Charles J. Dorman Nature Reviews Microbiology11,349–355(2013)
doi:10.1038/nrmicro3007

OPINION

Genome architecture and global gene regulation in bacteria: making progress towards a unified model?

Charles J. Dorman

Abstract | Data obtained with advanced imaging techniques, chromosome conformation capture methods, bioinformatics and molecular genetics, together with insights from polymer physics and mechanobiology, are helping to refine our understanding of the spatiotemporal organization of the bacterial nucleoid and its gene expression programmes. Here, I discuss the proposal that, in addition to DNA topology and nucleoid-associated proteins, gene regulation is an important organizing principle of nucleoid architecture.

The lack of a nuclear membrane is one of the defining features of prokaryotes. Instead of a nucleus, prokaryotes have a nucleoid, which is a macromolecular complex where the genetic material and its associated molecules are located. Although the bacterial nucleoid was first described more than 50 years ago[1], detailed information about its structure did not become available until recently, with significant progress being made in the past 3 years[2–10]. Advances in our understanding of nucleoid structure have relied historically on investigations by electron microscopy accompanied with biochemical and genetic studies. Powerful imaging technologies and physical analytical methods that were initially developed for the study of eukaryotic cells have recently been applied in bacterial research, and this has accelerated the rate of discovery of nucleoid structural detail.

It has been known for some time that the circular chromosome within the nucleoid is subjected to the compaction imposed by a combination of factors, including molecular crowding[11], DNA polymer dynamics[6], supercoiling of the DNA[12] and the interaction of the DNA with other molecules such as nucleoid-associated proteins (NAPs)[13]. In addition to underpinning nucleoid architecture, NAPs and DNA supercoiling also influence gene expression. Indeed, many NAPs were studied initially from the standpoint of gene regulation, and the impact of DNA supercoiling on transcription has been recognized for decades[14–16]. There is an intuitive appeal to the use of nucleoid-structuring features such as NAPs and DNA topology to influence gene expression, because this would integrate the process of gene expression with the very cellular structure within which it takes place. In this Opinion article, I discuss how our appreciation of the relationship between the nucleoid and gene expression has been deepened by recent findings, which suggest that the processes involved in efficient gene regulation themselves represent a nucleoid-structuring principle.

Nucleoid structure and superstructure

The chromosome of the model Gram-negative bacterium *Escherichia coli* is organized at a nanometre-scale structural level and at a micrometre-scale superstructural level[17]. When considering nucleoid organization, it is helpful to consider the dimensions of the container in which the nucleoid is found. During exponential growth, the *E. coli* cell measures approximately 2 µm in length by 1 µm in diameter. To put this in context, the *lac* operon, consisting of the three structural genes *lacZ*, *lacY* and *lacA* (encoding β-galactosidase, lactose permease and galactoside *O*-acetyltransferase, respectively), measures about 1.7 µm in length based on the number of nucleotides. The single, circular 4,639 kb chromosome of *E. coli* K-12, which was originally estimated to contain 4,288 protein-coding genes[18], is itself about 1.5 mm in circumference and, if opened fully, would have a diameter of 0.5 mm, which is 500 times the diameter of an *E. coli* cell. These dimensions illustrate the extent of the packaging problem associated with the bacterial nucleoid, a problem that is compounded by the need to arrange the DNA in ways that make it not just compact but also readily available for replication, segregation, gene expression and gene regulation.

Landmark genetic experiments published in 2004 by Boccard and colleagues revealed that the *E. coli* chromosome is composed of six distinguishable zones: four macrodomains (Ori, Left, Right and Ter) and two additional non-structured regions (NS-left and NS-right)[19] (FIG. 1a). This organization imposes certain restrictions on the permitted rearrangements to the linear-order sequence of the chromosome, perhaps even limiting its potential for further evolution in some respects[20].

The Ori macrodomain consists of the region around the origin of bidirectional chromosomal replication, *oriC*, and the Ter macrodomain is located at the opposite pole of the chromosome and contains the terminus of DNA replication[19,21]. The Ter macrodomain also contains the *dif* site, which is crucial for the resolution of chromosome dimers by the site-specific tyrosine recombinases XerC and XerD[22]. The four other regions make up the bulk of the left and right replichores, the two 'arms' of the chromosome along which DNA polymerase moves during bidirectional chromosome replication[19] (FIG. 1a). In *E. coli*, *oriC* is positioned at mid-cell, with the replichores to the left and right of this point and aligned with the long axis of the bacterium. This arrangement is dependent on MukBEF, which fulfils the structural maintenance of chromosomes (SMC) condensin function in *E. coli*[23]. The Ter macrodomain moves from cell pole to mid-cell in newborn cells and is maintained there through an interaction

between the Ter-binding protein MatP and the ZapB component of the cell division apparatus[24].

The DNA-binding protein MatP has a special relationship with the Ter macrodomain. In contrast to the widespread distribution of SeqA (a negative regulator of replication initiation) owing to the presence of binding sites around much of the chromosome, MatP binds to a sequence motif (*matS*) that is found uniquely within Ter, a macrodomain from which SeqA is excluded[24]. MatP has an important role in timing the separation of daughter chromosomes during cell division: it can hold two copies of the Ter macrodomain together, preventing premature chromosome segregation[25]. MatP also plays a part in the formation of loops in the Ter macrodomain, by holding copies of the *matS* sequence together and thus providing a degree of compaction to this macrodomain[26,27].

There is an interesting distribution of important cell cycle-associated DNA-binding proteins among the macrodomains, including SeqA, the nucleoid occlusion protein SlmA and MatP, and these proteins might contribute to macrodomain organization[21]. The intensively studied protein SeqA is involved in the timing of chromosome replication initiation and binds to hemimethylated 5′-GATC-3′ sites[28–30]. These sites are abundant at *oriC* but are also found elsewhere on the chromosome, although not in Ter[21,31,32].

SlmA has an important role in coordinating chromosome positioning within the dividing cell; it too has a distribution that corresponds with the macrodomain structure of the chromosome. It binds to a specific DNA sequence that is absent from the Ter macrodomain and infrequently found in the Left and Right domains. Thus, SlmA is mainly located in the Ori macrodomain and the flanking unstructured domains[21,33]. MatP, SeqA and SlmA differ from the NAPs in that they bind either exclusively inside (MatP) or exclusively outside (SeqA and SlmA) the Ter macrodomain; most NAPs do not display such strong macrodomain specificity in their binding preferences[13]. This might imply that the NAPs are involved in chromosome organization at a different level to the macrodomain-specific proteins. This conjecture is supported by several lines of evidence that link at least some NAPs to the maintenance of nucleoid structure at the microdomain level. Also, MatP and SlmA have not (so far) been reported to influence global gene expression, whereas many NAPs have.

Superimposed on its macrodomain structure is the organization of the chromosome at the level of looped DNA microdomains[34]. In the earliest studies of nucleoid structure, the number and size of the looped domains were under- and overestimated, respectively[35,36]. Much more accurate estimates were subsequently obtained using genetic methods that allowed the impact of looped-domain breakage on DNA supercoiling to be measured around the chromosome[37]. These data were combined with assessments of the influence of topological barriers at microdomain boundaries on site-specific recombination efficiency[38] and also a systematic examination of looped domains in electron microscopy images of chromosomes released from lysed cells[37]. It now seems that the chromosome of *E. coli* is divided into approximately 400 looped microdomains, each with an average circumference of 10–12 kb[37,38] (FIG. 1b). Microdomains might be both transient and predominantly a feature of chromosomes in exponentially growing bacteria[38]. Nucleoid structure seems to be more diffuse in slow-growing bacteria[10].

Several lines of evidence from physical and genetic studies indicate that NAPs, especially H-NS (histone-like, nucleoid-structuring protein) and Fis (factor-for-inversion stimulation), play a part in forming the boundaries of microdomains, where they act as insulators, or domainins[32]. Further support for H-NS as a domainin has come from super-resolution imaging combined

Glossary

Chromatin immunoprecipitation followed by microarray
(ChIP–chip). A method that allows the binding sites for a specific protein to be identified throughout a genome *in vivo*. The protein of interest is crosslinked to DNA in living bacteria with formaldehyde, and the genomic DNA is extracted and then sonicated to achieve a desired average DNA fragment length. An antibody specific for the protein of interest (or for an epitope tag that has been attached to the protein by genetic engineering) is used to precipitate the protein–DNA complex. The crosslinks are then reversed, the released DNA is fluorescently tagged, and its genomic location is identified using a DNA microarray.

Chromosome conformational capture
(3C). A technique that identifies physical interactions between parts of the genome (specifically, interactions that would not be predictable from a survey of the DNA sequence alone). Macromolecules are chemically crosslinked in living cells, and then the DNA is extracted, digested with a restriction enzyme and subjected to intramolecular ligation. PCR is used to detect novel junctions in the ligated DNA, which are predicted to arise from the close proximity of the now-joined sequences in the folded nucleoid. A chromatin immunoprecipitation step can be added to study novel interactions that depend on a specific protein, such as a nucleoid-associated protein.

Chromosome conformational capture carbon copy
(5C). A chromosome conformational capture (3C) library is first constructed, and then multiplex primers with universal primer extensions are annealed to the novel junctions in the library and ligated together. The 3C junctions serve as templates to guide the perfect ligation of the primers. These can then be used in microarrays or subjected to high-throughput sequencing to identify the DNA forming the junction.

Dps
(DNA protection during starvation). A nucleoid-associated protein that is expressed in stationary phase cultures (or in cultures experiencing oxidative stress) and is thought to protect the DNA from damage.

Fis
(Factor-for-inversion stimulation). A nucleoid-associated protein that is expressed in early exponential phase cultures, organizes the local DNA topology and modulates transcription.

H-NS
(Histone-like, nucleoid-structuring protein). A nucleoid-associated protein with a preference for binding to AT-rich DNA. H-NS is expressed at all stages of growth, silences the transcription of hundreds of genes and organizes nucleoid structure.

HU
A nucleoid-associated protein with a general DNA-binding and DNA-compacting activity.

IHF
(Integration host factor). A paralogue of HU with site-specific DNA-binding and DNA-bending activity.

Macrodomains
Genetically defined large-scale chromosomal segments that are unlikely to undergo recombination with each other because the resulting rearrangements are detrimental to the survival of the bacterium. The *Escherichia coli* chromosome has four macrodomains (Ori, Ter, Left and Right) and two so-called non-structured regions (NS-left and NS-right).

Microdomain
A topologically independent 10–12 kb loop that coexists with other microdomains within the macrodomain superstructure of the *Escherichia coli* genome. There are around 400 microdomain loops in the genome.

Nucleoid-associated proteins
(NAPs). Low-molecular-mass, abundant DNA-binding proteins that are thought to act as architectural components within the nucleoid and to modulate gene expression. *Escherichia coli* has at least 12 distinct NAPs.

Replichores
The two arms of the circular chromosome along which bidirectional DNA replication occurs. The right (or clockwise) replichore and the left (or anticlockwise) replichore each extend, in opposite directions, from the origin of chromosome replication (*oriC* in *Escherichia coli*) within the Ori macrodomain to the terminus of replication within the Ter macrodomain.

Topoisomerase
An enzyme that alters the linking number of the DNA by cutting, strand passage and religation.

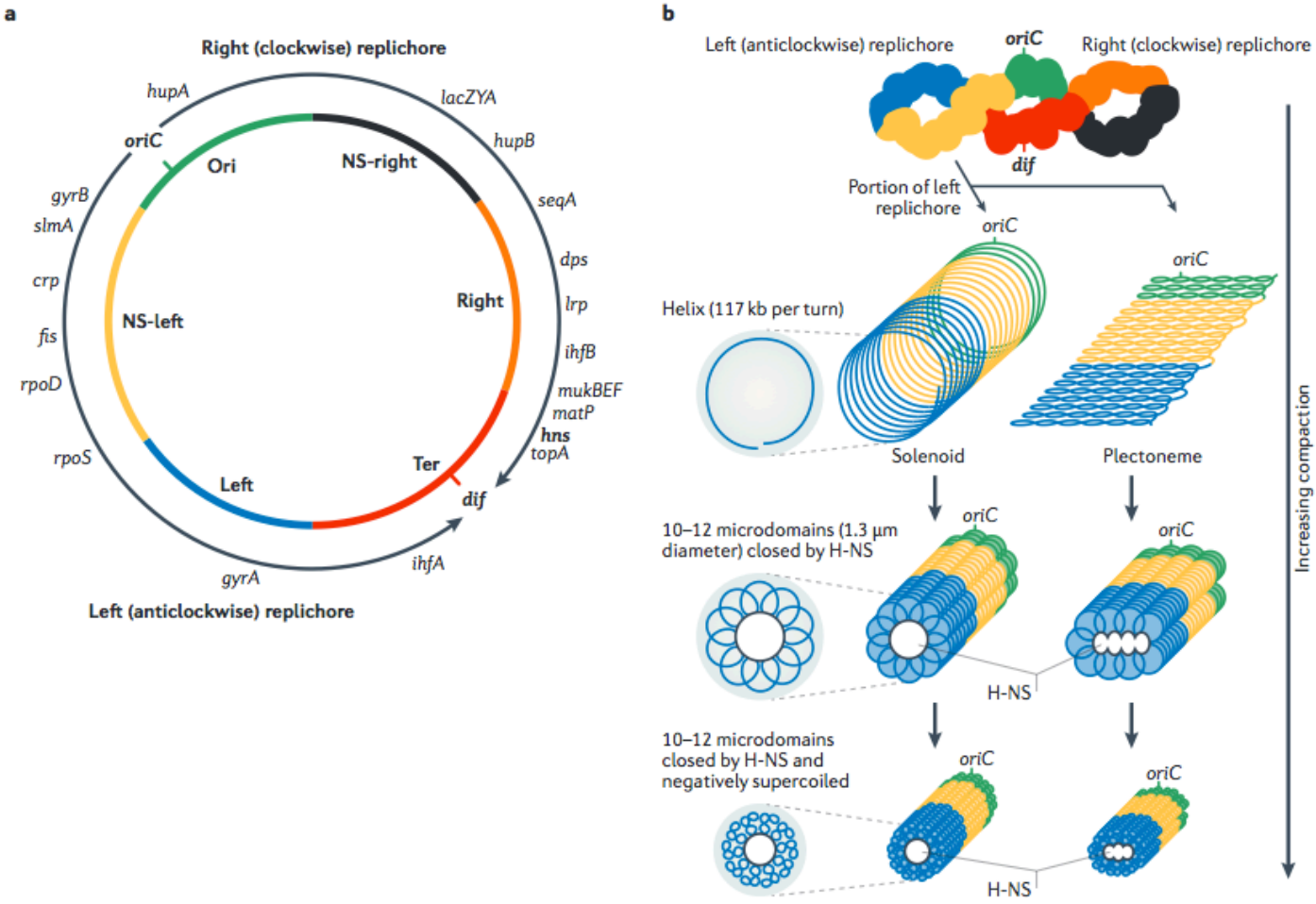

Figure 1 | **Organization of the *Escherichia coli* nucleoid.** **a** | The circular chromosome is shown with its macrodomains[19,21] indicated. The *oriC* locus is the origin of chromosome replication, and *dif* is the site where the XerC and XerD site-specific recombinases resolve chromosome dimers[22]. The directions of DNA replication are shown by the black arrows, and these arrows constitute the right (clockwise) and left (anticlockwise) replichores. The positions of key genes that encode proteins named in the main text are indicated. **b** | The chromosome is shown as a writhed structure (top), reflecting imaging data which suggest that it adopts a conformation of this type[2,10,64], at least in rapidly growing bacteria[10]. The thickness of this writhed DNA is indicative of the underlying layers of structure, as indicated below. A portion of the left replichore is illustrated as a solenoid and as a plectoneme, both of periodicity 117 kb. The DNA is next compacted by introducing 10–12 microdomains into each of its 117 kb units. These microdomain circles (each of 10–12 kb) have a diameter of approximately 1.3 µm, giving the nucleoid a cross-section of about 2.6 µm. Supercoiling these small circles compacts them approximately twofold[12]. The nucleoid-associated protein H-NS (histone-like, nucleoid-structuring protein)[4,13,32] is thought to have a core role within the two replichores, as it holds together the ends of the microdomain loops. *crp*, cyclic AMP receptor protein gene; *dps*, DNA protection during starvation; *fis*, factor-for-inversion stimulation; *hup*, HU subunit gene; *ihf*, integration host factor subunit; *gyr*, DNA gyrase subunit; *topA*, topoisomerase I.

with chromosome conformation capture (3C) technology[4]. The 3C data show that H-NS has the ability to enhance the frequency of colocalization of known H-NS-binding sites from different positions around the chromosome. The imaging data reveal that H-NS forms, on average, two prominent foci within the nucleoid that are consistent with the clustering of many microdomain boundaries[4] (FIG. 1b). The most straightforward interpretation of the imaging data is that H-NS is involved in organizing the DNA in each of the two replichores, which align with the long axis of the *E. coli* cell, like the H-NS foci (FIG. 1b). It might do this by bringing together hundreds of H-NS-binding sites along an H-NS scaffold. How robust this structure is and to what extent its integrity alters with the growth conditions and the growth phase is currently unclear. The 3C data suggest some preference for intra-replichore contacts among H-NS-binding sites, although there are also examples of contacts across replichores. This could mean that the H-NS-dependent DNA contacts introduce a degree of stochasticity into nucleoid architecture, perhaps generating variation in the gene expression patterns of individual members of the genetically identical bacterial population. This is because H-NS binds DNA at AT-rich sites that are stochastically distributed throughout the chromosome, and is itself a transcriptional repressor when it binds at a promoter (either excluding RNA polymerase from the promoter or trapping it there)[39,40]. It is interesting to note that StpA, a closely related paralogue of H-NS, does not form these large foci but is instead scattered throughout the nucleoid[4]. The significance of these differences in distribution is unknown, but they might reflect a preference on the part of StpA for binding to RNA and/or its preferential interaction with a portion of the H-NS population that is not involved in focus formation.

The negatively supercoiled nature of the chromosome itself represents an organizing influence, not least because it

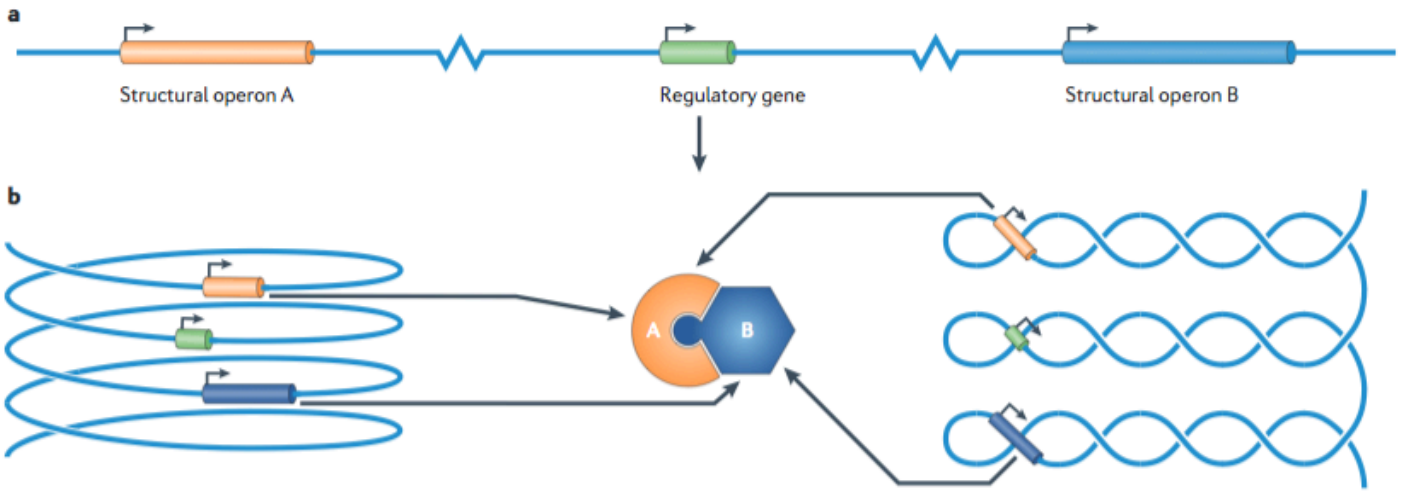

Figure 2 | **Nucleoid folding and gene regulation.** A simple regulon consisting of a regulatory gene and two structural operons, A and B, is illustrated in various conformations. **a** | When the chromosome is represented in a one-dimensional, linear form, the three genetic loci are separated by large distances in space. **b** | However, when the chromosome is reorganized as a solenoid (left) or as a plectoneme (right), the periodicity of these structures brings the three genes close together, facilitating communication between the regulatory gene and its two target operons. Moreover, the products of the A and B operons are produced in close proximity, favouring their interaction.

contributes to DNA compaction[12,41] (FIG. 1b). The bacterial type II topoisomerase DNA gyrase introduces negative supercoiling using the energy of ATP to drive the reaction[41–44]. In addition, the movement of the polymerases involved in transcription and DNA replication creates local domains of negatively supercoiled and relaxed DNA as the polymerases unwind the duplex[41,45]. DNA supercoiling affects transcription efficiency on several levels and so serves as a further integrating factor influencing gene expression and nucleoid structure[41–47]. Transcription has been proposed as a nucleoid-structuring principle in its own right, on the basis of observations that genes subject to high rates of transcription gather together to form foci[48]. Among the chromosomal genes reported to form foci are the ribosomal RNA (*rrn*) operons, and the promoters of these operons are controlled by NAPs, by negative supercoiling of the DNA and by guanosine tetraphosphate, the signalling molecule involved in the stringent response[16,48]. However, how these *rrn* operons are organized in the absence of high transcription rates is unclear, so it might be premature to conclude that high rates of transcription per se create *rrn* foci. It has also been reported that plasmids with a constitutively active promoter gather at the cell pole, but in the absence of transcription, these plasmids remain randomly distributed[49]. Although it must be noted that plasmids lie outside the nucleoid, this work does suggest that transcription can lead to gene repositioning and has the potential to influence nucleoid structure. Perhaps it would not be surprising if the need to regulate transcription had the potential to influence nucleoid structure.

Gene regulation then and now

Early gene regulation studies involving model bacterial systems (for example, the *lac* operon in *E. coli*) strongly informed opinion about the likely mechanisms used to regulate the other genes in the cell[50]. Although a variety of mechanisms became apparent reasonably quickly, the scene was dominated by the concepts of *trans*-acting protein-mediated repression or activation of transcription initiation, and the fact that regulatory genes were often located adjacent to the promoters that their protein products controlled. Placing two or more structural genes in an operon facilitates the regulation of these neighbouring genes by a single regulatory protein[51]. Work with the catabolite repressor cyclic AMP receptor protein (Crp) showed that one regulatory protein can affect the expression of a multitude of genes or operons and that these genes or operons can be located at many different chromosomal positions[52,53] (FIG. 2). Insights of this type were central to our understanding of the molecular processes underlying the concept of the 'regulon', a collective of (usually geographically dispersed) genes under the control of a common regulatory factor[51]. Crp is an abundant protein that binds to a vast number of potential sites in the chromosome, far more than can be accounted for by binding to Crp-dependent promoters alone[54]. This has led to the interesting suggestion that Crp has at least as much in common with NAPs as it does with conventional transcription factors[55]. However, it is also possible that this widespread binding of the abundant Crp protein reflects the need for high-level expression of those proteins that have geographically dispersed targets in the folded chromosome in the nucleoid.

Fluorescence *in situ* hybridization microscopy has been used to monitor the diffusion of labelled mRNA molecules expressed by the *groESL* operon and the crescentin (*creS*) gene in *Caulobacter crescentus* and by the *lacZ* gene in *E. coli*[3]. This analysis showed that mRNA translation occurs close to the DNA template. This discovery of de facto compartmentalization in bacteria has important consequences for our view of the efficacy of *trans*-acting factors in gene regulation, not least because it raises questions about the ability of geographically dispersed genes to communicate with each other. Upregulating *groESL* expression by heat shocking *C. crescentus* cells increases the dispersion of *groESL* mRNA, showing that transcripts from highly expressed genes can migrate from the site of their synthesis. In this context, it is interesting to note that a correlation exists between the level of expression of a regulatory gene and the number of genes that it controls[56]. Thus, one principle driving gene co-location in the nucleoid might be the need to bring regulatory genes within an effective range of their targets for physiologically meaningful regulation to occur [57]. (Perhaps small regulatory RNAs are more diffusible and thus provide a means of 'regulation-at-a-distance' that is more effective than protein-based mechanisms.)

It has been suggested that the periodic arrangements of genes along a solenoidal (that is, helically wound) chromosome provides a means of facilitating communication between genes and/or their products[58,59] (FIG. 2). A bioinformatic analysis of more than 100 bacterial genomes has identified statistically correlated gene pairs that tend to both co-occur and co-locate[60]. In *E. coli*, the genes in each gene pair are separated along the chromosome by multiples of 117 kb, leading to the suggestion that much of the chromosome is arranged in a helix-like structure with a 117 kb periodicity that facilitates the close alignment of these genes. Furthermore, these paired genes are associated with the most heavily transcribed regions of the genome[57]. Helical phasing of genes within each replichore would facilitate communication between genes and their products, and the need to accommodate this spatial co-location is likely to represent a strong organizing principle within the architecture of the nucleoid (FIG. 2). Such phasing could be achieved equally well by a solenoid-like structure or by a plectonemic (that is, braid-like) arrangement of the chromosomal DNA[7,61] (FIGS 1b,2).

Signal-processing analysis has been used on a transcriptomic scale to examine the co-expression of clusters of genes along the *E. coli* chromosome[62,63]. On the basis of data from one study, it has been proposed that there are three levels of transcriptional spatial organization: short range (up to 16 kb), medium range (100–125 kb) and long range (600–800 kb)[62]. It is interesting to note that these correspond in scale to the proposed sizes of, respectively, a chromosomal microdomain (10–12 kb), a helix with 117 kb periodicity and a macrodomain[37,38,58–60,62]. By contrast, another study detected a periodicity of 33 kb in addition to the 117 kb value[63]. These findings illustrate the importance of experimentally testing the *in vivo* importance of this periodicity.

Data obtained from imaging or from 3C or chromosome conformation capture carbon copy (5C) experiments have indicated that the arms of the nucleoid are interwound in *Bacillus subtilis*[64], *C. crescentus*[2] and *E. coli*[10]. Nucleoidal writhing can be expected to influence gene–gene proximity at a level above that of the periodic solenoidal structure (FIG. 1b). In *C. crescentus*, relocating *parS* (a sequence that is required for chromosome segregation in this species but is lacking in *E. coli*) changes the large-scale folding of the nucleoid without noticeably changing gene expression[2]. However, such alterations to the gross folding of the nucleoid might not have an impact at the small and intermediate scales, where gene–gene communication might influence nucleoid architecture and vice versa.

The order of the genes along the *E. coli* chromosome is remarkably similar to that in *Salmonella enterica*, even though these two species separated from their common ancestor about 100 million years ago[65]. Such conservation is indicative of an underlying structure–function imperative, and this conserved gene order has been considered in the context of the macrodomain structure of the *E. coli* nucleoid[5]. This analysis found that genes encoding *trans*-acting transcription factors are typically found in the same replichore as their targets, an observation that is in keeping with the need for co-location of regulatory genes and their regulatory subjects. By contrast, genes contributing to the same process (for example, ribosome production) are distributed in both replichores, but at comparable distances from *oriC*[5]. This might indicate a need to place genes at corresponding points along a putative DNA topological gradient in each replichore. Chromatin immunoprecipitation followed by microarray (ChIP–chip) data revealed that gyrase-binding sites are more abundant at the Ori pole than at the Ter pole of the chromosome and suggest that this results in a gradient of negative supercoiling extending from Ori to Ter[5,62]. However, it might simply reflect a need to maintain supercoiling at set points in zones of the chromosome that have different levels of transcriptionally induced supercoiling, as the average superhelical density around the chromosome is similar, at least under some growth conditions. On the one hand, the suggestion that the Ter region, which lies at the periphery of the nucleoid in bacteria growing in M9 growth medium[66], is in a more relaxed state is consistent with data from modelling, which indicate that Ter has a low level of topological complexity[7,57]. On the other hand, the condition of the Ter macrodomain might be a product of the growth conditions used in the experiments in which this macrodomain was analysed: rapidly growing bacteria have a more sharply delineated nucleoid structure than bacteria in stationary phase[10]. A systematic study of nucleoid architecture in the context of growth phase would be helpful in resolving this issue.

The fourth dimension

Bacterial physiology changes as the organism passes through successive growth stages. Following its introduction into fresh liquid medium in batch culture, the bacterial population spends a period of time in lag phase adjusting to its new environment. The population then enters the log phase of exponential growth, expanding at its maximal rate in the new environment until some vital component becomes limiting, causing a transition to stationary phase, when rapid growth ceases (FIG. 3). The cell composition and nucleoid architecture change throughout these growth stages. Patterns of gene regulation are dynamic, resulting in sequential changes to global gene expression.

In lag phase and stationary phase bacteria, the DNA has a lower superhelical density than it does in log phase bacteria[47], reflecting shifts in the ratio of the ATP/ADP concentration and the impact of this ratio

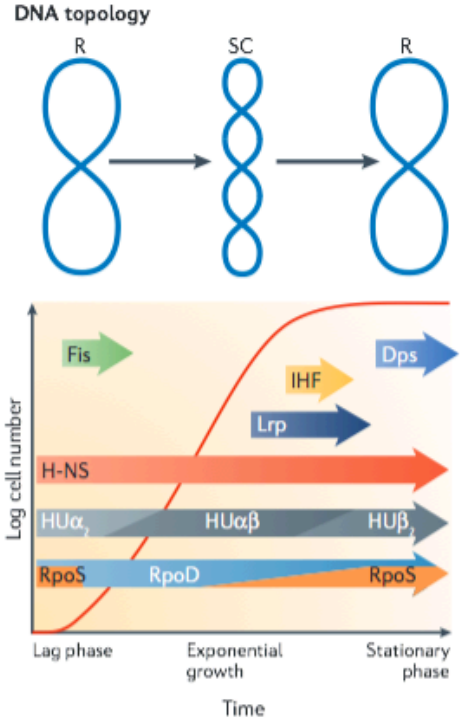

Figure 3 | **Growth phase and elements that affect nucleoid structure.** A typical growth curve for *Escherichia coli* growing in batch culture begins with a lag phase (while cells acclimatize), followed by the log phase of exponential growth and, finally, stationary phase (when the cells stop growing, usually owing to nutrient limitation). Important nucleoid-associated proteins are expressed at different times during the growth curve, as indicated, and the balance between the two main RNA polymerase σ-factors, RpoS and RpoD, also varies with growth stage. In addition, there are significant changes in DNA topology: DNA is negatively supercoiled (SC) in log phase cells, whereas it is more relaxed (R) in lag phase and stationary phase cells. The figure gives a purely qualitative impression of these events. Dps, DNA protection during starvation; Fis, factor-for-inversion stimulation; H-NS, histone-like, nucleoid-structuring protein; IHF, integration host factor.

on the ATP-dependent DNA-supercoiling activity of DNA gyrase[42,44]. Increased compaction of the nucleoid in stationary phase bacteria could be achieved by binding of NAPs, especially Dps (DNA protection during starvation)[67,68] (FIG. 3). The differential sensitivities of the two principal σ-factors of RNA polymerase to DNA supercoiling imposes a temporal control on their activities: RpoD (also known as σ^{70}) activity correlates with periods of relatively high chromosomal supercoiling, whereas RpoS (also known as σ^{S}) becomes dominant as the DNA relaxes[46] (FIG. 3).

The NAP population also shows a dynamic expression pattern[69] that is a function of the growth cycle[13,70] (FIG. 3). Fis and the HU α-subunit are expressed early; HU β-subunit and the two subunits of IHF (integration host factor) appear in exponential phase, with IHF peaking at the exponential-to-stationary phase transition; Lrp also peaks during this transition; Dps is maximally expressed in stationary phase; and H-NS is expressed at a constant ratio to chromosome copy number throughout growth[13] (FIG. 3). There is a rough correspondence between the proximity of NAP-encoding genes to *oriC* and the period at which they are expressed during growth[5]. As these NAPs influence the expression of many other genes, these determinants of nucleoid structure have a profound and widespread impact on the global gene expression profile of the cell. This is an integrating principle that links environment, gene regulation and nucleoid structure.

Implications of a unified model

The structure of the bacterial nucleoid is subjected to constraints at a number of levels because the chromosome has to be configured for optimal rates of accurate replication and segregation while accommodating the complex gene expression programmes that support the life of the cell, and because any chromosomal configuration must be compatible with the volume available to house it in the bacterium. Using DNA topology and NAPs simultaneously to modulate both gene expression and nucleoid architecture allows these two factors to be integrated. However, the associated folding of the genetic material in the nucleoid constrains the free movement of gene products, creating the need to co-locate certain genes to facilitate communication. Co-location can be achieved linearly or by exploiting the periodicity in nucleoid architecture to ensure that specific genes remain within regulatory range of each other. The levels and timing of gene expression, themselves subjected to regulation, can overcome the compartmentalization problem to some extent, allowing a given regulatory gene to exert influence at a distance.

To what extent is this apparently well-integrated nucleoid structure capable of further evolution? Bacterial chromosomes can undergo substantial rearrangement without incurring lethality[2,71], and the horizontal transfer and integration of novel genes is a routine event in many bacterial populations[72]. This suggests that nucleoid architecture has the scope to adapt to modifications, with the final arbiter of success being the manner in which the new form affects the competitive fitness of the bacterium.

Charles J. Dorman is at the Department of Microbiology, Moyne Institute of Preventive Medicine, Trinity College Dublin, Dublin 2, Ireland.
e-mail cjdorman@tcd.ie

doi:10.1038/nrmicro3007
Published online 3 April 2013
Corrected 5 April 2013

1. Kellenberger, E., Ryter, A. & Séchaud, J. Electron microscope study of DNA-containing plasms. II. Vegetative and mature phage DNA as compared with normal bacterial nucleoids in different physiological states. *J. Biophys. Biochem. Cytol.* **4**, 671–678 (1958).
2. Umbarger, M. A. *et al.* The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation. *Mol. Cell* **44**, 252–264 (2011).
3. Montero Llopis, P. *et al.* Spatial organization of the flow of genetic information in bacteria. *Nature* **466**, 77–82 (2010).
4. Wang, W., Li, G.-W., Chen, C., Xie, X. S. & Zhuang, X. Chromosome organization by a nucleoid-associated protein in live bacteria. *Science* **333**, 1445–1449 (2011).
5. Sobetzko, P., Travers, A. & Muskhelishvili, G. Gene order and chromosome dynamics coordinate spatiotemporal gene expression during the bacterial growth cycle. *Proc. Natl Acad. Sci. USA* **109**, E42–E50 (2012).
6. Pelletier, J. *et al.* Physical manipulation of the *Escherichia coli* chromosome reveals its soft nature. *Proc. Natl Acad. Sci. USA* **109**, E2649–E2656 (2012).
7. Wiggins, P. A., Cheveralls, K. C., Martin, J. S., Lintner, R. & Kondev, J. Strong intranucleoid interactions organize the *Escherichia coli* chromosome into a nucleoid filament. *Proc. Natl Acad. Sci. USA* **107**, 4991–4995 (2010).
8. Scolari, V. F., Bassetti, B., Sclavi, B. & Lagomarsino, M. C. Gene clusters reflecting macrodomain structure respond to nucleoid perturbations. *Mol. Biosyst.* **7**, 878–888 (2011).
9. Benza, V. G. *et al.* Physical descriptions of the bacterial nucleoid at large scales, and their biological implications. *Rep. Prog. Phys.* **75**, 076602 (2012).
10. Hadizadeh Yazdi, N., Guet, C. C., Johnson, R. C. & Marko, J. Variation of the folding and dynamics of the *Escherichia coli* chromosome with growth conditions. *Mol. Microbiol.* **86**, 1318–1333 (2012).
11. de Vries, R. DNA condensation in bacteria: interplay between macromolecular crowding and nucleoid proteins. *Biochimie* **92**, 1715–1721 (2010).
12. Trun, N. J. & Marko, J. Architecture of a bacterial chromosome. *ASM News* **64**, 276–283 (1998).
13. Dillon, S. C. & Dorman, C. J. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nature Rev. Microbiol.* **8**, 185–195 (2010).
14. Balke, V. L. & Gralla, J. D. Changes in the linking number of supercoiled DNA accompany growth transitions in *Escherichia coli*. *J. Bacteriol.* **169**, 4499–4506 (1987).
15. Dorman, C. J. & Corcoran, C. P. Bacterial DNA topology and infectious disease. *Nucleic Acids Res.* **37**, 672–678 (2009).
16. Travers, A. & Muskhelishvili, G. DNA supercoiling — a global transcriptional regulator for enterobacterial growth? *Nature Rev. Microbiol.* **3**, 157–169 (2005).
17. Dame, R. T., Espéli, O., Grainger, D. C. & Wiggins, P. A. Multidisciplinary perspectives on bacterial genome organization and dynamics. *Mol. Microbiol.* **86**, 1023–1030 (2012).
18. Blattner, F. R. *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1462 (1997).
19. Valens, M., Penaud, S., Rossignol, M., Cornet, F. & Boccard, F. Macrodomain organization of the *Escherichia coli* chromosome. *EMBO J.* **23**, 4330–4341 (2004).
20. Esnault, E., Valens, M., Espéli, O. & Boccard, F. Chromosome structuring limits genome plasticity in *Escherichia coli*. *PLoS Genet.* **3**, e226 (2007).
21. Dame, R. T., Kalmykowa, O. J. & Grainger, D. C. Chromosomal macrodomains and associated proteins: implications for DNA organization and replication in Gram-negative bacteria. *PLoS Genet.* **7**, e1002123 (2011).
22. Lesterlin, C., Barre, F. X. & Cornet, F. Genetic recombination and the cell cycle: what we have learned from chromosome dimers. *Mol. Microbiol.* **54**, 1151–1160 (2004).
23. Badrinarayanan, A., Lesterlin, C., Reyes-Lamothe, R. & Sherratt, D. The *Escherichia coli* SMC complex, MukBEF, shapes nucleoid organization independently of DNA replication. *J. Bacteriol.* **194**, 4669–4676 (2012).
24. Espéli, O. *et al.* A MatP-divisome interaction coordinates chromosome segregation with cell division in *E. coli*. *EMBO J.* **31**, 3198–3211 (2012).
25. Thiel, A., Valens, M., Vallet-Gely, I., Espéli, O. & Boccard, F. Long-range chromosome organization in *E. coli*: a site-specific system isolates the Ter macrodomain. *PLoS Genet.* **8**, e1002672 (2012).
26. Mercier, R. *et al.* The MatP/*matS* site-specific system organizes the terminus region of the *E. coli* chromosome into a macrodomain. *Cell* **135**, 475–485 (2008).
27. Dupaigne, P. *et al.* Molecular basis for a protein-mediated DNA-bridging mechanism that functions in condensation of the *E. coli* chromosome. *Mol. Cell* **48**, 560–571 (2012).
28. Bach, T., Krekling, M. A. & Skarstad, K. Excess SeqA prolongs sequestration of *oriC* and delays nucleoid segregation and cell division. *EMBO J.* **22**, 315–323 (2003).
29. Lu, M., Campbell, J. M., Boye, E. & Kleckner, N. SeqA: a negative modulator of replication initiation in *E. coli*. *Cell* **77**, 413–426 (1994).
30. von Freiesleben, U., Rasmussen, K. V. & Schaechter, M. SeqA limits DnaA activity in replication from *oriC* in *Escherichia coli*. *Mol. Microbiol.* **14**, 763–772 (1994).
31. Sánchez-Romero, M. A. *et al.* Dynamic distribution of SeqA protein across the chromosome of *Escherichia coli* K-12. *mBio* **1**, e00012-10 (2010).
32. Hardy, C. D. & Cozzarelli, N. R. A genetic selection for supercoiling mutants of *Escherichia coli* reveals proteins implicated in chromosome structure. *Mol. Microbiol.* **57**, 1636–1652 (2005).
33. Tonthat, N. K. *et al.* Molecular mechanism by which the nucleoid occlusion factor, SlmA, keeps cytokinesis in check. *EMBO J.* **30**, 154–164 (2011).
34. Pettijohn, D. E. Structure and properties of the bacterial nucleoid. *Cell* **30**, 667–669 (1982).
35. Sinden, R. R. & Pettijohn, D. E. Chromosomes in living *Escherichia coli* are segregated into domains of supercoiling. *Proc. Natl Acad. Sci. USA* **78**, 224–228 (1981).
36. Worcel, A. & Burgi, E. On the structure of the folded chromosome of *Escherichia coli*. *J. Mol. Biol.* **71**, 127–147 (1972).
37. Postow, L., Hardy, C. D., Arsuaga, J. & Cozzarelli, N. R. Topological domain structure of the *Escherichia coli* chromosome. *Genes Dev.* **18**, 1766–1779 (2004).
38. Deng, S., Stein, R. A. & Higgins, N. P. Organization of supercoil domains and their reorganization by transcription. *Mol. Microbiol.* **57**, 1511–1521 (2005).
39. Noom, M. C., Navarre, W. W., Oshima, T., Wuite, G. J. & Dame, R. T. H-NS promotes looped domain formation in the bacterial chromosome. *Curr. Biol.* **17**, R913–R914 (2007).
40. Dillon, S. C. *et al.* Genome-wide analysis of the H-NS and Sfh regulatory networks in *Salmonella* Typhimurium identifies a plasmid-encoded transcription silencing mechanism. *Mol. Microbiol.* **76**, 1250–1265 (2010).
41. Sinden, R. R. *DNA Structure and Function* (Academic Press, 1994).

42. Basu, A., Schoeffler, A. J., Berger, J. M. & Bryant, Z. ATP binding controls distinct structural transitions of *Escherichia coli* DNA gyrase in complex with DNA. *Nature Struct. Mol. Biol.* **19**, 538–546 (2012).
43. Rovinskiy, N., Agbleke, A. A., Chesnokova, O., Pang, Z. & Higgins, N. P. Rates of DNA gyrase supercoiling and transcription elongation control supercoil density in a bacterial chromosome. *PLoS Genet.* **8**, e1002845 (2012).
44. van Workum, M. *et al.* DNA supercoiling depends on the phosphorylation potential in *Escherichia coli*. *Mol. Microbiol.* **20**, 351–360 (1996).
45. Liu, L. F. & Wang, J. C. Supercoiling of the DNA template during transcription. *Proc. Natl Acad. Sci. USA* **84**, 7024–7027 (1987).
46. Bordes, P. *et al.* DNA supercoiling contributes to disconnect σ^S accumulation from σ^S-dependent transcription in *Escherichia coli*. *Mol. Microbiol.* **48**, 561–571 (2003).
47. Cameron, A. D. S. & Dorman, C. J. A fundamental regulatory mechanism operating through OmpR and DNA topology controls expression of *Salmonella* pathogenicity islands SPI-1 and SPI-2. *PLoS Genet.* **8**, e1002615 (2012).
48. Cabrera, J. E., Cagliero, C., Quan, S., Squires, C. L. & Jin, D. J. Active transcription of RNA operons condenses the nucleoid in *Escherichia coli*: examining the effect of transcription on nucleoid structure in the absence of transertion. *J. Bacteriol.* **191**, 4180–4185 (2009).
49. Sánchez-Romero, M. A., Lee, D. J., Sánchez-Morán, E. & Busby, S. J. Location and dynamics of an active promoter in *Escherichia coli* K-12. *Biochem. J.* **441**, 481–485 (2012).
50. Brock, T. D. *The Emergence of Bacterial Genetics* (Cold Spring Harbor Laboratory Press, 1990).
51. Maas, W. K. & McFall, E. Genetic aspects of metabolic control. *Annu. Rev. Microbiol.* **18**, 95–110 (1964).
52. Busby, S. J. in *Regulation of Gene Expression: 25 Years On* (eds Booth, I. R. & Higgins, C. F.) 51–77 (Cambridge Univ. Press, 1986).
53. Magasanik, B. & Neidhardt, F. C. in Escherichia coli *and* Salmonella, *Cellular and Molecular Biology* (ed. Neidhardt, F. C.) 1318–1325 (American Society for Microbiology Press, 1987).
54. Grainger, D. C., Hurd, D., Harrison, M., Holdstock, J. & Busby, S. J. Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc. Natl Acad. Sci. USA* **102**, 17693–17698 (2005).
55. Browning, D. F., Grainger, D. C. & Busby, S. J. Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr. Opin. Microbiol.* **13**, 773–780 (2010).
56. Janga, S. C., Salgado, H. & Martinez-Antonio, A. Transcriptional regulation shapes the organization of genes on bacterial chromosomes. *Nucleic Acids Res.* **37**, 3680–3688 (2009).
57. Fritsche, M., Li, S., Heermann, D. W. & Wiggins, P. A. A model for *Escherichia coli* chromosome packaging supports transcription factor-induced DNA domain formation. *Nucleic Acids Res.* **40**, 972–980 (2011).
58. Képès, F. Periodic transcriptional organization of the *E. coli* genome. *J. Mol. Biol.* **340**, 957–964 (2004).
59. Junier, I., Hérison, J. & Képès, F. Genomic organization of evolutionarily correlated genes in bacteria: limits and strategies. *J. Mol. Biol.* **419**, 369–386 (2012).
60. Wright, M. A., Kharchenko, P., Church, G. M. & Segrè, D. Chromosomal periodicity of evolutionarily conserved gene pairs. *Proc. Natl Acad. Sci. USA* **104**, 10559–10564 (2007).
61. Hong, S.-H. *et al. Caulobacter* chromosome *in vivo* configuration matches model predictions for a supercoiled polymer in a cell-like confinement. *Proc. Natl Acad. Sci. USA* **110**, 1674–1679 (2013).
62. Jeong, K. S., Ahn, J. & Khodursky, A. B. Spatial patterns of transcription activity in the chromosome of *Escherichia coli*. *Genome Biol.* **5**, R86 (2004).
63. Mathelier, A. & Carbone, A. Chromosomal periodicity and positional networks of genes in *Escherichia coli*. *Mol. Syst. Biol.* **6**, 366 (2010).
64. Berlatzky, I. A., Rouvinski, A. & Ben-Yehuda, S. Spatial organization of a replicating bacterial chromosome. *Proc. Natl Acad. Sci. USA* **105**, 14136–14140 (2008).
65. Groisman, E. A. & Ochman, H. How *Salmonella* became a pathogen. *Trends Microbiol.* **5**, 343–349 (1997).
66. Meile, J. C. *et al.* The terminal region of the *E. coli* chromosome localizes to the periphery of the nucleoid. *BMC Microbiol.* **11**, 1–28 (2011).
67. Frenkiel-Krispin, D. *et al.* Nucleoid restructuring in stationary phase bacteria. *Mol. Microbiol.* **2**, 395–405 (2004).
68. Kim, J. *et al.* Fundamental structural units of the *Escherichia coli* nucleoid revealed by atomic force microscopy. *Nucleic Acids Res.* **32**, 1982–1992 (2004).
69. Luijsterburg, M. S., Noom, M. C., Wuite, G. J. & Dame, R. T. The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J. Struct. Biol.* **156**, 262–272 (2006).
70. Martínez-Antonio, A., Lomnitz, J. G., Sandoval, S., Aldana, M. & Savageau, M. A. Regulatory design governing progression of population growth phases in bacteria. *PLoS ONE* **7**, e30654 (2012).
71. Andersson, D. I. & Hughes, D. Gene amplification and adaptive evolution in bacteria. *Annu. Rev. Genet.* **43**, 167–195 (2009).
72. Syvanen, M. Evolutionary implications of horizontal gene transfer. *Annu. Rev. Microbiol.* **46**, 341–358 (2012).

Acknowledgements
The author thanks R. L. Gourse and R. T. Dame for helpful discussions, and the three anonymous reviewers for insightful comments. The author also thanks N. Ní Bhriain for comments on the manuscript and M. J. Dorman for assistance with computation. This work was supported by a grant from Science Foundation Ireland.

Competing interests statement
The author declares no competing financial interests.

FURTHER INFORMATION
Charles J. Dorman's homepage: http://people.tcd.ie/cjdorman
ALL LINKS ARE ACTIVE IN THE ONLINE PDF