC library for NGS library preparation, a 3.0x volumetric

ratio of beads to DNA solution is used (step 1 of NGS library preparation). Larger DNA fragments bind to the beads at lower PEG concentrations. Thus, a 0.5x volumetric ratio of beads to DNA solution is used to separate long DNA fragments from the sequencing library during NGS library preparation (step 24 of NGS library preparation).

8. It is important to ensure that >750 ng of a purified Hi-C library is available to begin NGS library preparation. Depending on the restriction enzyme used for Hi-C, and the efficiencies of restriction digestion, biotin-labelled nucleotide fill-in, ligation, removal of biotin labels from unligated restriction fragment ends and coupling to magnetic beads, a significant fraction of the starting material is lost during the enrichment of biotin-labelled ligation junctions with streptavidin-coated magnetic beads (steps 15 and 16 of NGS library preparation). Hence, the effective quantity of the library is much lower. Small quantities of starting material increase the number of PCR cycles that are necessary to sufficiently amplify the library for Illumina® sequencing, hence, increasing the fraction of duplicate reads in the NGS dataset. While duplicates are filtered out during processing, the fraction of 'useful' Hi-C reads is minimized.
9. These are the recommended number of cycles for 750 – 500 ng of input DNA as measured in step 3 of NGS library preparation.
10. 2.0x SPRI-based clean-up of the adapter-ligated Hi-C library amplification reaction purifies fragments up to several kbs in length. Fragments longer than 600 bp interfere with cluster formation during NGS and reduce the quality of the sequencing run. These fragments are removed from the library with a 0.5x SPRI clean-up. This purification is performed in two steps as opposed to a single 0.5x SPRI clean-up to eliminate the interference of isostabilisers present in the 2X KAPA HiFi HotStart Ready Mix on the immobilization of DNA on beads.
11. On a 2% agarose gel, a purified, size-selected NGS library runs as a smear between 150 – 600 bp (Figure 2.2). A 'laddered' pattern in the smear or a strong DNA signal at ~120 bp indicate the presence of PCR duplicates and adapter dimers, respectively, both of which consume a significant proportion of the read depth and reduce the quality of the sequencing run.

Acknowledgements:
Haloquadratum walsbyi HBSQ001 used to optimize Hi-C in *H. walsbyi* was a gift from Dr. Henk Bolhuis (Microbiology & Biogeochemistry, NIOZ, The Netherlands)

References:

1. Condensin- and Replication-Mediated Bacterial Chromosome Folding and Origin Condensation Revealed by Hi-C and Super-resolution Imaging. *Mol. Cell*.
2. Le,T.B.K., Imakaev,M. V., Mirny,L.A. and Laub,M.T. (2013) High-resolution mapping of the spatial organization of a bacterial chromosome. *Science (80-.).*, 10.1126/science.1242059.
3. Lioy,V.S., Cournac,A., Marbouty,M., Duigou,S., Mozziconacci,J., Espéli,O., Boccard,F. and Koszul,R. (2018) Multiscale Structuring of the E. coli Chromosome by Nucleoid-Associated and Condensin Proteins. *Cell*, 10.1016/j.cell.2017.12.027.
4. Le,T.B. and Laub,M.T. (2016) Transcription rate and transcript length drive formation of chromosomal interaction domain boundaries. *EMBO J.*, 10.15252/embj.201593561.
5. Tran,N.T., Laub,M.T. and Le,T.B.K. (2017) SMC Progressively Aligns Chromosomal Arms in Caulobacter crescentus but Is Antagonized by Convergent Transcription. *Cell Rep.*, 10.1016/j.celrep.2017.08.026.
6. Wang,X., Brandão,H.B., Le,T.B.K., Laub,M.T. and Rudner,D.Z. (2017) Bacillus subtilis SMC complexes juxtapose chromosome arms as they travel from origin to terminus. *Science (80-.).*, 10.1126/science.aai8982.
7. Dame,R.T., Rashid,F.-Z.M. and Grainger,D.C. (2020) Chromosome organization in bacteria: mechanistic insights into genome structure and function. *Nat. Rev. Genet.*, 10.1038/s41576-019-0185-4.
8. Dekker,J., Rippe,K., Dekker,M. and Kleckner,N. (2002) Capturing chromosome conformation. *Science (80-.).*, 10.1126/science.1067799.
9. Lieberman-Aiden,E., Van Berkum,N.L., Williams,L., Imakaev,M., Ragoczy,T., Telling,A., Amit,I., Lajoie,B.R., Sabo,P.J., Dorschner,M.O., *et al.* (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science (80-.).*, 10.1126/science.1181369.
10. Takemata,N., Samson,R.Y. and Bell,S.D. (2019) Physical and Functional Compartmentalization of Archaeal Chromosomes. *Cell*, 10.1016/j.cell.2019.08.036.
11. Nagano,T., Lubling,Y., Várnai,C., Dudley,C., Leung,W., Baran,Y., Mendelson Cohen,N., Wingett,S., Fraser,P. and Tanay,A. (2017) Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature*, 10.1038/nature23001.
12. Nagano,T., Lubling,Y., Stevens,T.J., Schoenfelder,S., Yaffe,E., Dean,W., Laue,E.D., Tanay,A. and Fraser,P. (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*, 10.1038/nature12593.
13. Stevens,T.J., Lando,D., Basu,S., Atkinson,L.P., Cao,Y., Lee,S.F., Leeb,M., Wohlfahrt,K.J., Boucher,W., O'Shaughnessy-Kirwan,A., *et al.* (2017) 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature*, 10.1038/nature21429.
14. Crémazy,F.G., Rashid,F.Z.M., Haycocks,J.R., Lamberte,L.E., Grainger,D.C. and Dame,R.T. (2018) Determination of the 3D genome organization of bacteria using Hi-C. In *Methods in Molecular Biology*.
15. Hofmann, A. and Heermann, D.W., 2018. Processing and Analysis of Hi-C Data on Bacteria. In *Bacterial Chromatin* (pp. 19-31). Humana Press, New York, NY.
16. Hsieh,T.H.S., Fudenberg,G., Goloborodko,A. and Rando,O.J. (2016) Micro-C XL: Assaying chromosome conformation from the nucleosome to the entire genome. *Nat. Methods*, 10.1038/nmeth.4025.

Chapter 3:
Regulation of *proVWX* transcription by local chromatin re-modelling

Rashid F-Z.M., Chaurasiya, K.R., Crémazy, F.G.E., Prins, N., Herdtfelder, A., Brocken, D.J.W., and Dame, R.T. Regulation of *proVWX* transcription by local chromatin re-modelling. *In preparation for publication.*

Contributions statement:
F.M.R. and F.G.E.C. optimized the Chromosome conformation capture-based studies and prepared Hi-C libraries. F.M.R. performed the 3C-qPCR and RT-qPCR experiments. K.R.C. set-up the TIRF microscope. K.R.C., A.H., and D.J.W.B. performed the live-cell imaging experiments. K.R.C. and N.P. performed the smFRET studies. A.H. and F.M.R. engineered the strains.

Abstract

Nucleoid associated proteins maintain the architecture of the bacterial chromosome and regulate gene expression, hinting that their role as transcription factors may involve local three-dimensional chromosome re-modelling. Here, we provide the first evidence to support this hypothesis. We use ensemble RT-qPCR and 3C-qPCR, in addition to *in vivo* and *in vitro* single molecule fluorescence microscopy to show that the expression of the H-NS-regulated, osmosensitive *proVWX* operon of *Escherichia coli* involves structural re-modelling of the operon. The formation of a loop anchored between the P2 promoter of *proVWX* and the terminus of the operon represses its expression. Destabilization of the loop activates transcription. The model presented here provides clues for how H-NS and H-NS-like proteins may regulate the expression of other operons and genes within their regulons.

Introduction

Nucleoid Associated Proteins (NAPs) are architectural proteins that bind along the bacterial chromosome driving its compaction and organization. They do so by DNA bending, lateral filament formation along the DNA, or DNA bridging, the latter of which results in the formation of long- and short-range DNA loops that determine the global and local structural organization of the genome (1–10). NAP binding is sensitive to environmental changes such as fluctuations in temperature, pH, and osmolarity. Hence, NAPs organize the bacterial chromosome into a dynamic structure that is re-modelled in response to changes in the cell environment. NAPs also function as transcription factors that coordinate global gene regulation in response to environmental stimuli. Collectively, these characteristics consolidate a model proposing that NAPs regulate gene expression by reorganizing the local structure of individual operons in response to stimuli that modulate their architectural properties.

Histone-like Nucleoid Structing protein (H-NS) of *Escherichia coli* is a 137 amino acid long NAP that functions as a global regulator of gene expression (11) and as a xenogeneic silencer, repressing transcription from horizontally transferred DNA (12). It exists as a dimer in solution (13), formed via interaction between a pair of N-terminal dimerization domains (14–16). The H-NS dimer binds to AT-rich double stranded DNA by means of a C-terminal DNA binding domain (17). The interaction serves as a nucleation point for the co-operative multimerization of H-NS along the DNA to form a protein-DNA filament. Structurally, the H-NS dimers are held together via interactions between central dimer-dimer interaction domains (18, 19). The structural conformation of H-NS and, effectively, its global gene regulation is sensitive to environmental stimuli. Its osmo-sensitivity, in particular, is modulated by a K^+- and Mg^{2+}-sensitive α-helix, helix α3 (20, 21), that extends between the N-terminal dimerization domain and central dimer-dimer interaction domain (18). *In vitro* and *in silico* studies indicate that the stabilization of the helix α3 by Mg^{2+} 'opens' the H-NS dimer exposing a pair of C-terminal DNA-binding domains to promote the formation of DNA bridges (1, 20, 22, 23) – structures that impede transcription elongation *in vitro* (24). Destabilization of the helix α3 by K^+ 'closes' the H-NS molecule by folding one DNA-binding domain per dimer onto the body of the protein (20). The availability of only one DNA-binding domain per dimer favors the formation of an H-NS-DNA filament (20), a structure conducive to transcription (24). In contrast to eukaryotic systems where a causal link between chromatin architecture and transcription has been established (25), *in vivo* evidence highlighting an interplay between local

three-dimensional chromatin organization and gene expression in prokaryotes is lacking. The dual role of H-NS as a transcription factor and an architectural protein makes it an attractive target for addressing this challenge. To that end, we examined the structural regulation of the *proVWX* operon.

proVWX (*proU*) (26, 27) is an H-NS-regulated, osmo-sensitive operon activated by the rapid cytoplasmic accumulation of K^+ (counter-ion: glutamate) that ensues as a primary response to hyper-osmotic stress (28–34). The acute increase in cytoplasmic K^+ prevents cell dehydration by drawing water back from the environment (35). This, however, occurs at the expense of protein stability and function, hence, cellular processes (28). K^+ also behaves as a second messenger that activates the cell's secondary 'long-term' responses to osmotic stress, such as, the activation of the *proU* operon. *ProU* activation upregulates the expression of three structural proteins – *proV*, *proW*, and *proX* – that assemble into a transmembrane protein complex of the ATP-binding cassette (ABC) superfamily of transporters (36). The proU transporter imports osmoprotectants such as glycine-betaine, proline-betaine, and proline, among others, into the cell with high affinity (27, 29, 30, 37, 38). The osmoprotectants maintain a low osmotic potential in the cytoplasm without an adverse effect on cellular physiology, allowing K^+ to be expelled from the cell (28, 39).

The osmosensitivity of the *E. coli proU* operon is inherent in the sequences of its σ^{70}-dependent promoter (P2), and cryptic (40, 41) σ^S-dependent promoter (P1), positioned 60 and 250 bp upstream of the *proV* open reading frame (ORF), respectively (32, 42, 43) (Supplementary file 3.1; Figure 3.1). The osmosensitivity is also inherent in the *cis* regulatory elements, P1R and P2R, extending across the promoters (43). Indeed, *proU* is activated by K^+ in the absence of trans-acting factor. The addition of increasing concentrations of K-glutamate, but not of L-proline and glycine betaine, to a purified *in vitro* transcription system of RNA polymerase and nucleotides stimulates the transcription of *proU* cloned on a plasmid DNA template from both, the P1 and P2 promoters (32, 44, 45). Non-osmoregulated genes such as *bla*, *lac*, and *pepN* are repressed under identical conditions (45). The -10 sequence of P2 is a non-consensus sequence carrying three GC base pairs that interfere with dsDNA melting and open complex formation. Increasing concentrations of K-glutamate may cause microstructure changes to the promoter that either increase its accessibility to RNAP or favour the isomerisation of the RNAP-bound promoter to form the open promoter complex (45).

Owing to Rho-dependent termination of transcription from P1 under ordinary growth conditions, *proVWX* expression is primarily driven by P2 (40, 41). Expression therefrom is regulated by the negative regulatory element (NRE) — a region extending from ~300 bp upstream of the *proV* ORF to ~1100 bp into the *proV* ORF (43) (Supplementary file 3.1; Figure 3.1). As a gene spanning 1203 bp in length, truncates of *proV* double as truncates of the NRE. β-galactosidase assays performed by generating in-frame *lacZ* fusions to the 5' end of *proV* truncates indicate that the osmoresponse of *proVWX* weakens as the NRE is shortened from its downstream end (43). The NRE exerts its role by means of H-NS (Figure 3.1), and, as of yet, has been shown to confer osmosensitivity only to *proU* P2. Heterologous promoters cloned in place of *proU* P2 do not exhibit osmoresponsivity (42, 43, 46, 47). A pair of *cis* regulatory elements occurs within the broad region designated as NRE: the upstream regulatory element (URE) positioned at -229 to -47 of the P2 transcription start site (TSS), and the downstream regulatory element (DRE) that stretches across P2 from -40 to +177 (32, 42, 43, 48, 49) (Supplementary file 3.1; Figure 3.1). An intrinsic curvature in the structure of the URE (Supplementary file 3.1; Figure 3.1) contributes to the osmotic inducibility of P2. The insertion of spacer sequences between P2 and the upstream curved DNA sequence show that the two elements must be 'positioned stereospecifically', that is, the inserted spacer must comprise (multiples of) a full turn of DNA for the promoter to be fully activated at high osmolarity (50). The URE exerts its role by means of H-NS — a NAP that preferentially binds curved DNA (51–54) (Figure 3.1). The URE acts cooperatively with the DRE in an hns^{WT} strain to strengthen the repression of *proU* in low osmolarity and enhance its activation at higher osmolarity (42). Correspondingly, the DRE consists of a pair of high-affinity H-NS binding sites that function as nucleation sites for the formation of an H-NS—DNA nucleoprotein complex (48, 55) (Supplementary file 3.1; Figure 3.1). *In vitro* studies support a hypothesis according to which at low osmolarity (i.e. low intracellular K^+), the H-NS—DNA complex organises into a transcriptionally-repressive bridged conformation to silence *proVWX* (20, 24, 42, 48, 49), reminiscent of the role of H-NS at the *rrnB* P1, and *hdeAB* promoters (56, 57). Cellular influx of K^+ upon a switch to a higher osmolarity may drive a K^+-mediated reorganisation of the H-NS—DNA complex into a transcriptionally-conducive unbridged structure to relieve *proVWX* repression (20, 24, 32, 44).

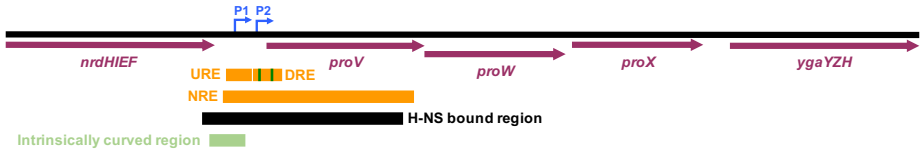


Figure 3.1: The regulatory elements of the *proVWX* operon. The *proVWX* operon comprises three ORFs: *proV*, *proW*, and *proX* (purple arrows), and is positioned between the *nrdHIEF* and *ygaGZH* operons (purple arrows) in the chromosome. The transcription of *proVWX* initates from the P1 and P2 (right angle arrows, blue) promoters recognised by σ^S and σ^{70}, respectively. The expression of the operon is regulated by the negative regulatory element (NRE, orange bar) that extends from ~300 bp upstream to ~1100 bp into the *proV* ORF. The NRE carries the upstream regulatory element (URE, orange bar) that forms part of the intrinsically curved region (light green bar), and the downstream regulatory element (DRE, orange bar). The green bars within the DRE illustrate the positions of high-affinity H-NS binding sites at which H-NS is predicted to bind and nucleate the formation of an H-NS—DNA nucleoprotein structure that spans the NRE (black bar).

In addition to H-NS, the NAPs IHF and HU also regulate *proU* (58–61). Genetic screening for mutants with decreased *proVWX* expression identified IHF and HU as positive regulators of *proVWX* expression (58, 59). IHF increases the expression of *proU* but does not affect the osmo-inducibility of the operon (58). IHF mediates its effect via a putative binding site positioned ~450 bp upstream of P2 (58), and the specific binding site of the protein in the -33 to +25 region around the P2 TSS (62). Due to limited study, it is unclear whether *proVWX* activation as a result of upstream IHF binding is affected through P1 or P2 (58, 59), however, the specific binding of IHF at P2 structurally modifies the promoter to partially relieve H-NS-mediated repression at low osmolarity (62). The precise molecular mechanism of the structural changes induced are unclear. The effect of *ihf* deletion on the expression of the *proVWX* operon was not observed in genome-wide studies (60), likely due to the modest decrease in the expression of the operon at low-osmolarity conditions in the absence of IHF (58). HU regulates expression from, and the osmoinducibilty of the proU P2 promoter (59). Δ*hupA*, Δ*hupB*, and Δ*hupAhupB* mutants show reduced proU-lac expression compared to the corresponding wild-type strain, and a reduced fold-change in expression from P2 between low and high osmolarity conditions (59). HU plays its role independently of H-NS as evidenced from Δ*hupB* Δ*hns* double mutants, where expression from P2 is realised as sum of the individual mutations rather than an epistatic output (59). The *proU*-lac expression results presented for Δ*hupA*, Δ*hupB*, and Δ*hupAhupB* in (59) are in conflict with genome-wide transcriptome data gathered in (61) which indicate that a Δ*hupA* mutation derepresses *proU* during the exponential phase of growth, Δ*hupB* derepresses the operon during transition-to-stationary phase, and Δ*hupAhupB* in stationary phase (61). However, the detection of *proU* expression at the transcriptional level from its native locus in the chromosome in (61) and

post-translationally from a plasmid-localised *proU-lacZ* gene in (59) may account for the discrepancy. The ectopic placement of a non-native *proU-lacZ* construct eliminates the effects of relevant regulatory elements and perturbs local topological states that impact *proU* expression either directly, or indirectly by affecting the binding of regulatory NAPs. StpA, an H-NS paralogue (58% sequence identity) (63, 64) that has been proposed to function as a 'molecular back-up' for H-NS (65) represses proV-lac 9-fold when expressed from multi-copy plasmids in *E. coli* Δ*hns* (66). Chromatin Immunoprecipitation (ChIP) of FLAG-tagged StpA shows that in wild-type *E. coli* and in *E. coli* Δ*hns*, StpA binds the *proVWX* regulatory region and its occupancy spans over P1 and P2 (67). StpA also binds the *proV* ORF, however, only as a hetero-oligomeric complex with H-NS (67). A ChIP signal of FLAG-tagged StpA is not observed at *proV* in *E. coli* Δ*hns* (67). Conversely, FLAG-tagged H-NS is detected at *proV* in *E. coli* Δ*stpA* (67).

The transcription of the *proVWX* operon is also regulated by genetic and chemical factors that modify the supercoiling density of the genome (43, 44, 68, 69). Mutations to DNA gyrase, and growth in media supplemented with non-lethal concentrations of novobiocin, a DNA gyrase inhibitor, reduce the levels of negative supercoiling in the cell and the expression of *proVWX* (68, 70). Mutations in *topA*, and growth in high osmolarity media increase negative supercoiling density and *proVWX* expression (68). However, extremely high levels of negative supercoiling observed in Δ*topA* mutants repress the operon (68). Furthermore, in a purified *in vitro* transcription system of RNA polymerase and nucleotides, transcription from the P2 promoter increases with increasing negative supercoiling density of the template (44). Increased negative supercoiling may favour transcription by facilitating the melting of the relatively GC-rich P2 promoter and promoting a structural reorganisation of the H-NS—NRE nucleoprotein complex (39).

Collectively, the coordinated role of an inherently osmo-sensitive promoter (32, 42, 43), supercoiling density, and architectural proteins – HU (59), IHF (58, 59, 62), StpA (66, 71), and the osmosensitive H-NS (20, 32, 42, 48, 49) – in the activation and repression of the *proU* operon from the P2 promoter implies that the regulation of *proVWX* may involve local three-dimensional reorganization of chromatin structure, analogous to eukaryotic chromatin remodelling. Here, we use 3C-qPCR and single-molecule TIRF microscopy to visualize the structural dynamics of *proVWX* in low- and high-salt conditions, and in response to a hyper-osmotic shock.

Materials and methods

Strains and plasmids

Bacterial strains used in this study were derived from *Escherichia coli* MG1655 using a one/two-step genome editing method based on the λ-red recombinase system (72) as described in (73). Chemocompetent bacteria were transformed with pKD46 (72), a temperature-sensitive plasmid encoding the λ-red proteins under an arabinose-inducible promoter. Electrocompetent pKD46+ cells were prepared from a culture grown at 30 °C in LB medium (1.0% bactotryptone (BD), 0.5% yeast extract (Alfa Aesar), 170 mM NaCl, pH 7.5) supplemented with 10 mM arabinose, for plasmid maintenance and to induce the expression of the λ-red proteins, respectively. The cells were transformed with a kanR-ccdB cassette (73) encoding kanamycin resistance and the ccdB toxin under a rhamnose-inducible promoter (72). The cassette was designed to carry ~1.5 kb extensions on either side homologous to the flanks of the genomic region to be edited. kanRccdB was amplified from plasmid pKD45 (72) and the homology regions from genomic DNA. The complete construct was assembled using overlap extension PCR (oePCR). Recombinants were selected for on LB agar (1.0% bactotryptone (BD), 0.5% yeast extract (Alfa Aesar), 170 mM NaCl, 1.5% bacteriological agar (Oxoid), pH 7.5) supplemented with 40 µg/mL kanamycin at 37 °C and verified using colony PCR and Sanger sequencing (BaseClear B.V.). The mutant allele construct with ~1.5kb homology regions was assembled into a plasmid using Gibson assembly (74). The allele was amplified and used to replace the genomic kanRccdB cassette via the λ-red recombinase system. Recombinants were selected for on M9 agar plates (42 mM Na_2HPO_4, 22 mM KH_2PO_4, 19 mM NH_4Cl, 8.5 mM NaCl, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 1.5% bacteriological agar (Oxoid)) supplemented with 1% rhamnose and verified with colony PCR and Sanger sequencing (BaseClear B.V.). The lac::CamR mutation was performed in a single step using a chloramphenicol resistance cassette carrying ends homologous to the left and right flanks of the *lac* operon. Selection was carried out on LB agar with 12.5 µg/mL chloramphenicol. All bacterial strains and intermediates were stored at -80 °C as glycerol stocks. Strains and plasmids used in this study are listed in Tables 3.1 and 3.2. Detailed protocols are available upon request.

Table 3.1: List of strains

Strain ID	Description	Reference
MG1655	*Escherichia coli* K-12	
NT331	MG1655 ΔendA	This chapter
NT453	MG1655 ΔendA ProU_terminus::lacO pProU::tetO-pProU	This chapter
NT455	MG1655 ΔendA ProU_terminus::lacO pProU::tetO-pProU pRD183	This chapter
NT606	MG1655 ΔendA; pKD46 (AmpR)	This chapter

NT607	MG1655 ΔendA lac::CamR	This chapter
NT617	MG1655 ΔendA lac::CamR; pKD46 (AmpR)	This chapter
NT618	MG1655 ΔendA lac::CamR ProU::kanRccdB	This chapter
NT620	MG1655 ΔendA lac::CamR ProU::tetO-proU-lacO	This chapter
NT623	MG1655 ΔendA lac::CamR ProU::kanRccdB; pRD418 (SpnR)	This chapter
NT624	MG1655 ΔendA lac::CamR ProU::kanRccdB; pRD419 (SpnR)	This chapter
NT625	MG1655 ΔendA lac::CamR ProU::tetO-pProU_GC-proU-lacO	This chapter
NT626	MG1655 ΔendA lac::CamR ProU::kanRccdB; pRD418 (SpnR) pRD310 (AmpR)	This chapter
NT627	MG1655 ΔendA lac::CamR ProU::kanRccdB; pRD418 (SpnR) pRD373 (AmpR)	This chapter
NT629	MG1655 ΔendA lac::CamR ProU::kanRccdB; pRD419 (SpnR) pRD310 (AmpR)	This chapter
NT630	MG1655 ΔendA lac::CamR ProU::kanRccdB; pRD419 (SpnR) pRD373 (AmpR)	This chapter
NT632	MG1655 ΔendA rnc::kanRccdB	This chapter
NT633	MG1655 ΔendA stpA::kanRccdB	This chapter
NT638	MG1655 ΔendA lac::CamR ProU_terminus::kanRccdB	This chapter
NT639	MG1655 ΔendA lac::CamR ProU_terminus::kanRccdB; pKD46 (AmpR)	This chapter
NT640	MG1655 ΔendA lac::CamR ProU_terminus::lacO	This chapter
NT641	MG1655 ΔendA lac::CamR ProU_terminus::lacO; pKD46 (AmpR)	This chapter
NT642	MG1655 ΔendA pProU::kanRccdB; pKD46 (AmpR)	This chapter
NT643	MG1655 ΔendA lac::CamR ProU_terminus::lacO pProU::kanRccdB; pKD46 (AmpR)	This chapter
NT644	MG1655 ΔendA pProU::pProU_GC	This chapter
NT645	MG1655 ΔendA lac::CamR ProU_terminus::lacO pProU::tetO-pProU	This chapter
NT646	MG1655 ΔendA lac::CamR ProU_terminus::lacO pProU::tetO-pProU_GC	This chapter
NT647	MG1655 ΔendA lac::CamR ProU_terminus::lacO pProU::tetO-pProU; pRD310 (AmpR)	This chapter
NT648	MG1655 ΔendA lac::CamR ProU_terminus::lacO pProU::tetO-pProU; pRD373 (AmpR)	This chapter
NT658	MG1655 ΔendA; pRD310 (AmpR)	This chapter
NT659	MG1655 ΔendA; pRD373 (AmpR)	This chapter
NT660	MG1655 ΔendA pProU::pProU_GC; pRD310 (AmpR)	This chapter
NT661	MG1655 ΔendA pProU::pProU_GC; pRD373 (AmpR)	This chapter
NT662	MG1655 ΔendA lac::CamR ProU_terminus::lacO pProU::tetO-pProU_GC; pRD310 (AmpR)	This chapter
NT663	MG1655 ΔendA lac::CamR ProU_terminus::lacO pProU::tetO-pProU_GC; pRD373 (AmpR)	This chapter

Table 3.2: List of plasmids

Plasmid ID/ Strain ID	Backbone	Insert	Resistance	Storage strain	Reference
pKD45/ XT198	pKD45	kanR P_{rhaB}-ccdB	Kanamycin	HCB1666	(72)
pKD46/ XT146	pKD46	araC P_{araB}-λ Red	Ampicillin	MG1655 ΔthyA	(72)
pTargetF/ XT205	pTarget	sgRNA (cadA)	Streptomycin, Spectinomycin	DH5α	Addgene #62226 (75)
pRD178/ NT279	pMK	6TetO hybrid array (tetO)	Kanamycin	KA797	This chapter

pRD179/ NT280	pMK	6LacO hybrid array (lacO)	Kanamycin	KA797	This chapter
pRD183/ NT304	pBAD24	LacI-mCherry-tetR-YFP	Ampicillin	DH5α	This chapter
pRD258/ NT397	pTarget	ProU_terminus-lacO	Streptomycin, Spectinomycin	DH5α	This chapter
pRD259/ NT398	pTarget	tetO-pProU	Streptomycin, Spectinomycin	DH5α	This chapter
pRD310/ NT477	pBAD24	J23101_promoter-LacI_mCherry-TetR_eYFP-rrnB_T1_terminator	Ampicillin	DH5α	This chapter
pRD359/ NT547	pTarget	pProU_GC	Streptomycin, Spectinomycin	DH5α	This chapter
pRD373/ NT574	pBAD24	J23101_promoter-LacI_mCherry-λtL3_terminator-J23101_promoter-TetR_eYFP-rrnB_T1_terminator	Ampicillin	DH5α	This chapter
pRD391/ NT591	pTarget	lac::CamR	Streptomycin, Spectinomycin, Chloramphenicol	DH5α	This chapter
pRD417/ XT235	pHA2	eGFP	Kanamycin, Chloramphenicol	DH5α	This chapter
pRD418/ NT621	pHA2	tetO-ProU-lacO	Streptomycin, Spectinomycin	KA797	This chapter
pRD419/ NT622	pHA2	tetO-pProU_GC-ProU-lacO	Streptomycin, Spectinomycin	KA797	This chapter

Media and growth conditions

For low-salt and hyper-osmotic shock conditions, a single bacterial colony from a freshly streaked plate was grown overnight at 37 ºC in low-salt LB medium (LS-LB: 1.0% bactotryptone (BD), 0.5% yeast extract (Alfa Aesar), 80 mM NaCl, pH 7.5) or low-salt M9 medium (LS-M9: 42 mM Na_2HPO_4, 22 mM KH_2PO_4, 19 mM NH_4Cl, 2.0 mM $MgSO_4$, 0.1 mM $CaCl_2$, 80 mM NaCl, 1X trace elements, 1% Bacto™ casamino acids (BD), 10 µg/mL Thiamine (Sigma-Aldrich), 0.4% glycerol (PanReac Applichem)). The overnight culture was used to inoculate fresh low-salt LB or M9 medium to a starting OD_{600} of 0.05 for up to four biological replicates and grown at 37 ºC to an OD_{600} of ~1.0. The culture was split into two aliquots. For hyper-osmotic shock, 5.0 M NaCl was added to one of the aliquots at a ratio of 46 µL per mL of culture. An equivalent volume of milliQ water was added to the second aliquot for the low salt condition. The cultures were grown for 10 minutes at 37 ºC, and then immediately harvested for RNA isolation, cell fixation for 3C-based experiments, and microscopy. For high-salt studies, a single bacterial colony was grown overnight in high-salt LB medium (LS-LB: 1.0% bactotryptone

(BD), 0.5% yeast extract (Alfa Aesar), 300 mM NaCl, pH 7.5) or high-salt M9 medium (HS-M9: 42 mM Na_2HPO_4, 22 mM KH_2PO_4, 19 mM NH_4Cl, 2.0 mM $MgSO_4$, 0.1 mM $CaCl_2$, 300 mM NaCl, 1X trace elements, 1% Bacto™ casamino acids (BD), 10 µg/mL Thiamine (Sigma-Aldrich), 0.4% glycerol (PanReac Applichem)). The overnight culture was used to inoculate fresh high-salt LB or M9 medium to an OD_{600} of 0.05 for up to four biological replicates. The cells were cultured at 37 °C to an OD_{600} of ~1.0, and for an additional 10 minutes. The cells were immediately harvested for RNA isolation, cell fixation, and microscopy.

Trace elements were prepared as a 100X stock solution of the following composition per 100 mL: 0.1 g $FeSO_4 \cdot 7H_2O$, 0.6 g $CaCl_2 \cdot 2H_2O$, 0.12 g $MnCl_2 \cdot 4H_2O$, 0.08 g $CoCl_2 \cdot 6H_2O$, 0.07 g $ZnSO_4 \cdot 7H_2O$, 0.03 g $CuCl_2 \cdot 2H_2O$, 2 mg H_3BO_3, and 0.5 g EDTA·Na_2.

RNA isolation and handling

Bacterial cells in 1 mL of culture (Media and growth conditions, Materials and methods) were collected by centrifugation at 13,000 xg for 2 minutes. The supernatant was removed and the pellet was resuspended in 200 µL of Max Bacterial Enhancement Reagent (TRIzol® Max™ Bacterial RNA Isolation Kit, Ambion®, life Technologies™) pre-heated to 95 °C. The lysate was incubated at 95 °C for 5 minutes. The preparation was treated with 1 mL of TRIzol® reagent (Ambion®, life Technologies™) and incubated at room temperature for 5 minutes. RNA isolation was paused at this step to accommodate cell fixation for 3C and Hi-C library preparation by flash-freezing the TRIzol®-lysate mix in liquid nitrogen and storing at -80 °C for up to 2 weeks. The TRIzol®-lysate mix was thawed at room temperature and RNA isolation was continued using the TRIzol® Max™ Bacterial RNA Isolation Kit (Ambion®, life Technologies™) according to the manufacturer's instructions. The concentration and purity (A_{260}/A_{280}) of RNA was measured with a NanoDrop™ 2000 spectrophotometer (Thermo Scientific™), and accordingly, RNA samples were diluted to a final concentration of 20 ng/µL with RNase-free water (Gibco®, life Technologies™). DNA contamination in the RNA samples was checked with RNase A (Qiagen) treatment and agarose gel electrophoresis. 50 µL of the RNA samples were transferred into wells of a 96-well RNase/DNase-free plate in triplicate to facilitate multi-channel pipetting for RT-qPCR experiments. RNA samples in the multi-well plate were stored at 4 °C and placed on ice during reaction set-up to avoid freeze-thaw cycles. Stock RNA samples were stored at -20 °C.

RT-qPCR

Primer design

Primers for RT-qPCR experiments were designed using the *Escherichia coli* K-12 MG1655 sequence (Accession number: NC_000913.3). *In silico* specificity screening was performed on SnapGene® Viewer 5.2. RT-qPCR primers were ordered as dried pellets from Sigma-Aldrich. The oligonucleotides were dissolved in 1X TE pH 8.0 to a final concentration of 100 µM and stored at 4 °C. The list of primers used for the RT-qPCR assay and corresponding amplicon details are provided in Supplementary file 3.2. Primer annealing sites and amplicon positions in the *proVWX* operon are provided in Supplementary files 3.1 and 3.2. The specificity of primer pairs was experimentally determined with Sanger sequencing (BaseClear B.V.) of the amplified PCR product, and melting curves.

Reaction set-up

RT-qPCR experiments were performed with the Luna® Universal One-Step RT-qPCR Kit (NEB, #E3005E) with a final reaction volume of 10 µL per well. The reaction composition is provided in Table 3.3. The reactions were set-up manually in Hard-Shell® 96-well PCR plates (Bio-Rad, #HSP9635) and run on a C1000 Touch™ thermal cycler heating block (Bio-Rad) with a CFX96™ Real-time system (Bio-Rad). The thermocycling parameters are provided in Table 3.4. NT331 genomic DNA (Supplementary methods) serially diluted in triplicate to 10X, 100X, 1000X, and 10000X dilution factors was used for the standard samples. Standard samples, test samples, and No Template Controls were processed in the same manner.

Table 3.3: RT-qPCR reaction composition per well. Modified from the manufacturer's protocol provided with the Luna® Universal One-Step RT-qPCR Kit.

Component	Volume in a 10.0 µL reaction	Final concentration
‡2X Luna Universal One-Step Reaction Mix (NEB)	5.0 µL	1X
‡20X Luna WarmStart® RT Enzyme Mix (NEB)	0.5 µL	1X
‡100 µM forward primer (Supplementary files 3.1 and 3.2)	0.04 µL	0.4 µM
‡100 µM reverse primer (Supplementary files 3.1 and 3.2)	0.04 µL	0.4 µM
‡Nuclease free water (Gibco®, life Technologies™)	3.92 µL	N/A
~20 ng/µL RNA sample	0.5 µL	~10 ng

‡ These components were added to each reaction as a master mix.

Table 3.4: RT-qPCR thermal cycling parameters. As per the manufacturer's instructions provided with the Luna® Universal One-Step RT-qPCR Kit.

No.	Step	Temperature	Duration (mm:ss)	Cycles
1	Reverse transcription	55 °C	10:00	1

2	Initial denaturation	95 °C	01:00	1
3	Denaturation	95 °C	00:10	45
4	Annealing/extension	60 °C	00:30	
5	Plate read (SYBR/FAM channel)	N/A	N/A	
6	Melt	60 °C	0:31	1
7	Ramp	60 °C	0:01	70
		+ 0.5 °C/cycle	+ 0:01/0.5 °C	
8	Plate read (SYBR/FAM channel)	N/A	N/A	

Cell fixation

6.0 mL of the bacterial culture in M9 medium (See media and growth conditions) was collected by centrifugation at 3000 xg for 5 minutes at 4 °C. The supernatant was discarded and the pellet was resuspended in 2 mL of LS-M9 medium for the low salt cultures or 2 mL of HS-M9 medium for the high salt and salt shock cultures. The cell suspension was treated with 8 mL of ice-cold methanol and incubated at 4 °C for 10 minutes. The cells were collected by centrifugation at 3000 xg for 5 minutes at 4 °C and washed with 20 mL of LS/HS-M9 medium. The pellet of washed, methanol-treated cells was resuspended in 3% formaldehyde in LS/HS-M9 medium. The reaction was incubated at 4 °C for 1 hour and subsequently quenched with a final concentration of 0.375 M glycine for 15 minutes at 4 °C. The fixed cells were collected by centrifugation and washed twice with 1X TE. The cell suspension was split into three aliquots and pelleted at 10000 xg at room temperature for 2 minutes. The supernatant was removed and the cell pellet was flash-frozen in liquid nitrogen for storage at -80 °C until use (up to 1 month).

Bacterial cultures grown in LB were fixed in the same manner, except that all wash steps and 3% formaldehyde preparation were performed with 1X LS-PBS (2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4), 1X HS-PBS (280 mM NaCl, 2.7 mM KCl, 10 mM mM Na_2HPO_4, 1.8 mM KH_2PO_4), or 1X PBS (137 mM NaCl, 2.7 mM KCl, 10 mM mM Na_2HPO_4, 1.8 mM KH_2PO_4) as required.

Hi-C and 3C library preparation

Hi-C libraries were prepared as described in (76). For 3C libraries, the Hi-C library preparation section in (76) was modified by eliminating 'Fill-in with biotin-labelled nucleotides' and followed up to step 40. NlaIII (NEB) was used as the restriction enzyme.

3C library handling

The concentration of the 3C libraries was measured with the Qubit™ dsDNA HS Assay Kit (Themo Fisher Scientific™, #Q32854) using the Qubit 2.0 fluorometer

(ThermoFisher Scientific™). The A_{260}/A_{280} ratio was not determined since the presence of glycogen in the library preparation interfered with absorbance at these wavelengths. The 3C libraries were diluted to a final concentration of 2 ng/µL with nuclease-free water (Invitrogen™) and transferred into wells of a 96-well RNase/DNase-free plate in triplicate to facilitate multi-channel pipetting for 3C-qPCR set-up. 3C libraries in the multi-well plate were stored at 4 ºC and placed on ice during reaction set-up to avoid freeze-thaw cycles. Library stocks were stored at -20 ºC.

3C-qPCR

The 3C-qPCR protocol was designed according to (77).

Primer and TaqMan probe (hydrolysis probe) design

Primers and TaqMan probes for 3C-qPCR experiments were designed using the Escherichia coli K-12 MG1655 sequence (Accession number: NC_000913.3) on SnapGene® Viewer 5.2. SnapGene® Viewer was also used for *in silico* specificity screening. For the fragments of interest for 3C-qPCR test reactions and cross-linking controls, primers and TaqMan probes were designed on opposite DNA strands within 100 bp of the downstream NlaIII site, with TaqMan probes positioned closer to the restriction site than the primers. All primers were designed to anneal to the genomic DNA with the same directionality so as to ensure that the signal observed during 3C-qPCR assays arises only as a result of re-ligation between the pair of fragments being tested and not due to incomplete restriction digestion. Furthermore, the hybridisation of the TaqMan probe to the non-primed strand of a fragment ensures that the hydrolysis of the probe, and hence, the detection of a fluorescence signal in qPCR only occurs if amplification crosses the ligation junction being tested.

For 3C-qPCR loading controls, primer pairs amplifying a ~100 bp region lacking an NlaIII digestion site were designed. For these amplicons, the TaqMan probe with the lowest predicted melting temperature for its secondary structures, as determined using OligoAnalyzer (IDT), was selected.

Double-quenched TaqMan probes (HPLC purified) with a 5' 6-FAM fluorophore, 3' Iowa Black™ Fluorescence Quencher, and an internal ZEN quencher positioned 9 bases from the fluorophore were ordered as dried pellets from IDT. Primers (desalted, dry) were ordered from IDT or Sigma-Aldrich. The oligonucleotides were dissolved in 1X TE pH 8.0 to a final concentration of 100 µM and stored at

4 °C. The lists of primers and probes used for 3C-qPCR, their positions within the *proU* operon and their annealing sites on the chromosome are provided in Supplementary files 3.1 and 3.3.

Control library preparation

Three sets of control libraries were prepared for 3C-qPCR: digested and randomly re-ligated NT331 genomic DNA, digested and randomly re-ligated NT644 gDNA, and a synthetic control template where chimeric fragments of interest were separately prepared by PCR and pooled in equimolar ratios. The control libraries were serially diluted to prepare the standard samples for 3C-qPCR.

Reaction set-up

3C-qPCR (77) was performed using the PrimeTime™ Gene Expression Master Mix Kit (Integrated DNA Technologies, #10557710) using 10 ng of the 3C libraries, corresponding to ~2 x 10^6 genome equivalents, as the template. A detailed reaction composition is provided in Table 3.5. Experiments were set-up manually in Hard-Shell® 96-well PCR plates (Bio-Rad, #HSP9635) and run on a C1000 Touch™ thermal cycler heating block (Bio-Rad) with a CFX96™ Real-time system (Bio-Rad) using the program outlined in Table 3.6. Standard, Test, and No Template Control samples were processed in the same manner.

Table 3.5: 3C-qPCR reaction composition per well. Modified from the manufacturer's protocol provided with the PrimeTime® Gene Expression Master Mix.

Component	Volume per reaction	Final concentration
‡2X PrimeTime® Gene Expression Master Mix (IDT)	5.0 µL	1X
‡100 µM TaqMan probe (IDT)	0.015 µL	0.15 µM
‡100 µM Constant primer (Supplementary file 3.3)	0.05 µL	0.5 µM
‡100 µM Test primer (Supplementary file 3.3)	0.05 µL	0.5 µM
~2 ng/µL 3C library	5.0 µL	~10 ng

‡ These components were added to each reaction as a master mix.

Table 3.6: 3C-qPCR thermal cycling parameters. As per the manufacturer's instructions provided with the PrimeTime® Gene Expression Master Mix.

No.	Step	Temperature	Duration (mm:ss)	Cycles
1	Polymerase activation	95 °C	03:00	1
3	Denaturation	95 °C	00:15	45
4	Annealing/extension	60 °C	01:00	
5	Plate read (SYBR/FAM channel)	N/A	N/A	

qPCR Data Analysis

RT- and 3C-qPCR data were processed using the Bio-Rad CFX Manager 3.1 program (Bio-Rad) and analysed with Microsoft Excel. All the raw data files, and

the exported .csv files are available in Supplementary folders 3.1-3.4. For all RT-qPCR experiments, the cycle of quantification (Cq) was determined using the Single Threshold option available on Bio-Rad CFX Manager 3.1. For 3C-qPCR, Cq was determined using the Regression function. Reactions with unreliable Cq values as a result of pipetting errors owing to the manual set-up of the experiment, or evaporation from improperly sealed wells, were eliminated from analysis. For the intercalator-based RT-qPCR experiment, melt curves were used to check the specificity of amplification, and hence, the validity of the experiment. Eliminated reactions and the justifications for each are provided in the Supplementary files 3.4-3.8.

The Cq values of the standard samples were used to plot a standard curve of Cq against the logarithm of the starting quantity of the template. Standard samples with pipetting errors were eliminated from the plot. To verify that quantification was performed in the linear dynamic range, it was ensured that the standard curve was linear (R^2 >0.95) and that the range of Cq values spanned by the standard curve encompassed the Cq values of the test reactions as much as possible. The slope, y-intercept, PCR efficiency, and R^2 value were extracted from the standard plot.

The Cq values of the technical replicates for each biological sample were used to determine the average Cq value (Cq_{avg}) and standard deviation of Cq (σCq) for each amplicon. Cq_{avg} was used to determine the relative expression level of the amplicons in RT-qPCR experiments and the relative interaction frequency in 3C-qPCR. σCq was used to determine the negative and positive errors for the biological samples. The calculations were done using the formula: 10^((Cq-Intercept)/Slope). This processing was applied for all Test amplicons and Internal control candidates (Supplementary files 3.4-3.8).

The relative expression levels and relative interaction frequencies of the Test amplicons were normalised to an Internal control amplicon to allow comparisons to be drawn between biological samples. This was done by assigning an arbitrary value of 100 to the relative expression level or interaction frequency of the selected Internal control. The values corresponding to the Test amplicons for all biological samples were re-calculated accordingly. Since all amplicons were quantified from standard curves plotted from the same standard samples, inter-assay comparisons were made possible.

Live cell FROS imaging

Slide preparation

800 µL of molten 3% (w/v) low-melt agarose dissolved in HS- or LS-M9 medium, as necessary, was applied on the surface of a microscope slide thoroughly cleaned with 70% ethanol. A pair of spacer slides were placed over the top and bottom edges of the microscope slide and a clean cover-slip was placed on top to mould the molten agarose into an agarose pad with a flattened surface. The spacer slides and the coverslips were removed after the agarose solidified and 5.0 µL of an *E. coli* culture concentrated to ~1/50th of its culture volume was pipetted onto the centre of the pad. The agarose pad was covered with a clean coverslip and sealed with nail polish.

Imaging

The cells were positioned at the centre of the imaging region of the TIRF microscopy set-up described in (78) using bright-field illumination, and brought into optimal focus for fluorescence imaging using the bottom half of the cell to focus the objective. The intensities of the excitation lasers were modulated to ~200 W/cm^2 with the shutters closed to prevent initial photobleaching. After opening of the shutters a continuous series of images was recorded until the fluorescent molecules bleached. ImageJ was used to identify the position of the centroids of fluorescent spots in the image series and link the positions to individual trajectories. The distance between the pair of FROS labels was determined every 4.2 seconds (frame rate: 14.3 frames/minute) for the recorded cells.

In vitro single molecule FRET

The 250 bp AT-rich construct and the set of 7 fragments for *proU* bridging experiments were obtained with PCR. A 685 bp (32% GC) molecule described in (79), and NT331 genomic DNA were used as the templates, respectively. The primers were fluorescently labelled with Cy3 or Cy5 and one primer of each pair was functionalized at the 5' end with BiotinTEG (Sigma-Aldrich). The amplified constructs were purified using the GFX^{TM} PCR DNA and Gel Band purification kit (GE Healthcare) and eluted in 1X TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA). The substrates for IHF-mediated DNA bending studies were prepared by annealing a pair of complementary single stranded oligonucleotides in annealing buffer (10 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM EDTA). For each pair one oligonucleotide was labelled with 5' Cy5 and the other with 5' Cy3 and 3' BiotinTEG. Primer

sequences are provided in Table S3.1. DNA substrates were serially diluted in annealing buffer to 25 pM for smFRET. SmFRET was performed using total internal reflection fluorescence (TIRF) microscopy using the set-up described in (78).

The fluorescent bead reference was used to calibrate the microscopy set-up prior to each experiment. The beads were excited with the 532 nm laser at 1 mW and 0 EM gain. The test sample was excited with the 532 nm laser at 20-50 mW and 700 EM gain. Movies of ~4000 frames (33 ms/frame) were recorded. A custom script written in IDL (78) was used to align the donor and acceptor images and to extract fluorescence intensities for single molecules of the test sample. A MatLab script (78) was used to validate the single-molecule traces from single-step photobleaching of the Cy3/Cy5 dyes, perform base-line corrections, and manually analyze the single-molecule traces. FRET efficiencies of selected traces were plotted in histograms using Igor Pro 7 (WaveMetrics, Inc., Lake Oswego, OR, USA), and fitted by single peak or multipeak fitting with gaussian distributions to determine the distribution of FRET states in the population.

Results

Degradation of *Escherichia coli* chromatin during 3C-based library preparation is overcome by the deletion of *endA*.

The *Escherichia coli* K-12 MG1655 strain closely resembles the genetic make-up of archetype *E. coli* and is used as a reference for genome-wide studies of NAP-binding profiles and transcription. The strain is, therefore, the optimal choice to study the interplay between three-dimensional chromatin organisation, NAP distribution, and gene expression. However, chromatin extracted from MG1655 underwent considerable degradation during the initial steps of 3-C and Hi-C library preparation. The degradation was not observed during the lysis and solubilisation steps (76, 80) that were carried out in a buffer with 1.0 mM EDTA, but occurred extensively once the cell lysate was diluted in a restriction digestion mix with a final EDTA concentration of 0.1 mM (Figure 3.2A). The dependence of chromatin degradation on the concentration of EDTA, and hence, the availability of divalent ions implied that the degradation was enzymatic.

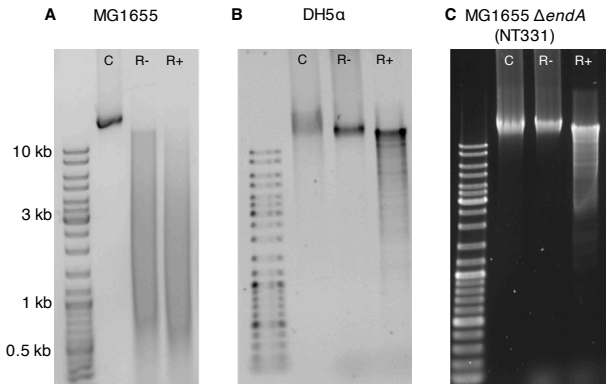


Figure 3.2: Degradation of *Escherichia coli* chromatin during 3C-based library preparation is overcome by *endA* deletion. Lane C: Chromatin preparation in 1X TE (EDTA concentration: 1.0 mM); **Lane R-:** Chromatin preparation after a 3-hour incubation in 1X restriction digestion buffer (EDTA concentration: 0.1 mM); **Lane R+:** Chromatin preparation after a 3-hour treatment with 0.4 U/µL of BglII in 1X restriction digestion buffer (EDTA concentration: 0.1 mM). **MG1655 (panel A)** chromatin undergoes extensive degradation in a restriction digestion buffer with a final EDTA concentration of 0.1 mM. The degradation is observed as a smear in lanes R- and R+ of panel A. Similar degradation is not observed in *endA* knock-out strains, **DH5α (panel B)** and **NT331 (panel C)**, where extracted chromatin (lane C) still runs as a heavy >10 kb band after a 3-hour incubation in a buffer with 0.1 mM EDTA (lane R-). Fixed chromatin extracted from the *endA*⁻ strains (panels B and C) can be digested by BglII (shown), PsuI (not shown), and NlaIII (not shown).

Endonuclease-I, a DNA-specific nuclease localised in the periplasm (81), digests dsDNA in a sequence independent manner and is responsible for the low quality of plasmid DNA preparations from *endA*$^+$ *E. coli* strains (82, 83). To investigate whether the enzyme also contributes to the degradation of chromatin in lysates of

formaldehyde-treated cells, the stability of fixed chromatin extracted from DH5α, an *endA*$^-$ strain of *E. coli* (83) during the initial steps of chromosome conformation capture was determined. Agarose gel electrophoresis showed that DH5α chromatin preparations do not degrade in the restriction digestion buffer with 0.1 mM EDTA (Figure 3.2B). Attempts to thermally denature endonuclease-I and overcome chromatin degradation were not pursued extensively since the conditions that reliably decreased degradation also promote reverse cross-linking of the chromatin and thereby interfere with proximity ligation in later steps of 3C-based protocols (Figure S3.1). Therefore, MG1655 Δ*endA* (henceforth referred to as NT331) was generated using the λ-red recombinase mediated gene replacement strategy (see: Materials and methods, Strains and plasmids) (72, 73). Chromatin preparations from fixed NT331 do not degrade when incubated in a buffer with a low concentration of EDTA (Figure 3.2C). Thus, all 3-C based experiments were carried out in a Δ*endA* background. It is noteworthy that using higher concentrations of formaldehyde for fixation, for instance, 7%, reduces chromatin degradation. *endA* deletion eliminates the requirement of using such concentrations and provides a wider window to optimize fixation conditions.

Proximity ligation with the insoluble fraction of digested, cross-linked chromatin, and methanol permeabilization of *Escherichia coli* prior to formaldehyde fixation improves the signal-to-noise ratio in chromosome contact maps.

Hi-C libraries of NT331 fixed with 3% formaldehyde have a low signal-to-noise ratio (Figure 3.3A), indicating inefficient formaldehyde-mediated cross-linking. Taking an earlier report of 3C-based studies in *E. coli* (84) into account, we raised the concentration of formaldehyde for fixation from 3% to 7%. However, the change did not contribute to a significant improvement in chromosome contact maps (Figure S3.2).

A low signal-to-noise ratio in chromosome contact maps may arise as a result of ligation between freely moving, non-crosslinked DNA molecules (85). This effect can be overcome by fractionating the digested, cross-linked chromatin into its soluble supernatant, and relatively insoluble pellet fractions by centrifugation (85, 86). The resulting enrichment of cross-linked DNA-protein complexes in the pellet and freely moving DNA molecules in the supernatant allows contact maps with a high signal-to-noise ratio to be generated when proximity ligation is carried out with only the pellet fraction (85, 86). In agreement with previous observations in E14.5 mouse embryos (86), *Saccharomyces cerevisiae* (85), and *S. pombe* (85),

incorporating fractionation and using only the pellet fraction for proximity ligation improved the signal-to-noise ratio of the *E. coli* NT331 contact map (Figure 3.3B). Fractionation was also incorporated in 3C-based studies of *Bacillus subtilis* (87).

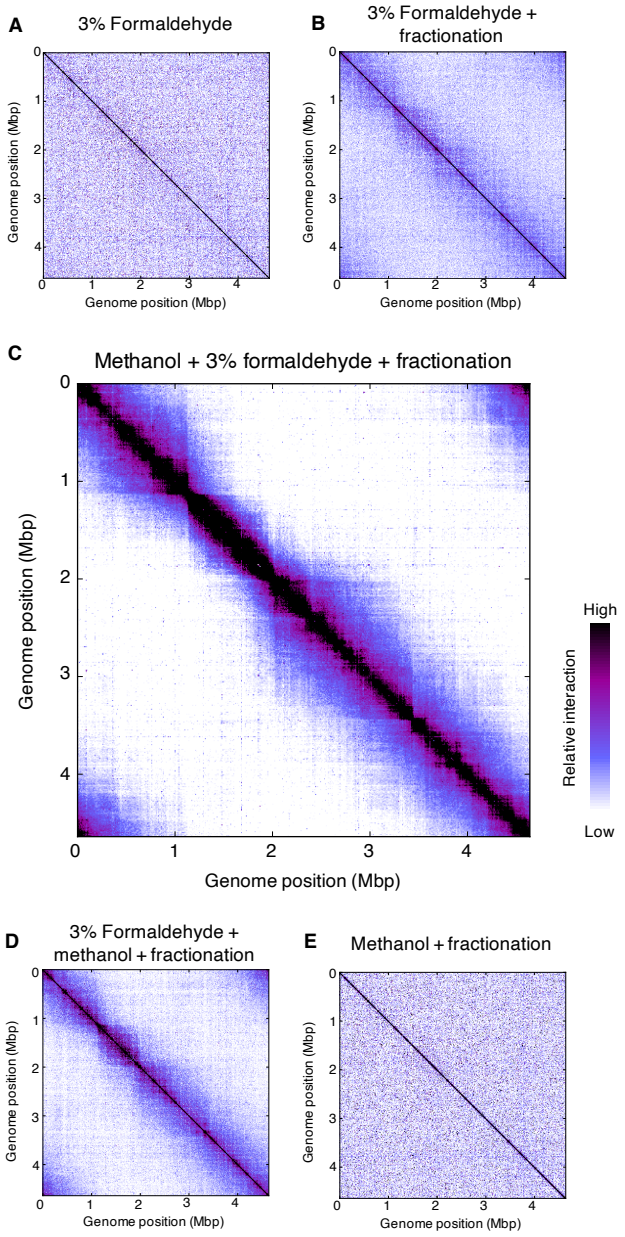


Figure 3.3: Proximity ligation with the insoluble fraction of digested, cross-linked chromatin, and methanol permeabilization of *Escherichia coli* prior to formaldehyde fixation improves the signal-to-noise ratio in chromosome contact maps. A: Hi-C libraries prepared from cells fixed with

3% formaldehyde have a low signal-to-noise ratio. **B, C:** The signal-to-noise ratio is improved by proximity ligation with the insoluble fraction of digested, cross-linked chromatin and methanol permeabilization of *E. coli* cells prior to formaldehyde fixation**. D:** Only marginal improvements in the signal-to-noise ratio are observed if methanol treatment is performed after formaldehyde fixation. **E:** *E. coli* cells permeabilised with methanol but not fixed with formaldehyde cannot be used to map chromosome structure. **Organism:** *Escherichia coli* MG1655 Δ*endA* (NT331); **3C-based study:** Hi-C; **Resolution:** 10 kb; **Growth conditions:** LB medium, 37 °C, exponential phase; **Fixation conditions: Panels A and B:** 3% formaldehyde, 1 hour, **Panel C:** 80% cold methanol for 10 minutes followed by 3% formaldehyde for 1 hour, **Panel D:** 3% formaldehyde for 1 hour followed by 80% cold methanol for 10 minutes, **Panel E:** 80% cold methanol for 10 minutes; **Restriction enzyme:** PsuI (ThermoFisher Scientific); **Fractionation: Panel A:** No, **Panels B-E:** Yes.

Formaldehyde is a commonly-used fixative in histology and (immuno–)histochemical studies where it is used either alone, or in combination with methanol (88, 89). Methanol dissolves lipids from cell membranes and coagulates proteins, thus, simultaneously permeabilizing and fixing histological preparations (89, 90). We extrapolated this to *E. coli*, and permeabilized the cells with 80% methanol prior to formaldehyde fixation. This treatment led to a significant improvement in the signal-to-noise ratio in chromosome contact maps (Figure 3.3C). Only a weak improvement was observed when methanol permeabilization followed formaldehyde fixation (Figure 3.3D). Methanol treatment alone could not be used to study chromosome conformation (Figure 3.3E).

Escherichia coli cells show global differences in the chromosome contact profiles during growth at different osmolarity conditions.

The binding of NAPs to DNA is sensitive to environmental conditions such as pH, temperature, and osmolarity. Consequently, changes to the ambient growth conditions of bacteria are reflected in an altered NAP binding profile of the chromosome, and hence, in the three-dimensional chromosome organization. We first examined the global differences in the chromosome contact profiles of NT331 during growth in a low-salt (0.08 M NaCl) medium, following a hyperosmotic shock (0.08 M -> 0.3 M NaCl), and in a high-salt (0.3 M NaCl) medium (Figures 3.4 and 3.5).

The NT331 chromosome contact maps exhibit a main diagonal of high interaction frequency that accounts for the physical proximity of chromosome regions close to each other in the primary genome sequence (black arrows, Figure 3.4). For NT331 cells growing in a low-salt medium (NT331 LS) high frequency interactions (in black) occur over distances of up to ~270 kb across the chromosome. The high frequency interactions occur over longer distances in the Ter macrodomain compared to the rest of the chromosome. This can be deduced from the width of the region of high interaction frequency around the main diagonal. This region is

thicker at the Ter macrodomain than the non-Ter regions of the chromosome. The width of this region is also indicative of local and global chromosome compaction. In NT331 LS, the chromosome is organized more compactly in the Ter macrodomain than it is in the non-Ter regions. Long-range, lower frequency interactions, observed in varying shades of purple, occur largely over distances of up to 1.1 Mb in non-Ter regions and ~670 kb in the Ter macrodomain (Figure 3.5A, left panel). Upon a hyperosmotic shock (NT331 SS), and in high salt media (NT331 HS), high frequency interactions still occur at distances of up to ~270-300 kb, however, the thickness of the diagonal indicates global decompaction of the chromosome. The changes are particularly evident in the non-structured right (NSR), non-structured left (NSL), and Ori macrodomains upon a hyperosmotic shock, and throughout the chromosome during growth in high-salt. For both conditions, long range, lower frequency interactions occur at distances of up to 1.3-1.6 Mb in non-Ter regions. In the Ter macrodomain, these interactions do not extend far beyond ~750 kb (Figures 3.5B and 3.5C, left panels).

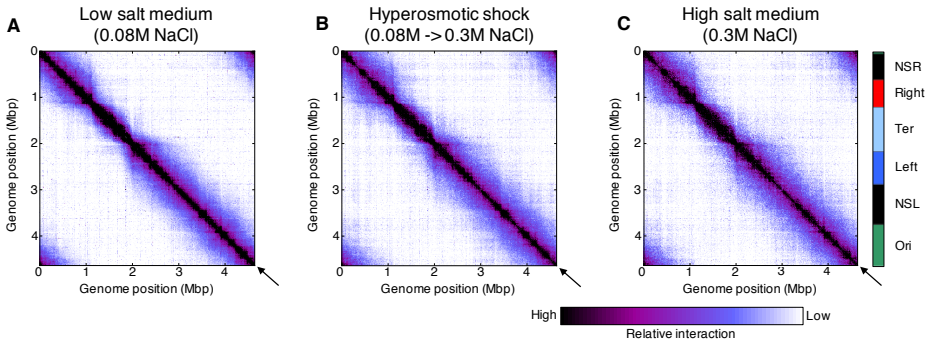


Figure 3.4: *Escherichia coli* cells show global differences in chromosome contact profiles during growth at different osmolarity conditions. Chromosome contact maps exhibit a main diagonal of high interaction frequency (black arrow). The thickness of this diagonal and the frequency of off-diagonal interactions varies with growth conditions, and along the length of the chromosome at the same growth condition. The panel on the far right illustrates the alignment of the macrodomains of the *E. coli* chromosome with the contact maps. **Organism:** *Escherichia coli* MG1655 Δ*endA* (NT331); **3C-based study:** Hi-C; **Resolution:** 10 kb; **Growth conditions: Panel A:** Low-salt LB medium (0.08 M NaCl), 37 ºC, exponential phase, **Panel B:** Low-salt LB medium (0.08 M NaCl) up to an OD_{600} of ~1.0 followed by a hyperosmotic shock with 5.0 M NaCl to a final concentration of 0.3 M for 10 minutes, 37 ºC, **Panel C:** High-salt LB medium (0.3 M NaCl), 37 ºC, exponential phase; **Fixation conditions:** 80% cold methanol for 10 minutes followed by 3% formaldehyde for 1 hour; **Restriction enzyme:** PsuI (ThermoFisher Scientific); **Fractionation:** Yes.

The structure of the chromosomal region in the vicinity of the *proVWX* operon reorganizes upon changes to the external environment. This region has been marked with a red triangle in the left panels of Figure 3.5 and is represented in the right panels of Figure 3.5. The position of *proVWX* is marked with white arrows

in the left panels, and with white '+' marks in the right panels (Figure 3.5). The chromosome around *proVWX* decompacts locally when *E. coli* cells in a low-salt medium are subjected to a hyperosmotic shock, while maintaining features of the finer chromosome organization, such as loops (marked with black arrows, Figures 3.5A and 3.5B, right panels) and the distinct 'flare' at 2.37 Mb (marked with white squares, Figures 3.5A and 3.5B, right panels). In a high salt environment, a condition that reflects the adaptation of *E. coli* to higher osmolarity following a hyperosmotic shock, chromosomal regions either decompact further such as the chromatin encompassing, and positioned locally downstream of *proVWX*, or show a stronger compaction compared to growth in a low-salt medium, for instance, the region encompassing the 'flare' at 2.37 Mb (marked with a white square, Figure 3.5C, right panel).

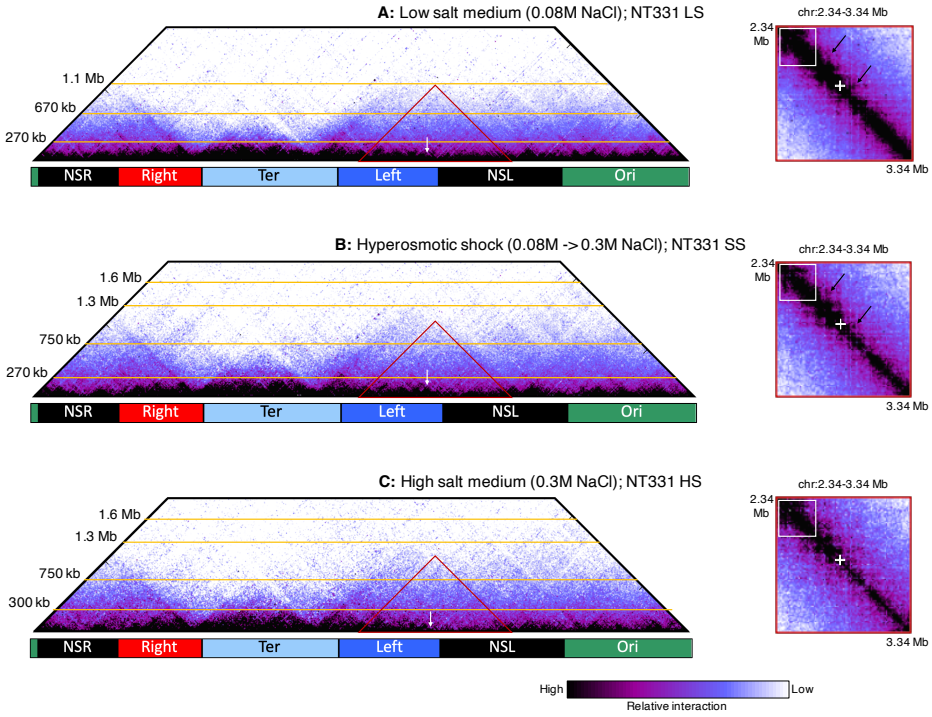


Figure 3.5: The *Escherichia coli* chromosome shows global and local changes to the chromosome contact profiles upon changes to the osmolarity of the growth medium, and during adaptation to altered osmolarity. A, left panel: High frequency interactions (black) in NT331 cells growing in low-salt LB medium occur over distances of up to ~270 kb across the chromosome. Long-range, lower frequency interactions, observed in varying shades of purple, occur largely over distances of up to 1.1 Mb in non-Ter regions and ~670 kb in the Ter macrodomain; **B & C, left panels:** Upon a hyperosmotic shock **(B)**, and in high salt media **(C)**, high frequency interactions still occur at distances of up to ~270-300 kb. Long range, lower frequency interactions occur at distances of up to 1.3-1.6 Mb in non-Ter regions. In the Ter macrodomain, these interactions do not extend far beyond ~750 kb; **A & B, right panels:** The chromosome around *proVWX* (also marked with a red triangle in the left panels; the

position of *proVWX* is marked with white arrows in the left panels and with white '+' marks in the right panels) decompacts locally when *E. coli* cells growing exponentially in a low-salt medium **(A)** are subjected to a hyperosmotic shock **(B)**. The local chromosome maintains features of the finer chromosome organization, such as loops (black arrows) and the 'flare' at 2.37 Mb (white squares). **C, right panel:** In a high salt environment, chromosomal regions are either decompacted compared to growth in a low-salt medium **(A)** such as the chromatin encompassing, and positioned locally downstream of *proVWX*, or show a stronger compaction, for example, at the region encompassing the 'flare' at 2.37 Mb (white square).

Hi-C in bacteria is limited by its resolution. Chromosomal gene density, the scarcity of intergenic DNA, and the short length of individual genes and operons means that an *E. coli* Hi-C map with, in our hands, a 10 kb resolution cannot be used to identify the underlying genetic features that encode global and finer structural changes to the chromosome, or the significance of these changes. Chromosome contact maps with a higher structural resolution assembled from 3C-Seq instead of Hi-C will be used to glean the contributing features.

The difference in local chromosome structure is associated with a change in the transcription profile of the *proVWX* operon

The global and finer structure of the chromosome is modified in response to changing osmolarity. Nucleoid associated proteins that regulate chromosome architecture also function as transcription factors, hence, the change in chromosome architecture may be reflected in transcription. We investigated the transcription profile of *proU* – an osmo-sensitive operon regulated by an osmo-sensitive NAP, H-NS – using RT-qPCR. As required by the MIQE guidelines (91), the details of RNA yield and purity (A_{260}/A_{280} measurements), and the results of RNA inhibition testing are provided in Supplementary file 3.1. Additionally, the results of a DNA contamination assessment of the RNA preparations showing no detectable contamination is provided in Figure S3.3.

Early studies examining the osmosensitivity of the *proU* operon involved the translational fusion of β-galactosidase to a truncate of ProV, and the subsequent detection of the specific activity of the enzyme in a variety of genetic backgrounds and osmolarity conditions as a measure of *proU* activation. While revolutionary, the technique is limited in that it relies on the post-translational detection of gene/operon activation, and it remains relatively ineffective in evaluating the transcriptional profile across an operon. To overcome this, we used RT-qPCR to study the osmo-sensitivity of *proU* at a transcriptional level. Eleven primer pairs were designed to evaluate the transcriptional profile across the operon and its flanking genes (Supplementary file 3.1; Figure 3.6). Primer specificity was determined with Sanger sequencing of the amplified product (Supplementary file

3.2; Supplementary folder 3.5). Identity of the amplicons was gauged with melting curves (Figure S3.4).

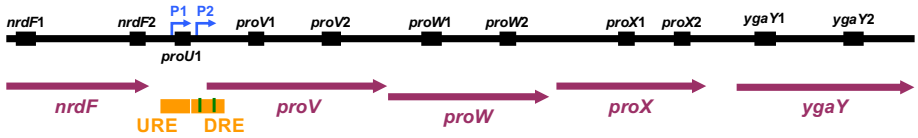


Figure 3.6: The positions of RT-qPCR amplicons (labelled, black) along *proVWX*, *nrdF*, and *ygaY* (purple arrows). P1 and P2 (right angle arrows, blue) designate the positions of the *proVWX* promoters recognised by σ^S and σ^{70}, respectively. URE and DRE (orange) mark the positions of the upstream and downstream regulatory elements of *proVWX*. The green bars within the DRE illustrate the positions of high-affinity H-NS binding sites.

Three internal controls, *rpoD*, *hcaT*, and *rrsA*, were selected for this study. *rpoD* encoding the RNA polymerase σ^{70} factor, is a house-keeping gene that has been validated as a stably-expressed mRNA suitable as an internal control in a diversity of bacterial species including *Klebsiella pneumoniae* (92), *Pseudomonas aeruginosa* (93), *Pseudomonas brassicacearum* GS20 (94), and *Gluconacetobacter diazotrophicus* (95). *hcaT* encodes a predicted 3-phenylpropionic transporter and has been validated as a suitable internal control in *Escherichia coli* (96, 97). *rrsA* codes for 16S rRNA and occurs in six copies in the *E. coli* MG1655 chromosome. It is accepted as a stably expressed house-keeping gene and was accepted as a validated internal control for *E. coli* RT-qPCR experiments (97). *rrsA* has been used to normalise relative expression levels in RT-qPCR studies of the *proU* operon (47). We tested the amplicon as a potential internal control, but, the low Cq values of *rrsA* in comparison to the Cq values of amplicons of the *proU* operon (Supplementary files 3.4 and 3.5) dissuaded the classification of *rrsA* as a reliable internal control. The ~2-fold higher expression of *rpoD* to *hcaT* motivated the use of only one internal control at a time for normalisation, since normalisation with a geometric average of the two amplicons biases the normalisation towards the higher expressed amplicon. The RT-qPCR results described here have been normalised using *rpoD* as the internal control. Equivalent data normalised with *hcaT* expression have been provided in the supplementary files for comparison.

The transcriptional profile of proU in NT331

The transcriptional profile across *proU* shows that the central *proW* gene has a lower relative expression level than the flanking *proV* and *proX* genes and that the terminal *proX* gene is expressed at a higher level than *proV*, in agreement with genome-wide RNA-seq studies (98). This trend is observed at 0.08 M NaCl, 0.3 M NaCl, and upon a hyperosmotic shock (Figures 3.7A-3.7C and S3.5A-S3.5C; Tables

S3.2 and S3.3). Mechanistically, the decline in relative expression between *proV* and *proW* may be accounted for by transcription termination at the junction of the two genes, while the sharp increase in expression between *proW* and *proX* may arise as a result of the presence of an internal promoter between the two genes. The increased transcript level in the downstream region of *proW* (amplicon *proW*2) compared to its upstream end (amplicon *proW*1) (Figures 3.7A-3.7C and S3.5A-S3.5C) implicates the presence of an additional *proX* promoter that lies within the *proW* ORF. Indeed, RNAP σ^{70} ChIP studies show an elevation in RNAP σ^{70} occupancy towards the downstream end of *proW*. However, a distinct RNAP σ^{70} peak within *proW* that may demarcate the additional internal promoter is not observed (Figure S3.6) (99). Nevertheless, a genome-wide study aimed at mapping RNA G-quadruplexes (rG4) in the *E. coli* transcriptome detected the presence of rG4 structures within the *proW* coding sequence (100), that occur between the *proW*1 and *proW*2 amplicons (Supplementary file 3.1). Using the *hemL* transcript that encodes glutamate-1-semialdehyde aminotransferase (aminomutase) as a model, rG4s were shown to play a role in the stabilisation of transcripts (100). K^+ stabilises rG4s when the ion occupies the centre of the quartet (101). Extrapolation of these findings to rG4s at *proW* hints that at higher intracellular concentrations of K^+, the increase in the relative expression level of *proX* transcripts initiated from the hypothetical internal promoter between *proW*1 and *proW*2 may be detected as a result of the stabilisation of rG4s at the 5' region of the transcript that, in turn, stabilises the transcript. The *proX* promoters are likely stringently-regulated non-canonical promoters rather than classical H-NS-repressed spurious promoters. H-NS ChIP shows no H-NS signal at *proX* (Figure S3.6) (102), and spurious transcripts have not been detected from *proX* in a Δ*hns* background (David C. Grainger, personal communication).

The *proVWX* operon is repressed in NT331 cells growing exponentially in a medium with an NaCl concentration of 0.08 M. Upon a hyper-osmotic shock that raises [NaCl] from 0.08 M to 0.3 M, *proVWX* expression increases up to ~8-fold (Figures 3.7D and S3.5D). The spike in expression may be mediated by the relief of H-NS-mediated repression as a result of the cytoplasmic influx of K^+ that occurs as an initial response to increased extracellular osmolarity (28–34). *In vitro* studies promote a model where the influx of K^+ drives a conformational change in the H-NS—DNA nucleoprotein structure at the NRE of *proVWX* from a transcriptionally-repressive DNA—H-NS—DNA bridge to a transcriptionally-conducive H-NS—DNA filament (20, 24). Exponentially-growing NT331 in a high-salt environment of 0.3 M NaCl, express *proVWX* ~1.5- to 2.5-fold higher than during exponential growth

at 0.08 M NaCl, and up to 4.5-fold lower than expression following a hyperosmotic shock (Figures 3.7E-3.7F and S3.5E-S3.5F). Exponential growth at 0.3 M NaCl reflects the adaptation of *E. coli* to increased extracellular osmolarity brought about by the import of osmoprotectants, the synthesis of trehalose, and the export of K^+ (28, 39). The export of K^+, in particular, re-instates repression. A higher expression of *proVWX* during exponential growth at 0.3 M NaCl compared to 0.08 M NaCl (Figures 3.7E and S3.5E) may account for the increased requirement of the ProU transporter at higher osmolarity.

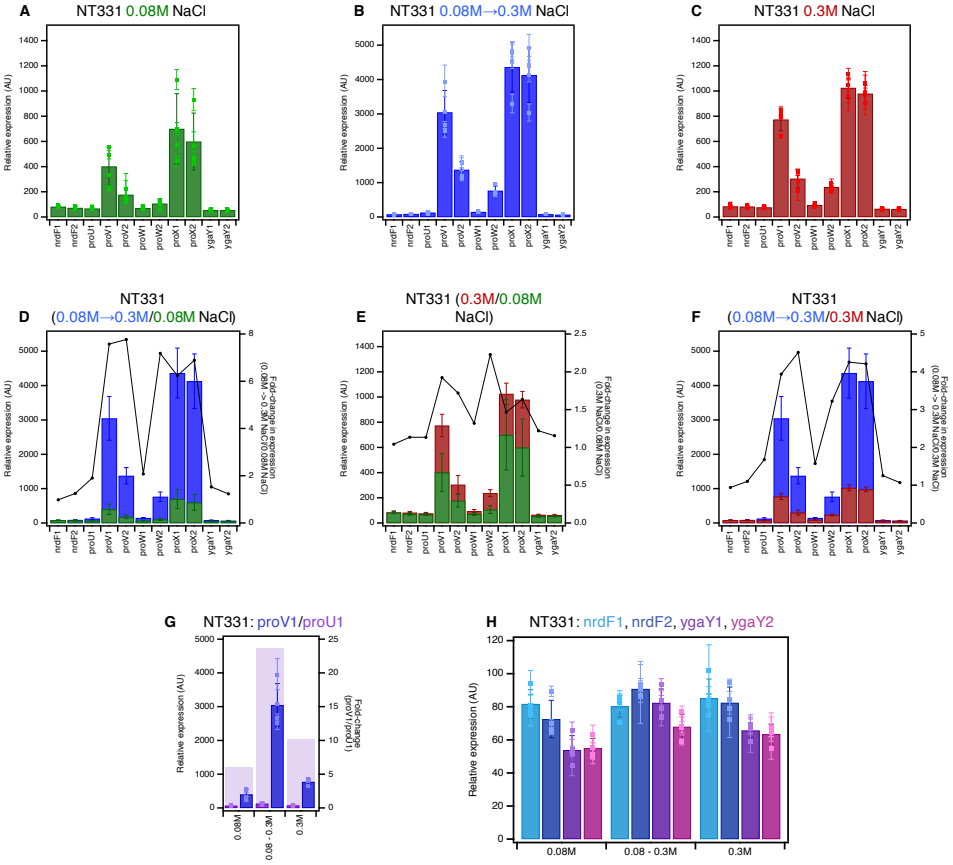


Figure 3.7: The transcriptional profile of *proVWX* and its flanking regions in NT331 during **(A)** exponential growth in M9 medium with 0.08 M NaCl, **(B)** hyperosmotic shock in M9 medium from 0.08 M to 0.3 M NaCl, and **(C)** exponential growth in M9 medium with 0.3 M NaCl. **The fold change in expression level of *proVWX* and its flanking regions** between **(D)** a hyperosmotic shock and exponential growth at 0.08 M NaCl, **(E)** exponential growth at 0.3 M NaCl and 0.08 M NaCl, and **(F)** a hyperosmotic shock and exponential growth at 0.3 M NaCl. **(G) The fold difference in expression at the *proV*1 amplicon compared to the *proU*1 amplicon** during exponential growth at 0.08 M NaCl, following a hyperosmotic shock, and during exponential growth at 0.3 M NaCl. **(H) The relative expression at amplicons flanking *proVWX*** during exponential growth at 0.08 M NaCl, following a hyperosmotic shock, and during exponential growth at 0.3 M NaCl. **Internal control: *rpoD***

The fold changes in the expression levels of the constituent genes of *proVWX* differ upon changes to ambient osmolarity. Following a hyperosmotic shock, the expression of *proV* increases ~7.5-fold, the P2 proximal amplicon of *proW* (amplicon *proW*1) increases ~2-fold, the P2 distal amplicon of *proW* (amplicon *proW*2) is expressed ~7-fold higher, and *proX* transcripts increase by a factor of ~6.5. In exponential growth at 0.3 M NaCl, in comparison to 0.08 M NaCl, the expression levels of *proV*, *proW*1, *proW*2, and *proX* increase by factors of ~1.8, ~1.3, ~2.2, and ~1.5, respectively (Figures 3.7D-3.7E, and S3.5D-S3.5E). Differences in the fold-induction or fold-repression of *proV*, *proW*, and *proX* in *E. coli* upon changes to growth conditions have also been observed in genome-wide studies (98). Collectively, the results show that *proV*, *proW*, and *proX* exhibit a degree of independent regulation despite occurring within the same operon. This provokes the hypothesis that the expression of co-regulated genes of an operon may be tweaked by additional factors that operate on individual genes. The *proU* operon encodes an ABC transporter – a member of the largest group of paralogous protein complexes (103). The components of ABC transporters are contained within operons, the earliest of which is predicted to have arisen before the divergence of bacteria and archaea (104, 105). The organization of ABC operons is evolutionarily conserved and is specific for each orthologous group (106, 107). In this light, the independent regulation of the constituent genes of *proU* suggests that other ABC transporter operons evolutionarily-related to *proVWX*, regardless of the host organism, may carry similar regulatory features.

Transcription of *proVWX* primarily initiates from its σ^{70}-dependent (P2) promoter positioned 60 bp upstream of the *proV* ORF. The *proVWX* σ^{S}-dependent (P1) promoter located at -250 bp of *proV* is cryptic and is activated by *rho* and *hns* mutations, and by cold stress (40). To record the contribution, if any, of P1 to the expression of *proVWX* in a wild-type NT331 background, the relative expression level of amplicon *proU*1 positioned between the P1 and P2 promoters was determined. At 0.08 M NaCl, *proU*1 was expressed ~6-fold lower than amplicon *proV*1 positioned 345 bp downstream of P2. Upon a hyperosmotic shock, the relative expression of *proU*1 increased ~2-fold, however, the expression was ~25-fold lower than that of *proV1*. During exponential growth at 0.3 M NaCl, the expression of *proU*1 was comparable to that at 0.08 M NaCl, and ~10-fold lower than *proV*1 (Figures 3.7G and S3.5G). This indicates that while the osmo-response of *proVWX* primarily arises from initiation at P2 (58, 59, 108), the activation of P1 contribution may ensue from the dismantling of the H-NS—DNA nucleoprotein

during the K^+ influx, to form a nucleoprotein structure that mimics an *hns* knock-down phenotype. Such a structure would alleviate Rho-dependent termination of transcripts initiated from P1 (40, 41, 98). Furthermore, P1 can contribute to the increased expression of the operon during a hyper-osmotic shock. Transcription from P1 across P2 may favour the relief of H-NS repression at the NRE (109, 110).

To detect cross-talk between the expression of *proU* and its flanking genes, the relative expression levels of *nrdF* and *ygaY* positioned upstream and downstream of *proU*, respectively, were determined. In the NT331 background, the expression of *ygaY* and *nrdF* was comparable for growth at 0.08 M NaCl, 0.3 M NaCl, and following a hyper-osmotic shock (Figures 3.7H and S3.5H; Tables S3.2 and S3.3), indicating insulation from flanking chromatin. It is important to note, however, that the expression of the *nrdF* amplicon proximal to *proU* (amplicon *nrdF*2) and both *ygaY* amplicons (amplicons *ygaY*1 and *ygaY*2) were detectably higher after a hyperosmotic shock. This may occur due to an increased local density of RNAP at this osmotic condition, or the decompaction of the local chromatin following the shock providing a more conducive environment for transcription.

The transcriptional profile of proVWX upon alleviation of H-NS-mediated repression at the DRE

H-NS-mediated *proU* repression was alleviated by introducing a series of point mutations in the AT-rich H-NS-binding regions that were experimentally detected by an *in vitro* DNase-I protection assay (49) in the *proU* DRE. This approach was chosen over *hns* deletion to prevent the pleiotropic effects associated with the latter. In this design, the -35 and -10 promoter elements of P2, the P2 TSS, and the Shine-Dalgarno sequence of *proVWX* were not edited. However, two point mutations were introduced in the high affinity H-NS binding site extending between -7 and +15 of the P2 TSS. AA>GC and TA>GC mutations were carried out at -3 to -2 and at +7 to +8 positions, respectively. AT-rich codons in H-NS binding sites in the *proV* ORF were switched out for GC-rich variants with the most similar codon usage ratio in the *E. coli* chromosome. Identical mutations, selected on the basis of the aforementioned criteria, were incorporated into both high-affinity H-NS binding sites occurring in the DRE (48). Following mutation, the AT-content of the DRE was lowered from 66.4% to 51.6% (Supplementary file 3.1, see Feature NT644_DRE). The strain was labelled NT644. The decrease in affinity of the mutated DRE for H-NS was validated *in vitro* using an electrophoretic mobility shift assay (Supplementary methods, Figure S3.8).

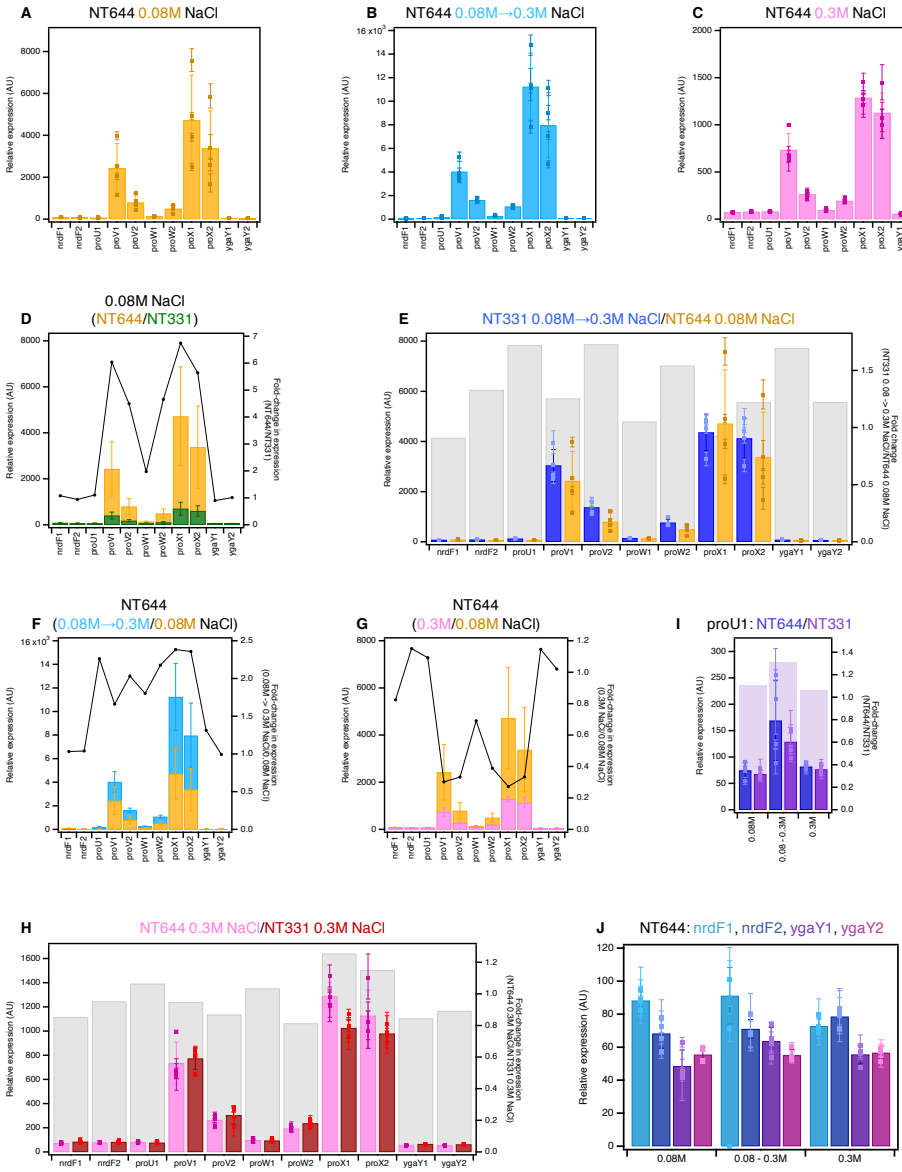


Figure 3.8: The transcriptional profile of *proVWX* and its flanking regions in NT644 during **(A)** exponential growth in M9 medium with 0.08 M NaCl, **(B)** hyperosmotic shock in M9 medium from 0.08 M to 0.3 M NaCl, and **(C)** exponential growth in M9 medium with 0.3 M NaCl. **The fold change in expression level of *proVWX* and its flanking regions** between **(D)** NT644 and NT331 during exponential growth at 0.08 M NaCl, **(E)** NT331 following a hyperosmotic shock and NT644 growing exponentially at 0.08 M NaCl, **(F)** a hyperosmotic shock and exponential growth at 0.08 M NaCl for NT644, **(G)** exponential growth at 0.3 M NaCl and 0.08 M NaCl for NT644, and **(H)** NT644 and NT331 growing exponentially at 0.3 M NaCl. **(I) The fold change in expression at the *proU*1 amplicon** between NT644 and NT331 during exponential growth at 0.08 M NaCl, following a hyperosmotic shock, and during exponential growth at 0.3 M NaCl. **(J) The relative expression level in NT644 at amplicons flanking *proVWX*** during exponential growth at 0.08 M NaCl, following a hyperosmotic shock, and during exponential growth at 0.3 M NaCl. **Internal control: *rpoD***

Alleviating H-NS binding at the *proU* operon activates the expression of its constituent genes at 0.08 M NaCl by up to ~6.5-fold (Figures 3.8A, 3.8D, S3.8A, and S3.8D; Tables S3.4 and S3.5). This is in line with observations by Nagarajavel *et al*., 2007, where, in a Δ*hns* background at 0.05 M, and 0.1 M NaCl, the β-galactosidase activity of a *proV-lacZ* fusion expressed from a construct cloned downstream of the P2 promoter, URE, and DRE on a plasmid was ~9- and ~4-fold higher than in the wild-type background (42). The relative expression of *proVWX* in NT644, a strain deficient in H-NS at the *proU* regulatory region, at 0.08 M NaCl is comparable to the relative expression in NT331, the wild-type strain, after a hyper-osmotic shock (0.08 M to 0.3 M NaCl) (Figures 3.8E and S3.8E; Tables S3.2-S3.5).

The *proVWX* osmoregulation studies performed in (42) show that the specific activity of β-galactosidase of a *proV-lacZ* fusion in a Δ*hns* background at 0.05 M, and 0.1 M NaCl is similar to the specific activity of proV-β-galactosidase in a wild-type background growing exponentially at 0.3 M NaCl, not wild-type cells during a hyperosmotic shock as in our observations. However, the post-translational detection of osmosensitivity in (42) may account for the differences between the results in (42) and those presented here. In this model, the up-shift of transcript levels following a hyper-osmotic shock rapidly establishes the required number of ProU transporters, while the decreased transcript levels in *E. coli* adapted to hyper-osmotic stress maintain the necessary levels of ProU.

We notice a disparity between our observations and those presented in (42) when examining the relative expression of *proU* in NT644. A hyper-osmotic shock to NT644, induces the expression of *proVWX* by ~2-fold (Figures 3.8B, 3.8F, S3.8B, and S3.8F), and upon adaptation to the hyperosmotic stress, represented by exponential growth at 0.3 M NaCl, *proVWX* expression levels decline to as low as up to ~25% of those at 0.08 M NaCl (Figures 3.8C, 3.8G, S3.8C, and S3.8G). Our results also show that NT644 (H-NS-deficient *proU* regulatory element) and NT331 (wild-type) have similar expression levels of *proU* at 0.3 M NaCl (Figures 3.8H and S3.8H; Tables S3.2-S3.5). This suggests that in *E. coli* cells adapted to higher osmolarities, the repression of *proU* is either mediated by factors other than H-NS occupancy at the *proU* promoter, or that H-NS occupancy at the mutated GC-rich *proU* promoter is favoured by one or more additional factors.

Expression from P1 is repressed in a Δ*hns* background (32, 43). The mutations at the *proVWX* regulatory element in NT644 are designed to disrupt H-NS binding to

the DRE (Supplementary file 3.1, see Feature NT644_DRE). H-NS occupancy at the URE, and effectively, at P1, (Figure S3.6) should remain unaffected. Indeed, the relative expression level of the *proU*1 amplicon was similar between NT331 and NT644 at the tested osmolarity conditions (Figures 3.8I and S3.8I).

Disruption of H-NS binding to the DRE did not affect the insulation of the *proVWX* operon. The relative expression levels of the flanking *nrdF* and *ygaY* genes were similar between NT331 and NT644 during growth at 0.08 M NaCl, 0.3 M NaCl, and following a hyperosmotic shock (Figures 3.7H, 3.8J, S3.5H, and S3.8J; Tables S3.2-S3.5).

The role of StpA in the regulation of the proU operon

StpA is a paralogue of H-NS that complements the protein in Δ*hns* strains (63–65). StpA stimulates the formation of DNA—H-NS—DNA bridges *in vitro* and stabilises the structure against changes in temperature and Mg^{2+} or K^{+} concentration (111). ChIP studies reveal that H-NS-bound regions of the chromosome contain StpA (67). At the *proVWX* operon, StpA occupies two sites within the NRE that overlap with H-NS-bound regions. The first site encompasses the P1 promoter and the second lies downstream of P2 and spans across the *proV* structural gene. For ease of communication, these sites are henceforth referred to as Site P1 and Site *proV* in this chapter. H-NS occupancy at *proVWX* appears to be unaffected in a Δ*stpA* background. In Δ*hns* strains, StpA only occupies Site P1 at *proVWX* – the nucleoprotein structure at Site *proV* is lost (67). This binding pattern may explain why the *proVWX* operon is only repressed by StpA when StpA is over-expressed in a Δ*hns* background (66). However, the role of StpA in *proVWX* regulation in an otherwise wild-type background is unclear. Therefore, we investigated the effect of *stpA* deletion on the osmoresponse of *proVWX* in NT331.

At 0.08 M NaCl, *stpA* deletion elevates the relative expression levels of all amplicons within the *proU* structural genes – with the exception of *proV*2 and *proW*1 – by ~4-fold (Figures 3.9A-3.9B, and S3.9A-S3.9B; Tables S3.2, S3.3, S3.6, and S3.7). *In vitro* observations of the architectural properties of StpA-deficient and StpA-supplemented H-NS—DNA structures (111) encourage the prediction that this rise in expression may be driven by an increased sensitivity of the repressive H-NS—DNA nucleoprotein at *proVWX* to intracellular K^{+} concentrations. The relative expression level of amplicon *proV*2, positioned on the *proV* ORF distal to the P1 and P2 promoters was ~7-fold higher upon *stpA* deletion, compared to the ~4-fold increase for amplicon *proV*1 that is positioned closer to the promoters (Figures 3.9A-3.9B, and S3.9A-S3.9B; Tables S3.2, S3.3,

S3.6, and S3.7). From an alternate perspective, the average relative expression level of *proV*2 at 0.08 M NaCl in NT331 was 44% of that of amplicon *proV*1; in NT331 Δ*stpA*, this value rose to 69%. This reflects an increased processivity of RNAP on the *proV* ORF in the absence of StpA, implying that StpA-deficient H-NS—DNA nucleoprotein complexes that form over the *proV* ORF, and by extension, on other sites throughout the chromosome, function as weaker transcription roadblocks than StpA-supplemented H-NS—DNA structures. The relative expression level of the *proW*1 amplicon was similar in both NT331 and NT331 Δ*stpA* (Figures 3.9A-3.9B, and S3.9A-S3.9B; Tables S3.2, S3.3, S3.6, and S3.7). This may be attributed to the strength of the terminator hypothesized to lie between *proV* and *proW*.

A hyperosmotic shock to NT331 Δ*stpA* from 0.08 M to 0.3 M NaCl increased the relative expression of *proV* and *proW* by ~14- to ~20-fold, and *proX* by ~10-fold (Figures 3.9C-3.9D, and S3.9C-S3.9D; Tables S3.6 and S3.7). In NT331, the increase stands at less than 8-fold (Figures 3.7B, 3.7D, S3.5B, and S3.5D; Tables S3.2 and S3.3). NT331 Δ*stpA* cells adapted to growth at 0.3 M NaCl express *proU* ~3- to ~4-fold higher than cells growing at 0.08 M NaCl and ~2.5- to ~4-fold lower than cells subjected to a hyperosmotic shock (Figures 3.9E-3.9G and S3.9E-S3.9G; Tables S3.6 and S3.7). These observations further support the model that StpA-deficient H-NS—DNA nucleoprotein filaments have a heightened osmosensitivity and lowered osmostability such that the structure is much more efficiently dismantled by K^+ (111).

Transcription initiation from P1 was lower in Δ*stpA* mutants compared to wild-type cells during exponential growth at 0.08 M and 0.3 M NaCl (Figures 3.9H and S3.9H; Tables S3.6 and S3.7). This is complementary to observations that expression from P1 is decreased in Δ*hns* strains (32, 43). How a generally repressive nucleoprotein complex of H-NS and StpA enhances expression from a promoter encapsulated within its structure is unclear. The puzzle is further complicated with the paradoxical observation that upon a hyperosmotic shock, a condition that dismantles H-NS—DNA complexes, expression from P1 increases and is similar between NT331, and NT331 Δ*stpA*.

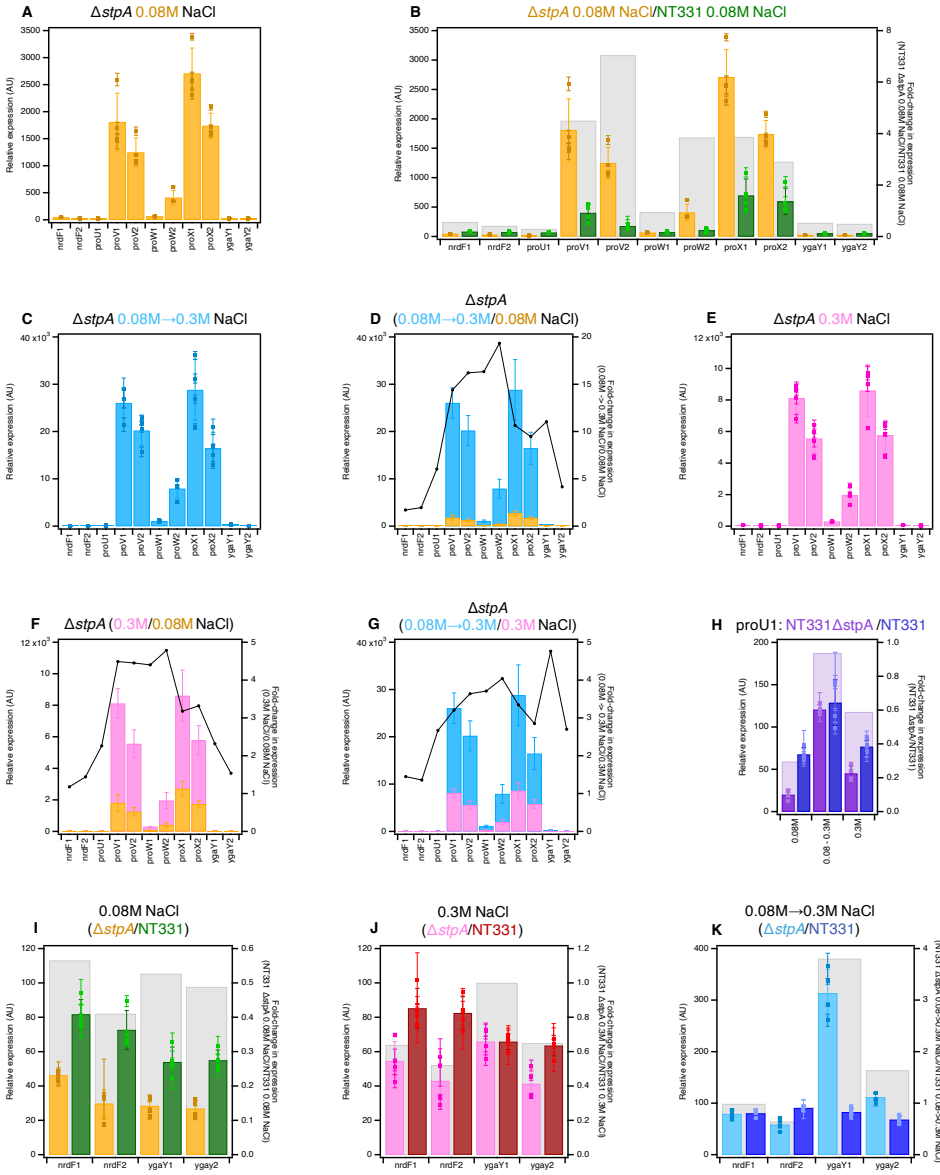


Figure 3.9: The transcriptional profile of the *proVWX* operon and its flanking regions in NT331 Δ*stpA* (A) during exponential growth at 0.08 M NaCl and **(B)** the fold-change in the expression levels of the amplicons compared to NT331, **(C)** upon a hyperosmotic shock from 0.08 M to 0.3 M NaCl, and **(D)** the fold-change in expression levels of the amplicons in comparison to exponential growth at 0.08 M NaCl, and **(E)** during exponential growth at 0.3 M NaCl, and the fold-change in expression levels of the amplicons with respect to **(F)** exponential growth at 0.08 M NaCl, and **(G)** a hyperosmotic shock. **(H) The fold-difference in expression level of amplicon *proU*1** between NT331 Δ*stpA* and NT331. **The fold-change in expression levels of the *nrdF* and *ygaY* amplicons between NT331 Δ*stpA* and NT331 (I)** during exponential growth at 0.08 M NaCl, **(J)** exponential growth at 0.3 M NaCl, and **(K)** following a hyper-osmotic shock. **Internal control: *rpoD***.

The expression of *nrdF* and *ygaY* genes in NT331 Δ*stpA* was affected by the osmotic condition. Both genes showed ~50% repression at 0.08 M NaCl (Figures 3.9I-3.9J, and S3.9I-S3.9J; Tables S3.2, S3.3, S3.6, and S3.7). A perceptible increase in expression at 0.3 M NaCl was observed, that was still maintained at ~50% of that in the wild-type background when expression was normalised relative to *rpoD* (Figures 3.9I-3.9J; Tables S3.2 and S3.6). When normalisation was performed using *hcaT* as an internal control, the expression was comparable to the wild-type background (Figures S3.9I and S3.9J; Tables S3.3 and S3.7). Upon a hyperosmotic shock to NT331 Δ*stpA*, the expression of *nrdF*, positioned upstream of *proU*, increased ~2-fold to a level comparable to that of NT331 under the same condition. Moreover, the expression of *ygaY*, positioned downstream of *proU*, increased 11-fold at the *ygaY*1 amplicon positioned proximal to the terminus of *proX* and 4-fold at the *ygaY*2 amplicon positioned further downstream, corresponding to an increase of 3.8-fold and 1.6-fold, respectively, compared to the wild-type strain (Figures 3.9K and S3.9K; Tables S3.2, S3.3, S3.6, and S3.7). This highlights that StpA insulates *proVWX* from its flanking operons, perhaps by constraining supercoils (65). Indeed, a weak ChIP signal for FLAG-tagged StpA is observed at the terminus of the *proVWX* operon (67). The decline in expression between *ygaY*1 and *ygaY*2 may be accounted for by the crypticity of *ygaY* (112, 113). Nine in-frame stop codons occur between the two amplicons.

The role of RNaseIII in regulating the expression of proU

RNaseIII, encoded by the *rnc* gene, regulates the expression of *proVWX* at a post-transcriptional level. In hypo-osmotic conditions and upon a hypo-osmotic shock, RNaseIII downregulates *proVWX* by cleaving a conserved secondary structure that forms at position +203 to +293 of the *proU* mRNA transcribed from P2. High osmolarity conditions inhibit the action of RNaseIII on the *proU* mRNA (47). The report in (47) was limited in that the regulatory role of RNaseIII on *proVWX* expression was studied using a plasmid-encoded construct of the *proU* regulatory elements and full-length *proV* in a strain carrying a chromosomal *ΔproVWX* mutation. Therefore, we built on this report and determined the effect of a Δ*rnc* mutation in an otherwise wild-type background on the regulation of chromosomal *proVWX* expression.

The *proVWX* operon was expressed perceptibly higher in NT331 Δ*rnc* compared to NT331 during exponential growth at 0.08 M NaCl (Figures 3.10A-3.10B, and S3.10A-S3.10B). The fold increase in the relative transcription level of amplicon *proV*2 – positioned distal to the P1 and P2 promoters– was higher than that of the

*proV*1 amplicon positioned closer to P1 and P2 (Figures 3.10B, and S3.10B). This indicates an increased stability in the absence of RNaseIII and highlights that in addition to the processing of the conserved hairpin located at +203 to +293 of the *proU* transcript, RNaseIII downregulates the expression of *proU* by degrading *proV* transcripts from the 3' end. Interestingly, the transcript level of *proX* in NT331 Δ*rnc* at 0.08 M NaCl was reduced to ~50% of that in NT331 (Figures 3.10B, and S3.10B). Since RNaseIII degrades RNA, the decrease in relative expression of *proX* may be triggered by decreased initiation from the predicted internal *proX* promoter. How the knock-out of an RNase could trigger such an effect is unclear.

A hyperosmotic shock from 0.08 M to 0.3 M NaCl, increased the expression of *proV* and *proX* by ~15-fold, *proW*1 by ~5-fold, and *proW*2 by ~11-fold (Figures 3.10C-3.10D, and S3.10C-S3.10D). In comparison to NT331, the relative expression levels of *proV* and *proW* were >2.4-fold higher in NT331 Δ*rnc* (Figures 3.10E, and S3.10E). In contrast, the relative expression of *proX* upon a hyperosmotic shock was similar for NT331 and the Δ*rnc* mutant (Figures 3.10E, and S3.10E). This indicates that during a hyper-osmotic shock RNaseIII specifically degrades transcripts initiating from P1 and P2, but does not interfere with transcripts expressed from the predicted internal *proX* promoter. This is in support of the inferences put forward in (47) that implicate a hairpin at the 5' end of *proU* transcripts in signalling RNaseIII-mediated processing, however, it does not explain the increased stability of 3' regions of *proU* transcripts in the Δ*rnc* mutant (Figures 3.10D-3.10E, and S3.10D-S3.10E).

During exponential growth at 0.3 M NaCl, *proV* and *proX* are induced 7.5- to 8.5-fold in comparison to growth at 0.08 M NaCl; *proW*1 is induced ~3-fold and *proW*2, ~5.5-fold (Figures 3.10F-3.10G, and S3.10F-S3.10G). The increase is in line with expectations of *proU* induction at higher osmolarities. The relative expression level of *proW* and *proX* in the Δ*rnc* mutant at 0.3 M NaCl is ~2-fold higher than in NT331. *proV*1 is ~5-fold more abundant and proV2, ~11-fold (Figures 3.10H, and S3.10H). This indicates that upon adaptation to higher osmolarity, RNaseIII contributes to downregulating the expression of *proU*. RNaseIII preferentially targets *proV* transcripts and destabilises the transcripts at the 3' end. The increase in the relative expression level of *proX* in NT331 Δ*rnc* at 0.3 M NaCl with respect to NT331 (Figures 3.10H, and S3.10H), implicates RNaseIII in *proX* processing. This is in contrast of *proX* expression at 0.08 M NaCl where RNaseIII knock-out represses the gene.

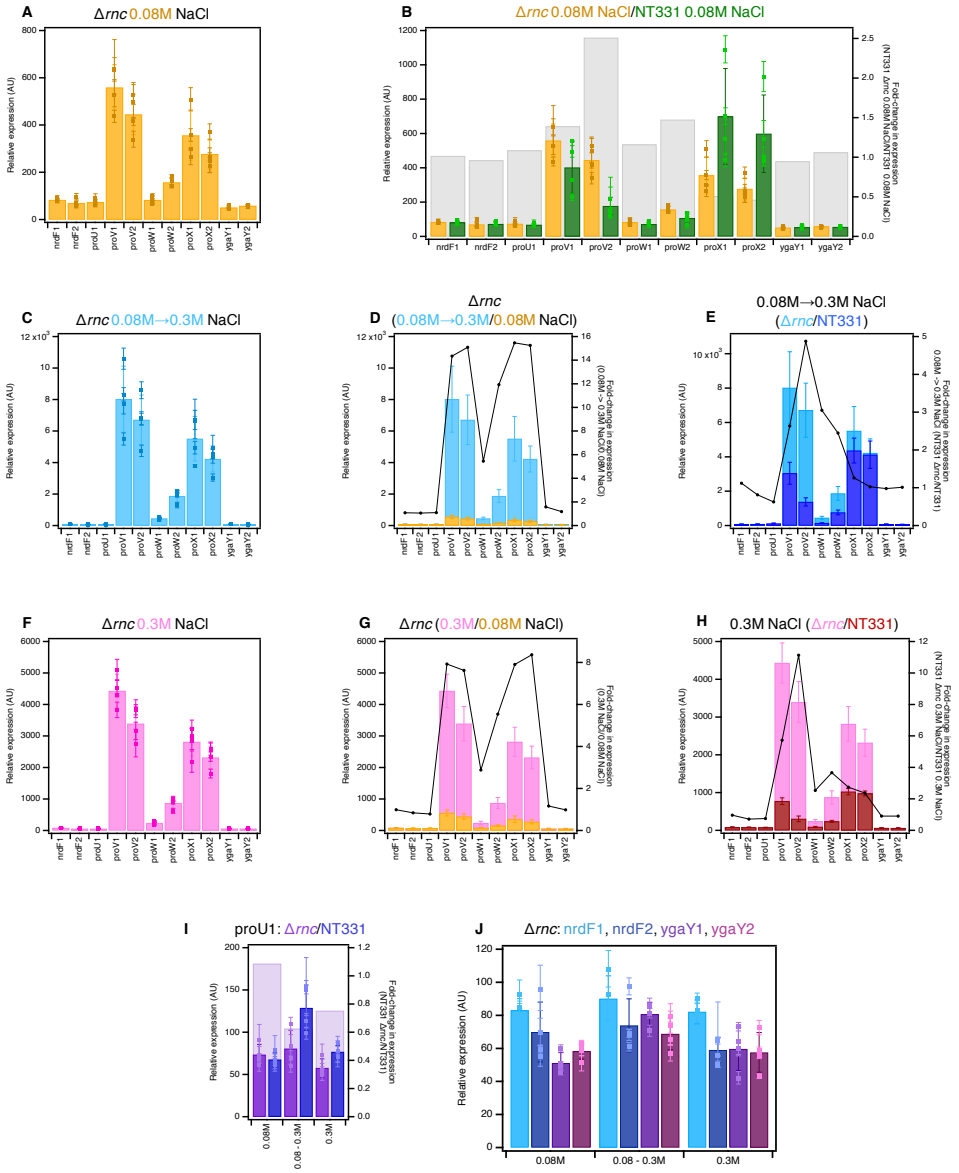


Figure 3.10: The transcriptional profile of *proVWX* and its flanking regions in NT331 Δ*rnc* during **(A)** exponential growth at 0.08 M NaCl, and **(B)** the fold-change in expression of the amplicons compared to NT331 growing exponentially at 0.08 M NaCl, upon **(C)** a hyperosmotic shock from 0.08 M NaCl to 0.3 M NaCl, and the fold-change in expression of the amplicons in comparison to **(D)** NT331 Δ*rnc* growing exponentially at 0.08 M NaCl, and **(E)** NT331 following a hyperosmotic shock. **The transcription profile of *proVWX* and its flanking regions in NT331 Δ*rnc*** during **(F)** exponential growth at 0.3 M NaCl, and the fold-change in expression of the amplicons relative to **(G)** NT331 Δ*rnc* growing exponentially at 0.08 M NaCl, and **(H)** NT331 growing exponentially at 0.3 M NaCl. **(I) The fold-difference in the relative expression of the *proU*1 amplicon** between NT331 Δ*rnc* and NT331. **(J) The relative expression level in NT331 Δ*rnc* at amplicons flanking *proVWX*** during exponential growth at 0.08 M NaCl, following a hyperosmotic shock, and during exponential growth at 0.3 M NaCl. **Internal control: *rpoD*.**

In NT331 Δ*rnc*, transcription from P1 is comparable to the wild-type background (Figures 3.10I, and S3.10I). *nrdF* and *ygaY* are insulated from *proU* and exhibit similar expression levels at high and low osmolarity that are comparable to the wild-type (Figures 3.10J, and S3.10J).

Structural analysis of the *proVWX* operon.

The expression of *proVWX* is regulated by H-NS – a NAP that responds to changes in osmolarity by modifying DNA architecture *in vitro* (20). H-NS may, therefore, regulate *proVWX* by modifying the local three-dimensional chromosome structure *in vivo*.

Live cell FROS imaging

To visualize the changes in the three-dimensional organization of *proVWX* in response to changing osmolarity, we used fluorescent-repressor operator system (FROS) labelling in live cells (114, 115). For this, the upstream and downstream ends of *proVWX* were labelled with Tet operator (TetO) and Lac operator (LacO) arrays, respectively (Supplementary file 3.1). Each array carries six repressor binding sites (Supplementary file 3.1) and accommodates a maximum of six dimers of fluorescently-labelled repressors. The arrays were flanked by terminators to account for terminator sequences (Supplementary file 3.1) that may have been disrupted by the insertion, and to ensure that transcription does not occur across the arrays to prevent interference with repressor binding. The fluorescently-labelled repressors (TetR-eYFP and LacI-mCherry) were expressed from under an arabinose-inducible promoter on a pBAD24 construct (pRD183).

A preliminary study was performed in NT455 (MG1655 *proVWX*::TetO-*proVWX*-LacO pRD183), a strain that does not carry a Δ*lac* mutation and, therefore, expresses an unlabelled variant of the Lac repressor. Furthermore, in this strain, LacI-mCherry can bind to the LacO sites in the *lac* operon. Nevertheless, the distance between the sites and the binding of a maximum of two LacI-mCherry molecules at each site means that the *lac* operon would not be detected above the background.

In NT455 growing exponentially in M9 medium with 0.08 M NaCl, the eYFP and mCherry foci that marked the ends of *proVWX* were observed in close proximity $d=94\pm1$ nm, $\sigma=9\pm1$ nm, $N=26$ cells (Figures 3.11A, and 3.11Da). The distance of ~100 nm lies at the limit of resolution of the TIRF microscopy set-up used in

this study. In this case, therefore, the foci can be considered to occur on top of each other. The observations imply that during exponential growth at 0.08 M NaCl – an osmotic condition at which *proVWX* is repressed (Figures 3.7A and S3.5A) – the operon forms a loop (Figure 3.11Db).

When an exponentially-growing culture of NT455 is subjected to a hyperosmotic shock in M9 medium that raises the concentration of NaCl from 0.08 M to 0.3 M, the eYFP and mCherry foci are detected at a distance of 436±7 nm, σ= 100±10 nm (N=45 cells) from each other (Figures 3.11B, and 3.11Ea). This indicates that the loop that forms between the ends of *proVWX* at 0.08 M NaCl disengages upon a hyperosmotic shock (Figure 3.11Eb) – a condition that is associated with a ~2- to ~8-fold increase in the expression of *proVWX* ORFs (Figures 3.7D, and S3.5D).

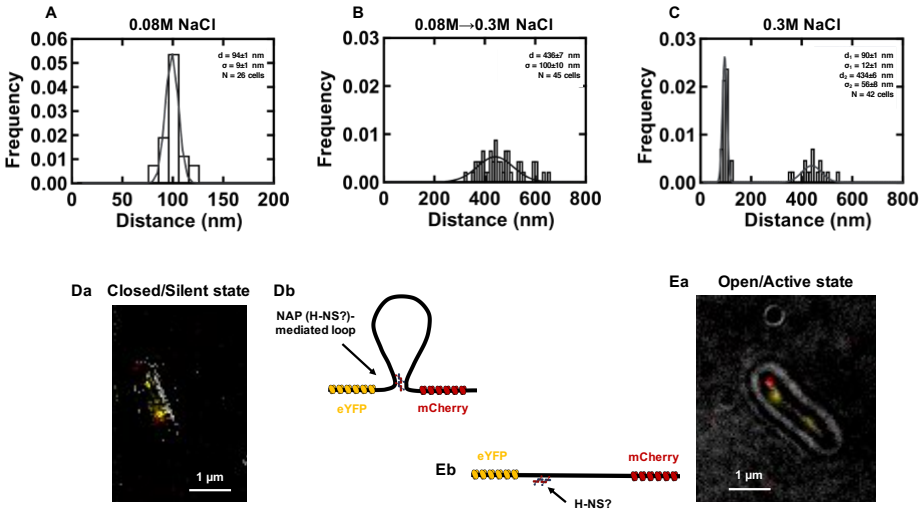


Figure 3.11: The three-dimensional organization of *proVWX* correlates with its expression state. **A:** In NT455 cells growing exponentially at osmolarity conditions where *proVWX* is repressed (0.08 M NaCl), the eYFP and mCherry foci that mark the ends of the operon lie in close proximity d=94±1 nm, σ=9±1 nm, N=26 cells. **B:** The foci diverge from each other following a hyperosmotic shock from 0.08 M to 0.3 M NaCl – an osmotic condition at which the expression of *proVWX* increases ~2- to ~8-fold; 436±7 nm, σ= 100±10 nm, N=45 cells. **C:** During exponential growth at 0.3 M NaCl, a condition at which the relative expression level of *proVWX* is between that of exponential growth at 0.08 M NaCl and that following a hyperosmotic shock, NT455 cells show two distinct populations with regards to the conformation of *proVWX*. A population in which the eYFP and mCherry foci lie in close proximity (d=90±1 nm, σ=12±1 nm, N=26), and another in which the foci diverge (434±6 nm, σ=56±8 nm, N=16). *D:* eYFP and mCherry foci occur in close proximity in cells where *proVWX* is repressed/silenced. This may arise as a result of the formation of a loop anchored between the ends of the operon. The loop may be mediated by H-NS. **E:** eYFP and mCherry foci lie ~400nm apart in cells where *proVWX* is activated. This may occur due to the destabilization of an H-NS-mediated loop between the ends of the operon.

During exponential growth at 0.3 M NaCl, an osmotic condition that signifies the adaptation of *E. coli* to a hyperosmotic shock, FROS signals in NT455 indicate the occurrence of two distinct populations of *E. coli* with regards to the conformation of *proVWX*: a population in which the operon occurs in its looped, silent conformation (d=90±1 nm, σ=12±1 nm, N=26), and another in which the operon exists in its open, active state (434±6 nm, σ=56±8 nm, N=16) (Figures 3.11C-3.11E).

The results show that the regulation of *proVWX* involves a structural aspect. To model the three-dimensional structure of *proVWX* at a higher resolution, we extended our study to ensemble chromosome conformation capture.

3C-qPCR

Hi-C probes the global conformation and contact profile of the chromosome (116). However, the Hi-C libraries prepared using PsuI as the restriction enzyme in our studies reliably afforded a resolution of only 10 kb. Since our region of interest in the chromosome that spans across the *proVWX* operon and includes its flanking regions is ~7 kb long (Supplementary file 3.1), the region is contained within a single pixel of the Hi-C maps (Figures 3.3-3.5). Hence, structural changes to the local chromatin cannot be resolved with Hi-C. Therefore, we considered 3C-qPCR (77) – a 'one-to-one' technique that probes the relative interaction frequency between only two loci at a time but affords a resolution of individual restriction digestion fragments. NlaIII was selected as the restriction enzyme for our study since the enzyme digests the regulatory region of *proVWX* in a way that separates the URE from the DRE, and the high-affinity H-NS binding sites in the DRE from each other (Figure 3.12, Supplementary file 3.1). In this experiment, fragment proU3_NlaIII (Figure 3.12, Supplementary file 3.1) that contains the *proVWX* P2 promoter and TSS was used as an anchor, and its relative interaction frequency with other restriction digestion fragments within the ~7kb region was investigated.

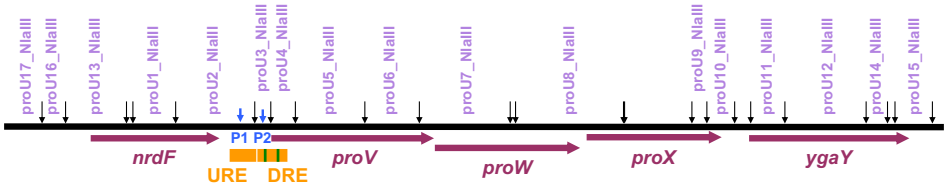


Figure 3.12: NlaIII restriction digestion sites (black arrows) and resulting chromatin fragments selected (purple labels) for the 3C-qPCR-based evaluation of local three-dimensional *proVWX*

structure. *proV*, *proW*, *proX*, *nrdF*, and *ygaY* (purple arrows) designate individual genes. P1 and P2 (blue arrows) mark the positions of the *proVWX* promoters. URE and DRE (orange) illustrate the upstream and downstream regulatory elements of *proVWX*. The green bars show the high-affinity H-NS binding sites within the DRE.

To quantify the relative interaction frequency of proU3_NlaIII with other fragments of *proVWX*, control libraries from digested and randomly re-ligated purified genomic DNA were prepared. The libraries were serially diluted and the hybrid ligation junctions were quantified at each dilution to plot a standard curve. However, only the undiluted and the 10X diluted control libraries produced a reliable fluorescent signal, thus, control libraries prepared from genomic DNA – as described in (77) – could not be used to plot standard curves. An alternative control library suggested in (77) requiring synthesis and equimolar pooling of each individual ligation junction to be probed was used instead. A preliminary test to evaluate the structure of *proVWX* in *E. coli* NT331 was performed with two biological replicates per osmotic condition. The test was carried out in LB medium.

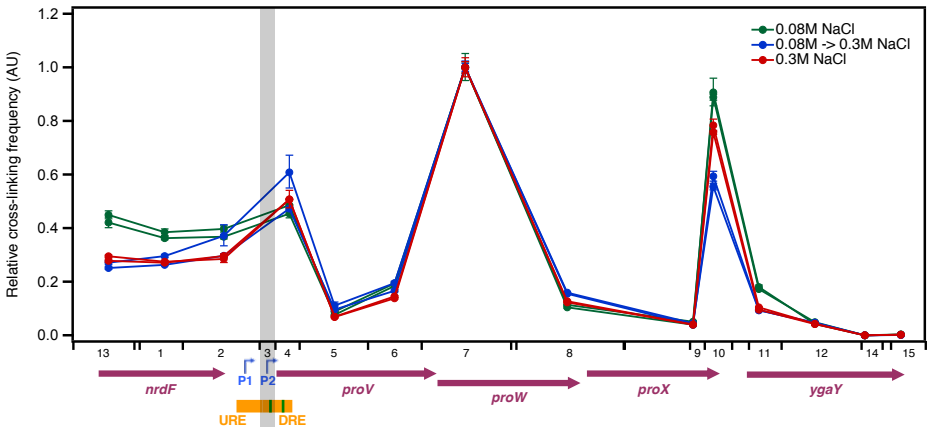


Figure 3.13: A 3C-qPCR study of the relative cross-linking frequency of proU3_NlaIII (highlighted in grey) with chromatin segments within and flanking the *proVWX* operon (marked along the horizontal axis) during exponential growth in LB medium with 0.08 M NaCl (green), following a hyperosmotic shock from 0.08 M to 0.3 M NaCl (blue), and during exponential growth at 0.3 M NaCl (red). Each curve represents the 3C profile of an independent biological replicate. The error bars represent the standard deviation of three quantitation values from the qPCR run. *proV*, *proW*, *proX*, *nrdF*, and *ygaY* (purple arrows) show individual genes. P1 and P2 (right-angled blue arrows) mark *proVWX* promoters. The upstream and downstream regulatory elements of *proVWX* (URE and DRE) are shown with orange bars. The green bars in the DRE designate high-affinity H-NS binding sites. **Internal cross-linking control:** proU3_NlaIII-proU7_NlaIII.

In NT331 cultures growing exponentially at 0.08 M NaCl, cultures subjected to a hyperosmotic shock, and cultures growing exponentially at 0.3 M NaCl, the proU3_NlaIII fragment exhibits three distinct interaction peaks. In order of

decreasing strength of the relative interaction frequency, proU3_NlaIII interacts with proU7_NlaIII which extends across the upstream half of the *proW* ORF, proU10_NlaIII which overlaps with the terminus of *proVWX*, and proU4_NlaIII that includes the downstream high-affinity H-NS binding site of the DRE (Figure 3.13; Tables S3.10-S3.12; Supplementary file 3.1).

The interaction between proU3_NlaIII and proU7_NlaIII (Figure 3.13; Tables S3.10-S3.12) sparks interest since it overlaps with the region that includes the *proW*1 amplicon (Supplementary file 3.1) that shows a dip in transcript levels in RT-qPCR studies (Figures 3.7-3.10, S3.5, and S3.8-S3.10; Tables S3.2-S3.9). The proU7_NlaIII fragment also lies just upstream of the *proW*2 RT-qPCR amplicon where the relative expression levels show a rise (Figures 3.7-3.10, S3.5, and S3.8-S3.10; Tables S3.2-S3.9). It is tempting to speculate that the proU3_NlaIII-proU7_NlaIII interaction functions as a transcription roadblock that contributes to the decreased expression at *proW*1 or even that it potentially plays a role in a coordinated regulation between P2 and the hypothetical internal promoter that may lie in this region. The proU3_NlaIII-proU7_NlaIII interaction is stable. Its stability is reflected in its reliability as an internal control. The *proVWX* interaction profiles of NT331 biological replicates at four separate osmolarity conditions tested in LB overlapped when the proU3_NlaIII-proU7_NlaIII interaction was used as an internal cross-linking control (Figure S3.11; Tables S3.10-S3.13).

NT331 *proVWX* 3C-qPCR profiles normalized to the proU3_NlaIII-proU7_NlaIII interaction show a distinctive feature at the proU10_NlaIII fragment (Figure 3.13; Tables S3.10-S3.12). At 0.08 M NaCl, when *proVWX* is repressed (Figures 3.7A, and S3.5A), the relative crosslinking frequency of proU3_NlaIII with proU10_NlaIII, a 195 bp region that includes the end of the *proX* ORF and the terminus of *proVWX* (Supplementary file 3.1), is ~0.9. This is ~4.5-fold higher than the proU3_NlaIII-proU11_NlaIII interaction and ~20-fold higher than that between proU3_NlaIII and proU9_NlaIII. This indicates that elements within fragment proU10_NlaIII form a bridge with the segment of the DRE that encompasses P2. This is in line with observations from live-cell FROS imaging that show that during exponential growth at 0.08 M NaCl, the ends of the *proVWX* operon lie in close proximity (Figure 3.11A). ChIP studies show that the DRE is bound by H-NS (Figure S3.6). Extrapolating the inferences from *in vitro* H-NS studies (20, 24), we propose a regulatory model on the basis of our ensemble 3C-qPCR and single-cell FROS microscopy observations that at low osmolarity conditions which are associated with a low intracellular concentration of K^+ (28),

H-NS adopts its bridging-capable conformation *in vivo*. The *proVWX* DRE forms an H-NS-mediated bridge with the terminus of the operon that functions as a transcription roadblock and represses *proVWX* (Figure 3.14).

A hyperosmotic shock to *E. coli* cells is associated with a rapid intracellular influx of K^+ (28). An osmotic condition that, *in vitro*, disrupts H-NS—DNA—H-NS bridges, thus, promoting the formation of transcriptionally-conducive H-NS—DNA filaments (20, 24). Following an increase in the NaCl concentration of the growth medium from 0.08 M to 0.3 M, the relative interaction frequency between proU3_NlaIII and proU10_NlaIII drops ~30% to ~0.6 (Figure 3.13; Tables S3.10 and S3.11). Indeed, live cell FROS imaging shows the divergence of the ends of *proVWX* following a hyperosmotic shock (Figure 3.11B). We expand our model to include the dismantling of the bridging interaction between the region of the DRE that carries the P2 promoter and the terminus of *proVWX* following a hyperosmotic shock to form an H-NS—DNA structure that does not impede *proVWX* expression (Figure 3.14).

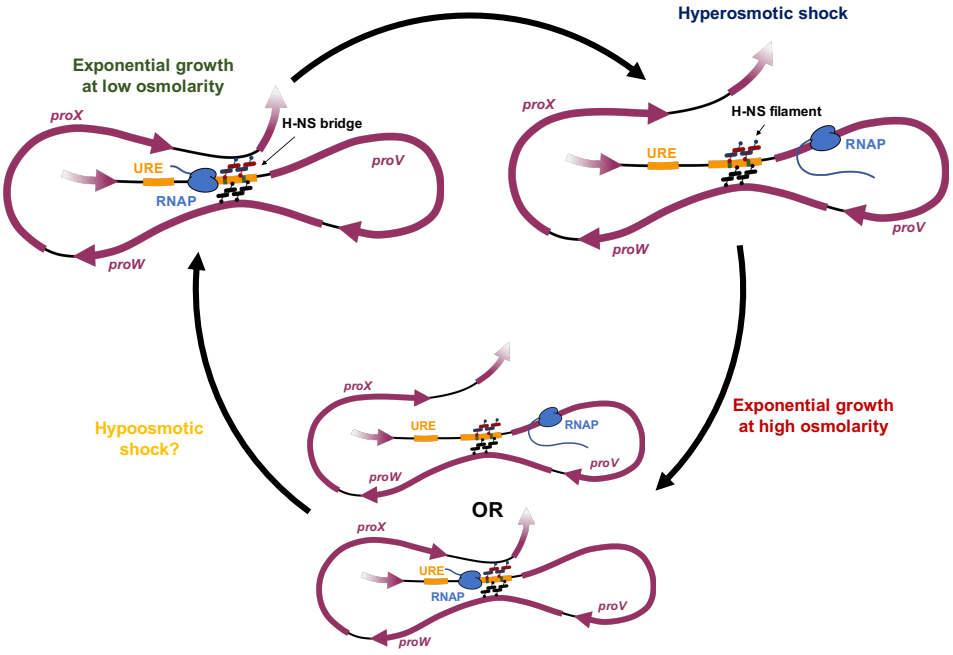


Figure 3.14: A model of the structural regulation of *proVWX* expression. Exponential growth at low osmolarity: During exponential growth at low osmolarity, a DNA—NAP—DNA bridge, likely mediated by H-NS, forms between the DRE and the terminus of *proVWX*. The bridge acts as a transcription roadblock that represses the operon. *In vitro* studies suggest that the formation of the H-NS-mediated bridge is promoted by the low cytoplasmic concentrations of K^+ that occur during growth in low osmolarity media. Decreased K^+ favors the bridging-capable conformation of H-NS (20). *In vitro*

experiments also indicate that the DNA—H-NS—DNA nucleoprotein structure functions as a transcription roadblock (24). **Hyperosmotic shock:** A hyperosmotic shock to *E. coli* drives a rapid cytoplasmic influx of K^+, that, under *in vitro* conditions, de-stabilises H-NS-mediated DNA bridges to form H-NS—DNA filaments that are conducive to transcription (20, 24). *In vivo*, a hyperosmotic shock dismantles the bridge between the DRE and the end of *proVWX* to form an H-NS—DNA filament across which RNA polymerase can transcribe. **Exponential growth at high osmolarity:** Adaptation to hyperosmotic stress is associated with the intracellular accumulation of osmoprotectants, and the decline of cytoplasmic K^+. The decrease in K^+ re-instates DNA bridging by H-NS that manifests as the re-establishment of a bridge between the DRE and the *proVWX* terminus. The bridge only forms in a fraction of cells where it silences the operon. In another population of *E. coli*, a decrease in the cytoplasmic concentration of osmoprotectants triggers the influx of K^+ that dismantles the bridge between the DRE and the terminus of *proVWX*, hence, activating the operon. **Hypoosmotic shock:** Preliminary experiments indicate that a hypoosmotic shock to *E. coli* rapidly establishes the bridge between the *proVWX* DRE and terminus to silence the operon.

Adaptation to hyperosmotic stress is associated with the intracellular accumulation of osmoprotectants, the decline of cytoplasmic K^+ (28), and, consequentially, the repression of *proVWX* (Figures 3.7C and S3.5C; Tables S3.2 and S3.3). During exponential growth at 0.3 M NaCl, the relative interaction frequency between proU3_NlaIII and proU10_NlaIII is increased compared to *E. coli* subjected to a hyperosmotic shock (Figure 3.13; Tables S3.11 and S3.12), but is still lower than exponential growth at 0.08 M NaCl (Figure 3.13; Tables S3.10 and S3.12). This observation from ensemble 3C-qPCR supports single-cell FROS microscopy studies that show that *proVWX* in *E. coli* cells in a medium with 0.3 M NaCl occurs either in its open, active conformation or in its closed, repressed state (Figure 3.11C). These results suggest that upon adaptation to high osmolarity, the decrease in intracellular K^+ favors the formation of H-NS—DNA—H-NS bridges. The formation of a bridge between proU3_NlaIII and proU10_NlaIII in a fraction of cells silences *proVWX*. In another population of *E. coli*, a decrease in the cytoplasmic concentration of osmoprotectants triggers the influx of K^+, the dismantling of the bridge between proU3_NlaIII and proU10_NlaIII, and hence, the expression of *proVWX* (Figure 3.14).

Fragment proU3_NlaIII also shows a high relative interaction frequency with fragment proU4_NlaIII that encompasses the downstream half of the DRE and contains the second high-affinity H-NS binding site (Figures 3.12-3.13; Supplementary file 3.1). Under the osmotic conditions investigated in this study, the frequency of the proU3_NlaIII-proU4_NlaIII interaction remains comparable (Figure 3.13), and may, therefore, be mediated by the physical proximity of the fragments rather than a regulatory purpose.

In NT331 cultures growing at 0.08 M NaCl, proU3_NlaIII shows a higher relative interaction frequency with proU13_NlaIII than in cultures growing at 0.3 M NaCl or cultures subjected to a hyperosmotic shock (Figure 3.13). Fragment proU13_NlaIII encompasses the end of *nrdE* and the beginning of *nrdF* (Figure 3.12, Supplementary file 3.1). Preliminary studies show that the proU3_NlaIII-proU13_NlaIII interaction in cultures subjected to a hypoosmotic shock from 0.3 M to 0.08 M NaCl is similar to cultures growing exponentially at 0.08 M NaCl (Figure S3.11). This implies the presence of a regulatory element within this region. Additional investigations will be required to identify the regulatory element and its role.

Single molecule Förster resonance energy transfer (smFRET) studies

3C-qPCR detected the non-dynamic, ensemble organization of *proVWX*. To study the single-molecule dynamics and the biophysical mechanism of the organization of *proVWX*, we visualized the structural dynamics of elements of *proVWX* immobilized on a glass surface using single-molecule Förster resonance energy transfer (smFRET) in combination with TIRF microscopy. The combinatorial approach affords a spatial resolution of 10-100Å, and a temporal resolution of 33 ms.

To verify the applicability of smFRET in visualizing structural dynamics of immobilized DNA molecules, a 250 bp-long DNA construct used in an earlier study by our laboratory to report on H-NS-mediated DNA bridging (20) was labelled with Cy3 and Cy5 as FRET donor and acceptor fluorophores, respectively, and with Biotin-TEG for immobilization onto a streptavidin-coated glass surface (Table S3.1). Henceforth, constructs with the above described features will be referred to as smFRET-compatible. FRET between the donor and acceptor fluorophores was not observed in the absence of H-NS (Figure 3.15A). In the presence of 200 nM H-NS, in a buffer supplemented with 10 mM Mg^{2+} to favor H-NS-mediated bridging (20), two structural configurations of the DNA substrate with FRET efficiencies of 0.16±0.07 and 0.50±0.12 were detected, corresponding to distances of 79Å and 60Å, respectively, between the fluorophores (Figures 3.15B-3.15D). These observations show intra-molecular bridging by H-NS, and highlight that the structural change associated with the formation of the bridges and the structural state adopted by an individual DNA molecule can be detected by smFRET.

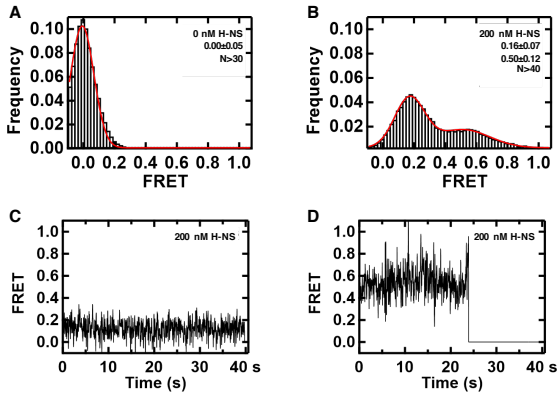


Figure 3.15: H-NS forms intramolecular bridges within a 250 bp-long, AT-rich DNA substrate. A: In the absence of H-NS, the FRET efficiency recorded for 250bp AT-rich DNA substrates is 0.00±0.05 (N>30). **B:** At 200 nM H-NS, the DNA substrates are folded into two structural configurations with FRET efficiencies of 0.16±0.07 and 0.50±0.12 (N>40). **C, D:** Representative FRET traces of single DNA molecules in the presence of 200 nM H-NS.

H-NS does not mediate the formation of a bridge between the proVWX URE and DRE

The *proVWX* URE and DRE function in concert to coordinate the expression of *proVWX* (42). Whether this requires a structural interaction between the two regulatory elements, such as the formation of a loop, is unclear. The 3C-qPCR studies presented earlier do not provide conclusive evidence that structural organization contributes to the coordinated role of the URE and DRE. Therefore, we investigated if the formation of bridge between the URE and DRE mediated by H-NS is possible.

A 460 bp smFRET-compatible DNA construct spanning the regulatory elements was used for this purpose. Given that the persistence length of DNA is 50 nm – approximately 150 bp – the 460 bp construct covers three persistence lengths and is, therefore, capable of forming an end-to-end loop. Despite extensive data acquisition in a Mg^{2+}-supplemented buffer (20) at H-NS concentrations ranging from 10-200 nM, only few reliable FRET traces representing H-NS-mediated bridge formation were identified. This points to the rarity of the event, indicating that either *proVWX* regulation by the URE and DRE may not involve the formation of a loop between the pair of elements, or that a NAP-mediated bridge cannot form within the 460 bp URE-DRE construct. To account for the latter, the smFRET experiments were repeated with Rok and MvaT – H-NS family proteins in *Bacillus sp*. and *Pseudomonas sp*., respectively (21, 117). Rok folds the DNA construct to form intramolecular bridges with FRET efficiencies of 0.19±0.09 and 0.42±0.12 (Figures 3.16A-3.16C). A mutated variant of MvaT ($MvaT_{F36D,M44D}$) that cannot

oligomerize, and, therefore, forms DNA—MvaT—DNA bridges but not DNA—MvaT filaments (21), also folds the 460 bp construct into structures with FRET efficiencies between ~0.2 and ~0.5 (Figures 3.16D-3.16F). Collectively, the results show that a NAP-mediated bridged complex within a DNA fragment comprising of the *proVWX* URE and DRE can be formed. However, the formation of the intramolecular bridge by H-NS is a rare occurrence and may, therefore, not carry a biological significance.

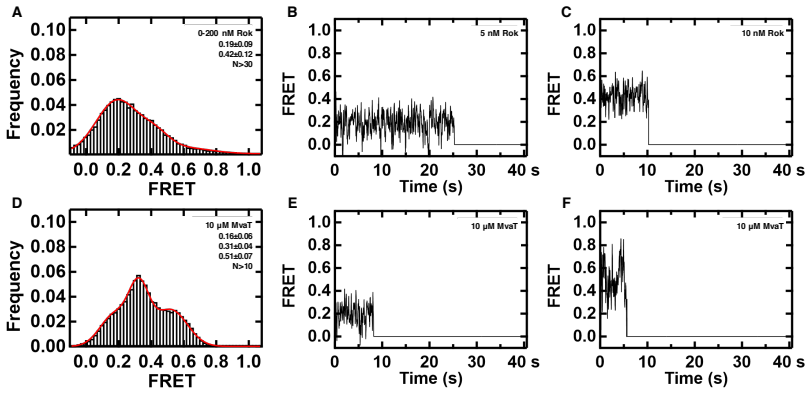


Figure 3.16: A bridged nucleoprotein complex between the *proVWX* URE and DRE can be mediated by Rok and MvaT$_{F36D,M44D}$. A: Rok folds the URE-DRE DNA construct to form intramolecular bridges with FRET efficiencies of 0.19±0.09 and 0.42±0.12 (N>30). **B, C:** Representative FRET traces of Rok-structured DNA constructs. **D:** MvaT$_{F36D,M44D}$ folds the DNA into structures with FRET efficiencies between ~0.2 and ~0.5. **E, F:** Representative traces of MvaT-structured DNA molecules.

The interaction between the proVWX P2 and proW

3C-qPCR of *proVWX* shows that the chromatin fragment proU3_NlaIII that carries the P2 promoter interacts with fragment proU7_NlaIII that spans the upstream half of *proW*. The resolution afforded by NlaIII as the restriction enzyme in the 3C study cannot pinpoint the precise location of the interaction. To elucidate this, a preliminary *in vitro* smFRET study was performed. Six smFRET-compatible DNA constructs – ProW1 to ProW6 – were designed. The Biotin-TEG- and Cy3-labelled ends of the constructs were positioned at -46 relative to P2, and the Cy5-labelled ends terminated at various positions within the proU7_NlaIII fragment, ~100 bp apart (Figure 3.17). FRET efficiencies for the constructs were determined at 0 nM and 200 nM H-NS in a buffer supplemented with 10 mM Mg^{2+}. H-NS was selected for the study since the protein is a bridging-capable NAP that shows a strong ChIP signal in the proU3_NlaIII fragment (67, 102).

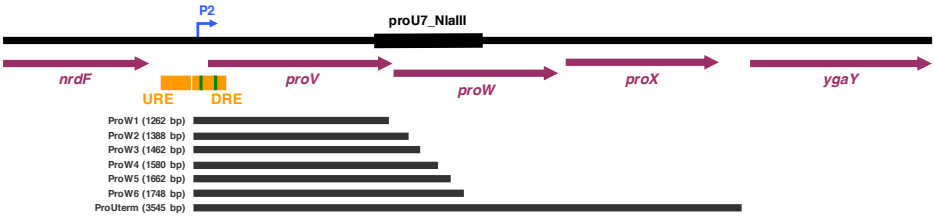


Figure 3.17: smFRET constructs (grey) designed to study loop formation between the *proVWX* P2 promoter (right-angle arrow, blue), and the proU7_NlaIII fragment (black bar) and terminus of the operon. Purple arrows represent ORFs. URE and DRE (orange) mark the positions of the upstream and downstream regulatory elements of *proVWX*. The green bars within the DRE illustrate the positions of high-affinity H-NS binding sites.

FRET was not observed for ProW1 and ProW2 in the absence or presence of H-NS, indicating that H-NS does not form a bridge between P2 and the ends of the ProW1 and ProW2 constructs under our experimental conditions (Figures 3.18A-3.18B). It is important to note that our experimental design is only sensitive to the formation of a bridge between the ends of an smFRET compatible substrate. Nucleoprotein structures that form internally within the substrate, or between a labelled end and an unlabelled section of the construct cannot be detected.

Constructs ProW3 and ProW4 both show a low efficiency FRET peak, in addition to a peak at 0 FRET at 0 nM H-NS (Figures 3.18C-3.18D, left panels). This suggests that the constructs spend a significant time fraction exploring a bent conformation even in the absence of protein. Indeed, when fragments ProW3 and ProW4 are evaluated with bend.it® (118), a server that predicts the curvature of a DNA molecule from its sequence using the BEND algorithm (119), 28.4% of the ProW3 sequence and 29.2% of ProW4 are predicted to have an intrinsic curvature of >5 degrees per helical turn when the bend angle is calculated over three turns considering an ideal B-DNA structure. An intrinsic curvature of >7.5 degrees per helical turn is predicted for ~9.5% of the two sequences (Predicted curvature window size: 31 bp, curvature parameter: derived from DNaseI digestion, length of helical turn: 10.5 bp). Nevertheless, the bends may not be sufficiently in phase to bring the ends of the DNA substrate in proximity for FRET. The low efficiency FRET state may also be accounted for by DNA duplex pairing (120). The phenomenon suggests that sequence of bases in a DNA duplex generate a backbone twist that doubles as a fingerprint of the molecule. The twist generates a 'coat' of alternating positive and negative charges (in the major groove, and at the phosphate backbone, respectively) around the duplex. The electrostatic 'coat' is only in register with that of a homologous sequence and drives a preferential pairing of the helices. In the presence of Mg^{2+}, a cation that specifically interacts

with DNA, the positive charge of the groove is strengthened and can overcome the repulsion between dsDNA molecules that is mediated by the negatively-charged phosphate backbones (120). Weak homology between sequences at the ends of the smFRET-compatible constructs used in our study may favour a bent conformation, hence a low FRET signal, even in the absence of H-NS. In follow-up studies, atomic force microscopy visualizations of the ProW3 and ProW4 constructs will provide clues regarding preferential conformations that may be adopted by the molecules (121, 122). With 200 nM H-NS, construct ProW3, but not ProW4, shows an additional peak at 0.41±0.15 FRET, corresponding to a distance of 54 Å (Figures 3.18C-3.18D, right panels). This indicates that the end of the ProW3 fragment may form an H-NS-mediated interaction with P2.

Similar to constructs ProW1 and ProW2, construct ProW5 does not show FRET in the absence of H-NS (Figure 3.18E, left panel). At 200 nM H-NS, however, a peak at 0.65±0.08 FRET becomes apparent (Figure 3.18E, right panel). The rarity of this event means that additional data acquisition will be required to validate its significance.

The smFRET observations for construct ProW6 are remarkable. At both, 0 nM and 200 nM H-NS, events with a FRET efficiency of ~0.8 corresponding to a distance of ~47Å are observed (Figure 3.18F). Furthermore, the probability of observing the high FRET state increases in the presence of H-NS. However, traces of this state do not show single-step photobleaching and can only be considered artefactual. To validate the high FRET states by single-step photobleaching, the event will be stimulated in follow-up studies by increasing laser power and measurement duration, and by reducing the concentrations of the oxygen scavenger and anti-blinking/anti-bleaching agent used in the experimental set-up (78). At 200 nM H-NS, events with a FRET efficiency of 0.48±0.06 (61Å) that were validated with single-step photobleaching were observed, indicating that the end of construct ProW6 forms an H-NS-mediated interaction with P2 (Figure 3.18F, right panel). This may account for the high frequency interaction observed between fragments proU3_NlaIII and proU7_NlaIII in 3C-qPCR studies. Interestingly, the additional 86 bp in the ProW6 construct compared to ProW5 lie between the *proW*1 and *proW*2 amplicons investigated with RT-qPCR. On the basis of our RT-qPCR results, we hypothesized that an internal promoter lies in this region.

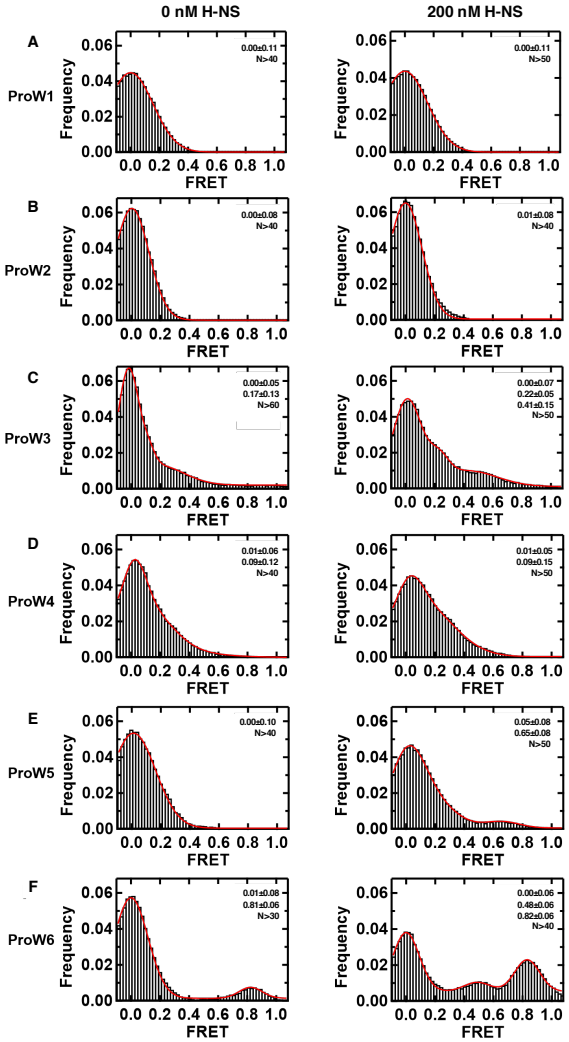


Figure 3.18: H-NS stabilises *proVWX* DRE-ProW3, and DRE-ProW6 loops. FRET distributions of ProW1-ProW6 (A-F) fragments with 0 nM (left) and 200 nM (right) H-NS.

The interaction between proVWX P2 and terminator.

3C-qPCR shows that the segment of the DRE that carries P2 forms an osmo-sensitive interaction with the terminus of *proVWX* that may play a role in the regulation of the operon. We tested the role of H-NS in mediating this interaction using an smFRET-compatible construct extending from -46 to +3499 relative to P2. At 0 nM H-NS no FRET was observed for the constructs with the exception of a single molecule with a FRET efficiency of 0.81±0.09 (Figure 3.19A). Similar to the ~0.8 FRET state detected for the ProW6 constructs, this molecule did not

photobleach and its observation is likely artefactual. At 200 nM H-NS, artefactual high FRET states at 0.95±0.05 were observed (Figure 3.19B). Nevertheless, two additional FRET states with efficiencies of 0.23±0.07 and 0.47±0.05 (N=9) were detected, supporting the occurrence of a bridge between the DRE and the terminus of *proVWX* (Figure 3.19B).

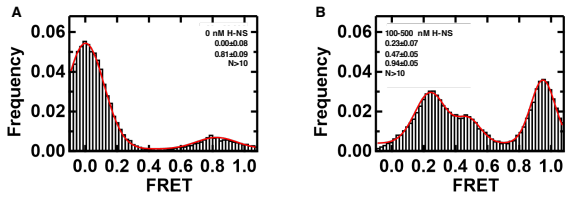


Figure 3.19: H-NS can form a bridge between the *proVWX* DRE and the terminus of the operon. **A:** FRET was not observed for the ProUterm (Figure 3.12) construct at 0 nM H-NS (N>10) with the exception of a single molecule exhibiting an (artefactual) FRET efficiency of 0.81±0.09. **B:** At 200 nM H-NS, FRET states with efficiencies of 0.23±0.07 and 0.47±0.05 (N=9) were detected, in addition to an artefactual FRET state at 0.95±0.05.

Conclusions and perspectives

Nucleoid associated proteins are environmentally sensitive architectural proteins of the bacterial chromosome that drive its compaction and organisation. NAPs also function as transcription factors and regulate the expression of genes in response to environmental stimuli (1–3). We have provided the first *in vivo* evidence that NAPs regulate gene expression by re-modelling local chromatin using the *Escherichia coli proVWX* operon as a model system.

Using a modified 3C-based library preparation technique that includes methanol fixation prior to formaldehyde fixation, and proximity ligation with the insoluble fraction of digested, cross-linked chromatin to improve signal-to-noise ratio, we show that the *E. coli* chromosome undergoes global rearrangements during growth at different osmotic conditions. Ensemble RT-qPCR experiments highlight that conditions which modify chromosome conformation also affect gene expression — *proVWX* in our study — and reveal that despite the occurrence of *proV*, *proW*, and *proX* within a single operon, individual genes show a degree of independent regulation. The role of NAPs in *proVWX* regulation was verified by manipulating H-NS binding profiles at the URE and DRE of the operon and in a Δ*stpA* strain, both of which showed altered gene expression profiles.

The role of StpA in the regulation of *proVWX* is intriguing. RT-qPCR studies show that the deletion of StpA is associated with an up to ~8-fold increase in relative

expression in *proVWX*, an increase in the processivity of RNA polymerase across *proV*, and the loss of insulation of *proVWX* from flanking operons. Collectively, this underlines that StpA has specific functions in transcription regulation that are distinct from those of H-NS and that the deletion of *stpA* from *E. coli* that still expresses H-NS comes at a physiological cost.

By using TIRF microscopy to observe FROS-labelled *proVWX* in single cells, and ensemble 3C-qPCR to model the local three-dimensional conformation of the operon, we show that conditions that repress *proVWX* are associated with the formation of a loop that is anchored between the *proVWX* P2 promoter region and the terminus of the operon. Destabilisation of the loop during hyperosmotic stress is associated with the activation of *proVWX*. The 3C-qPCR studies also reveal an interaction between P2 and the region spanning the end of *nrdE* and the beginning of *nrdF* that is evidently stronger at hypoosmotic stress than hyperosmotic stress. This suggests that the regulatory elements of individual operons may occur within other transcriptional units beyond their mapped regulatory regions, illustrating the importance of studying the regulation of transcriptional units at native chromosomal loci.

It is noteworthy that in *proVWX*, an H-NS-regulated operon, low K^+ concentrations that favour the bridging competent form of H-NS *in vitro* also facilitate the formation of a loop between the ends of the operon *in vivo*. Similarly, higher K^+ concentrations that destabilise the bridging competent structure of H-NS *in vitro* also destabilise the loop between the ends of *proVWX in vivo*. Furthermore, smFRET studies reported here indicate that H-NS can form a loop between the ends of *proVWX*. It is curious that while ChIP studies of the H-NS binding profile show the occurrence of a DNA—H-NS nucleoprotein structure at *proVWX* P2, a corresponding signal is not observed at the terminus of the operon.

The results presented in this chapter show that transcription of an H-NS-regulated operon correlates with its three-dimensional structure. H-NS and H-NS-like proteins are wide-spread amongst bacteria, boasting large regulons in these organisms (117, 123–126). This implies that the structural regulation of gene expression may be a typical feature of transcriptional control in bacteria.

References:

1. Dame,R.T. (2005) The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol. Microbiol.*, 10.1111/j.1365-2958.2005.04598.x.
2. Dame,R.T., Kalmykowa,O.J. and Grainger,D.C. (2011) Chromosomal macrodomains and associated proteins: Implications for DNA organization and replication in gram negative bacteria. *PLoS Genet.*, 10.1371/journal.pgen.1002123.
3. Dame,R.T., Rashid,F.Z.M. and Grainger,D.C. (2020) Chromosome organization in bacteria: mechanistic insights into genome structure and function. *Nat. Rev. Genet.*, **21**, 227–242.
4. Dillon,S.C. and Dorman,C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.*, 10.1038/nrmicro2261.
5. Dorman,C.J. (2013) Genome architecture and global gene regulation in bacteria: making progress towards a unified model? *Nat. Rev. Microbiol.*, **11**, 349–55.
6. Rimsky,S. and Travers,A. (2011) Pervasive regulation of nucleoid structure and function by nucleoid-associated proteins. *Curr. Opin. Microbiol.*, 10.1016/j.mib.2011.01.003.
7. Luijsterburg,M.S., White,M.F., Van Driel,R. and Th. Dame,R. (2008) The major architects of chromatin: Architectural proteins in bacteria, archaea and eukaryotes. *Crit. Rev. Biochem. Mol. Biol.*, 10.1080/10409230802528488.
8. Dame,R.T. and Tark-Dame,M. (2016) Bacterial chromatin: Converging views at different scales. *Curr. Opin. Cell Biol.*, **40**, 60–65.
9. Noom,M.C., Navarre,W.W., Oshima,T., Wuite,G.J.L. and Dame,R.T. (2007) H-NS promotes looped domain formation in the bacterial chromosome. *Curr. Biol.*, 10.1016/j.cub.2007.09.005.
10. van der Valk,R.A., Vreede,J., Crémazy,F. and Dame,R.T. (2015) Genomic Looping: A Key Principle of Chromatin Organization. *J. Mol. Microbiol. Biotechnol.*, **24**, 344–359.
11. Atlung,T. and Ingmer,H. (1997) H-NS: A modulator of environmentally regulated gene expression. *Mol. Microbiol.*, 10.1046/j.1365-2958.1997.3151679.x.
12. 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 (80-.).*, 10.1126/science.1128794.
13. Falconl,M., Gualtierl,M.T., La Teana,A., Losso,M.A. and Pon,C.L. (1988) Proteins from the prokaryotic nucleoid: primary and quaternary structure of the 15-kD Escherichia coli DNA binding protein H-NS. *Mol. Microbiol.*, 10.1111/j.1365-2958.1988.tb00035.x.
14. Bloch,V., Yang,Y., Margeat,E., Chavanieu,A., Augé,M.T., Robert,B., Arold,S., Rimsky,S. and Kochoyan,M. (2003) The H-NS dimerization domain defines a new fold contributing to DNA recognition. *Nat. Struct. Biol.*, 10.1038/nsb904.
15. Esposito,D., Petrovic,A., Harris,R., Ono,S., Eccleston,J.F., Mbabaali,A., Haq,I., Higgins,C.F., Hinton,J.C.D., Driscoll,P.C., *et al.* (2002) H-NS oligomerization domain structure reveals the mechanism for high order self-association of the intact protein. *J. Mol. Biol.*, 10.1016/S0022-2836(02)01141-5.
16. Ueguchi,C., Suzuki,T., Yoshida,T., Tanaka,K.I. and Mizuno,T. (1996) Systematic mutational analysis revealing the functional domain organization of Escherichia coli nucleoid protein H-NS. *J. Mol. Biol.*, 10.1006/jmbi.1996.0566.
17. Shindo,H., Iwaki,T., Ieda,R., Kurumizaka,H., Ueguchi,C., Mizuno,T., Morikawa,S., Nakamura,H. and Kuboniwa,H. (1995) Solution structure of the DNA binding domain of a nucleoid-associated protein, H-NS, from Escherichia coli. *FEBS Lett.*, 10.1016/0014-5793(95)00079-O.
18. Arold,S.T., Leonard,P.G., Parkinson,G.N. and Ladbury,J.E. (2010) H-NS forms a superhelical protein scaffold for DNA condensation. *Proc. Natl. Acad. Sci. U. S. A.*, 10.1073/pnas.1006966107.
19. 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.*, 10.1111/j.1365-2958.2009.06754.x.
20. 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*, 10.7554/eLife.27369.
21. Qin,L., Bdira,F. Ben, Sterckx,Y.G.J., Volkov,A.N., Vreede,J., Giachin,G., Van Schaik,P., Ubbink,M. and Dame,R.T. (2020) Structural basis for osmotic regulation of the DNA binding properties of H-NS proteins. *Nucleic Acids Res.*, 10.1093/nar/gkz1226.
22. Dame,R.T., Wyman,C. and Goosen,N. (2000) H-NS mediated compaction of DNA visualised by

atomic force microscopy. *Nucleic Acids Res.*, 10.1093/nar/28.18.3504.
23. Dame,R.T., Noom,M.C. and Wuite,G.J.L. (2006) Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature*, **444**, 387–390.
24. 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*, 10.7554/eLife.04970.
25. Sexton,T. and Cavalli,G. (2015) The role of chromosome domains in shaping the functional genome. *Cell*, **160**, 1049–1059.
26. Csonka,L.N. (1982) A third L-proline permease in Salmonella typhimurium which functions in media of elevated osmotic strength. *J. Bacteriol.*, **151**, 1433–1443.
27. Gowrishankar,J. (1985) Identification of osmoresponsive genes in Escherichia coli: Evidence for participation of potassium and proline transport systems in osmoregulation. *J. Bacteriol.*, 10.1128/jb.164.1.434-445.1985.
28. Sleator,R.D. and Hill,C. (2002) Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. *FEMS Microbiol. Rev.*, 10.1111/j.1574-6976.2002.tb00598.x.
29. Cairney,J., Booth,I.R. and Higgins,C.F. (1985) Osmoregulation of gene expression in Salmonella typhimurium: proU encodes an osmotically induced betaine transport system. *J. Bacteriol.*, 10.1128/jb.164.3.1224-1232.1985.
30. Barron,A., May,G., Bremer,E. and Villarejo,M. (1986) Regulation of envelope protein composition during adaptation to osmotic stress in Escherichia coli. *J. Bacteriol.*, 10.1128/jb.167.2.433-438.1986.
31. Sutherland,L., Cairney,J., Elmore,M.J., Booth,I.R. and Higgins,C.F. (1986) Osmotic regulation of transcription: Induction of the proU betaine transport gene is dependent on accumulation of intracellular potassium. *J. Bacteriol.*, 10.1128/jb.168.2.805-814.1986.
32. Rajkumari,K., Kusano,S., Ishihama,A., Mizuno,T. and Gowrishankar,J. (1996) Effects of H-NS and potassium glutamate on σ(S)- and σ70-directed transcription in vitro from osmotically regulated P1 and P2 promoters of proU in Escherichia coli. *J. Bacteriol.*, 10.1128/jb.178.14.4176-4181.1996.
33. Jovanovich,S.B., Record,M.T. and Burgess,R.R. (1989) In an Escherichia coli coupled transcription-translation system, expression of the osmoregulated gene proU is stimulated at elevated potassium concentrations and by an extract from cells grown at high osmolality. *J. Biol. Chem.*
34. Ramirez,R.M. and Villarejo,M. (1991) Osmotic signal transduction to proU is independent of DNA supercoiling in Escherichia coli. *J. Bacteriol.*, 10.1128/jb.173.2.879-885.1991.
35. Dinnbier,U., Limpinsel,E., Schmid,R. and Bakker,E.P. (1988) Transient accumulation of potassium glutamate and its replacement by trehalose during adaptation of growing cells of Escherichia coli K-12 to elevated sodium chloride concentrations. *Arch. Microbiol.*, 10.1007/BF00408306.
36. Gowrishankar,J. (1989) Nucleotide sequence of the osmoregulatory proU operon of Escherichia coli. *J. Bacteriol.*, 10.1128/jb.171.4.1923-1931.1989.
37. May,G., Faatz,E., Villarejo,M. and Bremer,E. (1986) Binding protein dependent transport of glycine betaine and its osmotic regulation in Escherichia coli K12. *MGG Mol. Gen. Genet.*, 10.1007/BF00430432.
38. Lucht,J.M. and Bremer,E. (1994) Adaptation of Escherichia coli to high osmolarity environments: Osmoregulation of the high-affinity glycine betaine transport system ProU. *FEMS Microbiol. Rev.*, 10.1111/j.1574-6976.1994.tb00067.x.
39. Gowrishankar,J. and Manna,D. (1996) How is osmotic regulation of transcription of the Escherichia coli proU operon achieved? A review and a model. *Genetica*, 10.1007/BF00055322.
40. Rajkumari,K. and Gowrishankar,J. (2001) In vivo expression from the RpoS-dependent P1 promoter of the osmotically regulated proU operon in Escherichia coli and Salmonella enterica serovar typhimurium: Activation by rho and hns mutations and by cold stress. *J. Bacteriol.*, 10.1128/JB.183.22.6543-6550.2001.
41. Gowrishankar,J. (1999) Ploughing a lonely furrow: The curious case of the P1 promoter in the osmotically regulated proU operon of Escherichia coli. *J. Indian Inst. Sci.*
42. Nagarajavel,V., Madhusudan,S., Dole,S., Rahmouni,A.R. and Schnetz,K. (2007) Repression by binding of H-NS within the transcription unit. *J. Biol. Chem.*, 10.1074/jbc.M702753200.

43. Dattananda,C.S., Rajkumari,K. and Gowrishankar,J. (1991) Multiple mechanisms contribute to osmotic inducibility of proU operon expression in Escherichia coli: Demonstration of two osmoresponsive promoters and of a negative regulatory element within the first structural gene. *J. Bacteriol.*, 10.1128/jb.173.23.7481-7490.1991.
44. Ueguchi,C. and Mizuno,T. (1993) The Escherichia coli nucleoid protein H-NS functions directly as a transcriptional repressor. *EMBO J.*, 10.1002/j.1460-2075.1993.tb05745.x.
45. Prince,W.S. and Villarejo,M.R. (1990) Osmotic control of proU transcription is mediated through direct action of potassium glutamate on the transcription complex. *J. Biol. Chem.*
46. Overdier,D.G. and Csonka,L.N. (1992) A transcriptional silencer downstream of the promoter in the osmotically controlled proU operon of Salmonella typhimurium. *Proc. Natl. Acad. Sci. U. S. A.*, 10.1073/pnas.89.7.3140.
47. Kavalchuk,K., Srinivasan,M. and Schnetz,K. (2012) RNase III initiates rapid degradation of proU mRNA upon hypo-osmotic stress in Escherichia coli. *RNA Biol.*, 10.4161/rna.9.1.18228.
48. 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.*, 10.1038/nsmb1233.
49. Lucht,J.M., Dersch,P., Kempf,B. and Bremer,E. (1994) Interactions of the nucleoid-associated DNA-binding protein H-NS with the regulatory region of the osmotically controlled proU operon of Escherichia coli. *J. Biol. Chem.*
50. Tanaka,K. ichi, Ueguchi,C. and Mizuno,T. (1994) Importance of Stereospecific Positioning of the Upstream cis-Acting DNA Element Containing a Curved DNA Structure for the Functioning of the Escherichia coli proV Promoter. *Biosci. Biotechnol. Biochem.*, 10.1271/bbb.58.1097.
51. Dame,R.T., Wyman,C. and Goosen,N. (2001) Structural basis for preferential binding of H-NS to curved DNA. *Biochimie*, 10.1016/S0300-9084(00)01213-X.
52. Yamada,H., Muramatsu,S. and Mizuno,T. (1990) An Escherichia coli protein that preferentially binds to sharply curved DNA. *J. Biochem.*, 10.1093/oxfordjournals.jbchem.a123216.
53. Yamada,H., Yoshida,T., Tanaka,K. ichi, 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. *MGG Mol. Gen. Genet.*, 10.1007/BF00290685.
54. Owen-Hughes,T.A., Pavitt,G.D., Santos,D.S., Sidebotham,J.M., Hulton,C.S.J., Hinton,J.C.D. and Higgins,C.F. (1992) The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. *Cell*, 10.1016/0092-8674(92)90354-F.
55. Tanaka,K. ichi, Muramatsu,S., Yamada,H. and Mizuno,T. (1991) Systematic characterization of curved DNA segments randomly cloned from Escherichia coli and their functional significance. *MGG Mol. Gen. Genet.*, 10.1007/BF00260648.
56. Dame,R.T., Wyman,C., Wurm,R., Wagner,R. and Goosen,N. (2002) Structural basis for H-NS-mediated trapping of RNA polymerase in the open initiation complex at the rrnB P1. *J. Biol. Chem.*, 10.1074/jbc.C100603200.
57. Shin,M., Song,M., Joon,H.R., Hong,Y., Kim,Y.J., Seok,Y.J., Ha,K.S., Jung,S.H. and Choy,H.E. (2005) DNA looping-mediated repression by histone-like protein H-NS: Specific requirement of Eσ70 as a cofactor for looping. *Genes Dev.*, 10.1101/gad.1316305.
58. Lucht,J.M. and Bremer,E. (1991) Characterization of mutations affecting the osmoregulated proU promoter of Escherichia coli and identification of 5' sequences required for high-level expression. *J. Bacteriol.*, 10.1128/jb.173.2.801-809.1991.
59. Manna,D. and Gowrishankar,J. (1994) Evidence for involvement of proteins HU and RpoS in transcription of the osmoresponsive proU operon in Escherichia coli. *J. Bacteriol.*, 10.1128/jb.176.17.5378-5384.1994.
60. Prieto,A.I., Kahramanoglou,C., Ali,R.M., Fraser,G.M., Seshasayee,A.S.N. and Luscombe,N.M. (2012) Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in Escherichia coli K12. *Nucleic Acids Res.*, **40**, 3524–3537.
61. Oberto,J., Nabti,S., Jooste,V., Mignot,H. and Rouviere-Yaniv,J. (2009) The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. *PLoS One*, 10.1371/journal.pone.0004367.
62. Khodr,A., Fairweather,V., Bouffartigues,E. and Rimsky,S. (2015) IHF is a trans-acting factor implicated in the regulation of the proU P2 promoter. *FEMS Microbiol. Lett.*, 10.1093/femsle/fnu049.
63. Shi,X. and Bennett,G.N. (1994) Plasmids bearing hfq and the hns-like gene stpA complement hns

mutants in modulating arginine decarboxylase gene expression in Escherichia coli. *J. Bacteriol.*, 10.1128/jb.176.21.6769-6775.1994.
64. Zhang,A. and Belfort,M. (1992) Nucleotide sequence of a newly-identified Escherichia coli gene, stpA, encoding an H-NS-like protein. *Nucleic Acids Res.*, 10.1093/nar/20.24.6735.
65. Zhang,A., Rimsky,S., Reaban,M.E., Buc,H. and Belfort,M. (1996) Escherichia coli protein analogs StpA and H-NS: Regulatory loops, similar and disparate effects on nucleic acid dynamics. *EMBO J.*, 10.1002/j.1460-2075.1996.tb00476.x.
66. Williams,R.M., Rimsky,S. and Buc,H. (1996) Probing the structure, function, and interactions of the Escherichia coli H-NS and StpA proteins by using dominant negative derivatives. *J. Bacteriol.*, 10.1128/jb.178.15.4335-4343.1996.
67. Uyar,E., Kurokawa,K., Yoshimura,M., Ishikawa,S., Ogasawara,N. and Oshima,T. (2009) Differential binding profiles of StpA in wild-type and hns mutant cells: A comparative analysis of cooperative partners by chromatin immunoprecipitation- microarray analysis. *J. Bacteriol.*, 10.1128/JB.01594-08.
68. Higgins,C.F., Dorman,C.J., Stirling,D.A., Waddell,L., Booth,I.R., May,G. and Bremer,E. (1988) A physiological role for DNA supercoiling in the osmotic regulation of gene expression in S. typhimurium and E. coli. *Cell*, 10.1016/0092-8674(88)90470-9.
69. Bhriain,N.N., Dorman,C.J. and Higgins,C.F. (1989) An overlap between osmotic and anaerobic stress responses: a potential role for DNA supercoiling in the coordinate regulation of gene expression. *Mol. Microbiol.*, 10.1111/j.1365-2958.1989.tb00243.x.
70. Karem,K. and Foster,J.W. (1993) The influence of DNA topology on the environmental regulation of a pH-regulated locus in Salmonella typhimurium. *Mol. Microbiol.*, 10.1111/j.1365-2958.1993.tb00905.x.
71. Fletcher,S.A. and Csonka,L.N. (1995) Fine-structure deletion analysis of the transcriptional silencer of the proU operon of Salmonella typhimurium. *J. Bacteriol.*, 10.1128/jb.177.15.4508-4513.1995.
72. Datsenko,K.A. and Wanner,B.L. (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.*, 10.1073/pnas.120163297.
73. Kolmsee,T. and Hengge,R. (2011) Rare codons play a positive role in the expression of the stationary phase sigma factor RpoS (σS) in Escherichia coli. *RNA Biol.*, 10.4161/rna.8.5.16265.
74. Gibson,D.G., Young,L., Chuang,R.Y., Venter,J.C., Hutchison,C.A. and Smith,H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, **6**, 343–345.
75. Jiang,Y., Chen,B., Duan,C., Sun,B., Yang,J. and Yang,S. (2015) Multigene editing in the Escherichia coli genome via the CRISPR-Cas9 system. *Appl. Environ. Microbiol.*, 10.1128/AEM.04023-14.
76. Rashid,F.Z.M. (2021) Chapter 2: Hi-C in bacteria and archaea.
77. Hagege,H., Klous,P., Braem,C., Splinter,E., Dekker,J., Cathala,G., De Laat,W. and Forne,T. (2007) Quantitative analysis of chromosome conformation capture assays (3c-qpcr). *Nat. Protoc.*, 10.1038/nprot.2007.243.
78. Chaurasiya,K.R. and Dame,R.T. (2018) Single molecule FRET analysis of DNA binding proteins. In *Methods in Molecular Biology*.
79. Laurens,N., Driessen,R.P.C., Heller,I., Vorselen,D., Noom,M.C., Hol,F.J.H., White,M.F., Dame,R.T. and Wuite,G.J.L. (2012) Alba shapes the archaeal genome using a delicate balance of bridging and stiffening the DNA. *Nat. Commun.*, 10.1038/ncomms2330.
80. Crémazy,F.G., Rashid,F.Z.M., Haycocks,J.R., Lamberte,L.E., Grainger,D.C. and Dame,R.T. (2018) Determination of the 3D genome organization of bacteria using Hi-C. In *Methods in Molecular Biology*.
81. Cordonnier,C. and Bernardi,G. (1965) Localization of E. coli endonuclease I. *Biochem. Biophys. Res. Commun.*, 10.1016/0006-291X(65)90434-1.
82. Taylor,R.G., Walker,D.C. and Mclnnes,R.R. (1993) E.coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Res.*, 10.1093/nar/21.7.1677.
83. Hanahan D. (1985) Techniques for transformation of Escherichia coli. *En DNA Cloning a Pract. Approach (ed. Glover, D.M.). Oxford. IRL Press.*

84. Lioy,V.S., Cournac,A., Marbouty,M., Duigou,S., Mozziconacci,J., Espéli,O., Boccard,F. and Koszul,R. (2018) Multiscale Structuring of the E. coli Chromosome by Nucleoid-Associated and Condensin Proteins. *Cell*, 10.1016/j.cell.2017.12.027.
85. Hsieh,T.H.S., Fudenberg,G., Goloborodko,A. and Rando,O.J. (2016) Micro-C XL: Assaying chromosome conformation from the nucleosome to the entire genome. *Nat. Methods*, **13**, 1009–1011.
86. Gavrilov,A.A., Gushchanskaya,E.S., Strelkova,O., Zhironkina,O., Kireev,I.I., Iarovaia,O. V. and Razin,S. V. (2013) Disclosure of a structural milieu for the proximity ligation reveals the elusive nature of an active chromatin hub. *Nucleic Acids Res.*, 10.1093/nar/gkt067.
87. Condensin- and Replication-Mediated Bacterial Chromosome Folding and Origin Condensation Revealed by Hi-C and Super-resolution Imaging. *Mol. Cell*.
88. Kiernan,J.A. (1981) Histological and histochemical methods: theory and practice Oxford, UK.
89. Jamur,M.C. and Oliver,C. (2010) Cell fixatives for immunostaining. *Methods Mol. Biol.*, 10.1007/978-1-59745-324-0_8.
90. Jamur,M.C. and Oliver,C. (2010) Permeabilization of cell membranes. *Methods Mol. Biol.*, 10.1007/978-1-59745-324-0_9.
91. Bustin,S.A., Benes,V., Garson,J.A., Hellemans,J., Huggett,J., Kubista,M., Mueller,R., Nolan,T., Pfaffl,M.W., Shipley,G.L., *et al.* (2009) The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.*, 10.1373/clinchem.2008.112797.
92. Gomes,A.É.I., Stuchi,L.P., Siqueira,N.M.G., Henrique,J.B., Vicentini,R., Ribeiro,M.L., Darrieux,M. and Ferraz,L.F.C. (2018) Selection and validation of reference genes for gene expression studies in Klebsiella pneumoniae using Reverse Transcription Quantitative real-time PCR. *Sci. Rep.*, 10.1038/s41598-018-27420-2.
93. Savli,H., Karadenizli,A., Kolayli,F., Gundes,S., Ozbek,U. and Vahaboglu,H. (2003) Expression stability of six housekeeping genes: A proposal for resistance gene quantification studies of Pseudomonas aeruginosa by real-time quantitative RT-PCR. *J. Med. Microbiol.*, 10.1099/jmm.0.05132-0.
94. Bai,B., Ren,J., Bai,F. and Hao,L. (2020) Selection and validation of reference genes for gene expression studies in Pseudomonas brassicacearum GS20 using real-time quantitative reverse transcription PCR. *PLoS One*, 10.1371/journal.pone.0227927.
95. Galisa,P.S., da Silva,H.A.P., Macedo,A.V.M., Reis,V.M., Vidal,M.S., Baldani,J.I. and Simes-Araújo,J.L. (2012) Identification and validation of reference genes to study the gene expression in Gluconacetobacter diazotrophicus grown in different carbon sources using RT-qPCR. *J. Microbiol. Methods*, 10.1016/j.mimet.2012.07.005.
96. Zhou,K., Zhou,L., Lim,Q., Zou,R., Stephanopoulos,G. and Too,H.P. (2011) Novel reference genes for quantifying transcriptional responses of Escherichia coli to protein overexpression by quantitative PCR. *BMC Mol. Biol.*, 10.1186/1471-2199-12-18.
97. Peng,S., Stephan,R., Hummerjohann,J. and Tasara,T. (2014) Evaluation of three reference genes of Escherichia coli for mRNA expression level normalization in view of salt and organic acid stress exposure in food. *FEMS Microbiol. Lett.*, 10.1111/1574-6968.12447.
98. 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 Dev.*, **26**, 2621–2633.
99. Myers,K.S., Yan,H., Ong,I.M., Chung,D., Liang,K., Tran,F., Keleş,S., Landick,R. and Kiley,P.J. (2013) Genome-scale Analysis of Escherichia coli FNR Reveals Complex Features of Transcription Factor Binding. *PLoS Genet.*, 10.1371/journal.pgen.1003565.
100. Shao,X., Zhang,W., Umar,M.I., Wong,H.Y., Seng,Z., Xie,Y., Zhang,Y., Yang,L., Kwok,C.K. and Deng,X. (2020) RNA G-quadruplex structures mediate gene regulation in bacteria. *MBio*, 10.1128/mBio.02926-19.
101. Huppert,J.L., Bugaut,A., Kumari,S. and Balasubramanian,S. (2008) G-quadruplexes: The beginning and end of UTRs. *Nucleic Acids Res.*, 10.1093/nar/gkn511.
102. Kahramanoglou,C., Seshasayee,A.S.N., 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 Res.*, 10.1093/nar/gkq934.
103. Tatusov,R.L., Mushegian,A.R., Bork,P., Brown,N.P., Hayes,W.S., Borodovsky,M., Rudd,K.E. and Koonin,E. V. (1996) Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli. *Curr. Biol.*, **6**, 279–291.

104. Tomii,K. and Kanehisa,M. (1998) A comparative analysis of ABC transporters in complete microbial genomes. *Genome Res.*, **8**, 1048–1059.
105. Higgins,C.F. (1992) ABC Transporters: From microorganisms to man. *Annu. Rev. Cell Biol.*, **8**, 67–113.
106. Tam,R. and Saier,M.H. (1993) Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol. Rev.*, **57**, 320–346.
107. Saurin,W. and Dassa,E. (1994) Sequence relationships between integral inner membrane proteins of binding protein-dependent transport systems: Evolution by recurrent gene duplications. *Protein Sci.*, **3**, 325–344.
108. Mellies,J., Brems,R. and Villarejo,M. (1994) The Escherichia coli proU promoter element and its contribution to osmotically signaled transcription activation. *J. Bacteriol.*, **176**, 3638–3645.
109. Wade,J.T. and Grainger,D.C. (2018) Waking the neighbours: disruption of H-NS repression by overlapping transcription. *Mol. Microbiol.*, 10.1111/mmi.13939.
110. Rangarajan,A.A. and Schnetz,K. (2018) Interference of transcription across H-NS binding sites and repression by H-NS. *Mol. Microbiol.*, 10.1111/mmi.13926.
111. 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 Res.*, 10.1093/nar/gky265.
112. Adhya,S. and Gottesman,M. (1978) Control of transcription termination. *Annu. Rev. Biochem.*, **47**, 967–996.
113. Adhya,S., Gottesman,M. and De Crombrugghe,B. (1974) Termination and antitermination in transcription: control of gene expression. *Basic Life Sci.*, **3**, 213–221.
114. Robinett,C.C., Straight,A., Li,G., Willhelm,C., Sudlow,G., Murray,A. and Belmont,A.S. (1996) In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J. Cell Biol.*, 10.1083/jcb.135.6.1685.
115. Lau,I.F., Filipe,S.R., Søballe,B., Økstad,O.A., Barre,F.X. and Sherratt,D.J. (2003) Spatial and temporal organization of replicating Escherichia coli chromosomes. *Mol. Microbiol.*, **49**, 731–743.
116. van Berkum,N.L., Lieberman-Aiden,E., Williams,L., Imakaev,M., Gnirke,A., Mirny,L.A., Dekker,J. and Lander,E.S. (2010) Hi-C: a method to study the three-dimensional architecture of genomes. *J. Vis. Exp.*, 10.3791/1869.
117. Qin,L., Erkelens,A.M., Ben Bdira,F. and Dame,R.T. (2019) The architects of bacterial DNA bridges: A structurally and functionally conserved family of proteins. *Open Biol.*, 10.1098/rsob.190223.
118. Vlahoviček,K., Kaján,L. and Pongor,S. (2003) DNA analysis servers: Plot.it, bend.it, model.it and IS. *Nucleic Acids Res.*, 10.1093/nar/gkg559.
119. Goodsell,D.S. and Dickerson,R.E. (1994) Bending and curvature calculations in B-DNA. *Nucleic Acids Res.*, 10.1093/nar/22.24.5497.
120. Kornyshev,A.A. and Leikin,S. (2001) Sequence recognition in the pairing of DNA duplexes. *Phys. Rev. Lett.*, 10.1103/PhysRevLett.86.3666.
121. Zuccheri,G., Scipioni,A., Cavaliere,V., Gargiulo,G., De Santis,P. and Samori,B. (2001) Mapping the intrinsic curvature and flexibility along the DNA chain. *Proc. Natl. Acad. Sci. U. S. A.*, 10.1073/pnas.051631198.
122. Scipioni,A., Anselmi,C., Zuccheri,G., Samori,B. and De Santis,P. (2002) Sequence-dependent DNA curvature and flexibility from scanning force microscopy images. *Biophys. J.*, 10.1016/S0006-3495(02)75254-5.
123. Colangeli,R., Helb,D., Vilchèze,C., Hazbón,M.H., Lee,C.G., Safi,H., Sayers,B., Sardone,I., Jones,M.B., Fleischmann,R.D., *et al.* (2007) Transcriptional regulation of multi-drug tolerance and antibiotic-induced responses by the histone-like protein Lsr2 in M. tuberculosis. *PLoS Pathog.*, 10.1371/journal.ppat.0030087.
124. Castang,S., McManus,H.R., Turner,K.H. and Dove,S.L. (2008) H-NS family members function coordinately in an opportunistic pathogen. *Proc. Natl. Acad. Sci. U. S. A.*, 10.1073/pnas.0808215105.
125. Castang,S. and Dove,S.L. (2010) High-order oligomerization is required for the function of the H-NS family member MvaT in Pseudomonas aeruginosa. *Mol. Microbiol.*, 10.1111/j.1365-2958.2010.07378.x.
126. Gordon,B.R.G., Li,Y., Wang,L., Sintsova,A., Van Bakel,H., Tian,S., Navarre,W.W., Xia,B. and

Liu,J. (2010) Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in Mycobacterium tuberculosis. *Proc. Natl. Acad. Sci. U. S. A.*, 10.1073/pnas.0913551107.

Supplementary information

Supplementary methods

Genomic DNA preparation

An isolated *Escherichia coli* colony was cultured to stationary phase in LB medium at 37ºC with shaking at 200 rpm. Cells in 1.5 mL of the culture were collected by centrifugation at 10000xg at room temperature and resuspended in 400 µL of TES buffer (50 mM Tris-HCl, 10 mM NaCl, 10 mM EDTA, pH 7.5). The cell suspension was incubated with 1.0% Sarkosyl, 100 µg/mL RNase A (Qiagen®), and 100 µg/mL Proteinase K (Qiagen®) at 65ºC until the solution cleared. 400 µL of 4M NH_4OAc was added to the lysate, and the solution was extracted twice with 25:24:1 Phenol:Chloroform:Isoamyl alcohol (Sigma-Aldrich®) and once with chloroform (Sigma-Aldrich®). Genomic DNA was precipitated with an equal volume of isopropanol (Sigma-Aldrich®) for 10 minutes at room temperature, and pelleted by centrifuged for 20 minutes at 25000xg at 4ºC. The pellet was dissolved in 400 µL of 0.1 M NaOAc, pH 6.0, and re-precipitated with 800 µL of cold 100% ethanol (Sigma-Aldrich®) for 15 minutes at room temperature. The precipitated DNA was collected by centrifugation at 25000xg for 15 minutes at 4ºC. The pellet was washed with cold 70% ethanol (Sigma-Aldrich®) and air dried. Genomic DNA was dissolved in 1X TE (Sigma-Aldrich®) (10 mM Tris, 0.1 mM EDTA, pH 8.0).

H-NS purification

Escherichia coli BL21 Δ*hns* pLysE pRD18 was cultured in LB medium (1.0% bactotryptone (BD), 0.5% yeast extract (Alfa Aesar), 170 mM NaCl, pH 7.5) at 37ºC to an OD_{600} of 0.4. Expression of H-NS from pRD18 was induced with a final concentration of 120 µg/mL IPTG for 3 hours. The cells were collected by centrifugation at 7000 xg for 15 minutes at 25ºC. Cell pellets were resuspended in 20 mL low-salt H-NS buffer (130 mM NaCl, 20 mM Tris, 10% glycerol, 8 mM β-mercaptoethanol, and 3 mM benzamidine, pH 7.2) containing 100 mM NH_4Cl with 1 µg/mL DNase, 100 µg/mL lysozyme, and 1 mM PMSF. The cells were lysed by sonication and the soluble fraction was loaded onto a P11 column. H-NS was eluted in low-salt H-NS buffer with NH_4Cl using a gradient of 100 mM to 1.0 M NH_4Cl in 100 mL at a rate of 1 mL/min. Fractions with high concentrations of H-NS were pooled and dialysed overnight at 4ºC into low-salt H-NS buffer. The dialysed sample was loaded onto a 1 mL HisTrap™ HP column (Amersham Biosciences) and eluted in H-NS buffer (20 mM Tris, 10% glycerol, 8 mM β-mercaptoethanol, and 3 mM benzamidine, pH 7.2) using a gradient of 130 mM to

1M NaCl in 20 mL at a rate of 1 mL/min. The fractions with high H-NS concentrations were pooled and dialysed at 4ºC overnight into low-salt H-NS buffer. The H-NS was loaded onto a 1 mL Resource-Q column and eluted into a highly concentrated and pure peak with a block elution in H-NS buffer from 200 mM to 500 mM NaCl. The purified H-NS was dialysed at 4ºC overnight into H-NS buffer with 300 mM KCl. H-NS was separated into 50.0 µL aliquots, flash-frozen in liquid nitrogen, and stored at -80ºC until use.

Electrophoretic mobility shift assay

Purified H-NS was serially diluted in low-salt binding buffer (130 mM NaCl, 20 mM Tris, 10% glycerol, 8 mM β-mercaptoethanol, and 3 mM benzamidine, pH 7.2). The dilutions were mixed with an equivalent volume of 400 nM DRE^{mut} or DRE^{wt} (Supplementary file 3.1) in nuclease-free water (Ambion) and incubated for 20 minutes at 25 ºC followed by 10 minutes at 4 ºC. The samples were resolved on a polyacrylamide gel (Mini-Protean® TGX™ precast gels, 4-15%, Bio-Rad) at 30 V at 4 ºC. Experiments were performed in triplicate. The polyacrylamide gels were stained for 45 minutes in a solution of 10X GelRed (Biotium) and thereafter imaged with GelDoc™ XR+ (Bio-Rad) using Bio-Rad's ImageLab software.

Supplementary figures

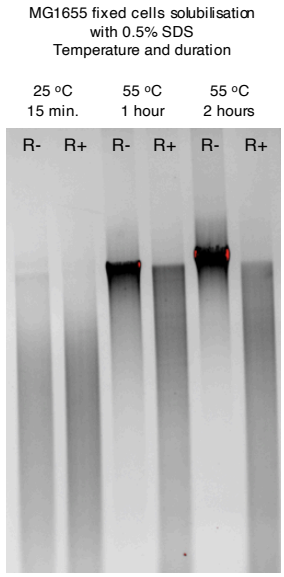


Figure S3.1: Cell solubilisation conditions that reliably decrease degradation of MG1655 chromatin in low [EDTA] buffers also promote de-crosslinking of formaldehyde-mediated crosslinks. Raising the temperature and increasing the duration of 0.5% SDS treatment during lysis and solubilisation of fixed cells increases the stability of the extracted chromatin in low [EDTA] buffer.

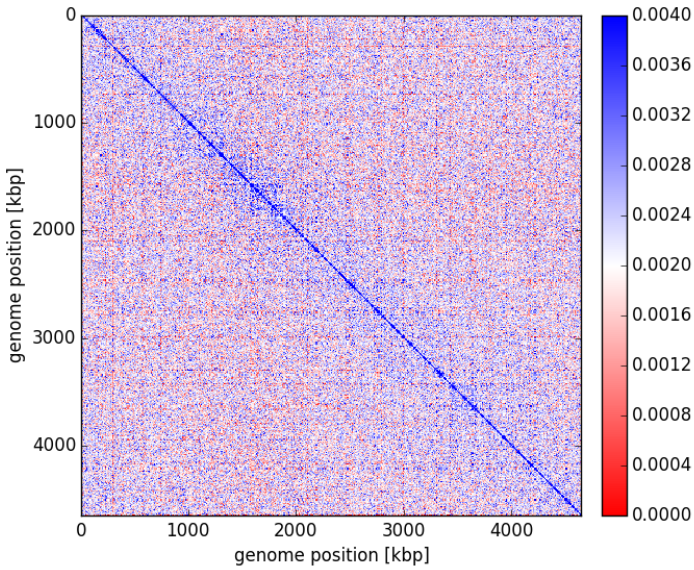


Figure S3.2: Fixation of *E. coli* cells with higher concentrations of formaldehyde does not significantly improve the signal-to-noise ratio in chromosome contact maps. Chromosome contact

maps generated from 3C-based libraries prepared with *E. coli* cells fixed using 7% formaldehyde do not show a significant improvement in the signal-to-noise ratio compared to maps generated from *E. coli* cells fixed with 3% formaldehyde (Figure 3.2A). **Organism:** *Escherichia coli* MG1655 Δ*endA* (NT331); **3C-based study:** 3C-Seq; **Growth conditions:** LB medium, 37 ºC, exponential phase; **Fixation conditions:** 7% formaldehyde, 1 hour; **Restriction enzyme:** HpaII (NEB); **Fractionation:** Not performed.

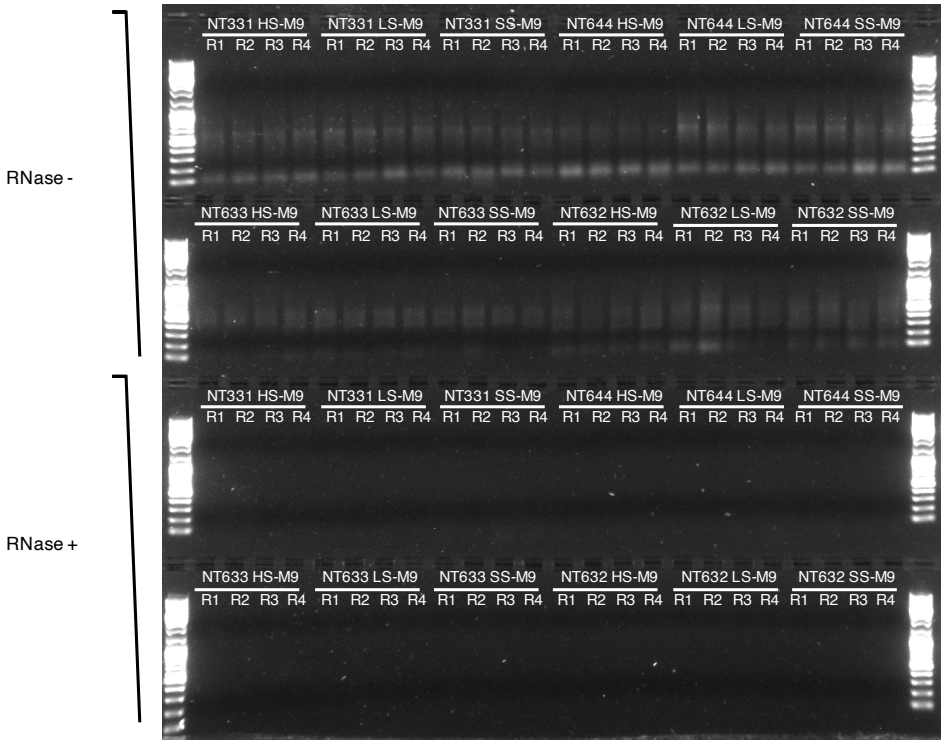


Figure S3.3: DNA contamination assessment of RNA preparations. ~100 ng of RNA with (RNase+) or without (RNase-) RNase treatment were visualised on a 1.2% agarose gel pre-stained with 1X GelRed (Sigma-Aldrich). The absence of a nucleic acid signal in the RNase+ wells indicates undetectable DNA contamination.

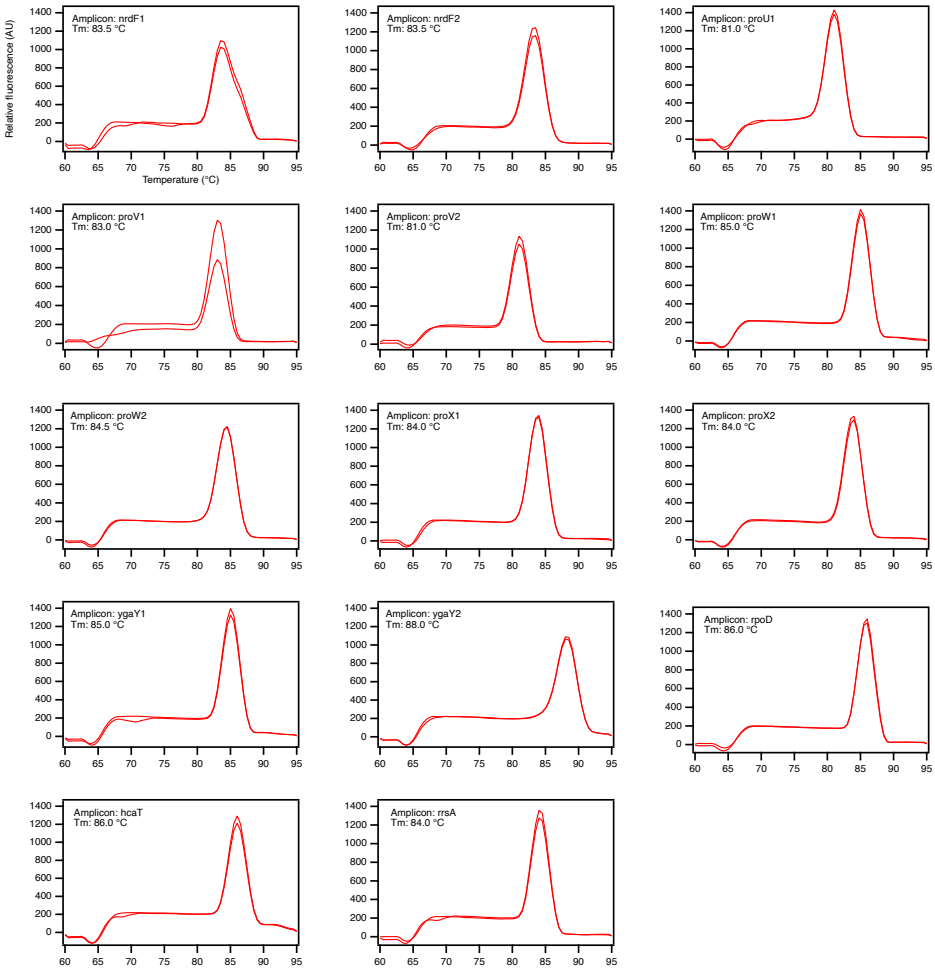


Figure S3.4: RT-qPCR amplicon melting curves. The specificity of amplification in RT-qPCR reactions was gauged from the melting curve of the fragment amplified in each well. The melting curve and the melting temperature (Tm) of the amplicons reported on in this chapter are shown here. The sequences of the fragments used for this experiment were verified with Sanger sequencing (Supplementary file 3.2).

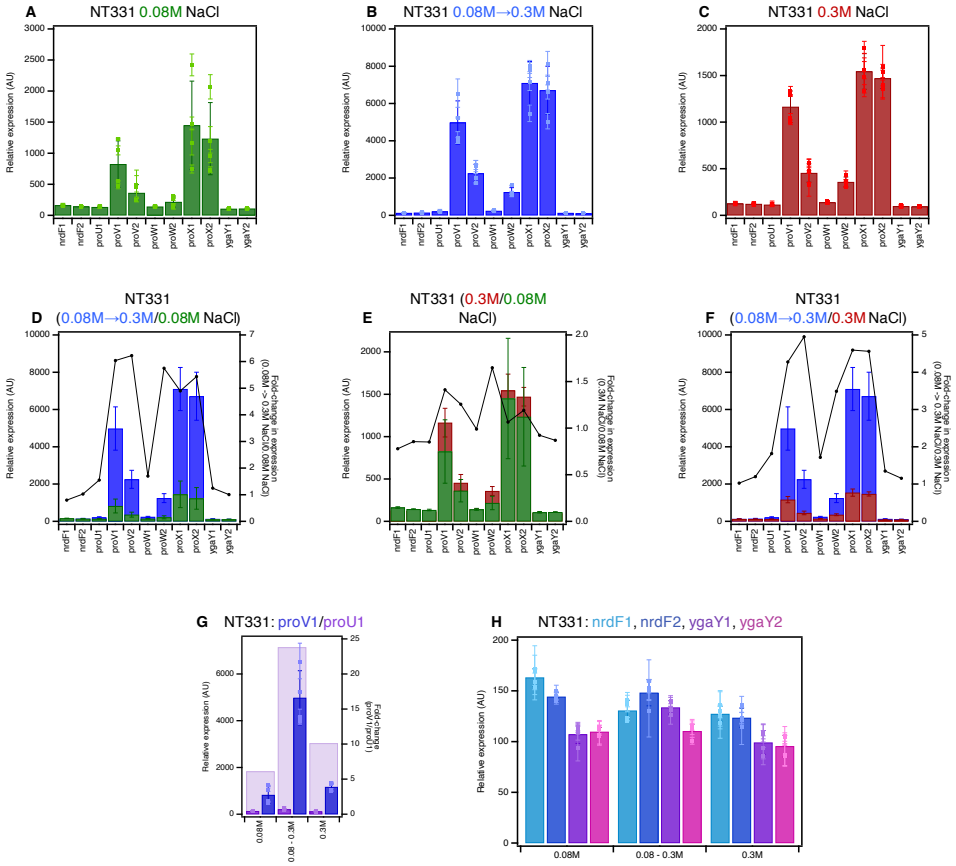


Figure S3.5: The transcriptional profile of *proVWX* and its flanking regions in NT331 during **(A)** exponential growth in M9 medium with 0.08 M NaCl, **(B)** hyperosmotic shock in M9 medium from 0.08 M to 0.3 M NaCl, and **(C)** exponential growth in M9 medium with 0.3 M NaCl. **The fold change in expression level of *proVWX* and its flanking regions** between **(D)** a hyperosmotic shock and exponential growth at 0.08 M NaCl, **(E)** exponential growth at 0.3 M NaCl and 0.08 M NaCl, and **(F)** a hyperosmotic shock and exponential growth at 0.3 M NaCl. **(G) The fold difference in expression at the *proV*1 amplicon compared to the *proU*1 amplicon** during exponential growth at 0.08 M NaCl, following a hyperosmotic shock, and during exponential growth at 0.3 M NaCl. **(H) The relative expression at amplicons flanking *proVWX*** during exponential growth at 0.08 M NaCl, following a hyperosmotic shock, and during exponential growth at 0.3 M NaCl. **Internal control: *hcaT***.

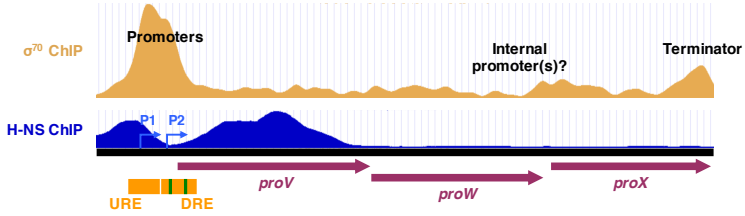
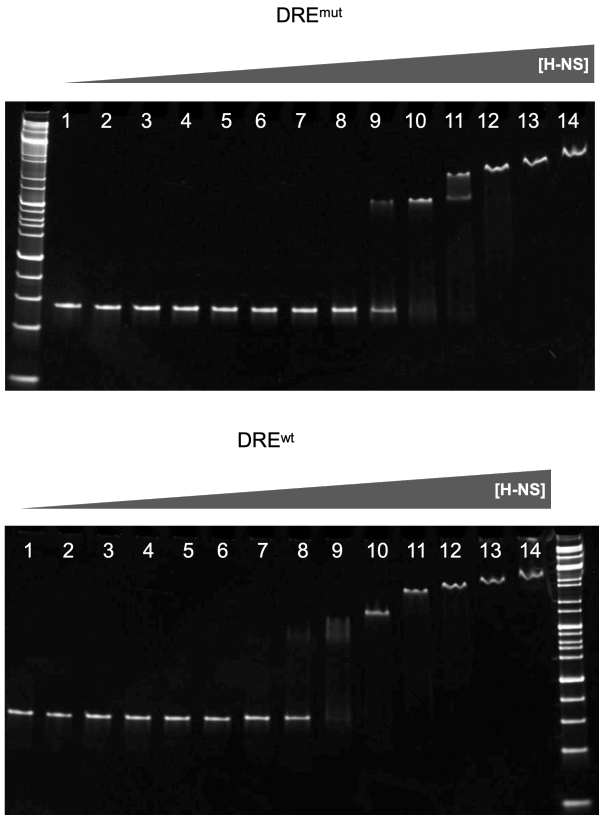


Figure S3.6: The chromatin immunoprecipitation (ChIP) profiles of RNA Polymerase σ^{70} (yellow) (99) **and H-NS (blue)** (102) **at *proVWX*.** The purple arrows represent the *proV*, *proW*, and *proX* ORFs. The orange bars mark the *proVWX* upstream and downstream regulatory elements (URE and DRE). The green bars within the DRE designate high-affinity H-NS binding sites. The P1 and P2 *proVWX* promoters and their transcriptional directionalities are shown with blue right-angle arrows.



Lane	[H-NS] (nM)
1	31.7
2	52.8
3	88.0
4	146.6
5	244.4
6	407.3
7	678.8
8	1131
9	1885
10	3142
11	5238
12	8730
13	14550
14	24250

Figure S3.7: The mutated DRE (top) has a lower affinity for H-NS than the wild-type DRE (bottom)

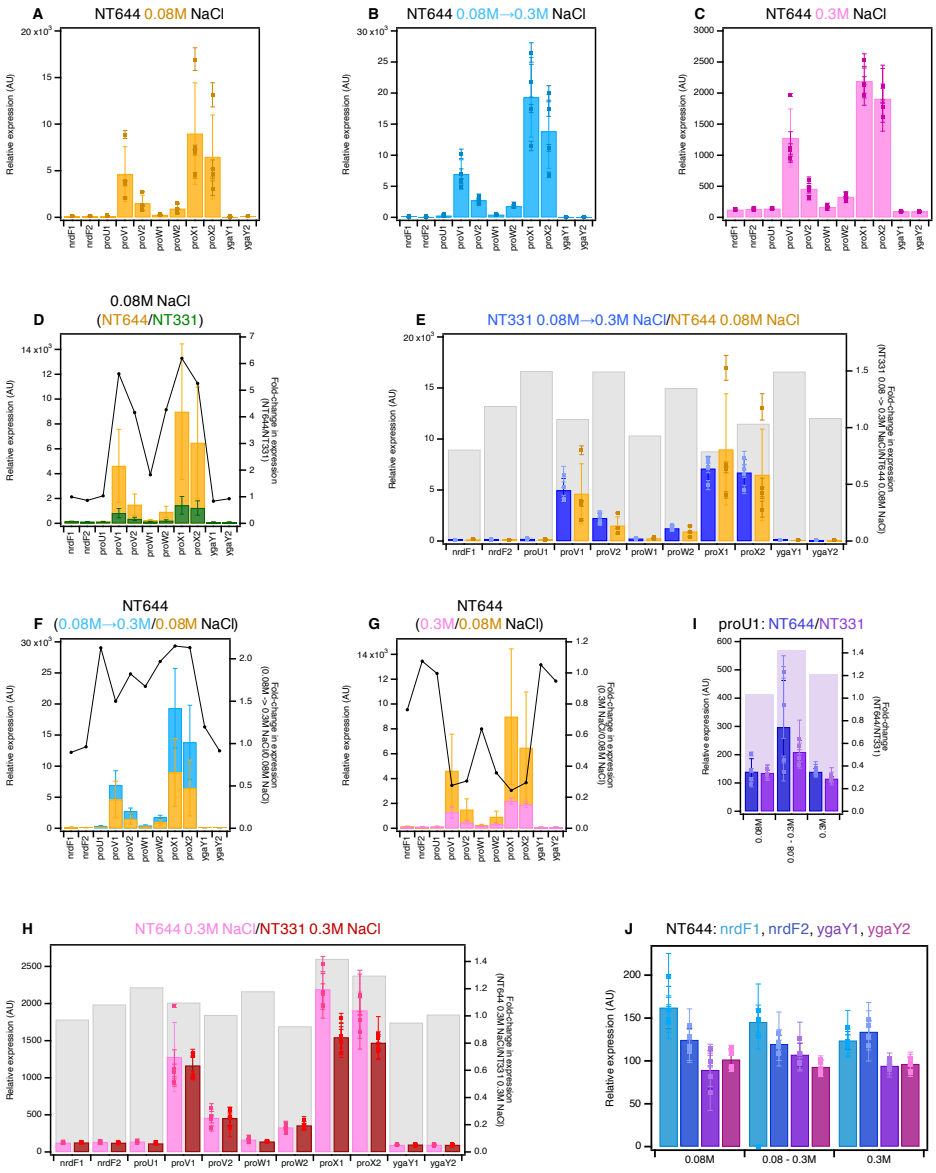


Figure S3.8: The transcriptional profile of *proVWX* and its flanking regions in NT644 during **(A)** exponential growth in M9 medium with 0.08 M NaCl, **(B)** hyperosmotic shock in M9 medium from 0.08 M to 0.3 M NaCl, and **(C)** exponential growth in M9 medium with 0.3 M NaCl. **The fold change in expression level of *proVWX* and its flanking regions** between **(D)** NT644 and NT331 during exponential growth at 0.08 M NaCl, **(E)** NT331 following a hyperosmotic shock and NT644 growing exponentially at 0.08 M NaCl, **(F)** a hyperosmotic shock and exponential growth at 0.08 M NaCl for NT644, **(G)** exponential growth at 0.3 M NaCl and 0.08 M NaCl for NT644, and **(H)** NT644 and NT331 growing exponentially at 0.3 M NaCl. **(I) The fold change in expression at the *proU*1 amplicon** between NT644 and NT331 during exponential growth at 0.08 M NaCl, following a hyperosmotic shock, and during exponential growth at 0.3 M NaCl. **(J) The relative expression level in NT644 at amplicons flanking *proVWX*** during exponential growth at 0.08 M NaCl, following a hyperosmotic shock, and during exponential growth at 0.3 M NaCl. **Internal control: *hcaT***

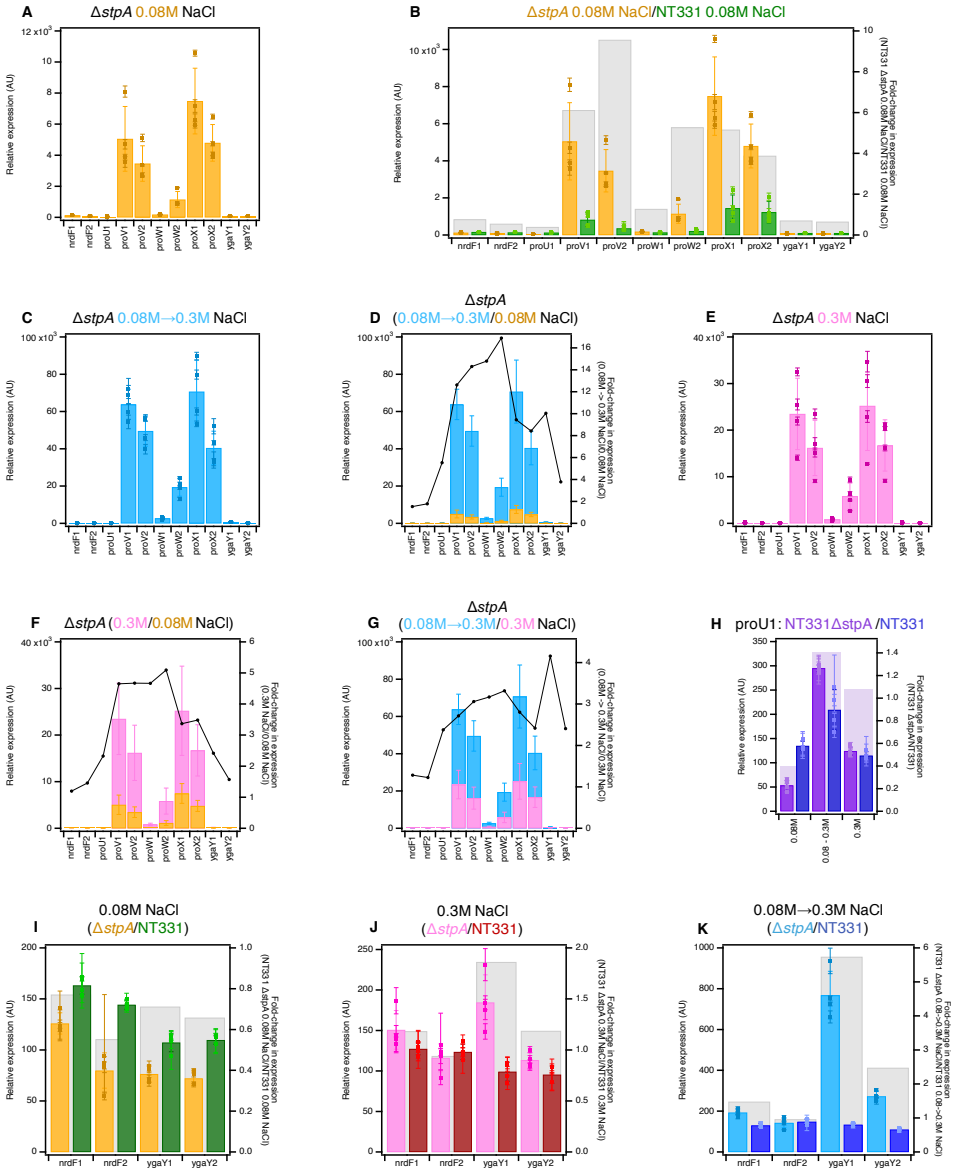


Figure S3.9: The transcriptional profile of the *proVWX* operon and its flanking regions in NT331 Δ*stpA* (A) during exponential growth at 0.08 M NaCl and **(B)** the fold-change in the expression levels of the amplicons compared to NT331, **(C)** upon a hyperosmotic shock from 0.08 M to 0.3 M NaCl, and **(D)** the fold-change in expression levels of the amplicons in comparison to exponential growth at 0.08 M NaCl, and **(E)** during exponential growth at 0.3 M NaCl, and the fold-change in expression levels of the amplicons with respect to **(F)** exponential growth at 0.08 M NaCl, and **(G)** a hyperosmotic shock. **(H) The fold-difference in expression level of amplicon *proU*1** between NT331 ΔstpA and NT331. **The fold-change in expression levels of the *nrdF* and *ygaY* amplicons between NT331 Δ*stpA* and NT331 (I)** during exponential growth at 0.08 M NaCl, **(J)** exponential growth at 0.3 M NaCl, and **(K)** following a hyper-osmotic shock. **Internal control: *hcaT***.

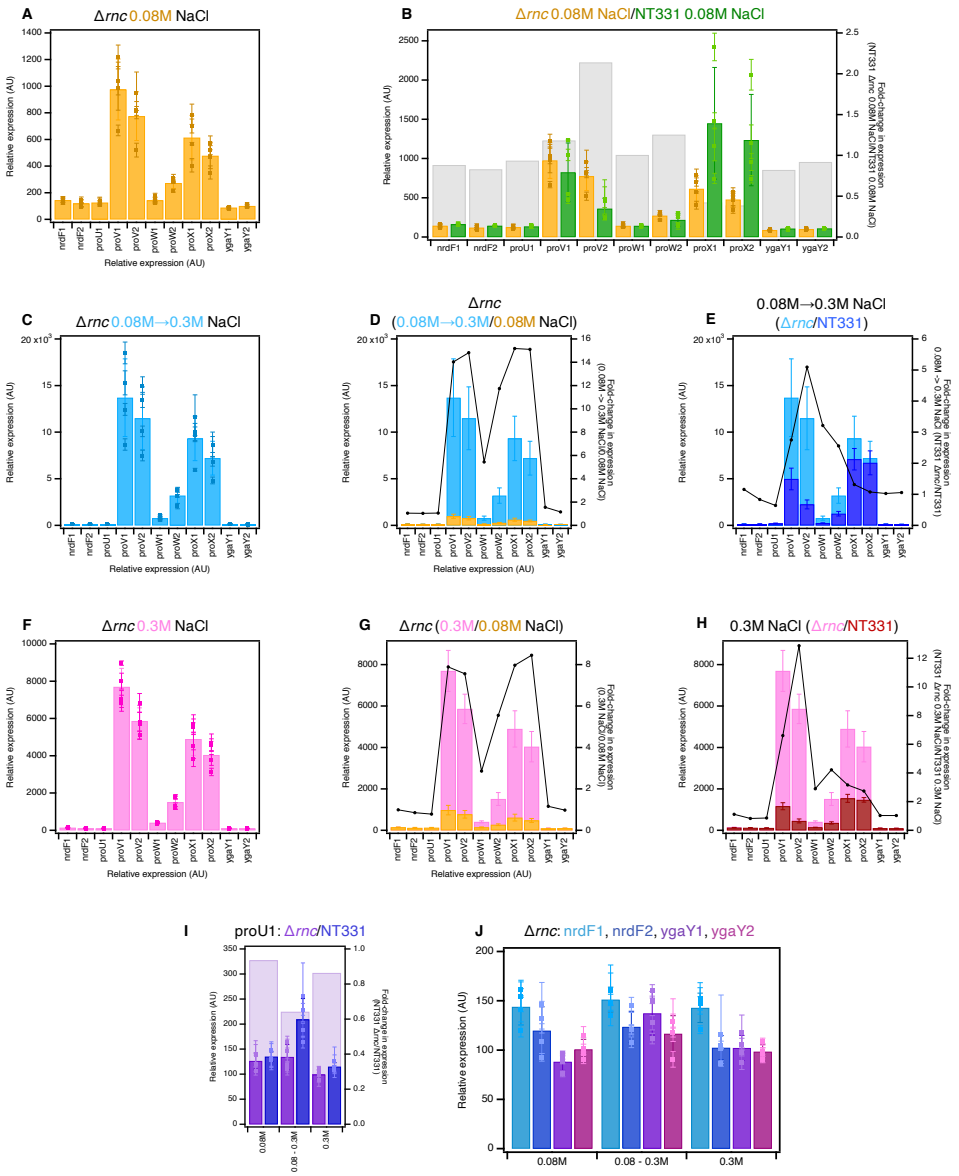


Figure S3.10: The transcriptional profile of *proVWX* and its flanking regions in NT331 Δ*rnc* during **(A)** exponential growth at 0.08 M NaCl, and **(B)** the fold-change in expression of the amplicons compared to NT331 growing exponentially at 0.08 M NaCl, upon **(C)** a hyperosmotic shock from 0.08 M NaCl to 0.3 M NaCl, and the fold-change in expression of the amplicons in comparison to **(D)** NT331 Δ*rnc* growing exponentially at 0.08 M NaCl, and **(E)** NT331 following a hyperosmotic shock. **The transcription profile of *proVWX* and its flanking regions in NT331 Δ*rnc*** during **(F)** exponential growth at 0.3 M NaCl, and the fold-change in expression of the amplicons relative to **(G)** NT331 Δ*rnc* growing exponentially at 0.08 M NaCl, and **(H)** NT331 growing exponentially at 0.3 M NaCl. **(I) The fold-difference in the relative expression of the *proU*1** amplicon between NT331 Δ*rnc* and NT331. **(J) The relative expression level in NT331 Δ*rnc* at amplicons flanking *proVWX*** during exponential

growth at 0.08 M NaCl, following a hyperosmotic shock, and during exponential growth at 0.3 M NaCl. **Internal control: *hcaT***.

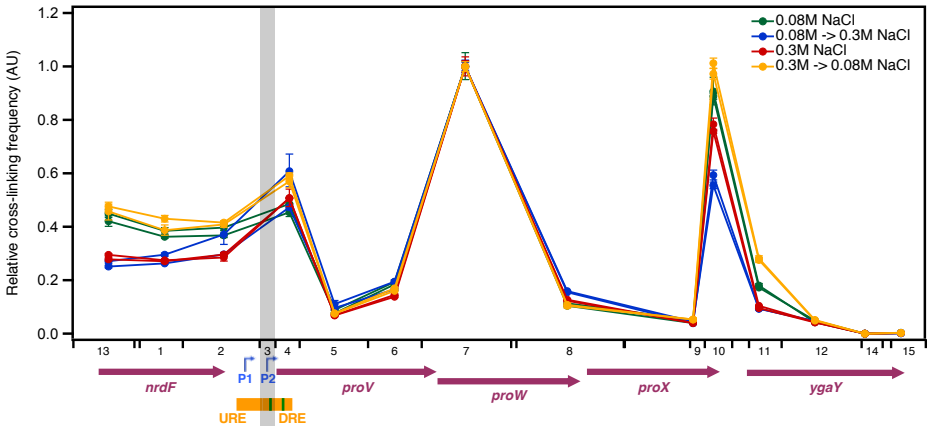


Figure S3.11: A 3C-qPCR study of the relative cross-linking frequency of proU3_NlaIII (highlighted in grey) with chromatin segments within and flanking *proVWX* (marked along the horizontal axis) during exponential growth in LB medium with 0.08 M NaCl (green), following a hyperosmotic shock from 0.08 M to 0.3 M NaCl (blue), during exponential growth at 0.3 M NaCl (red), and following a hypoosmotic shock from 0.3 M to 0.08 M NaCl (yellow). Each curve represents the 3C profile of an independent biological replicate. The error bars represent the standard deviation of three quantitation values from the qPCR run. *proV*, *proW*, *proX*, *nrdF*, and *ygaY* (purple arrows) show individual genes. P1 and P2 (right-angled blue arrows) mark proVWX promoters. The upstream and downstream regulatory elements of *proVWX* (URE and DRE) are shown with orange bars. The green bars in the DRE designate high-affinity H-NS binding sites. **Internal cross-linking control:** proU3_NlaIII-proU7_NlaIII.

Supplementary tables

Table S3.1: List of primers for smFRET studies.

Primer name	Sequence
HNS_bridge_fwd	BtnTEG-ATACATCCy3dTGTAACTTGAACGGCGTAAAAGAGG
HNS_bridge_rev	Cy5-TGAATTCTTCGATATGAGTCCTTAGTAACTTACCC
URE-DRE_rev	BtnTEG-TTTTTCy3dTCCAGACTGGCGTCTTTTACG
URE-DRE_fwd	Cy5-TAATTCCCGCCAAATAGCTT
DRE_fwd	BtnTEG-GCCATCAGGGGTTGCCTCAGATTCTCAGTATGTTAGGGTAGAAAAAAGCy3dTGACTATTTCC
ProW1_rev	Cy5-CTTTCGAAATGATGCCGACATACTGTTGGTCC
ProW2_rev	Cy5-AGTCGTCGGTGTACCCCAGGCGTCTGCGGATTG
ProW3_rev	Cy5-AGAATATTAAAATGCTCGACGTTTGGCGCAGGCGTA
ProW4_rev	Cy5-GATATAATCAACCGGAACGCGCACGCCCTGGAAGAC
ProW5_rev	Cy5-AAATCTGCCAGGCGATGAGAGCGAAAACGATAATCG
ProW6_rev	Cy5-CAGAGTCACCATTGCCTGCGACCAGGCACCGATTGC
ProUterm_rev	Cy5-GGGATGAAAGATAGATTGTTGAAATGTGCGATGTGG
IHF_60bp_fwd	Cy5-TTGGCATTATAAAAAAGCATTGCTTATCAATTTGTTGCAACGAACAGGTCACGTAATCGT
IHF_60bp_rev	Cy3-ACGATTACGTGACCTGTTCGTTGCAACAAATTGATAAGCAATGCTTTTTTATAATGCCAATTT-BtnTEG
IHF_35bp_A_fwd	Cy5-GATCTCAGGACGTAATACCAGTCCAGTACGACTTT
IHF_35bp_A_rev	Cy3-AAAGTCGTACTGGACTGGTATTACGTCCTGAGATCTTT-BtnTEG
IHF_35bp_B_fwd	Cy5-GGCCAAAAAAGCATTGCTTATCAATTTGTTGCACC
IHF_35bp_B_rev	Cy3-GGTGCAACAAATTGATAAGCAATGCTTTTTTGGCCTTT-BtnTEG
IHF_35bp_C_fwd	Cy5-TCAGGGGTTGCCTCAGATTCTCAGTATGTTAGGGT
IHF_35bp_C_rev	Cy3-ACCCTAACATACTGAGAATCTGAGGCAACCCCTGATTT-BtnTEG

Table S3.2: Relative expression values of amplicons within, and flanking the *proVWX* operon in NT331. Internal control: *rpoD*

Amplicon	NT331 0.08 M NaCl	NT331 Hyperosmotic shock	NT331 0.3 M NaCl
nrdF1	81.76±8.58	80.31±6.67	85.26±11.58
nrdF2	72.65±11.31	90.84±2.98	82.50±9.49
proU1	67.60±7.90	128.77±27.04	76.72±7.43
proV1	402.16±150.56	3047.84±637.44	773.91±89.45
proV2	177.21±50.60	1376.57±238.83	304.88±73.74
proW1	71.52±11.27	148.65±13.91	94.40±12.63
proW2	106.71±31.29	766.58±137.55	237.86±27.42
proX1	699.64±279.30	4365.00±726.11	1025.76±84.89
proX2	598.77±226.02	4125.85±795.50	979.43±65.30
ygaY1	53.92±8.70	82.41±8.39	65.80±4.53
ygaY2	54.98±5.93	67.00±7.56	63.53±6.03

Table S3.3: Relative expression values of amplicons within, and flanking the *proVWX* operon in NT331. Internal control: *hcaT*

Amplicon	NT331 0.08 M NaCl	NT331 Hyperosmotic shock	NT331 0.3 M NaCl
nrdF1	163.32±8.61	130.70±10.10	127.36±7.29
nrdF2	144.37±4.35	148.28±12.99	123.56±9.44
proU1	135.01±9.47	209.23±41.67	115.02±7.97
proV1	825.06±374.20	4982.62±1159.38	1165.23±168.78

proV2	362.27±132.65	2255.64±482.54	455.49±98.55
proW1	142.41±11.69	242.24±25.71	141.03±7.61
proW2	217.29±79.75	1250.22±241.01	358.28±53.15
proX1	1449.73±710.92	7101.04±1153.99	1545.50±192.71
proX2	1234.74±581.37	6713.35±1288.12	1471.58±109.99
ygaY1	107.28±9.05	133.79±6.59	99.17±11.73
ygaY2	109.74±4.21	110.35±6.71	95.51±10.66

Table S3.4: Relative expression values of amplicons within, and flanking the *proVWX* operon in NT644. Internal control: *rpoD*

Amplicon	NT644 0.08 M NaCl	NT644 Hyperosmotic shock	NT644 0.3 M NaCl
nrdF1	88.27±5.12	91.17±17.10	72.81±4.94
nrdF2	68.34±8.97	71.09±5.69	78.61±5.15
proU1	74.79±15.20	169.18±76.33	81.69±3.77
proV1	2428.61±1177.42	4032.33±868.78	734.54±175.17
proV2	796.10±341.97	1618.03±183.18	264.97±35.89
proW1	141.29±21.40	254.71±41.71	97.67±14.01
proW2	496.96±194.66	1081.64±127.27	193.62±19.97
proX1	4717.25±2145.92	11245.74±2835.34	1287.77±115.42
proX2	3378.13±1791.64	7965.02±2759.61	1128.00±210.66
ygaY1	48.54±9.77	63.73±8.35	55.60±5.71
ygaY2	55.56±4.41	55.20±3.46	56.70±3.78

Table S3.5: Relative expression values of amplicons within, and flanking the *proVWX* operon in NT644. Internal control: *hcaT*

Amplicon	NT644 0.08 M NaCl	NT644 Hyperosmotic shock	NT644 0.3 M NaCl
nrdF1	162.28±24.75	145.43±15.62	123.93±10.31
nrdF2	124.49±13.55	119.73±10.06	133.98±13.44
proU1	139.59±46.30	297.39±165.21	139.22±12.29
proV1	4632.45±2942.73	6948.43±2328.41	1279.27±465.74
proV2	1508.89±862.37	2748.93±522.37	458.92±116.22
proW1	260.22±58.78	436.10±120.87	166.54±27.97
proW2	927.22±448.16	1826.21±278.49	330.78±47.92
proX1	8984.84±5455.14	19333.58±6401.31	2193.00±238.20
proX2	6489.04±4488.92	13841.45±5963.81	1907.92±247.87
ygaY1	89.55±22.12	107.22±13.53	94.29±5.71
ygaY2	101.75±12.28	93.06±7.94	96.49±7.63

Table S3.6: Relative expression values of amplicons within, and flanking the *proVWX* operon in NT331 Δ*stpA* (NT633). Internal control: *rpoD*

Amplicon	NT633 0.08 M NaCl	NT633 Hyperosmotic shock	NT633 0.3 M NaCl
nrdF1	46.25±2.53	79.10±7.59	54.52±11.33
nrdF2	29.79±8.16	58.43±11.39	43.00±13.54
proU1	19.98±5.58	120.60±5.80	45.17±8.90
proV1	1808.97±533.26	26048.07±3254.38	8119.52±931.61
proV2	1248.25±267.61	20226.51±3164.81	5557.57±880.61
proW1	67.33±7.67	1100.01±205.74	296.37±28.02
proW2	410.52±136.89	7934.30±1992.17	1963.72±512.42
proX1	2708.60±468.16	28803.66±6405.98	8606.95±1628.96
proX2	1739.62±233.27	16464.36±3391.55	5773.60±938.28
ygaY1	28.40±5.41	313.55±46.54	65.85±7.87

ygaY2	26.85±5.16	111.52±9.33	41.30±8.68

Table S3.7: Relative expression values of amplicons within, and flanking the *proVWX* operon in NT331 Δ*stpA* (NT633). Internal control: *hcaT*

Amplicon	NT633 0.08 M NaCl	NT633 Hyperosmotic shock	NT633 0.3 M NaCl
nrdF1	125.97±10.43	193.56±14.99	150.77±23.88
nrdF2	79.80±17.35	142.70±24.42	116.26±17.29
proU1	53.39±10.33	295.26±6.41	124.22±8.26
proV1	5052.77±2083.77	63812.32±8172.40	23504.72±7658.21
proV2	3465.18±1140.43	49570.11±8135.08	16194.99±5930.20
proW1	182.25±3.84	2697.16±538.20	850.69±237.53
proW2	1147.82±517.95	19404.52±4834.93	5847.17±2784.05
proX1	7488.12±2115.74	70727.28±16904.09	25217.37±9585.28
proX2	4800.36±1181.15	40439.05±9029.73	16724.05±5495.00
ygaY1	76.42±7.41	768.07±117.33	184.66±34.12
ygaY2	72.25±6.94	273.21±23.27	113.45±9.30

Table S3.8: Relative expression values of amplicons within, and flanking the *proVWX* operon in NT331 Δ*rnc* (NT632). Internal control: *rpoD*

Amplicon	NT632 0.08 M NaCl	NT632 Hyperosmotic shock	NT632 0.3 M NaCl
nrdF1	83.09±6.94	90.11±13.79	82.23±7.34
nrdF2	69.94±18.13	73.93±16.09	59.03±7.40
proU1	73.48±12.17	80.54±21.69	57.72±10.85
proV1	559.26±96.36	8023.76±2095.41	4434.21±530.35
proV2	444.98±85.59	6715.55±1571.80	3394.43±544.03
proW1	83.09±14.12	453.45±83.01	239.58±53.78
proW2	157.33±19.22	1876.76±403.74	872.33±175.58
proX1	356.34±107.62	5511.92±1414.30	2818.13±460.38
proX2	277.14±63.50	4227.19±823.92	2319.72±359.24
ygaY1	51.17±6.31	80.78±7.12	59.70±13.03
ygaY2	58.37±4.84	68.84±9.64	57.59±11.98

Table S3.9: Relative expression values of amplicons within, and flanking the *proVWX* operon in NT331 Δ*rnc* (NT632). Internal control: *hcaT*

Amplicon	NT632 0.08 M NaCl	NT632 Hyperosmotic shock	NT632 0.3 M NaCl
nrdF1	143.94±18.40	151.32±12.66	142.88±14.94
nrdF2	119.79±24.13	123.63±15.83	102.38±12.24
proU1	126.33±15.92	134.26±24.93	99.27±9.15
proV1	975.93±228.81	13700.86±4167.98	7699.76±991.49
proV2	775.22±183.41	11494.71±3369.42	5862.75±710.33
proW1	143.06±22.42	779.19±221.60	409.41±48.64
proW2	272.42±40.24	3200.35±827.86	1512.10±312.42
proX1	614.23±166.78	9329.09±2384.86	4897.48±877.88
proX2	477.58±95.71	7217.89±1818.43	4037.97±734.01
ygaY1	88.16±8.82	137.41±22.07	102.15±12.71
ygaY2	100.83±10.11	116.59±18.27	98.49±7.54

Table S3.10: Relative cross-linking frequency of proU3_NlaIII with fragments within, and flanking the *proVWX* operon in NT331 during exponential growth at 0.08 M NaCl. Cross-linking control: proU3_NlaIII-proU7_NlaIII

	Replicate 1	Replicate 2

Cross-link fragment	Interaction frequency	Positive error	Negative error	Interaction frequency	Positive error	Negative error
proU13_NlaIII	0.47635	0.02006	0.01915	0.44966	0.01503	0.01455
proU1_NlaIII	0.43016	0.00613	0.00603	0.38473	0.01239	0.01200
proU2_NlaIII	0.41590	0.00820	0.00802	0.39729	0.01191	0.01156
proU4_NlaIII	0.59093	0.01647	0.01590	0.48582	0.00206	0.00206
proU5_NlaIII	0.07650	0.00201	0.00204	0.08878	0.00019	0.00018
proU6_NlaIII	0.16982	0.00540	0.00525	0.19367	0.00667	0.00645
proU7_NlaIII	1	0.05170	0.04916	1	0.02068	0.02026
proU8_NlaIII	0.10778	0.00391	0.00377	0.11385	0.00107	0.00106
proU9_NlaIII	0.05314	0.00072	0.00071	0.05070	0.00272	0.00258
proU10_NlaIII	1.01213	0.05303	0.05010	0.88884	0.01121	0.01107
proU11_NlaIII	0.28176	0.00764	0.00733	0.17288	0.00093	0.00092
proU12_NlaIII	0.05133	0.00222	0.00212	0.04904	0.00192	0.00185
proU14_NlaIII	0.00012	0.000005	0.000005	0.00012	0.000007	0.000007
proU15_NlaIII	0.00307	0.00013	0.00012	0.00243	0.00015	0.00014

Table S3.11: Relative cross-linking frequency of proU3_NlaIII with fragments within, and flanking the *proVWX* operon in NT331 following a hyperosmotic shock from 0.08 M to 0.3 M NaCl. Cross-linking control: proU3_NlaIII-proU7_NlaIII

Cross-link fragment	Replicate 1			Replicate 2		
	Interaction frequency	Positive error	Negative error	Interaction frequency	Positive error	Negative error
proU13_NlaIII	0.25154	0.00515	0.00505	0.27086	0.00779	0.00758
proU1_NlaIII	0.26281	0.00279	0.00276	0.29538	0.00621	0.00608
proU2_NlaIII	0.29662	0.00692	0.00676	0.37169	0.04124	0.03712
proU4_NlaIII	0.47300	0.00721	0.00710	0.60827	0.06438	0.05822
proU5_NlaIII	0.09636	0.00262	0.00255	0.11098	0.01334	0.01191
proU6_NlaIII	0.16592	0.00424	0.00413	0.19467	0.00717	0.00692
proU7_NlaIII	1	0.02407	0.02351	1	0.01493	0.01471
proU8_NlaIII	0.15494	0.00273	0.00268	0.15920	0.00116	0.00115
proU9_NlaIII	0.03905	0.00151	0.00145	0.04484	0.00167	0.00161
proU10_NlaIII	0.59386	0.01880	0.01823	0.55523	0.01225	0.01199
proU11_NlaIII	0.09405	0.00475	0.00452	0.09333	0.00179	0.00175
proU12_NlaIII	0.04729	0.00104	0.00102	0.04630	0.00504	0.00454
proU14_NlaIII	0.00008	0.000004	0.000004	0.00008	0.000006	0.000006
proU15_NlaIII	0.00200	0.00116	0.00073	0.00140	0.00016	0.00014

Table S3.12: Relative cross-linking frequency of proU3_NlaIII with fragments within, and flanking the *proVWX* operon in NT331 during exponential growth at 0.3 M NaCl. Cross-linking control: proU3_NlaIII-proU7_NlaIII

Cross-link fragment	Replicate 1			Replicate 2		
	Interaction frequency	Positive error	Negative error	Interaction frequency	Positive error	Negative error
proU13_NlaIII	0.29507	0.00471	0.004632	0.27786	0.00622	0.00609
proU1_NlaIII	0.27475	0.00504	0.00495	0.27088	0.00657	0.00642
proU2_NlaIII	0.28512	0.01430	0.01362	0.29622	0.00739	0.00721
proU4_NlaIII	0.50758	0.03359	0.03150	0.50704	0.00138	0.00137
proU5_NlaIII	0.06922	0.00202	0.00196	0.06809	0.00520	0.00483
proU6_NlaIII	0.14520	0.00312	0.00305	0.13839	0.00289	0.00283
proU7_NlaIII	1	0.03612	0.03486	1	0.02139	0.02095
proU8_NlaIII	0.12685	0.00294	0.00287	0.12323	0.00208	0.00205
proU9_NlaIII	0.04040	0.00152	0.00147	0.03948	0.00108	0.00105
proU10_NlaIII	0.78385	0.02326	0.02259	0.75720	0.00780	0.00772
proU11_NlaIII	0.10415	0.00185	0.00181	0.09542	0.00104	0.00103

proU12_NlaIII	0.04305	0.00094	0.00092	0.04197	0.00293	0.00274
proU14_NlaIII	0.00008	0.000006	0.000005	0.00007	0.000003	0.000003
proU15_NlaIII	0.00225	0.00033	0.00028	0.00217	0.00001	0.00001

Table S3.13: Relative cross-linking frequency of proU3_NlaIII with fragments within, and flanking the *proVWX* operon in NT331 following a hypoosmotic shock from 0.3 M to 0.08 M NaCl. Cross-linking control: proU3_NlaIII-proU7_NlaIII

Cross-link fragment	**Replicate 1**			**Replicate 2**		
	Interaction frequency	Positive error	Negative error	Interaction frequency	Positive error	Negative error
proU13_NlaIII	0.47635	0.00854	0.00839	0.45766	0.03432	0.03193
proU1_NlaIII	0.43016	0.01236	0.01202	0.38738	0.01975	0.01880
proU2_NlaIII	0.41590	0.00708	0.00697	0.40856	0.01089	0.01060
proU4_NlaIII	0.59094	0.01113	0.01093	0.57023	0.02499	0.02394
proU5_NlaIII	0.07650	0.00031	0.00030	0.07484	0.00075	0.00074
proU6_NlaIII	0.16982	0.00218	0.00215	0.16048	0.00292	0.00287
proU7_NlaIII	1	0.01818	0.01785	1	0.00620	0.00616
proU8_NlaIII	0.10779	0.00681	0.00641	0.10431	0.00942	0.00864
proU9_NlaIII	0.05315	0.00162	0.00157	0.05294	0.00619	0.00554
proU10_NlaIII	1.01214	0.01928	0.01892	0.97268	0.02005	0.01965
proU11_NlaIII	0.28176	0.00971	0.00938	0.27600	0.00942	0.00911
proU12_NlaIII	0.05133	0.00075	0.00074	0.04911	0.00134	0.00130
proU14_NlaIII	0.00012	0.00001	0.00001	0.00011	0.000003	0.000003
proU15_NlaIII	0.00307	0.00005	0.00005	0.00320	0.00008	0.00008

Chapter 4:
HI-NESS: A family of genetically-encoded DNA labels based on a bacterial nucleoid associated protein

Rashid, F-Z.M., Mahlandt, E., Van der Vaart, M., Solari, J., Henneman, B., Brocken, D.J.W., Voskamp, P., Blok, A., Shimizu, T., Meijer, A.H., Goedhart, J., Crémazy, F.G.E., and Dame, R.T. HI-NESS: A family of DNA labelling dyes constructed by fluorescent protein fusion to the DNA binding domain of a bacterial nucleoid associated protein. *Manuscript in preparation*.

Contributions statement:
R.T.D conceived and supervised the project. F.M.R., F.G.E.C., B.H., D.J.W.B, cloned the HI-NESS expression vectors. F.M.R. and M.vd.V performed HI-NESS imaging in *Escherichia coli*. E.M. imaged HI-NESS in eukaryotic cell lines. M.vd.V carried out imaging experiments in zebrafish larvae. F.M.R. analysed imaging data captured for *E. coli*, eukaryotic cell lines, and zebrafish larvae. E.M. analysed imaging data for eukaryotic cell lines. J.S. performed PALM single-particle tracking studies. P.V., A.B., and F.M.R. over-expressed and purified HI-NESS. F.M.R. performed HI-NESS *in vitro* experiments. F.M.R. wrote the manuscript. E.M., M.vd.V, T.S., A.H.M., J.G., F.G.E.C., and R.T.D. edited the manuscript.

Abstract

The interplay between three-dimensional chromosome organisation and genomic processes such as replication and transcription necessitates *in vivo* studies of chromosome dynamics. Fluorescent organic dyes are often used for chromosome labelling *in vivo*. The mode of binding of these dyes to DNA cause its distortion, elongation, and partial unwinding. The structural changes induce DNA damage and interfere with the binding dynamics of chromatin-associated proteins, consequently perturbing gene expression, genome replication, and cell cycle progression. We have developed a minimally-perturbing fluorescent DNA label by translationally fusing a (photo-switchable) fluorescent protein to the DNA binding domain of H-NS – a bacterial nucleoid-associated protein. We show that this DNA label, abbreviated as HI-NESS (H-NS-based indicator for nucleic acid stainings), is minimally-perturbing to genomic processes and labels chromosomes in bacteria, eukaryotic cells in culture, and in zebrafish embryos.

Introduction

Developments in the field of chromosome biology have highlighted an intricate interplay between the spatiotemporal organisation of the chromosome and its activities such as transcription, replication, and segregation (1–3). Investigations of such time-resolved structural dynamics of chromosomes in live cells can be performed by fluorescence microscopy. Fluorescent organic dyes are often the first choice for labelling chromosomes. Bisbenzimides such as 6-diamidino-2-phenylindole (DAPI) and Hoechst insert into the minor groove of A-T rich double-stranded DNA (dsDNA), resulting in an enhanced fluorescence emission in the blue range of the visible spectrum (4–6). Hoechst 33258 has also been conjugated to other fluorescent dyes including, but not limited to, IR-786, fluorescein, and silicon-rhodamine (SiR), that shift its spectral properties from the phototoxic UV/blue range towards longer wavelengths of the spectrum (7–11). Cyanine dyes such as those of the TOTO, TO-PRO, and SYTOX families intercalate between DNA bases and exhibit fluorescence emission enhancement (12–14). The amplified fluorescence of organic dyes upon DNA binding allows DNA labelling with a high signal-to-noise ratio. Due to the reversible binding of the fluorophores to DNA, this feature also allows the detection of single binding events that enables super-resolution imaging by binding-activated localization microscopy (BALM) and fluctuation-assisted BALM (15, 16), and point accumulation for imaging in nanoscale topography (PAINT) (17, 18). Stochastic blinking of fluorophores such as silicon-rhodamine and carboxyrhodamine conjugated to Hoechst allows super-resolution imaging with stimulated emission detection (STED) microscopy (10, 11). Furthermore, bisbenzimides undergo stochastic photoconversion upon UV exposure that shifts the excitation and emission spectra of the dyes from the UV/blue to the blue/green and green/red ranges (19–21). While being problematic for multicolour fluorescence microscopy studies, this property is exploited in single molecule localisation microscopy (SMLM) to image chromatin in intact cell nuclei at a resolution of up to ~40 nm (22, 23).

Despite their broad use, the binding mode of organic DNA labels results in distortion, elongation, and/or partial unwinding of the double-helix (24–27). The structural changes interfere with the binding of DNA processing enzymes and affect their enzymatic activity (28–32). Fluorescent organic dyes can also generate single and double strand breaks in the DNA owing to reactive oxygen species that are produced when the excited fluorophore reacts with molecular oxygen (33, 34). DNA intercalation may also trigger DNA damage signalling and cell cycle arrest in the absence of laser excitation (35). Thus, fluorescent organic DNA dyes are

generally cytotoxic, especially during prolonged incubation periods necessary for time-lapse experiments (35–38). Moreover, some fluorescent organic DNA dyes are unable to permeate the membrane of live cells, requiring cell fixation and the use of permeabilisation agents such as SDS and Triton X-100 for DNA staining. The dyes may strongly bind to RNA as well, calling for RNase A treatment prior to imaging (39).

An alternative approach that relies on covalent labelling of DNA with organic fluorophores has also been introduced. Covalent binding of azide-functionalized fluorophores such as tetramethylrhodamine azide (TAMRA) and Alexa Fluor-azide to alkyne-functionalized nucleotide analogues – EdU (5-ethynyl-20-deoxyuridine) and EdC (5-ethynyl-20-deoxycytidine) – with click chemistry (40) has been used to localize DNA in mammalian cells (41–43), plant tissues (44), and *Escherichia coli* (45–47). While affording resolutions of <20nm with super-resolution microscopy techniques (41, 45), this approach is of limited applicability in live cell imaging. The incorporation of EdU and EdC into DNA triggers DNA damage signalling, interferes with cell cycle progression, and induces apoptosis (48, 49). Fluorescent labelling of the nucleotide analogues requires cell fixation and permeabilisation, and Cu^{+} to catalyse the click reaction (40–47). Moreover, the covalent attachment of large fluorophores to chromosomal DNA would predictably be cytotoxic.

Chromosomes can also be visualised by the expression of fluorescent proteins fused to DNA binding proteins. For eukaryotic systems, the histones H2B, H3 and H4 tagged with (photoactivatable) fluorescent proteins have been shown to incorporate into functional nucleosomes, and to accurately describe chromosome structure throughout the cell cycle (50–52). Such fusions have been used to obtain constitutive DNA staining in mice, nematodes, zebrafish, drosophila and arabidopsis transgenic strains (53–57). mEos2- and mEos3-tagged Heterochromatin protein 1 (HP1) have been used to study the distribution of heterochromatin in human embryonic stem cells at super-resolution (58). In bacterial systems, the distributions of fluorescently labelled nucleoid associated proteins (NAPs) such as GFP-labelled α and β subunits of *E. coli* HU (HupA and HupB, respectively) (59) and Fis (60), and GFP/mCherry-labelled HBsu(61, 62) – the *Bacillus subtilis* homologue of HU — have been shown to overlap with that of DAPI indicating that such protein fusions may be used as alternative DNA labels. Indeed, HupA-mCherry, and GFP-Fis have been used to follow the structural changes to the *E. coli* nucleoid during growth and cell division (63, 64), HupB-

EGFP was used to stain the chromosome in *Mycoplasma smegmatis* to study the distribution of fluorescently-labelled Lsr2 and $Lsr2_{\Delta NTD}$ in single cells (65), and GFP-HBsu has been used to image the nucleoid in *B. subtilis* using 3D-structured illumination microscopy (3D-SIM) to visualise high-density regions in the chromosome (66). However, fluorescent protein fusions to native cell proteins can impair protein function by interfering with protein folding and the equilibria of the protein's interactions with other macromolecules (61, 67–69). Fluorescent protein fusions also require titration of expression levels to match that of the native protein. Furthermore, certain fusions have a limited applicability, for instance, fluorescently labelled histone proteins can only be used to visualise the chromosome in eukaryotic cells.

Collectively, this creates a need for a universal, minimally perturbing DNA label for visualizing chromosomes in live cells. To that end, we have designed a fusion protein that exploits the spectral properties of (photoactivatable) fluorescent proteins, and the DNA binding properties of H-NS – a bacterial nucleoid-associated protein. We show that the DNA label, termed HI-NESS (H-NS-based indicator for nucleic acid stainings), is minimally perturbing to cells, accurately describes chromosome structure in bacteria, eukaryotic cells in culture and live zebrafish, and is customizable with regards to the fluorescent protein used for its visualization.

Materials and methods

Cloning HI-NESS expression vectors

The HI-NESS constructs were assembled and cloned into bacterial pBAD33 and eukaryotic pcDNA3.1(+) expression vectors in a single step using Gibson Assembly (70). The constructs were verified by Sanger sequencing and archived as DH5α glycerol stocks. A complete list of the template plasmids used for Gibson Assembly and the HI-NESS vectors assembled therefrom is provided in Table S4.1. The plasmids designed in this study are deposited on Addgene.

HI-NESS imaging in *Escherichia coli*

E. coli cells (MG1655 and MG1655 Δhns) were grown at 37°C in M9 or H1 medium (71) with the appropriate antibiotics (Table S4.1). The medium was supplemented with 0.1% w/v arabinose (Sigma-Aldrich) to induce expression from the pBAD33 vector. For experiments that required DAPI labelling, the dye (DAPI, Sigma-Aldrich) was added to a final concentration of 10 µg/mL to cultures at an OD_{600} of ~0.1. The cells were harvested at an OD_{600} of 0.2 by centrifugation at 3000 xg for 5 minutes and resuspended in M9 or H1 medium to an OD_{600} of ~2.0. 4.0 µL of the culture was pipetted onto a 1 mm-thick, 1.5% agarose pad prepared on a microscope slide. A cover-slip was placed over the cells and the slide was sealed with nail polish. Extremely clean cover-slips, prepared as described in (72), were used for PALM experiments.

For diffraction-limited imaging, *E. coli* cells were visualised using a Leica TCS SPE or SP8 confocal microscope with a 64X oil immersion objective (NA = 1.4) and excited by 405, 488, or 532 nm laser lines. PALM single particle tracking experiments (73, 74) were carried out using the PALM imaging set-up described in (72) with an exposure time of 15 ms and an inter frame interval of 65 ms. Fluorophores were tracked using the u-track package (75), and the diffusion coefficient of each tracked molecule was calculated using a covariance-based estimator method (76). The tracks were sorted into bound and unbound fractions by fitting a histogram of the diffusion coefficients with the sum of two Gaussian functions. The distribution (mean ± standard deviation) of the unbound molecules was extracted as the apparent diffusion coefficient.

HI-NESS imaging in eukaryotic cell lines: cell culture and sample preparation

HeLa (CCL-2, American Tissue Culture Collection; Manassas, VA, USA) and U-2OS (HTB-96, American Tissue Culture Collection, Manassas, VA, USA) cells were cultured in Dulbecco's modified Eagle's medium + GlutaMAX™-I (Giboc) with

10% fetal calf serum (Giboc) (DMEM + FCS) at 37˚C in 7% CO_2. For transfection 25 000 to 50 000 cells were seeded on 24 mm ø cover-slip (Menzel, Thermo Fisher Scientific) in a 6 well plate with 2 ml DMEM + FCS and cultured for 24 hours. A transfection mix containing 0.5 to 1 µg plasmid (Table S4.1), linear polyethylenimine (PEI, pH 7.3, Polysciences) with a concentration of 1 mg/mL per 100 ng DNA, and 200 µl OptiMEM (Thermo Fisher Scientific) was added to each well. 24 h after transfection, cells were incubated with 2 mM thymidine (CAS: 50-89-5, Sigma-Aldrich) in DMEM + FCS for 18 h to increase the percentage of dividing cells. Thereafter, cells were washed twice with DMEM and incubated for another 5 h before imaging. For SiR-Hoechst labeling, the cells were incubated with 500 nM SiR-DNA (SC007, SpiroChrome Probes for Bioimaging) in DMEM, 4 h prior to imaging. HeLa and U-2OS cells were imaged between 24 to 48 h after transfection in an Attofluor cell chamber (Thermo Fischer Scientific) in 1 ml of Microscopy medium (20 mM HEPES (pH=7.4),137 mM NaCl, 5.4 mM KCl, 1.8 mM $CaCl_2$, 0.8 mM $MgCl_2$ and 20 mM Glucose) at 37˚C.

Blood outgrowth endothelial cells (BOEC) were cultivated from healthy adult donor blood as described previously (77) and cultured in Endothelial Cell Growth Medium-2 BulletKit (CC-3162, Lonza) with 100 U/mL Penicillin (Thermo Fisher Scientific) and 100 µg/mL Streptomycin (Thermo Fisher Scientific), and 20% fetal calf serum (Giboc) (EGM +) at 37˚C in 5% CO_2. Culture dishes and cover-slips were coated with 0.1% gelatin (CAS 9000-70-8, Merck) in phosphate-buffered saline 30 min prior to cell seeding. Transfection was performed with 2 µg endotoxin free plasmid, using the Neon™ Electroporation Transfection system (MPK5000, Invitrogen) with the associated Neon™ Transfection System 100 µl Kit (MPK10096, Invitrogen) generating a single pulse of 1300 V for 30 ms. Cells were seeded on 24 mm ø cover-slip in a 6 well plate with 2 ml EGM +. BOECs were imaged between 24 to 48 h after microporation in an Attofluor cell chamber (Thermo Fischer Scientific) in 1 ml EGM + at 37˚C and 5% CO_2.

HI-NESS imaging in eukaryotic cell lines: spinning disk microscopy

Cells were imaged with a Nikon Ti-E microscope equipped with a Yokogawa CSU X-1 spinning disk unit, a 60x objective (Plan Apo, VC, oil, DIC, NA 1.4), a 100x objective (Apo, TIRF, oil, DIC, N2), Perfect Focus System, and the Nikon NIS elements software. Images were acquired with an Andor iXon 897 CCD camera. mTurquoise2 was imaged using a 440 nm laser line, a triple dichroic mirror (suitable for 440, 514, 561 nm laser) and a 460 – 500 nm emission filter. mEos3.2 was imaged using a 488 nm laser line, a triple dichroic mirror (suitable for 405,

488, 561 nm laser) and a 500 nm long pass emission filter. mEos3.2 was photo-converted with a 405 nm laser line and imaged using 561 nm laser line, a triple dichroic mirror (suitable for 405, 488, 561 nm laser) and a 600 – 660 nm emission filter.

HI-NESS imaging in eukaryotic cell lines: wide-field microscopy

Dividing cells were imaged on a Nikon Ti-E widefield microscope, equipped with a 60x objective (Plan Apo λ, 60x, oil) and a 20x air objective (Plan Apo, VC, DIC, N2), a Lumencor Spectra X light source, the Perfect Focus System, a Hamamatsu C11440-22C camera (SN:100256), and Nikon NIS elements software. For overnight time lapse movies, HeLa cells were imaged in DMEM + FCS at 37˚C and 5% CO_2 in an Attofluor cell chamber (Thermo Fischer Scientific) in a humidified environment. mTurquoise2 was imaged with an excitation wavelength of 440/20 nm and emission light was detected at 459-499 nm with an emission filter in combination with a dichroic mirror (455-491, 523-557, 590-800 nm transmission bands). mScarlet-I was imaged with an excitation wavelength of 550/15 nm and emission light was detected at 570–616 nm with an emission filter in combination with a dichroic mirror (411-452, 485-541, 567-621, 656-793 nm transmission bands). Phase contrast images were acquired with the phase contrast condenser PH3.

HI-NESS imaging in eukaryotic cell lines: confocal microscopy

Confocal microscopy images were obtained with a Leica SP8 equipped with a 63x objective (HC PL Apo, C2S, NA 1.40, oil), the pinhole was set to 1 Airy unit, using line scan, 4x frame averaging, at a scan speed of 40 Hz. SiR-Hoechst was imaged using a 633 nm laser line, emission light was detected between 642 – 788 nm with a HyD detector and the gain was get to 50V. mTurquoise2 was imaged using a 442 nm laser line, emission light was detected between 452 – 598 nm with a HyD detector and the gain was set to 40V. mEos3.2 was imaged using a 488 nm laser line, emission light was detected between 495 – 554 nm with a HyD detector and the gain was set to 110 V.

Line scan analysis

Line scans were performed using the ‘Plot profile’ function in Fiji (78). All profiles were normalized to the most intense pixel of the line scan that was assigned an arbitrary intensity value of 1000.

HI-NESS imaging in zebrafish embryos and larvae

Zebrafish lines used in this study (AB/TL wild types) were handled in compliance with local animal welfare regulations, as overseen by the Animal Welfare Body of Leiden University (License number: 10612) and maintained according to standard protocols (http://zfin.org/). All experiments were done on embryos or larvae up to 5 days post-fertilization (dpf), that had not yet reached the free-feeding stage. Embryos/larvae were kept in egg water (60 μg/ml Instant Ocean sea salts) at 28.5°C and anesthetized with 0.02% ethyl 3-aminobenzoate methanesulfonate (Tricaine, Sigma-Aldrich) before imaging and fixation.

To achieve mosaic expression of HI-NESS and the mEos3.2 control, 25 – 50 pg of pRD188 or pRD190 (Table S4.1) in 1x Danieau buffer (58 mM NaCl, 0.7 mM KCl, 0.4 mM $MgSO_4$, 0.6 mM $Ca(NO_3)_2$ and 5.0 mM HEPES, pH 7.6.) was microinjected into zebrafish embryos at the one-cell stage. After 24 hours, the embryos were screened for fluorescence using a Leica MZ16FA stereo fluorescence microscope. For co-staining with DAPI, larvae expressing HI-NESS or the mEos3.2 control were fixed with 4% PFA in 1x PBS at 4°C overnight. Fixed larvae were washed with 1x PBS and stained with DAPI (Sigma-Aldrich) at a final concentration of 100 μg/ml in 1x PBS. Fixed or live embryos of 2 dpf were mounted with 1.5% low melting point agarose (SERVA) in egg water and imaged using a Leica TCS SPE or SP8 confocal microscope with a 40X water immersion objective (NA = 0.8) and excited by 405, 488, or 532 nm laser lines.

Results and Discussion

HI-NESS can be displaced by native H-NS from chromosomal DNA

We engineered HI-NESS by translationally fusing mEos3.2, a photo-switchable fluorescent protein, to the N-terminus of an *Escherichia coli* H-NS truncation comprising residues 80-137 of the wild-type protein. In this construct, mEos3.2 allows visualisation of the DNA-labelling protein in diffraction-limited and super-resolution microscopy; residues 96-137 of H-NS fold into a DNA-binding domain, and H-NS residues 80-95 form a linker that separates the aforementioned moieties to prevent steric clashes and protein misfolding. In this construct, mEos3.2 was fused to the DNA binding domain of H-NS as opposed to full length H-NS, to achieve a high DNA dissociation constant (79), and to prevent the potential multimerization of HI-NESS that may arise from the presence of an oligomerisation domain (65). These features are expected to make the DNA labelling protein less perturbing to genomic transactions in the cell. To verify this, the distribution of HI-NESS was tested in *E. coli*, where the protein was ectopically expressed from a plasmid. HI-NESS was expected to be outcompeted by native H-NS in binding to chromosomal DNA, while DNA labelling with a high signal-to-noise ratio was expected in the absence of endogenous H-NS. Indeed, HI-NESS distributes relatively homogeneously in wild-type *E. coli* with poor colocalisation with DAPI (Figure 4.1A). In comparison, the distribution of HI-NESS correlates well with the DAPI signal in *E. coli* Δ*hns* (Figure 4.1B).

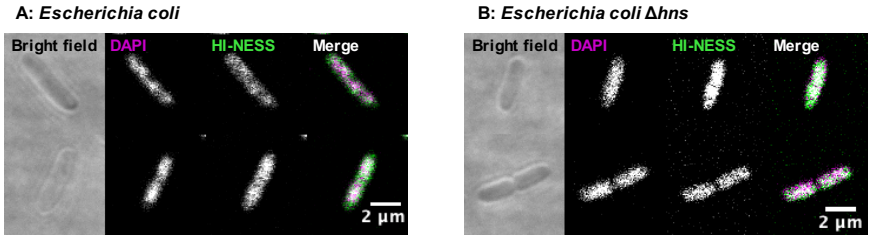


Figure 4.1: HI-NESS distribution in wild-type *Escherichia coli* and *E. coli* Δ*hns*. A: HI-NESS (green) distributes homogeneously in *E. coli* and correlates poorly with the DAPI signal (magenta). **B:** The HI-NESS signal correlates well with the distribution of the DAPI signal in *E. coli* Δ*hns*. A white signal in the Merge images represents colocalization.

To further verify the higher dissociation constant of HI-NESS compared to mEos3.2-H-NS, we used PALM single-particle tracking (73, 74) to characterise the mobility of mEos3.2, HI-NESS, and mEos3.2-H-NS in wild-type *E. coli*. We determined the apparent diffusion coefficients (D_{app}) of these proteins to be 0.44±0.33 µm^2/s (n=1320), 0.31±0.26 µm^2/s (n=3768) and 0.08±0.15 µm^2/s (n=3559), respectively (Figure 4.2). The D_{app} for mEos3.2 in our measurements is an order of magnitude lower than previously reported (74). This is due to longer

camera/laser exposure times and inter frame intervals used in our study. An exposure time of 15 ms in our study – versus 1 ms in (74) – increased the motion blur effect (80) limiting the detection of the population of fast-diffusing mEos3.2 fluorophores. Additionally, our inter frame interval of 65 ms limited the resolution of the sub-diffusive trajectory of the fluorophore in the cytoplasm, hence D_{app} was underestimated (74). Nevertheless, all our single particle tracking experiments were performed with the same imaging parameters ensuring that the D_{app} values determined in our study can be compared.

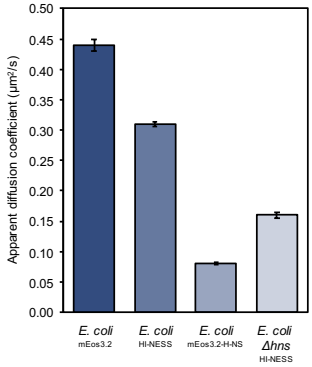


Figure 4.2: Measurements of the apparent diffusion coefficients (D_{app}) of mEos3.2, HI-NESS, and mEos3.2-H-NS with PALM single-particle tracking. mEos3.2 in *E. coli*: D_{app} = 0.44±0.33 μm²/s, n=1320; **HI-NESS in *E. coli*:** D_{app} = 0.31±0.26 μm²/s, n=3768; **mEos3.2-H-NS in *E. coli*:** D_{app} = 0.08±0.15 μm²/s, n=3559; **HI-NESS in *E.coli Δhns*:** D_{app} = 0.16±0.22 μm²/s, n=1796. The errors for D_{app} values in the figure legend are the standard deviations of the apparent diffusion coefficients of the fluorophores and, as such, represent the distribution of diffusion coefficients of unbound fluorescent molecules as determined by PALM single-particle tracking. The errorbars in the graph represent standard error of the mean.

Apparent diffusion coefficients of proteins in the *E. coli* cytoplasm that do not exhibit specific interactions with cytosolic elements, follow the Einstein-Stoke's equation for the diffusion of spherical particles in a classical fluid. Apparent diffusion coefficients are predictable for molecules up to a size of ~110 kDa, with a 30% increase in molecular weight accounting for a 10% decrease in the apparent diffusion coefficient (81). Larger decreases indicate drag due to increased interaction with cytosolic contents (81). HI-NESS (32.1 kDa) has a ~30% lower D_{app} than the fluorophore itself (25.7 kDa). With a molecular weight ~26% larger than that of mEos3.2, the size of HI-NESS only accounts for ~10% of the decrease (81). The difference in D_{app} values may indicate DNA binding by HI-NESS, however, it may also be accounted for by the protein structured as a pair of spheres held together by a flexible linker, rather than an ideal spherical particle. The ~4-fold higher D_{app} of HI-NESS (32.1 kDa) compared to that of mEos3.2-H-NS (41.3 kDa, ~30% larger than HI-NESS) highlights that HI-NESS has a higher

dissociation constant, and may thus be minimally perturbing to genome transactions. Furthermore, in *E. coli* Δ*hns*, HI-NESS has a D_{app} value of 0.16±0.22 $\mu m^2/s$ (n=1796) – ~50% lower than that observed in wild-type *E. coli* (Figure 4.2). This indicates that in wild-type *E. coli*, HI-NESS can be displaced by native H-NS from binding to the chromosome. Together, these results imply that HI-NESS is not a suitable DNA stain for use in *E. coli*.

HI-NESS labels chromosomes in eukaryotic cells

The DNA binding properties of HI-NESS make it an excellent candidate for chromosome labelling in eukaryotic cells. A pcDNA3.1(+) vector was used to express HI-NESS flanked by a pair of SV40 T-antigen derived nuclear localisation signals (NLSs) in HeLa, BOEC, and U2OS cells. The fluorescence signal appeared as several discrete and dense foci in nuclei (Figure 4.3 and S4.1) that overlapped with the SiR-Hoechst signal (Figure 4.3). Such foci were not visible when the cells expressed NLS-mEos3.2-NLS lacking the H-NS DNA binding domain (Figure S4.2). HI-NESS also accumulated in nucleoli – structures that tend to be devoid of DNA, but enriched in RNA and protein (Figure 4.3 and S4.1). The accumulation is expected to be largely non-specific and to occur due to high levels of HI-NESS in the nucleus since HI-NESS lacks detectable nucleolar localisation/retention signals (NoRSs) (82, 83), and the SV40 T-antigen derived NLS cannot drive nucleolar accumulation of proteins (84). However, the affinity of HI-NESS for RNA in *in vitro* assays indicates that the accumulation may partly be due to RNA labelling (Supplementary methods; Table S4.2; Figure S4.3).

We then evaluated the experimental advantage of using HI-NESS over fluorescently tagged histone proteins to label eukaryotic chromosomes. H2A-mTurquoise2 (H2A-mTq2) (85) was expressed in HeLa cells and the nuclei were co-stained with SiR-Hoechst. H2A-mTq2 exhibited extensive nucleolar accumulation that drowned the fluorescence signal over the rest of the nucleus, consequently, reducing the signal-to-noise ratio for chromosome visualisation (Figure 4.4, Nuclei 1&2). Nucleolar accumulation of histones has also been observed for fluorescently-labelled H2B, driven by the presence of a NoRS in the protein's nuclear localisation signal (86). NoRS tend to be enriched in positively-charged (basic) amino acids that facilitate electrostatic interactions with the negatively-charged (acidic) contents of the nucleolus (86). In the case of H2A-mTq2, nucleolar accumulation occurs in the absence of a detectable NoRS in the construct (82, 83), and may be promoted by the inherent basicity of the protein. Occasionally, we observed cells with minimal, if any, nucleolar accumulation of

H2A-mTq2 (Figure 4.4, Nuclei 3&4). These cells expressed low levels of H2A-mTq2 as inferred from low fluorescence in the mTurquoise2 channel. In these cases, line scans across nuclei show that the H2A-mTq2 signal recapitulates the SiR-Hoechst signal (Figure 4.4, Nuclei 3&4). A similar observation was made for nuclei that lacked visible nucleoli (Figure 4.4, Nucleus 5).

We co-expressed HI-NESS and H2A-mTq2 in HeLa cells to compare the distribution of the DNA labels in the same nuclei. The study reproduced the extensive nucleolar accumulation of H2A-mTq2 (Figure 4.5). In comparison, chromosomes were visualised with a high signal-to-noise ratio in the HI-NESS channel, with the DNA label exhibiting decreased nucleolar accumulation (Figure 4.5) compared to cells in which only HI-NESS was ectopically expressed (Figure 4.3). This indicates that H2A-mTq2 has a higher propensity for nucleolar retention than HI-NESS and excludes HI-NESS from the sub-nuclear compartment.

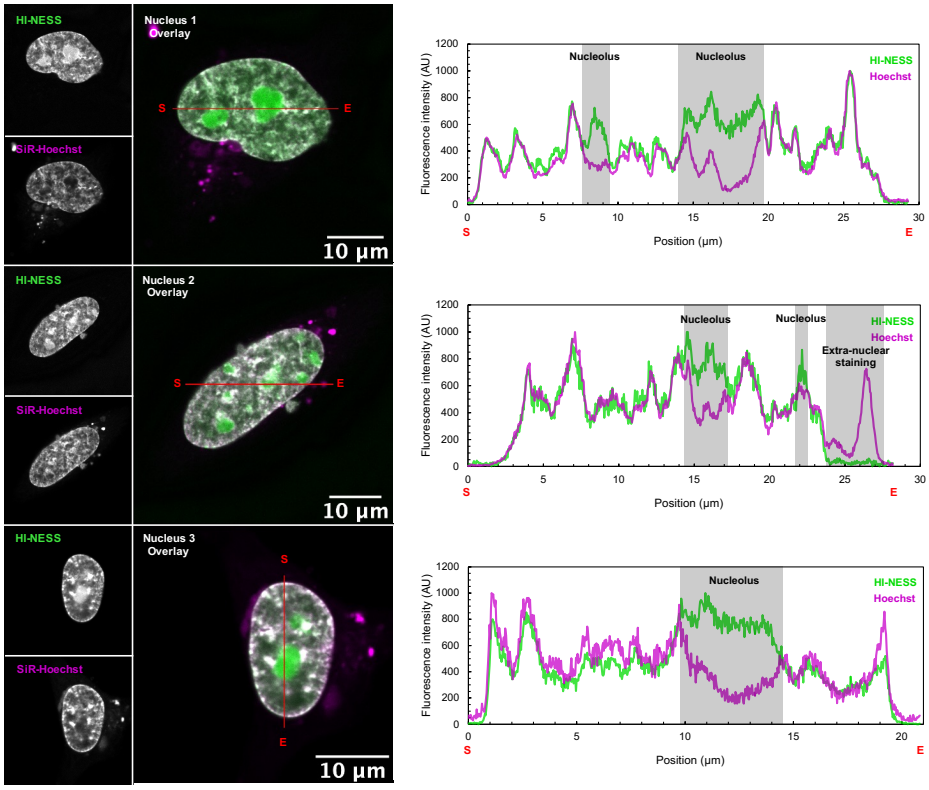


Figure 4.3: HI-NESS labels chromosomes in HeLa cells in culture. Line scans (marked in red with start and end positions indicated with S and E, respectively) across nuclei of HeLa cells co-stained with HI-NESS (green) and SiR-Hoechst (magenta) highlight the overlap between the distributions of the two DNA labels. However, HI-NESS also accumulates in nucleoli due to high levels of the protein in the nucleus and its affinity for RNA as detected in *in vitro* studies (Figure S4.3).

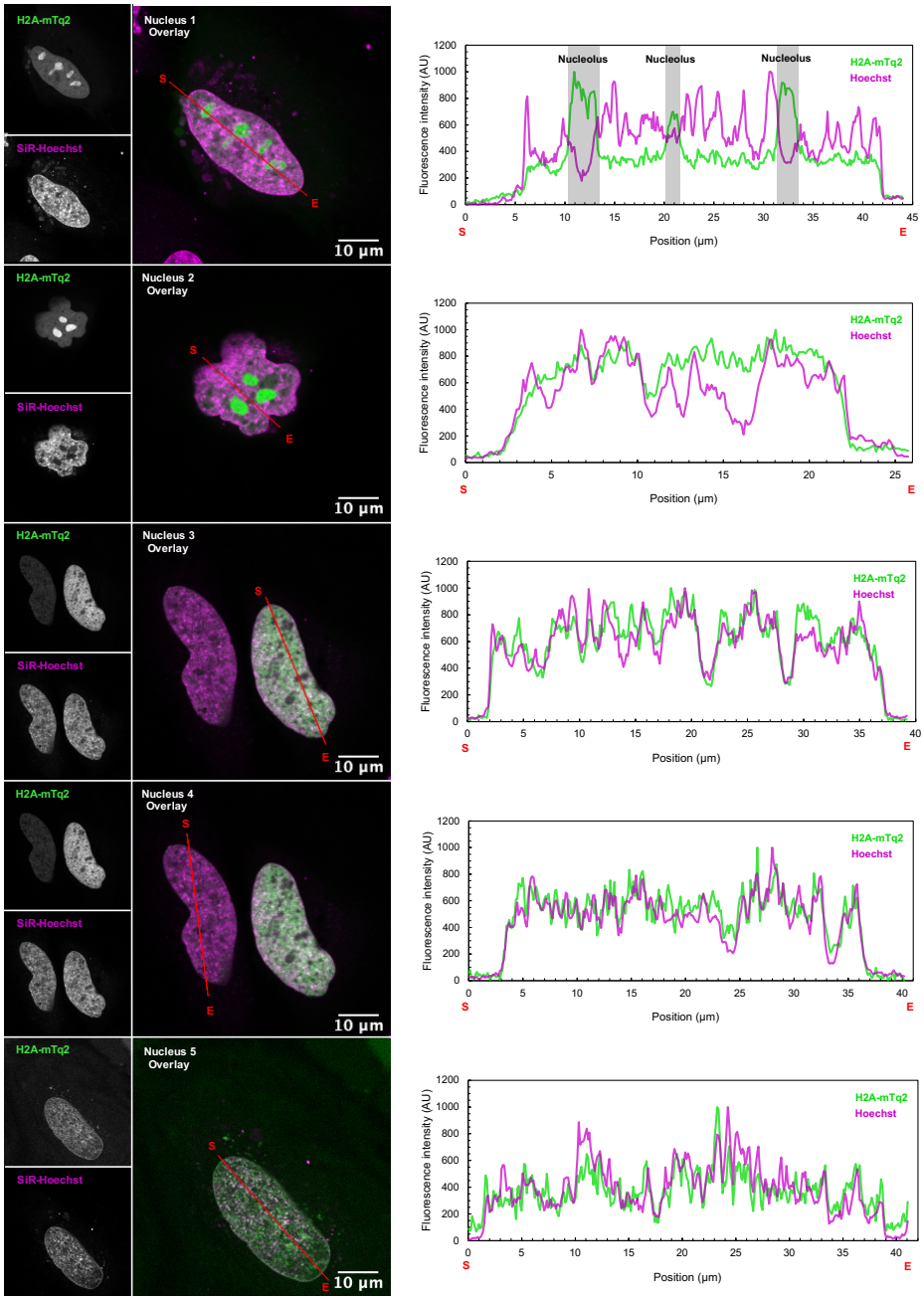


Figure 4.4: The overlap between H2A-mTurquoise2 (green) and SiR-Hoechst (magenta) signals in HeLa cells. Nuclei 1 and 2: Extensive nucleolar accumulation of H2A-mTurquoise2 (H2A-mTq2) drowns its fluorescent signal over the rest of the nucleus. Line scans (marked in red with start and end positions indicated with S and E, respectively) across such nuclei show that the H2A-mTq2 signal only recapitulates that of SiR-Hoechst when the line scan does not cross a nucleolus. **Nuclei 3, 4, and 5:** In cells expressing low levels of H2A-mTq2 (nuclei 3, and 4), and in nuclei with no visible nucleoli (nucleus 5) the distribution of the SiR-Hoechst and H2A-mTq2 signals are comparable.

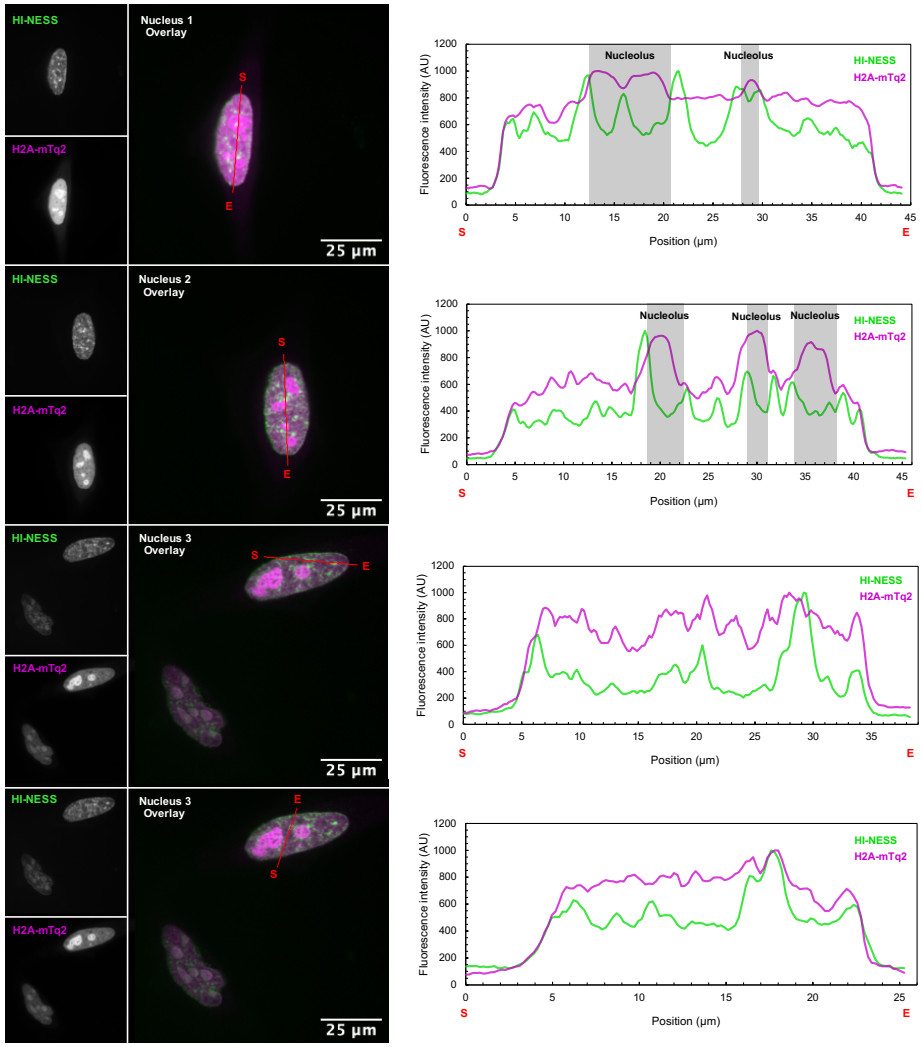


Figure 4.5: HI-NESS (green) labels the chromosomes of HeLa cells in culture with a higher signal-to-noise ratio than H2A-mTq2 (magenta). In HeLa cells co-expressing HI-NESS and H2A-mTq2, the fluorescently-tagged histone exhibits higher nucleolar retention than HI-NESS and, consequently, stains the chromosomes with a lower signal-to-noise ratio. Line scans (marked in red with start and end positions indicated with S and E, respectively) across these nuclei highlight the effect of nucleolar retention on the signal over the rest of the nucleus and show a similarity in the distribution of the DNA labelling proteins in non-nucleolar regions.

HI-NESS as a DNA labelling protein in zebrafish

The feasibility of chromosomal DNA staining in whole organisms using HI-NESS was investigated in zebrafish. HI-NESS and the NLS-mEos3.2-NLS control were expressed in zebrafish from pcDNA3.1(+) vectors microinjected into the embryos at the 1-cell stage. The distribution of the fluorophores was visualised using confocal laser scanning microscopy in fixed zebrafish larvae co-stained with DAPI.

Predictably, HI-NESS was detected in cell nuclei where its distribution overlapped with that of DAPI (Figure 4.6, Movies S4.1-S4.5). Interestingly, nucleolar accumulation of HI-NESS was not observed. This may be due to a lower expression of HI-NESS, however, it also indicates that while HI-NESS may exhibit a comparable affinity for RNA and DNA *in vitro* (Figure S4.3), HI-NESS preferentially binds to DNA in an *in vivo* system. In zebrafish larvae expressing NLS-mEos3.2-NLS lacking the H-NS DNA binding domain, the fluorescent protein accumulated in nucleoli and distributed uniformly over the non-nucleolar regions of the nucleus. Evidently, the mEos3.2 signal did not recapitulate the distribution of DAPI (Figure S4.4, Movies S4.6-S4.8).

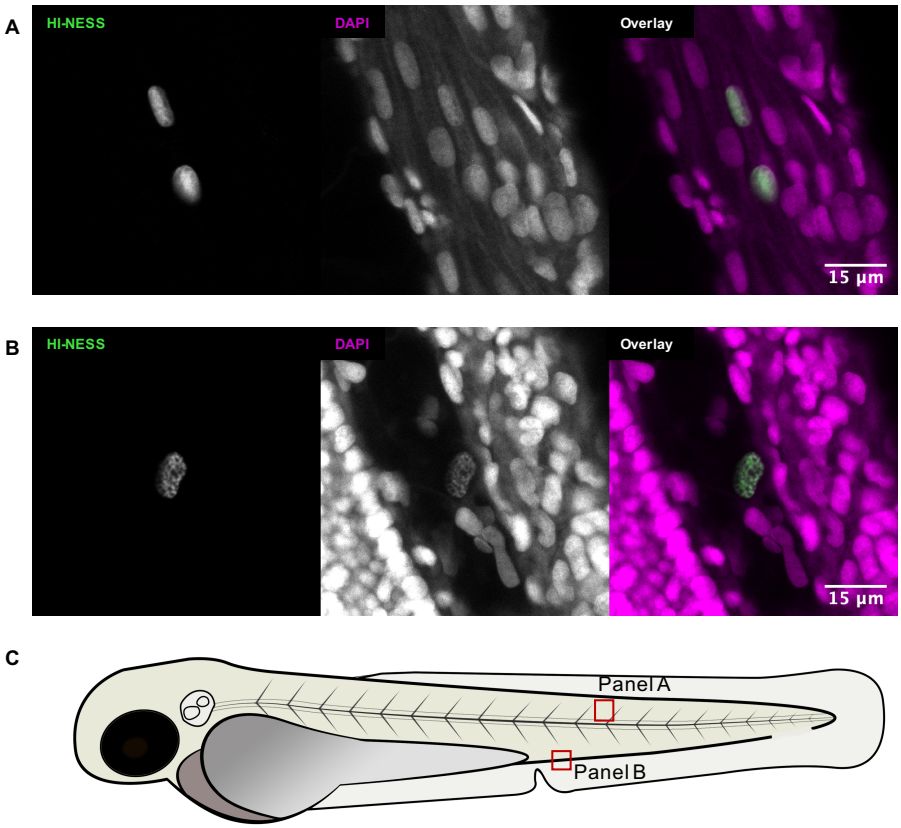


Figure 4.6: HI-NESS labels chromosomes in zebrafish larvae. The distribution of the mEos3.2 signal (green) in nuclei of zebrafish larvae overlapped with that of DAPI (magenta). Nucleolar accumulation of HI-NESS was not observed. **A:** Striated muscle cells (Trunk); **B:** from left to right: trunk, blood vessel, yolk extension. The HI-NESS labelled cell is within a blood vessel. The approximate locations in the zebrafish embryo at which **A** and **B** were imaged is provided in **C**.

Zebrafish larvae exhibited a mosaic expression of HI-NESS and NLS-mEos3.2-NLS in our experiments (Figure S4.5) owing to the microinjection of the pcDNA3.1(+)

vector into embryos at the one-cell stage. For constitutive HI-NESS expression in an animal model, the Tol2 transposon-based gene insertion system (reviewed in (87, 88)) may be used to integrate the HI-NESS gene into the genome. Constructs with the HI-NESS gene placed downstream of a cell-type-specific promoter may also be designed to selectively label nuclei in a live animal model.

HI-NESS is customisable

HI-NESS is a modular protein comprised of fluorescent, DNA-binding, and, optionally, organelle-targeting segments. The protein domains forming these segments can potentially be switched out for others exhibiting similar properties. We have generated a range of HI-NESS labels where mEos3.2 has been swapped for non-photoswitchable fluorescent proteins (Table S4.1). In eukaryotes, HI-NESS can be used to visualise chromosome dynamics during the cell cycle at a high spatial and temporal resolution (Movie S4.9), and follow the movements of nuclei in a live animal model (Movie S4.10). The addition of organelle-targeting/localisation signals can be exploited to specifically label nuclear, mitochondrial, or chloroplast DNA. In prokaryotes, we predict that HI-NESS can be used to visualise the chromosome in bacteria naturally lacking H-NS and H-NS-like proteins. The H-NS-based DNA-binding module can also be replaced with archaeal DNA-binding domains or stable variants evolved from the H-NS-based module to study chromosome dynamics in extremophiles.

Conclusion

We have designed a minimally-perturbing, DNA label to visualise chromosomes in bacteria, eukaryotic cells in culture, and in live animal models. We have shown that in these systems, the distribution of the label (HI-NESS) overlaps with that of traditional DNA labels such as DAPI and SiR-Hoechst. In eukaryotes, HI-NESS offers an alternative to fluorescently labelled histone proteins that tend to accumulate in nucleoli. Furthermore, HI-NESS staining closely mimics that of organic DNA dyes. It is, therefore, a more effective choice for eukaryotic DNA labelling than histone markers.

Availability

Fiji/ImageJ are an open source image processing programs made available by the National Institutes of Health (NIH), Bethesda, USA.

Prism8 is an analysis and graphing software developed by GraphPad Software, San Diego, USA.

References:

1. Dame,R.T., Rashid,F.Z.M. and Grainger,D.C. (2020) Chromosome organization in bacteria: mechanistic insights into genome structure and function. *Nat. Rev. Genet.*, **21**, 227–242.
2. Dame,R.T. and Tark-Dame,M. (2016) Bacterial chromatin: Converging views at different scales. *Curr. Opin. Cell Biol.*, 10.1016/j.ceb.2016.02.015.
3. Yu,M. and Ren,B. (2017) The Three-Dimensional Organization of Mammalian Genomes. *Annu. Rev. Cell Dev. Biol.*, 10.1146/annurev-cellbio-100616-060531.
4. Banerjee,D. and Pal,S.K. (2008) Dynamics in the DNA recognition by DAPI: Exploration of the various binding modes. *J. Phys. Chem. B*, 10.1021/jp077090f.
5. Trotta,E., D'Ambrosio,E., Ravagnan,G. and Paci,M. (1996) Simultaneous and different binding mechanisms of 4',6-diamidino-2- phenylindole to DNA hexamer (d(CGATCG))2: A 1H NMR study. *J. Biol. Chem.*, 10.1074/jbc.271.44.27608.
6. Bailly,C., Colson,P., Hénichart,J. pierre and Houssier,C. (1993) The different binding modes of hoechst 33258 to DNA studied by electirc linear dichroism. *Nucleic Acids Res.*, 10.1093/nar/21.16.3705.
7. Nakayama,A., Bianco,A.C., Zhang,C.Y., Lowell,B.B. and Frangioni,J. V. (2003) Quantification of brown adipose tissue perfusion in transgenic mice using near-infrared fluorescence imaging. *Mol. Imaging*, 10.1162/153535003765276273.
8. Nakamura,A., Takigawa,K., Kurishita,Y., Kuwata,K., Ishida,M., Shimoda,Y., Hamachi,I. and Tsukiji,S. (2014) Hoechst tagging: A modular strategy to design synthetic fluorescent probes for live-cell nucleus imaging. *Chem. Commun.*, 10.1039/c4cc01753f.
9. Dasari,M., Lee,S., Sy,J., Kim,D., Lee,S., Brown,M., Davis,M. and Murthy,N. (2010) Hoechst-IR: An imaging agent that detects necrotic tissue in vivo by binding extracellular DNA. *Org. Lett.*, 10.1021/ol100923d.
10. Lukinavičius,G., Blaukopf,C., Pershagen,E., Schena,A., Reymond,L., Derivery,E., Gonzalez-Gaitan,M., D'Este,E., Hell,S.W., Gerlich,D.W., *et al.* (2015) SiR-Hoechst is a far-red DNA stain for live-cell nanoscopy. *Nat. Commun.*, 10.1038/ncomms9497.
11. Bucevičius,J., Keller-Findeisen,J., Gilat,T., Hell,S.W. and Lukinavičius,G. (2019) Rhodamine-Hoechst positional isomers for highly efficient staining of heterochromatin. *Chem. Sci.*, 10.1039/c8sc05082a.
12. Rye,H.S., Yue,S., Wemmer,D.E., Quesada,M.A., Haugland,R.P., Mathies,R.A. and Glazer,A.N. (1992) Stable fluorescent complexes of double-stranded DNA with bis-intercalating asymmetric cyanine dyes: Properties and applications. *Nucleic Acids Res.*, 10.1093/nar/20.11.2803.
13. Van Hooijdonk,C.A.E.M., Glade,C.P. and Van Erp,P.E.J. (1994) TO-PRO-3 iodide: A novel HeNe laser-excitable DNA stain as an alternative for propidium iodide in multiparameter flow cytometry. *Cytometry*, 10.1002/cyto.990170212.
14. Glazer,A.N. and Rye,H.S. (1992) Stable dye-DNA intercalation complexes as reagents for high-sensitivity fluorescence detection. *Nature*, 10.1038/359859a0.
15. Schoen,I., Ries,J., Klotzsch,E., Ewers,H. and Vogel,V. (2011) Binding-activated localization microscopy of DNA l. *Nano Lett.*, 10.1021/nl2025954.
16. Szczurek,A., Klewes,L., Xing,J., Gourram,A., Birk,U., Knecht,H., Dobrucki,J.W., Mai,S. and Cremer,C. (2017) Imaging chromatin nanostructure with binding-activated localization microscopy based on DNA structure fluctuations. *Nucleic Acids Res.*, 10.1093/nar/gkw1301.
17. Legant,W.R., Shao,L., Grimm,J.B., Brown,T.A., Milkie,D.E., Avants,B.B., Lavis,L.D. and Betzig,E. (2016) High-density three-dimensional localization microscopy across large volumes. *Nat. Methods*, 10.1038/nmeth.3797.
18. Spahn,C.K., Glaesmann,M., Grimm,J.B., Ayala,A.X., Lavis,L.D. and Heilemann,M. (2018) A toolbox for multiplexed super-resolution imaging of the E. coli nucleoid and membrane using novel PAINT labels. *Sci. Rep.*, 10.1038/s41598-018-33052-3.
19. Karg,T.J. and Golic,K.G. (2018) Photoconversion of DAPI and Hoechst dyes to green and red-emitting forms after exposure to UV excitation. *Chromosoma*, 10.1007/s00412-017-0654-5.
20. Zurek-Biesiada,D., Kedracka-Krok,S. and Dobrucki,J.W. (2013) UV-activated conversion of Hoechst 33258, DAPI, and Vybrant DyeCycle fluorescent dyes into blue-excited, green-emitting protonated forms. *Cytom. Part A*, 10.1002/cyto.a.22260.
21. Piterburg,M., Panet,H. and Weiss,A. (2012) Photoconversion of DAPI following UV or violet excitation can cause DAPI to fluoresce with blue or cyan excitation. *J. Microsc.*, 10.1111/j.1365-2818.2011.03591.x.

22. Szczurek,A.T., Prakash,K., Lee,H.K., Zurek-Biesiada,D.J., Best,G., Hagmann,M., Dobrucki,J.W., Cremer,C. and Birk,U. (2014) Single molecule localization microscopy of the distribution of chromatin using hoechst and DAPI fluorescent probes. *Nucl. (United States)*, 10.4161/nucl.29564.
23. Zurek-Biesiada,D., Szczurek,A.T., Prakash,K., Best,G., Mohana,G.K., Lee,H.K., Roignant,J.Y., Dobrucki,J.W., Cremer,C. and Birk,U. (2016) Quantitative super-resolution localization microscopy of DNA in situ using Vybrant® DyeCycle™ Violet fluorescent probe. *Data Br.*, 10.1016/j.dib.2016.01.041.
24. Spielmann,H.P. (1998) Dynamics of a bis-intercalator DNA complex by 1H-detected natural abundance 13C NMR spectroscopy. *Biochemistry*, 10.1021/bi980789e.
25. Spielmann,H.P., Wemmer,D.E. and Jacobsen,J.P. (1995) Solution Structure of a DNA Complex with the Fluorescent Bis-Intercalator TOTO Determined by NMR Spectroscopy. *Biochemistry*, 10.1021/bi00027a004.
26. Kamitori,S. and Takusagawa,F. (1992) Crystal structure of the 2:1 complex between d(GAAGCTTC) and the anticancer drug actinomycin D. *J. Mol. Biol.*, 10.1016/0022-2836(92)90931-9.
27. Günther,K., Mertig,M. and Seidel,R. (2010) Mechanical and structural properties of YOYO-1 complexed DNA. *Nucleic Acids Res.*, 10.1093/nar/gkq434.
28. Chiang,S.Y., Welch,J., Beerman,T.A. and Rauscher,F.J. (1994) Effects of Minor Groove Binding Drugs on the Interaction of TATA Box Binding Protein and TFIIA with DNA. *Biochemistry*, 10.1021/bi00189a003.
29. Störl,K., Störl,J., Zimmer,C. and Lown,J.W. (1993) Minor-groove binders are inhibitors of the catalytic activity of DNA gyrases. *FEBS Lett.*, 10.1016/0014-5793(93)81513-Y.
30. Straney,D.C. and Crothers,D.M. (1987) Effect of Drug–DNA Interactions upon Transcription Initiation at the lac Promoter. *Biochemistry*, 10.1021/bi00381a031.
31. Woynarowski,J.M., McHugh,M., Sigmund,R.D. and Beerman,T.A. (1989) Modulation of topoisomerase II catalytic activity by DNA minor groove binding agents distamycin, Hoechst 33258, and 4',6-diamidine-2-phenylindole. *Mol. Pharmacol.*
32. Parolin,C., Montecucco,A., Ciarrocchi,G., Pedrali-Noy,G., Valisena,S., Palumbo,M. and Palu,G. (1990) The effect of the minor groove binding agent DAPI (2-amidino-diphenyl-indole) on DNA-directed enzymes: an attempt to explain inhibition of plasmid expression in Escherichia coli. *FEMS Microbiol. Lett.*, 10.1016/S0378-1097(05)80065-5.
33. Åkerman,B. and Tuite,E. (1996) Single- and double-strand photocleavage of DNA by YO, YOYO and TOTO. *Nucleic Acids Res.*, 10.1093/nar/24.6.1080.
34. Tycon,M.A., Dial,C.F., Faison,K., Melvin,W. and Fecko,C.J. (2012) Quantification of dye-mediated photodamage during single-molecule DNA imaging. *Anal. Biochem.*, 10.1016/j.ab.2012.03.021.
35. Sen,O., Saurin,A.T. and Higgins,J.M.G. (2018) The live cell DNA stain SiR-Hoechst induces DNA damage responses and impairs cell cycle progression. *Sci. Rep.*, 10.1038/s41598-018-26307-6.
36. Bielawski,K., Woczyski,S. and Bielawska,A. (2001) DNA-binding activity and cytotoxicity of the extended diphenylfuran bisamidines in breast cancer MCF-7 cells. *Biol. Pharm. Bull.*, 10.1248/bpb.24.704.
37. Durand,R.E. and Olive,P.L. (1982) Cytotoxicity, mutagenicity and DNA damage by Hoechst 33342. *J. Histochem. Cytochem.*, 10.1177/30.2.7061816.
38. Haraguchi,T., Ding,D.Q., Yamamoto,A., Kaneda,T., Koujin,T. and Hiraoka,Y. (1999) Multiple-color fluorescence imaging of chromosomes and microtubules in living cells. In *Cell Structure and Function*.
39. Martin,R.M., Leonhardt,H. and Cardoso,M.C. (2005) DNA labeling in living cells. *Cytom. Part A*, 10.1002/cyto.a.20172.
40. Kolb,H.C., Finn,M.G. and Sharpless,K.B. (2001) Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chemie Int. Ed.*, 10.1002/1521-3773(20010601)40:11<2004::aid-anie2004>3.3.co;2-x.
41. Zessin,P.J.M., Finan,K. and Heilemann,M. (2012) Super-resolution fluorescence imaging of chromosomal DNA. *J. Struct. Biol.*, 10.1016/j.jsb.2011.12.015.
42. Salic,A. and Mitchison,T.J. (2008) A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proc. Natl. Acad. Sci. U. S. A.*, 10.1073/pnas.0712168105.
43. Qu,D., Wang,G., Wang,Z., Zhou,L., Chi,W., Cong,S., Ren,X., Liang,P. and Zhang,B. (2011) 5-Ethynyl-20-deoxycytidine as a new agent for DNA labeling: Detection of proliferating cells. *Anal.*

Biochem., 10.1016/j.ab.2011.05.037.
44. Kotogány,E., Dudits,D., Horváth,G. V. and Ayaydin,F. (2010) A rapid and robust assay for detection of S-phase cell cycle progression in plant cells and tissues by using ethynyl deoxyuridine. *Plant Methods*, 10.1186/1746-4811-6-5.
45. Spahn,C., Endesfelder,U. and Heilemann,M. (2014) Super-resolution imaging of Escherichia coli nucleoids reveals highly structured and asymmetric segregation during fast growth. *J. Struct. Biol.*, 10.1016/j.jsb.2014.01.007.
46. Spahn,C., Cella-Zannacchi,F., Endesfelder,U. and Heilemann,M. (2015) Correlative super-resolution imaging of RNA polymerase distribution and dynamics, bacterial membrane and chromosomal structure in Escherichia coli. *Methods Appl. Fluoresc.*, 10.1088/2050-6120/3/1/014005.
47. Ferullo,D.J., Cooper,D.L., Moore,H.R. and Lovett,S.T. (2009) Cell cycle synchronization of Escherichia coli using the stringent response, with fluorescence labeling assays for DNA content and replication. *Methods*, 10.1016/j.ymeth.2009.02.010.
48. Zhao,H., Halicka,H.D., Li,J., Biela,E., Berniak,K., Dobrucki,J. and Darzynkiewicz,Z. (2013) DNA damage signaling, impairment of cell cycle progression, and apoptosis triggered by 5-ethynyl-2′-deoxyuridine incorporated into DNA. *Cytom. Part A*, 10.1002/cyto.a.22396.
49. Ligasová,A., Liboska,R., Friedecký,D., Mičová,K., Adam,T., Ozdian,T., Rosenberg,I. and Koberna,K. (2016) Dr Jekyll and Mr Hyde: A strange case of 5-ethynyl-2′-deoxyuridine and 5-ethynyl-2′-deoxycytidine. *Open Biol.*, 10.1098/rsob.150172.
50. Kanda,T., Sullivan,K.F. and Wahl,G.M. (1998) Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. *Curr. Biol.*, 10.1016/S0960-9822(98)70156-3.
51. Kimura,H. and Cook,P.R. (2001) Kinetics of core histones in living human cells: Little exchange of H3 and H4 and some rapid exchange of H2B. *J. Cell Biol.*, 10.1083/jcb.153.7.1341.
52. McKinney,S.A., Murphy,C.S., Hazelwood,K.L., Davidson,M.W. and Looger,L.L. (2009) A bright and photostable photoconvertible fluorescent protein. *Nat. Methods*, 10.1038/nmeth.1296.
53. Köster,R.W. and Fraser,S.E. (2001) Tracing transgene expression in living zebrafish embryos. *Dev. Biol.*, 10.1006/dbio.2001.0242.
54. Fraser,S.T., Hadjantonakis,A.K., Sahr,K.E., Willey,S., Kelly,O.G., Jones,E.A.V., Dickinson,M.E. and Baron,M.H. (2005) Using a histone yellow fluorescent protein fusion for tagging and tracking endothelial cells in ES cells and mice. *Genesis*, 10.1002/gene.20139.
55. Das,T., Payer,B., Cayouette,M. and Harris,W.A. (2003) In vivo time-lapse imaging of cell divisions during neurogenesis in the developing zebrafish retina. *Neuron*, 10.1016/S0896-6273(03)00066-7.
56. Savoian,M.S. and Rieder,C.L. (2002) Mitosis in primary cultures of Drosophila melanogaster larval neuroblasts. *J. Cell Sci.*
57. Boisnard-Lorig,C., Colon-Carmona,A., Bauch,M., Hodge,S., Doerner,P., Bancharel,E., Dumas,C., Haseloff,J. and Berger,F. (2001) Dynamic analyses of the expression of the histone::YFP fusion protein in Arabidopsis show that syncytial endosperm is divided in mitotic domains. *Plant Cell*, 10.1105/tpc.13.3.495.
58. Hu,Y.S., Zhu,Q., Elkins,K., Tse,K., Li,Y., Fitzpatrick,J.A.J., Verma,I.M. and Cang,H. (2013) Light-sheet Bayesian microscopy enables deepcell super-resolution imaging of heterochromatin in live human embryonic stem cells. *Opt. Nanoscopy*, 10.1186/2192-2853-2-7.
59. Wery,M., Woldringh,C.L. and Rouviere-Yaniv,J. (2001) HU-GFP and DAPI co-localize on the Escherichia coli nucleoid. *Biochimie*, 10.1016/S0300-9084(01)01254-8.
60. Hadizadeh,N., Johnson,R.C. and Marko,J.F. (2016) Facilitated dissociation of a nucleoid protein from the bacterial chromosome. *J. Bacteriol.*, 10.1128/JB.00225-16.
61. Köhler,P. and Marahiel,M.A. (1997) Association of the histone-like protein HBsu with the nucleoid of Bacillus subtilis. *J. Bacteriol.*, 10.1128/jb.179.6.2060-2064.1997.
62. Smits,W.K. and Grossman,A.D. (2010) The transcriptional regulator Rok binds A+T-rich DNA and is involved in repression of a mobile genetic element in Bacillus subtilis. *PLoS Genet.*, 10.1371/journal.pgen.1001207.
63. Hadizadeh Yazdi,N., Guet,C.C., Johnson,R.C. and Marko,J.F. (2012) Variation of the folding and dynamics of the Escherichia coli chromosome with growth conditions. *Mol. Microbiol.*, 10.1111/mmi.12071.
64. Fisher,J.K., Bourniquel,A., Witz,G., Weiner,B., Prentiss,M. and Kleckner,N. (2013) Four-

dimensional imaging of E. coli nucleoid organization and dynamics in living cells. *Cell*, 10.1016/j.cell.2013.04.006.
65. Kołodziej,M., Trojanowski,D., Bury,K., Hołówka,J., Paściak,M., Matysik,W., Kąkolewska,H., Feddersen,H., Giacomelli,G., Bramkamp,M., *et al.* (2020) Lsr2 is a nucleoid-associated protein that exerts pleiotropic effects on mycobacterial cellular processes. *bioRxiv*, 10.1101/2020.04.27.063487.
66. Marbouty,M., Le Gall,A., Cattoni,D.I., Cournac,A., Koh,A., Fiche,J.B., Mozziconacci,J., Murray,H., Koszul,R. and Nollmann,M. (2015) Condensin- and Replication-Mediated Bacterial Chromosome Folding and Origin Condensation Revealed by Hi-C and Super-resolution Imaging. *Mol. Cell*, 10.1016/j.molcel.2015.07.020.
67. Heo,M., Nord,A.L., Chamousset,D., Van Rijn,E., Beaumont,H.J.E. and Pedaci,F. (2017) Impact of fluorescent protein fusions on the bacterial flagellar motor. *Sci. Rep.*, 10.1038/s41598-017-11241-w.
68. Zhang,F., Moniz,H.A., Walcott,B., Moremen,K.W., Wang,L. and Linhardt,R.J. (2014) Probing the impact of GFP tagging on Robo1-heparin interaction. *Glycoconj. J.*, 10.1007/s10719-014-9522-1.
69. Swulius,M.T. and Jensen,G.J. (2012) The helical mreb cytoskeleton in Escherichia coli MC1000/pLE7 is an artifact of the N-terminal yellow fluorescent protein tag. *J. Bacteriol.*, 10.1128/JB.00505-12.
70. Gibson,D.G., Young,L., Chuang,R.Y., Venter,J.C., Hutchison,C.A. and Smith,H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, 10.1038/nmeth.1318.
71. Hazelbauer,G.L., Mesibov,R.E. and Adler,J. (1969) Escherichia coli mutants defective in chemotaxis toward specific chemicals. *Proc. Natl. Acad. Sci. U. S. A.*, 10.1073/pnas.64.4.1300.
72. Solari,J., Anquez,F., Scherer,K.M. and Shimizu,T.S. (2018) Bacterial chemoreceptor imaging at high spatiotemporal resolution using photoconvertible fluorescent proteins. In *Methods in Molecular Biology*.
73. Manley,S., Gillette,J.M., Patterson,G.H., Shroff,H., Hess,H.F., Betzig,E. and Lippincott-Schwartz,J. (2008) High-density mapping of single-molecule trajectories with photoactivated localization microscopy. *Nat. Methods*, 10.1038/nmeth.1176.
74. English,B.P., Hauryliuk,V., Sanamrad,A., Tankov,S., Dekker,N.H. and Elf,J. (2011) Single-molecule investigations of the stringent response machinery in living bacterial cells. *Proc. Natl. Acad. Sci. U. S. A.*, 10.1073/pnas.1102255108.
75. Jaqaman,K., Loerke,D., Mettlen,M., Kuwata,H., Grinstein,S., Schmid,S.L. and Danuser,G. (2008) Robust single-particle tracking in live-cell time-lapse sequences. *Nat. Methods*, 10.1038/nmeth.1237.
76. Vestergaard,C.L., Blainey,P.C. and Flyvbjerg,H. (2014) Optimal estimation of diffusion coefficients from single-particle trajectories. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, 10.1103/PhysRevE.89.022726.
77. Martin-Ramirez,J., Hofman,M., Van Den Biggelaar,M., Hebbel,R.P. and Voorberg,J. (2012) Establishment of outgrowth endothelial cells from peripheral blood. *Nat. Protoc.*, 10.1038/nprot.2012.093.
78. Schindelin,J., Arganda-Carreras,I., Frise,E., Kaynig,V., Longair,M., Pietzsch,T., Preibisch,S., Rueden,C., Saalfeld,S., Schmid,B., *et al.* (2012) Fiji: An open-source platform for biological-image analysis. *Nat. Methods*, 10.1038/nmeth.2019.
79. Dame,R.T., Noom,M.C. and Wuite,G.J.L. (2006) Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature*, 10.1038/nature05283.
80. Di Paolo,D., Afanzar,O., Armitage,J.P. and Berry,R.M. (2016) Single-molecule imaging of electroporated dye-labelled chey in live escherichia coli. *Philos. Trans. R. Soc. B Biol. Sci.*, **371**.
81. Nenninger,A., Mastroianni,G. and Mullineaux,C.W. (2010) Size dependence of protein diffusion in the cytoplasm of Escherichia coli. *J. Bacteriol.*, 10.1128/JB.00284-10.
82. Scott,M.S., Boisvert,F.M., McDowall,M.D., Lamond,A.I. and Barton,G.J. (2010) Characterization and prediction of protein nucleolar localization sequences. *Nucleic Acids Res.*, 10.1093/nar/gkq653.
83. Scott,M.S., Troshin,P. V. and Barton,G.J. (2011) NoD: A Nucleolar localization sequence detector for eukaryotic and viral proteins. *BMC Bioinformatics*, 10.1186/1471-2105-12-317.
84. Martin,R.M., Ter-Avetisyan,G., Herce,H.D., Ludwig,A.K., Lättig-Tünnemann,G. and Cardoso,M.C.

(2015) Principles of protein targeting to the nucleolus. *Nucleus*, 10.1080/19491034.2015.1079680.
85. Goedhart,J., Von Stetten,D., Noirclerc-Savoye,M., Lelimousin,M., Joosen,L., Hink,M.A., Van Weeren,L., Gadella,T.W.J. and Royant,A. (2012) Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. *Nat. Commun.*, 10.1038/ncomms1738.
86. Musinova,Y.R., Lisitsyna,O.M., Golyshev,S.A., Tuzhikov,A.I., Polyakov,V.Y. and Sheval,E. V. (2011) Nucleolar localization/retention signal is responsible for transient accumulation of histone H2B in the nucleolus through electrostatic interactions. *Biochim. Biophys. Acta - Mol. Cell Res.*, 10.1016/j.bbamcr.2010.11.003.
87. Kawakami,K. (2005) Transposon tools and methods in zebrafish. *Dev. Dyn.*, 10.1002/dvdy.20516.
88. Kawakami,K. (2007) Tol2: A versatile gene transfer vector in vertebrates. *Genome Biol.*, 10.1186/gb-2007-8-s1-s7.
89. Guzman,L.M., Belin,D., Carson,M.J. and Beckwith,J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose P(BAD) promoter. *J. Bacteriol.*, 10.1128/jb.177.14.4121-4130.1995.
90. Zhang,M., Chang,H., Zhang,Y., Yu,J., Wu,L., Ji,W., Chen,J., Liu,B., Lu,J., Liu,Y., *et al.* (2012) Rational design of true monomeric and bright photoactivatable fluorescent proteins. *Nat. Methods*, 10.1038/nmeth.2021.
91. Lau,I.F., Filipe,S.R., Søballe,B., Økstad,O.A., Barre,F.X. and Sherratt,D.J. (2003) Spatial and temporal organization of replicating Escherichia coli chromosomes. *Mol. Microbiol.*, 10.1046/j.1365-2958.2003.03640.x.
92. Bindels,D.S., Haarbosch,L., Van Weeren,L., Postma,M., Wiese,K.E., Mastop,M., Aumonier,S., Gotthard,G., Royant,A., Hink,M.A., *et al.* (2016) MScarlet: A bright monomeric red fluorescent protein for cellular imaging. *Nat. Methods*, 10.1038/nmeth.4074.
93. Chertkova,A.O., Mastop,M., Postma,M., Bommel,N. van, Niet,S. van der, Batenburg,K.L., Joosen,L., Gadella,T.W.J., Okada,Y. and Goedhart,J. (2017) Robust and Bright Genetically Encoded Fluorescent Markers for Highlighting Structures and Compartments in Mammalian Cells. *bioRxiv*, 10.1101/160374.

Supplementary Information

Supplementary methods:

Cloning HI-NESS overproduction vectors:

For protein overproduction, the HI-NESS gene encoding an N-terminal His-tag and a TEV-cleavable linker was cloned into a pET28a vector using Gibson Assembly (70). The plasmid was checked by Sanger sequencing and archived in a DH5α glycerol stock.

HI-NESS overproduction and purification

Plasmid pRD421 for the expression of N-terminally His-tagged HI-NESS with a TEV cleavable linker was transformed into BL21 (DE3) pLysS cells. 5 mL of an overnight culture of a single colony was used to inoculate 500 mL of LB medium. The culture was grown at 37 °C to an OD_{600} of 0.2 and thereafter, induced for 6 hours with 1 mM IPTG. The cells were pelleted by centrifugation and resuspended in 10 mL of low-salt binding buffer (130 mM NaCl, 20 mM Tris, 10% glycerol, 8 mM β-mercaptoethanol, and 3 mM benzamidine, pH 7.2) with 1 µg/mL DNase, 100 µg/mL lysozyme, and 1 mM PMSF. The cells were lysed by sonication and the soluble fraction was loaded on a 1 mL HisTrap™ HP column (Amersham Biosciences). The column was washed with 5 mL of high-salt binding buffer (500 mM NaCl, 20 mM Tris, 10% glycerol, 8 mM β-mercaptoethanol, and 3 mM benzamidine, pH 7.2) and, subsequently, His-tagged HI-NESS was eluted with 100% elution buffer (130 mM NaCl, 20 mM Tris, 10% glycerol, 8 mM β-mercaptoethanol, 3 mM benzamidine, and 500 mM Imidazole, pH 7.2). The peak fractions were exchanged back to low-salt binding buffer with a PD-10 Desalting column (Amersham Biosciences). The protein was cleaved for 48 hours at 4°C with 1 mg of TEV in the low-salt buffer supplemented with 10 mM DTT, and 2 mM EDTA to produce HI-NESS with a single glycine scar at the N-terminal. The cleavage reaction was loaded on a 1 mL HisTrap™ HP column (Amersham Biosciences) and washed with high-salt binding buffer to elute HI-NESS. HI-NESS was immediately exchanged to low-salt binding buffer with a PD-10 Desalting column (Amersham Biosciences). Protein concentration was determined using the Molar extinction co-efficient of mEos3.2 at 507 nm. 10 µL aliquots of the protein were flash-frozen in liquid nitrogen and stored at -80 °C until use.

Electrophoretic mobility shift assay

Purified HI-NESS was serially diluted in low-salt binding buffer (130 mM NaCl, 20 mM Tris, 10% glycerol, 8 mM β-mercaptoethanol, and 3 mM benzamidine, pH

7.2). The dilutions were mixed with an equivalent volume of 1 µM AT- or GC-rich DNA or RNA (Table S4.2) in nucleic acid buffer (10 mM Tris, 1 mM EDTA, and 50 mM NaCl, pH 7.5) and incubated for 20 minutes at 25 °C followed by 10 minutes at 4 °C. The samples were resolved on a polyacrylamide gel (Mini-Protean® TGX™ precast gels, 4-15%, Bio-Rad) at 30 V at 4 °C. Experiments were performed in triplicate. The polyacrylamide gels were stained for 45 minutes in a solution of 10X GelRed (Biotium) and thereafter imaged with GelDoc™ XR+ (Bio-Rad) using Bio-Rad's ImageLab software. The 'Plot profile' function in Fiji was used to quantify the unbound nucleic acid for each HI-NESS dilution. The percentage of bound nucleic acid was calculated therefrom. The dissociation constant (Kd) of HI-NESS for AT- and GC-rich DNA and RNA was extracted from a non-linear fit of the binding curves assuming specific binding with a Hill slope using the Prism8 software.

Supplementary figures:

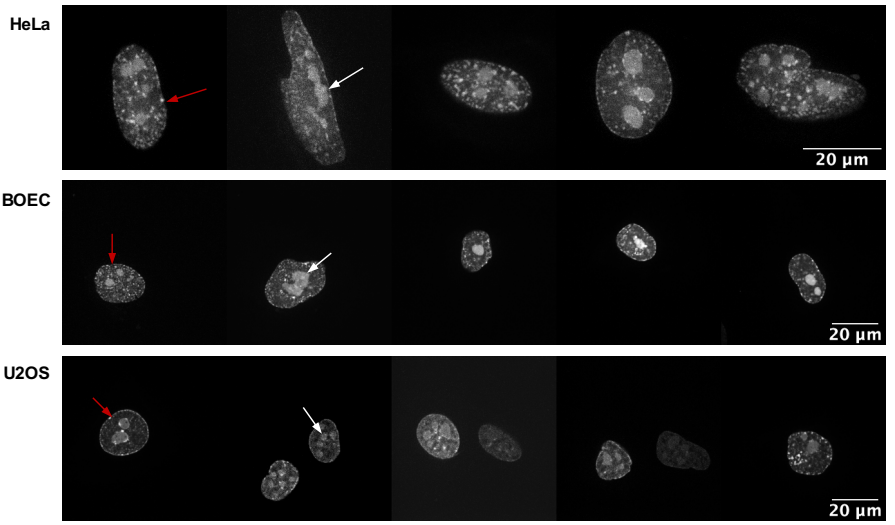


Figure S4.1: HI-NESS labels chromosomes in eukaryotic cells in culture. The HI-NESS signal appears as discrete and dense foci in nuclei (red arrows) and accumulates in nucleoli (white arrows).

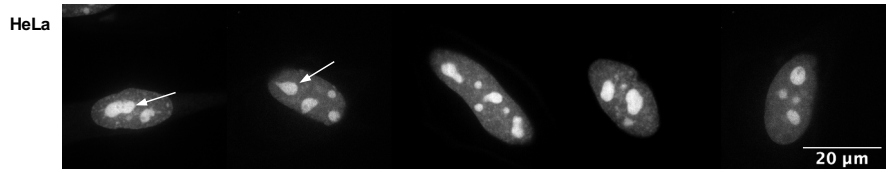


Figure S4.2: NLS-mEos3.2-NLS accumulates in the nucleoli. The mEos3.2 signal in HeLa cells expressing NLS-mEos3.2-NLS is observed primarily in the nucleoli (white arrows). The nuclei do not exhibit the distinct 'speckled' pattern that is observed when the chromosomes are stained with HI-NESS or SiR-Hoechst (Figure 4.2 and S4.1).

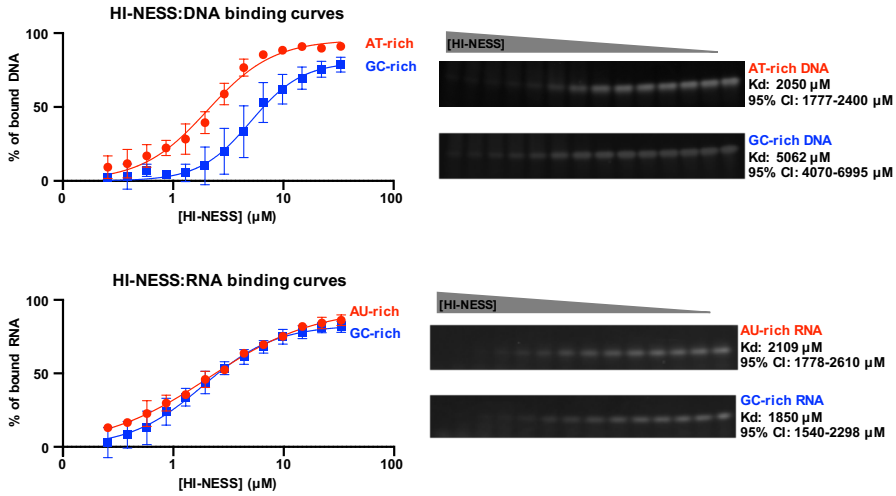


Figure S4.3: *In vitro* binding affinities of HI-NESS for DNA and RNA. The binding affinities of HI-NESS for AT/AU- and GC-rich DNA and RNA were estimated using an electrophoretic mobility shift assay (see supplementary methods). Representative results are shown on the right panel. **AT-rich DNA:** Kd = 2050 μM, 95% CI = 1777-2400 μM; **GC-rich DNA:** Kd = 5062 μM, 95% CI = 4070-6995 μM; **AU-rich RNA:** Kd = 2109 μM, 95% CI = 1778-2610 μM; **GC-rich RNA:** Kd = 1850 μM, 95% CI = 1540-2298 μM. HI-NESS displays a high affinity for RNA *in vitro*, however, the scarce nucleolar accumulation of HI-NESS in cells of zebrafish larvae (Figure S4.5; Movies S4.1-S4.5, S4.10) indicate that this affinity does not extrapolate *in vivo*.

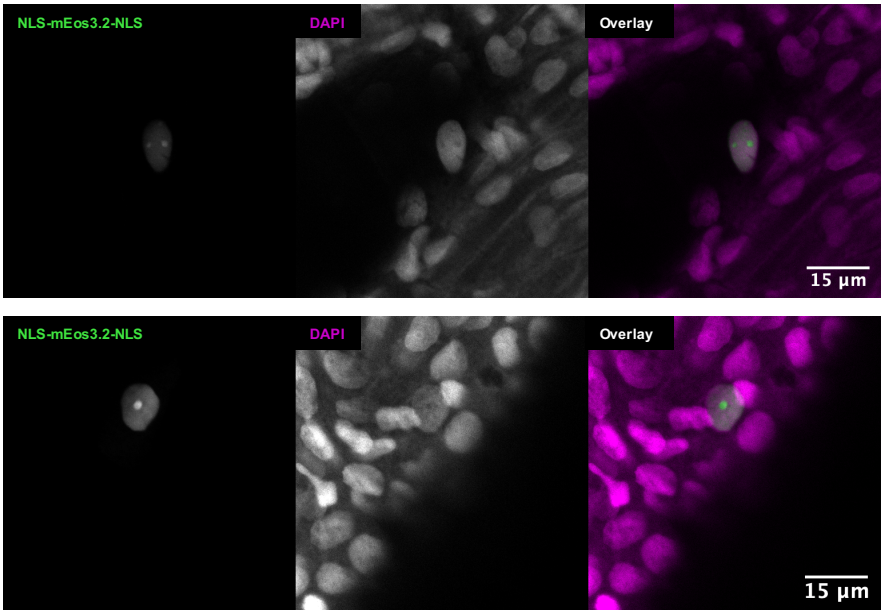


Figure S4.4: NLS-mEos3.2-NLS (green) does not recapitulate the distribution of DAPI (magenta) in nuclei of zebrafish larvae.

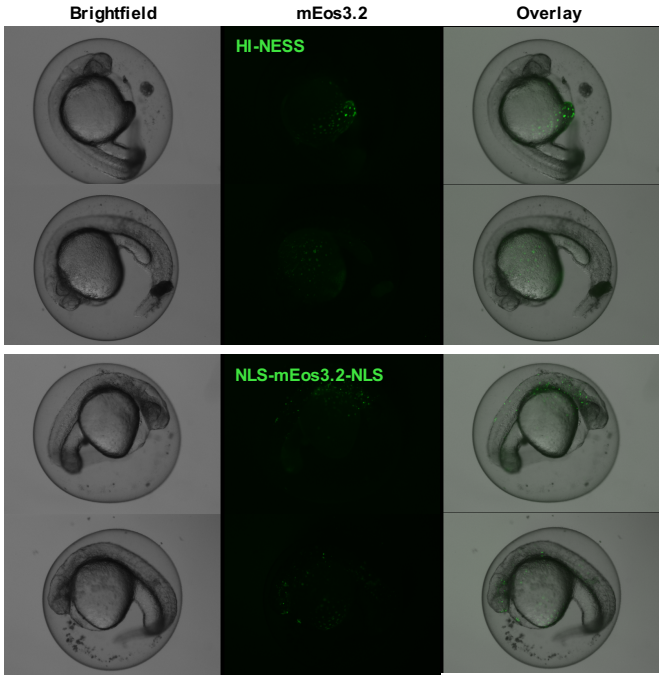


Figure S4.5: The mosaic expression of HI-NESS and NLS-mEos3.2-NLS in zebrafish larvae.

Supplementary tables:

Table S4.1:List of plasmids used in this study

Plasmid	Plasmid backbone	Insert	Resistance marker	Reference
pBAD33	pBAD33	N/A	Chloramphenicol	(89)
pcDNA3.1(+)	pcDNA3.1(+)	N/A	Ampicillin; Neomycin	Invitrogen
pRD128	pBAD33	mTurquoise2_H-NSdbd	Chloramphenicol	This chapter
pRD129	pBAD33	mEos3.2_H-NSdbd	Chloramphenicol	This chapter
pRD188	pcDNA3.1(+)	NLS_mEos3.2_H-NSdbd_NLS	Ampicillin; Neomycin	This chapter
pRD190	pcDNA3.1(+)	NLS_mEos3.2_NLS	Ampicillin; Neomycin	This chapter
pRD198*	pBAD33	mEGFP_H-NSdbd	Chloramphenicol	This chapter
pRD395*	pBAD33	mCherry_H-NSdbd	Chloramphenicol	This chapter
pRD396	pBAD33	eYFP_H-NSdbd	Chloramphenicol	This chapter
pRD397	pcDNA3.1(+)	NLS_mTurquoise2_H-NSdbd_NLS	Ampicillin; Neomycin	This chapter
pRD398*	pcDNA3.1(+)	NLS_mCherry_H-NSdbd_NLS	Ampicillin; Neomycin	This chapter
pRD399*	pcDNA3.1(+)	NLS_mEGFP_H-NSdbd_NLS	Ampicillin; Neomycin	This chapter
pRD400	pcDNA3.1(+)	NLS_eYFP_H-NSdbd_NLS	Ampicillin; Neomycin	This chapter
pRD403	pBAD33	mScarlet-I_H-NSdbd	Chloramphenicol	This chapter
pRD404	pBAD33	mNeonGreen_H-NSdbd	Chloramphenicol	This chapter
pRD405	pcDNA3.1(+)	NLS_mScarlet-I_H-NSdbd_NLS	Ampicillin; Neomycin	This chapter
pRD406	pcDNA3.1(+)	NLS_mNeonGreen_H-NSdbd_NLS	Ampicillin; Neomycin	This chapter
pRD421	pET28a	His-tag_TEV-cleavable linker_mEos3.2_H-NSdbd	Kanamycin	This chapter
pmTurquoise2_C1	pEGFP-C1	mTurquoise2	Kanamycin; Neomycin	(85)
mEos3.2-C1	mEos3.2-C1	mEos3.2	Kanamycin; Neomycin	Addgene #54550 (90)
pLau53	pBAD24	eYFP	Ampicillin	(91)
pmScarlet-i_C1	pC1	mScarlet-I	Kanamycin	From TWJG Addgene #85044 (92)
LifeAct-mNeonGreen	pEGFP-N1	mNeonGreen	Kanamycin	From TWJG Addgene #98877 (93)

* Sequences for mCherry, and mEGFP were prepared with PCR and Gibson assembly.

Table S4.2: List of nucleic acids (Sigma-Aldrich) used for the electrophoretic mobility shift assays.

Nucleic acid	Sequence
AT-rich dsDNA	5'-ATGGCAATTAAATTAGAAATTAAAAATCTTTATAAAATATTTGGC-3'
GC-rich dsDNA	5'-ATGGCAATCAAACTCGAGATCAAGAACCTCTACAAGATCTTCGGC-3'
AU-rich ssRNA	5'-AUGGCAAUUAAAUUAGAAAUUAAAAAUCUUUAUAAAAUAUUUGGC-3'
GC-rich ssRNA	5'-AUGGCAAUCAAACUCGAGAUCAAGAACCUCUACAAGAUCUUCGGC-3'

Supplementary movies:
Movies S4.1-S4.10 are available in the attached electronic supplement:
2021_Rashid_PhD_Thesis/Electronic_Supplement/Chapter 4/

Movies S4.1-S4.5: HI-NESS labels chromosomes in Zebrafish larvae. The distribution of the mEos3.2 signal (green) in nuclei of zebrafish larvae overlapped with that of DAPI (magenta). Nucleolar accumulation of HI-NESS was not observed. **S4.1:** Epithelial cells (tail fin); **S4.2:** Muscle cells (trunk); **S4.3:** Yolk syncytial cells (yolk syncytial layer); **S4.4:** From left to right: trunk, blood vessel, yolk extension; **S4.5:** Yolk syncytial cells (yolk syncytial layer).

Movies S4.6-S4.8: NLS-mEos3.2-NLS does not label the chromosomes in zebrafish larvae. In cells of zebrafish larvae expressing NLS-mEos3.2-NLS (green), the fluorophore accumulates in nucleoli and distributes uniformly over the non-nucleolar regions of the nucleus. The distribution does not correlate with that of DAPI (magenta). **S4.6:** Muscle cells (trunk); **S4.7:** Left: muscle cells (trunk), right: epithelial cells (skin); **S4.8:** Yolk syncytial cells (yolk syncytial layer).

Movie S4.9: HI-NESS can be used to follow chromosome dynamics in eukaryotic cells in culture. The time-lapse (in min:sec) shows cell division in a HeLa cell. From left to right: Phase contrast image, Lck-mTurquoise2 membrane label, HI-NESS (mScarlet-I).

Movie S4.10: HI-NESS (mScarlet-I) can be used to follow the movements of nuclei in a live animal model.

Chapter 5:
Outlook

Chromosome architecture is closely intertwined with gene transcription. In *Caulobacter crescentus* and *Bacillus subtilis*, the boundaries between self-interacting chromosome interaction domains (CIDs) are mediated by highly-transcribed genes (1, 2). In preliminary studies, the chromosome contact map of *Escherichia coli* treated with Rifampicin to limit transcription shows a loss of chromosome architecture (Figure 5.1). A similar change was also observed for the euryarchaeum *Haloferax volcanii* following treatment with the transcription inhibitory drug actinomycin D (3). In addition, the first chromosome contact maps generated for *Haloquadratum walsbyi*, a slow growing archaeal species that populates hypersaline lakes (4–6) and has low levels of transcription (7), show limited chromosome structure (Figure 5.2).

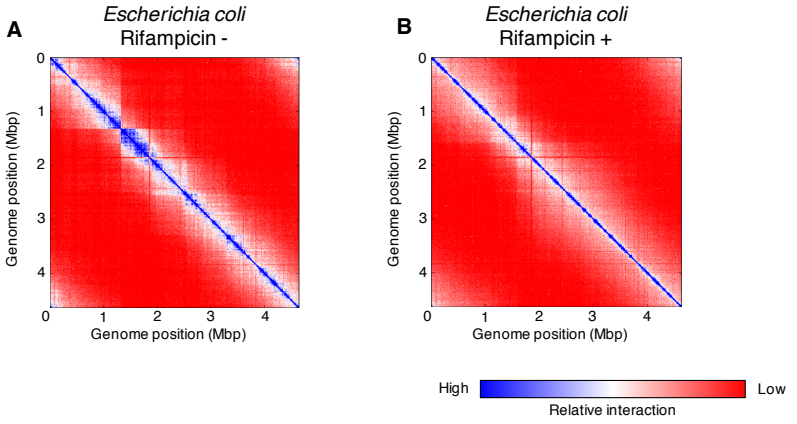


Figure 5.1: Global chromosome organisation in *Escherichia coli* is dependent on active transcription. The chromosome contact profiles of exponentially growing *E. coli* **(A)** show a loss of three-dimensional chromosome architecture upon treatment with Rifampicin **(B)** – a drug that inhibits transcription. **Organism:** *Escherichia coli* MG1655 Δ*endA* (NT331); **3C-based study:** Hi-C (8); **Resolution:** 10 kb; **Growth conditions: Panel A:** LB medium, 37 ºC, exponential phase, **Panel B:** LB medium, 37 ºC, exponential phase followed by a 2-hour long treatment with 100 µg/mL Rifampicin; **Fixation conditions:** 80% cold methanol for 10 minutes followed by 3% formaldehyde for 1 hour; **Restriction enzyme:** PsuI (ThermoFisher Scientific); **Fractionation:** Yes.

The work presented in this thesis builds on the observations of a global interplay between chromosome structure and transcription. It provides evidence of an interplay between local three-dimensional chromosome architecture and gene expression using the *proVWX* operon of *Escherichia coli* that encodes the ProU osmoprotectant transporter as a model system. At *proVWX*, a bridge between the σ^{70}-dependent promoter of the operon and the terminator is associated with a repressed operon. The disassembly of the bridge in response to hyperosmotic stress correlates with an active operon (Chapter 3, this thesis). *proVWX* encodes an ABC transporter (9). ABC transporters form part of the largest group of paralogous

operons/genes that evolved before the divide between archaea and bacteria (10, 11). Evidently, ABC transporters are prevalent in archaea, bacteria, and eukaryotes (11). The model of the interplay between the chromosome structure and gene expression presented for *proVWX* may, therefore, provide clues to potentially conserved regulatory features of several operons and genes across all domains of life. The prevalence of ABC transporters makes the operons and genes that encode these structures an ideal model system to study the evolution of the interplay between local chromatin structure and gene expression.

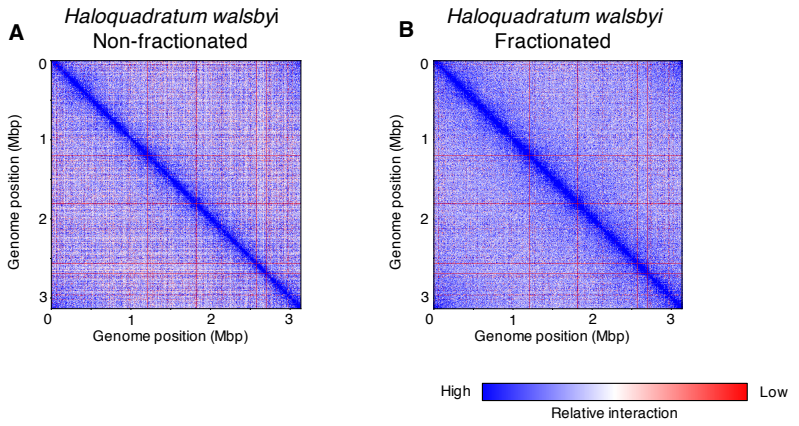


Figure 5.2: The chromosome contact profiles of *Haloquadratum walsbyi*, a slow-growing archaeon with reduced transcription shows limited chromosome architecture. Organism: *Haloquadratum walsbyi* HBSQ001; **3C-based study:** Hi-C; **Resolution:** 10 kb; **Growth conditions:** 37 ºC, stationary phase, standing culture; **Fixation conditions:** 5% formaldehyde, 25 ºC, 3 hours; **Restriction enzyme:** PsuI (ThermoFisher Scientific); **Fractionation: Panel A:** No, **Panel B:** Yes.

ProVWX orthologs are encoded in the genomes of *Salmonella enterica* serovars and *Shigella* species, the latter of which are evolved from *Escherichia* (12). The three genera share identity in the *proVWX* nucleotide sequence, encode H-NS, and exhibit an activation of *proVWX* in response to hyperosmotic stress (13–15). However, *Salmonellae*, in particular, differ in the expression of the operon. Due to limited studies, a similar statement cannot be made for *Shigella*, yet. Indeed, in *Salmonella enterica* serovar Typhimurium, relative expression decreases across the operon from *proV* to *proW* and to *proX* (16), in contrast to the transcriptional profile across *E. coli proVWX* (Chapter 3, this thesis). In *Salmonella enterica* serovar Typhi, a -1 frameshift in *proV* has rendered ProV non-functional (17), nevertheless, the compensation of the absence of ProV by its paralogues in Δ*proV* mutants of *Shigella sonnei* (14) hints that the transport of osmoprotectants by the ProWX in *Salmonella* Typhi may not be entirely abolished. *Salmonella* Typhi could,

therefore, carry a partially cryptic *proVWX* operon. The detailed transcriptional profile of *proVWX*, and the three-dimensional structure of the operon at various osmotic stress conditions in *Salmonella* and *Shigella* – as described in this thesis for *Escherichia coli* – remains to be investigated. Such a study is expected to provide a starting point to highlight principles of H-NS-mediated structural regulation of osmosensitive operons. The ATP-binding cassette (the ATPase) is the most evolutionarily conserved of the ATPase, transmembrane transporter, and substrate binding subunits of ABC transporters (18). A bioinformatic survey of the cross-domain distribution of the ProV ATPase can be used to assemble a phylogenetic tree of the distribution of ProV and ProVWX transporter orthologs. Such a survey will reveal targets to systematically investigate the structural regulation of *proVWX* in organisms that may or may not carry H-NS-like proteins – the integral regulatory NAP of the operon (19–22). The study may lead to the discovery of other architectural transcription regulators that directly detect and respond to osmotic stress.

Salmonella enterica encodes three osmoprotectant transporters: ProP (23), ProU (15), and OsmU (24). ProP is a proton symporter of the major facilitator superfamily of permeases (25). ProU and OsmU are ABC transporters (9, 24). In contrast to *proVWX*, the operon encoding OsmU is arranged as *osmYXWV*, where *osmV* encodes the ATPase, *osmW* and *osmY* encode the polypeptides for the heterodimeric transmembrane transporter, and *osmX* encodes the periplasmic substrate binding protein (24). A functional *osmYXWV* system allows *Salmonella* to survive hyperosmotic stress in a *proP- proVWX-* background, following a lag phase that may last up to 12 hours (24). The *osmYXWV* operon triggers interest owing to its conservation in several species of the *Enterobacteriaceae* family at a nucleotide level and members of the phylum Euryarchaeota at a translational level (24). Furthermore, *Chromobacterium violaceum* of the *Neisseriaceae* also carries an *osmYXWV* ortholog that shares nucleotide sequence similarity with the *osmYXWV* of *Enterobacteriaceae* (24). However, the operon occurs in a chromosomal region that shares no similarity with the genetic neighbourhood of *osmYXWV* in *Enterobacteriaceae*, indicating that the operon may have been acquired by horizontal transfer in *C. violaceum* (24). A comparison of the transcriptional (RNA-Sequencing, and RT-qPCR), structural (3C-based studies), and NAP binding (Chromatin immunoprecipitation) profiles of archaeal and bacterial *osmYXWV* operons will reveal conserved features of the structural regulation of genes/operons during osmotic stress response.

The interdependency of chromosome structure and gene expression may play an evolutionary role by providing a means of 'sampling' repressed regions of the chromosome during replication. The progress of the replication machinery over the chromosome requires local dismantling of the template chromatin and its reassembly on replicated DNA. This may give rise to a time frame during which the repressed promoter of a cryptic gene or that of a horizontally transferred gene is left unattended. This time frame may be affected by stress, for instance, during the stringent response – a condition that affects replication, transcription, and metabolism (Reviewed in (26)). The binding of RNA Polymerase to such a promoter, and the expression of the repressed unit allows the cell to 'sample' cryptic or horizontally transferred genes. Genes that provide an advantageous characteristic may undergo gradual de-silencing and become incorporated into the transcriptome and proteome. A cross-domain, multi-species study of active and cryptic ABC operons and genes will provide clues to the validity of this hypothesis.

The *Pyrococci* of the Euryarchaeota encode cellobiose/β-glucoside ABC transporters (27–30). The gene cluster is upregulated in *Pyrococcus furiosus* growing in the presence of cellobiose in a medium with limited carbohydrate availability (30). However, in *P. abyssi* and *P. horikoshii* the cellobiose/β-glucoside ABC transporter cluster is not activated under similar conditions, and evidently, the cells are unable to grow in carbohydrate limiting conditions with cellobiose (30). A comparison of the nucleotide sequences, NAP binding profiles, and local three-dimensional architecture of active, cryptic, encrypted (active gene clusters inactivated by mutagenesis), and decrypted (inactive gene clusters activated by mutagenesis) cellobiose/β-glucoside ABC transporter clusters of the *Pyrococci* will enable the investigation of 'gene sampling' in comparable genetic backgrounds. A phylogenetic study of the distribution of orthologous cellobiose/β-glucoside uptake systems and the identification of cryptic and active clusters will allow a step-wise investigation of the role of local chromosome structure in evolution.

In addition to the silencing of transcriptional units mediated by local repressive chromatin, the higher order structure of chromatin assembled from the concerted binding of multiple NAPs may play an additional evolutionary role. NAPs are environmental sensors. The proteins detect and respond to changes in osmolarity, temperature, and pH by modifying their architectural properties, effectively altering three-dimensional chromosome architecture and, hence, regulating genome transactions (31–33). The binding of a diversity of NAPs along the genome and their combined function enables each segment of the chromosome to

respond differently to environmental stimuli. Therefore, the higher-order chromatin structure in the neighbourhood of a transcriptional unit may place a bias on its evolution. For instance, genes encoding elements that alleviate osmotic stress may preferentially be activated when positioned within osmoresponsive chromatin, but may remain cryptic when positioned at pH-sensitive sites. A database of paralogs and orthologs of osmoprotectant ABC transporter operons/genes such as the *proVWX* and *osmYXWV* in which the entries are sorted according to their cryptic/active state, and predicted chromatin environment deduced from NAP binding profiles and transcriptional profiles, will enable validation of this hypothesis.

Acknowledgements:
Haloquadratum walsbyi HBSQ001 was a gift from Dr. Henk Bolhuis (Microbiology & Biogeochemistry, NIOZ, The Netherlands)

References:

1. Le,T.B.K., Imakaev,M. V, Mirny,L.A. and Laub,M.T. (2013) High-resolution mapping of the spatial organization of a bacterial chromosome. *Science*, **342**, 731–4.
2. Condensin- and Replication-Mediated Bacterial Chromosome Folding and Origin Condensation Revealed by Hi-C and Super-resolution Imaging. *Mol. Cell*.
3. Cockram,C., Thierry,A., Gorlas,A., Lestini,R. and Koszul,R. (2020) Euryarchaeal genomes are folded into SMC-dependent loops and domains, but lack transcription-mediated compartmentalization. *Mol. Cell*, 10.1016/j.molcel.2020.12.013.
4. Antón,J., Llobet-Brossa,E., Rodríguez-Valera,F. and Amann,R. (1999) Fluorescence in situ hybridization analysis of the prokaryotic community inhabiting crystallizer ponds. *Environ. Microbiol.*, 10.1046/j.1462-2920.1999.00065.x.
5. Benlloch,S., López-López,A., Casamayor,E.O., Øvreås,L., Goddard,V., Daae,F.L., Smerdon,G., Massana,R., Joint,I., Thingstad,F., *et al.* (2002) Prokaryotic genetic diversity throughout the salinity gradient of a coastal solar saltern. *Environ. Microbiol.*, 10.1046/j.1462-2920.2002.00306.x.
6. Walsby,A.E. (1980) A square bacterium [16]. *Nature*, 10.1038/283069a0.
7. Bolhuis,H., Martín-Cuadrado,A.B., Rosselli,R., Pašić,L. and Rodriguez-Valera,F. (2017) Transcriptome analysis of Haloquadratum walsbyi: Vanity is but the surface. *BMC Genomics*, 10.1186/s12864-017-3892-2.
8. Rashid,F.Z.M. (2021) Chapter 2: Hi-C in bacteria and archaea.
9. Gowrishankar,J. (1989) Nucleotide sequence of the osmoregulatory proU operon of Escherichia coli. *J. Bacteriol.*, 10.1128/jb.171.4.1923-1931.1989.
10. Tomii,K. and Kanehisa,M. (1998) A comparative analysis of ABC transporters in complete microbial genomes. *Genome Res.*, **8**, 1048–1059.
11. Higgins,C.F. (1992) ABC Transporters: From microorganisms to man. *Annu. Rev. Cell Biol.*, **8**, 67–113.
12. Peng,J., Zhang,X., Yang,J., Wang,J., Yang,E., Bin,W., Wei,C., Sun,M. and Jin,Q. (2006) The use of comparative genomic hybridization to characterize genome dynamics and diversity among the serotypes of Shigella. *BMC Genomics*, 10.1186/1471-2164-7-218.
13. Gowrishankar,J. and Manna,D. (1996) How is osmotic regulation of transcription of the Escherichia coli proU operon achieved? A review and a model. *Genetica*, 10.1007/BF00055322.
14. Mahmoud,R.Y., Li,W., Eldomany,R.A., Emara,M. and Yu,J. (2017) The Shigella ProU system is required for osmotic tolerance and virulence. *Virulence*, 10.1080/21505594.2016.1227906.
15. Cairney,J., Booth,I.R. and Higgins,C.F. (1985) Osmoregulation of gene expression in Salmonella typhimurium: proU encodes an osmotically induced betaine transport system. *J. Bacteriol.*, 10.1128/jb.164.3.1224-1232.1985.
16. Maserati,A., Fink,R.C., Lourenco,A., Julius,M.L. and Diez-Gonzalez,F. (2017) General response of Salmonella enterica serovar Typhimurium to desiccation: A new role for the virulence factors sopD and sseD in survival. *PLoS One*, 10.1371/journal.pone.0187692.
17. Parkhill,J., Dougan,G., James,K.D., Thomson,N.R., Pickard,D., Wain,J., Churcher,C., Mungall,K.L., Bentley,S.D., Holden,M.T.G., *et al.* (2001) Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18. *Nature*, 10.1038/35101607.
18. Davidson,A.L., Dassa,E., Orelle,C. and Chen,J. (2008) Structure, Function, and Evolution of Bacterial ATP-Binding Cassette Systems. *Microbiol. Mol. Biol. Rev.*, 10.1128/mmbr.00031-07.
19. Lucht,J.M., Dersch,P., Kempf,B. and Bremer,E. (1994) Interactions of the nucleoid-associated DNA-binding protein H-NS with the regulatory region of the osmotically controlled proU operon of Escherichia coli. *J. Biol. Chem.*
20. Rajkumari,K., Kusano,S., Ishihama,A., Mizuno,T. and Gowrishankar,J. (1996) Effects of H-NS and potassium glutamate on σ(S)- and σ70-directed transcription in vitro from osmotically regulated P1 and P2 promoters of proU in Escherichia coli. *J. Bacteriol.*, 10.1128/jb.178.14.4176-4181.1996.
21. Nagarajavel,V., Madhusudan,S., Dole,S., Rahmouni,A.R. and Schnetz,K. (2007) Repression by binding of H-NS within the transcription unit. *J. Biol. Chem.*, 10.1074/jbc.M702753200.
22. 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.*, 10.1038/nsmb1233.

23. Cairney,J., Booth,I.R. and Higgins,C.F. (1985) Salmonella typhimurium proP gene encodes a transport system for the osmoprotectant betaine. *J. Bacteriol.*, 10.1128/jb.164.3.1218-1223.1985.
24. Frossard,S.M., Khan,A.A., Warrick,E.C., Gately,J.M., Hanson,A.D., Oldham,M.L., Sanders,D.A. and Csonka,L.N. (2012) Identification of a third osmoprotectant transport system, the osmU system, in Salmonella enterica. *J. Bacteriol.*, 10.1128/JB.00495-12.
25. MacMillan,S. V., Alexander,D.A., Culham,D.E., Kunte,H.J., Marshall,E. V., Rochon,D. and Wood,J.M. (1999) The ion coupling and organic substrate specificities of osmoregulatory transporter ProP in Escherichia coli. *Biochim. Biophys. Acta - Biomembr.*, 10.1016/S0005-2736(99)00085-1.
26. Irving,S.E., Choudhury,N.R. and Corrigan,R.M. (2020) The stringent response and physiological roles of (pp)pGpp in bacteria. *Nat. Rev. Microbiol.*, 10.1038/s41579-020-00470-y.
27. Kawarabayasi,Y., Sawada,M., Horikawa,H., Haikawa,Y., Hino,Y., Yamamoto,S., Sekine,M., Baba,S.I., Kosugi,H., Hosoyama,A., *et al.* (1998) Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, Pyrococcus horikoshii OT3. *DNA Res.*, 10.1093/dnares/5.2.55.
28. Robb,F.T., Maeder,D.L., Brown,J.R., DiRuggiero,J., Stump,M.D., Yeh,R.K., Weiss,R.B. and Dunn,D.M. (2001) Genomic sequence of hyperthermophile, Pyrococcus furiosus: Implications for physiology and enzymology. *Methods Enzymol.*, 10.1016/S0076-6879(01)30372-5.
29. Albers,S.V., Koning,S.M., Konings,W.N. and Driessen,A.J.M. (2004) Insights into ABC Transport in Archaea. *J. Bioenerg. Biomembr.*, 10.1023/B:JOBB.0000019593.84933.e6.
30. Koning,S.M., Elferink,M.G.L., Konings,W.N. and Driessen,A.J.M. (2001) Cellobiose uptake in the hyperthermophilic archaeon Pyrococcus furiosus is mediated by an inducible, high-affinity ABC transporter. *J. Bacteriol.*, 10.1128/JB.183.17.4979-4984.2001.
31. Dame,R.T. (2005) The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol. Microbiol.*, 10.1111/j.1365-2958.2005.04598.x.
32. Dame,R.T., Kalmykowa,O.J. and Grainger,D.C. (2011) Chromosomal macrodomains and associated proteins: Implications for DNA organization and replication in gram negative bacteria. *PLoS Genet.*, 10.1371/journal.pgen.1002123.
33. Dame,R.T., Rashid,F.Z.M. and Grainger,D.C. (2020) Chromosome organization in bacteria: mechanistic insights into genome structure and function. *Nat. Rev. Genet.*, **21**, 227–242.

Summary

When an *Escherichia coli* cell pellet is lysed in a final volume up to 50 times that of the pellet, it produces a solution so viscous that jelly-like droplets can be picked with a pair of tweezers. It is fascinating to envision that this viscous material, concentrated by a factor of 50, is encapsulated within a membrane, and that this capsule is alive. The blueprint of its life is contained within a circular genome 1.5 mm in length that is folded by nucleoid associated proteins (NAPs) and packaged into a phase-separated structure with the cell. The cell faces a delicate challenge: the blueprint of its life must remain compactly packaged to facilitate its segregation into daughter cells prior to cell division, and, at the same time, every locus must remain accessible for the cell to rapidly respond to favourable and stressful environmental stimuli. The NAPs that address this challenge must, therefore, have an architectural property to compact the genome, and be environmentally-sensitive to facilitate genomic response to stimuli. Advances in fluorescence microscopy, chromosome conformation capture (3C)-based techniques, and *in vitro* genome-in-a-box approaches have begun to shed light on the principles of chromosome organisation and dynamics. In this thesis, single-cell fluorescence microscopy, ensemble reverse transcriptase quantitative PCR (RT-qPCR) and 3C-qPCR, and single-molecule Förster resonance energy transfer (smFRET) have been used to provide insights into the intriguing interplay between three-dimensional chromatin structure and gene expression. An experimentally-supported model that consolidates the architectural properties and regulatory functions of NAPs has also been described.

An up-to-date review on the interplay between chromosome organisation and genomic transactions including replication, transcription, and segregation in bacteria has been provided in **Chapter 1**. Chromosome folding and compaction at a molecular level mediated by NAPs, and the modulation of the structure as a result of cross-talk between NAPs, and proteins that regulate NAP function has been described. The hierarchical organisation of bacterial chromosomes has also been addressed in **Chapter 1** and an overview of the impact of architectural NAPs and genome transactions on local and global patterns of chromosome organisation, and how this may be affected by environmental conditions has been provided.

The three-dimensional structure of the prokaryotic chromosome is encoded in its sequence as binding sites of varying strength for the diversity of NAPs expressed by the cell, and as the positions of transcriptional units and associated regulatory elements. This compels that chromosome architecture be studied in the context of

genomic sequence. Hi-C enables it. Hi-C is a chromosome conformation capture based technique that combines proximity ligation with next generation sequencing to detect the physical proximity of chromosomal loci in three-dimensional space. A detailed protocol for Hi-C in prokaryotes has been provided in **Chapter 2**.

In **Chapter 3** the H-NS-regulated, osmosensitive *proVWX* operon of *Escherichia coli* has been used as a model system to demonstrate, for the first time in prokaryotes, that the expression of a transcriptional unit is correlated with its three-dimensional architecture. RT-qPCR was used to report on the transcriptional profile of *proVWX* and verify the roles of H-NS and StpA in regulating the expression of the operon. 3C-qPCR, and live cell FROS were used to show the formation of a loop anchored between the promoter and terminator of *proVWX* under conditions at which the operon is repressed, and the destabilisation of the loop at osmolarity conditions that activate *proVWX*. The results that have been presented in **Chapter 3** also reveal that StpA, often considered a molecular back-up for H-NS, plays physiological roles in a wild-type cell that are distinct from H-NS.

The dynamic organisation of the chromosome and its association with genome processes creates a need for visualizing the genome in a non-perturbing manner. In **Chapter 4**, HI-NESS (H-NS-based indicator for nucleic acid stainings) – a minimally-perturbing DNA labelling dye designed by translationally fusing the DNA binding domain of H-NS to a fluorescent protein – has been described. HI-NESS labels the chromosome in *Escherichia coli* cells that lack H-NS, eukaryotic cells in culture, and in zebrafish embryos.

How the work reported in this thesis forms a foundation for investigating chromosome organisation and dynamics in other bacteria, and in archaea – a domain that is generally regarded as the missing link between bacteria and eukaryotes – has been discussed in **Chapter 5**. In the chapter, a model of how the interplay between chromosome structure and genome transactions may have driven evolution has been proposed.

Samenvatting

Wanneer een *Escherichia coli* celpellet gelyseerd wordt in een eindvolume dat 50 keer kleiner is dan het pellet zelf, is het resultaat een oplossing die zo viskeus is dat er met een pincet geleiachtige druppels uit kunnen worden gehaald. Het is fascinerend om te bedenken dat dit viskeuze materiaal, dat 50 keer geconcentreerd is, ingekapseld door een membraan een cel vormt die leeft. De blauwdruk van het leven van deze cel staat geschreven in het circulaire genoom dat een diameter heeft van ongeveer 1.5 mm. Het genoom wordt gevouwen door "nucleoid associated proteins" (NAPs) en bestaat in de vorm van een fase-gescheiden structuur in de cel. De cel staat voor een grote uitdaging: de blauwdruk van zijn leven moet opeengepakt blijven om voor de celdeling verdeling over dochtercellen mogelijk te maken, terwijl tegelijkertijd individuele genen toegankelijk moeten blijven voor de cel om snel te kunnen reageren op zowel gunstige als ongunstige omgevingsfactoren. De NAPs die hier zorg voor dragen moeten daarom beschikken over architecturele eigenschappen om het genoom compact te maken én moeten gevoelig zijn voor signalen uit de omgeving om een respons vanuit het genoom mogelijk te maken. Door vooruitgangen in fluorescentiemicroscopie, "chromosome conformation capture" (3C) gebaseerde technieken en *in vitro* "genome-in-a-box" benaderingen is het mogelijk geworden een begin te maken om de principes van chromosoomorganisatie en -dynamiek te ontrafelen. In dit proefschrift worden "single-cell" fluorescentiemicroscopie, "ensemble reverse transcriptase quantitative PCR" (RT-qPCR) in combinatie met 3C-qPCR en "single-molecule Förster resonance energy transfer" (smFRET) gebruikt om inzicht te krijgen in de fascinerende wisselwerking tussen driedimensionale chromatinestructuur en genexpressie. Op basis van de observaties met deze technieken wordt een genregulatiemodel voorgesteld waarin de architecturele eigenschappen en regulatoire functies van NAPs samen komen.

Een actueel overzicht van de wisselwerking tussen chromosoomorganisatie en genoomtransacties, inclusief replicatie, transcriptie en segregatie in bacteriën wordt gegeven in **Hoofdstuk 1**. Chromosoomvouwing en -compactie op moleculair niveau ten gevolge van de binding van NAPs wordt beschreven, evenals de wisselwerking tussen NAPs en de eiwitten die NAPs reguleren waardoor chromatinestructuur wordt gemoduleerd. De hiërarchische organisatie van bacteriële chromosomen wordt ook behandeld in **Hoofdstuk 1**. Tenslotte wordt er een overzicht gegeven van de invloed van NAPs met architecturele eigenschappen en genoomtransacties op lokale en globale patronen van chromosoomorganisatie, en hoe dit beïnvloed kan worden door omgevingsfactoren

De driedimensionale structuur van het prokaryote chromosoom is gecodeerd in zijn sequentie als bindingsplaatsen met verschillende sterkte voor de verschillende NAPs die tot expressie komen in de cel, en in de vorm van de locatie van transcriptie-eenheden en geassocieerde regulerende elementen. Dit vraagt om het bestuderen van chromosoomarchitectuur in relatie tot genoomsequentie. Hi-C maakt dit mogelijk. Hi-C is een “chromosome conformation capture” gebaseerde techniek die op nabijheid gebaseerde ligatie combineert met “next generation sequencing” om de fysieke nabijheid van chromosomale loci in de driedimensionale ruimte te detecteren. Een gedetailleerd protocol voor Hi-C in prokaryoten wordt gegeven in **Hoofdstuk 2**.

In **Hoofdstuk 3** wordt het door H-NS gereguleerde osmosensitieve operon *proVWX* van *Escherichia coli* gebruikt als modelsysteem en wordt - voor het eerst in bacteriën - aangetoond dat de activiteit van een transcriptie-eenheid gecorreleerd is met driedimensionale genoomarchitectuur. RT-qPCR werd gebruikt om het transcriptieprofiel van het *proVWX* operon in kaart te brengen. Specifiek, werd de rol van H-NS en StpA in de regulatie van expressie van het operon bestudeerd. 3C-qPCR en “live cell” FROS werden gebruikt om de vorming van een lus tussen de promoter en de terminator van *proVWX* aan te tonen onder condities waarbij het operon niet actief is, dan wel de destabilisatie van de lus bij osmolaire condities die *proVWX* activeren. De resultaten gepresenteerd in **Hoofdstuk 3** laten ook zien dat StpA, dat vaak wordt gezien als moleculaire “back-up” van H-NS, eigen fysiologische rollen heeft in wildtype cellen naast H-NS.

De dynamiek in chromosoomorganisatie en de associatie met genomische processen creëren een noodzaak om het genoom op een niet-invasieve, natuurlijke manier in beeld te brengen. In **Hoofdstuk 4** wordt HI-NESS (H-NS-based indicator for nucleic acid stainings) beschreven. HI-NESS is een minimaal invasieve kleurstof die DNA aankleurt en is gecreëerd door het DNA bindende domein van H-NS te fuseren met een fluorescent eiwit. HI-NESS labelt het chromosoom in *Escherichia coli* cellen die H-NS missen, eukaryote cellen in cultuur en in embryos van zebravissen.

Hoe het gerapporteerde werk in dit proefschrift een basis vormt voor het onderzoeken van chromosoomorganisatie en dynamiek in bacteriën en archaea – een domein dat algemeen wordt gezien als de ontbrekende schakel tussen bacteriën en eukaryoten – wordt bediscussieerd in **Hoofdstuk 5**. In dit hoofdstuk

wordt een model voorgesteld waarin de wisselwerking tussen chromosoomstructuur en genoomtransacties een drijvende kracht in evolutie is.