Architecture of a Bacterial Chromosome

A built-in loop-domain structure, proteins, and other macromolecules fold and compact this otherwise unwieldy molecule of DNA

Nancy J. Trun and John F. Marko

C ells expend considerable effort to ensure that their chromosomes function and are inherited properly—precisely controlling how, when, and where these genetic materials are replicated, repaired, recombined, moved around the growing cell, and folded to fit in their allotted space. The penalty for errors in any of these processes is chromosome damage or loss, leading to mutations or cell death.

Bacterial Chromosomes Are Huge and Unwieldy

The dimensions of bacterial chromosomes present striking challenges to cells. For example, in *Escherichia coli*, the chromosome is a circular DNA molecule of about 4.7 million base pairs. In the B form, this molecule has a total length (circumference) of 1.5 mm, while its thickness is only 2 nm, or 2×10^{-6} mm. Thus, *E. coli* cells must accommodate a molecule that is 750,000 times longer than it is wide! Moreover, those are the dimensions for an unreplicating chromosome in a cell with a generation time more than 60 minutes. Faster-growing cells contain even more DNA—as they race to produce the material that their daughter cells will inherit.

The problems cells encounter dealing with these unwieldy molecules are compounded by the small space in which the DNA must fit. A typical *E. coli* cell is about 1 μm by 0.5 μm, and DNA appears to occupy about half this space (Fig. 1A). Thus, the 1.5-mm-by-2 $\times 10^{-6}$-mm chromosome must fit in an approximately cylindrical space that is 5×10^{-4} mm long with a diameter of 5×10^{-4} mm. The DNA-occupied space within the *E. coli* cell is called the bacterial nucleoid.

Fitting the *E. coli* chromosome in its allotted space seems a monumental undertaking. If this large, circular molecule were folded into the most compact geometrical configuration possible, more than 2,500 folds would be needed for it to fit lengthwise within the cell. However, if folded in this manner, the chromosomal DNA would occupy only about 5% of the volume of the nucleoid. This simple analysis suggests that the majority of the nucleoid space may be occupied by something other than DNA and that the chromosome compaction problem entails more than mere folding. Indeed, the cell needs somehow to maintain this long, thin molecule in an active, functional configuration.

Not All Folding Is Random for the *E. coli* Chromosome

The disparity between the 1.5-mm length of the *E. coli* chromosomal DNA and the 1-μm cell length dramatizes the packaging problem. Of course, long DNA molecules that are suspended in liquid are not straight but, instead, are bent by Brownian motion into random coils (Fig. 2A). DNA molecules remain reasonably straight only over 50-nm (150-bp) stretches, a length called the thermal persistence length, or A. The persistence length characterizes the bending stiffness of the double helix, or its inclination to resist those distorting Brownian forces. Particular base pair sequences generate permanent bends along the double helix, further randomizing the shape of DNA in solution.

For a DNA molecule that is much longer than $A = 50$ nm, the bends generated by Brownian motion reduce its overall length, L, to the much smaller geometrical mean of the persistence length and its total length, or $\sqrt{AL}$. The pro-

Nancy J. Trun is an investigator at the Laboratory of Molecular Biology, National Cancer Institute, Bethesda, Md., and John F. Marko is an assistant professor at the Department of Physics, The University of Illinois at Chicago.

FIGURE 1

A

B

1 µm
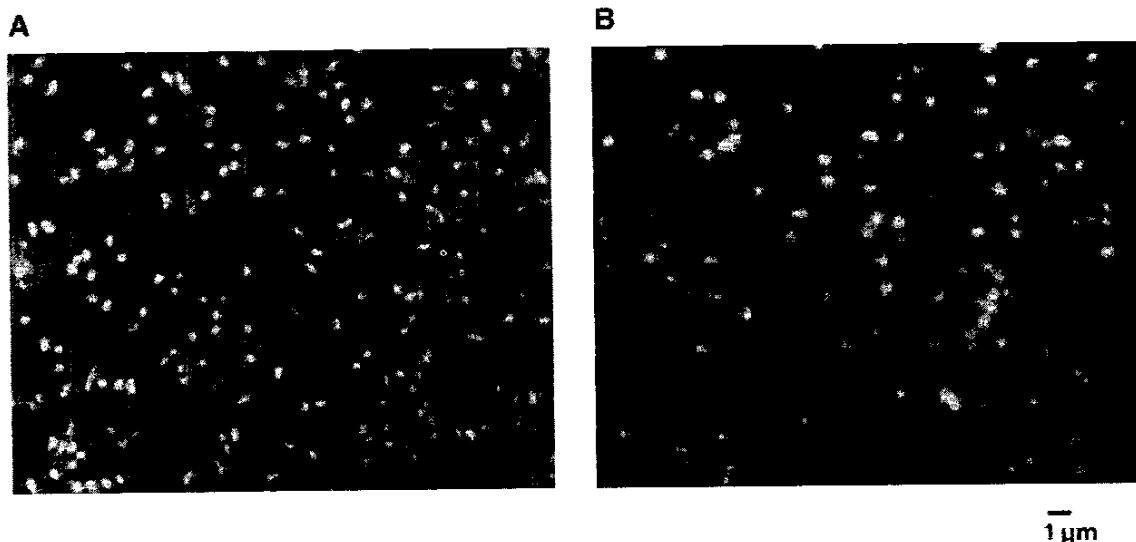
E. coli chromosomes, or nucleoids, in the cell (A) and after their release into solution (B). The DNA is visualized using the fluorescent DNA-specific dye DAPI.

portionality of the molecule size to the square root of its total length is characteristic of a "random walk" shape. Many flexible polymers in dilute solution (e.g., polystyrene in toluene) take on fluctuating "random coil" conformations, for which random walks, or shapes generated by taking a series of fixed-length steps in random directions, provide a useful mathematical model for analyzing their effective, folded size in solution.

Assuming that A = 50 nm and L = 1.5 mm, the random walk-based estimate of size indicates that Brownian motion randomly reduces the effective length of the *E. coli* chromosome to a mere 10 µm (Fig. 2A). However, although considerably condensed, this length is still much too large for the molecule to fit within the *E. coli* cell.

A far worse problem is that such a molecule would be completely disorganized: in a random coil, there is no control over what sequences are close to what other sequences, or of how the molecule is entangled with itself. Furthermore, a random coil will continually change its organization as its different parts are shuffled. Although important to physical properties and biochemical functions, random bending is not viable as a scheme for folding a chromosome within a cell.

Large-Scale Structure Is Built into Chromosomal DNA

Chromosomes have permanent, large-scale structural features built into them. In the *E. coli* chromosome, the largest of the known structural features, called the loop domain structure, entails the apparent fastening of the chromosome to itself at roughly 50-kb intervals (Fig. 2B). Estimates for the number of such domains range from 30 to 200. According to direct microscopy studies, these domains yield an overall flower structure when chromosomes are appropriately isolated (Fig. 2C).

The elements that form the anchors of the loops in the flower structure are not known. Some researchers suggest that repetitive sequence elements (REP) in the chromosome constitute the DNA part of the anchor. In the *E. coli* genome, there are 314 REP elements containing between 1 and 12 tandem copies of a 40-bp palindromic sequence. REPs are dispersed throughout the entire genome, usually between genes. Although these elements bind DNA gyrase in vitro, further experiments are necessary to determine if the REP sequences are part of the anchors.

FEATURES ★

DNA Supercoiling Plays a Vital Role in *E. coli* Cells

Eubacterial DNA is generally underwound (untwisted), and consequently supercoiled, or wrapped around itself in the manner of a twisted telephone cord. Most circular plasmids are supercoiled because about 4% of their B-form helix turns are removed.

The *E. coli* chromosome is similarly underwound, and also has roughly 4% of the DNA turns removed relative to the relaxed B-form. In the *E. coli* chromosome, the degree of supercoiling of the different domains is independent—supercoils apparently cannot move between adjacent domains (Fig. 2C). This feature indicates that the loop anchors must torsionally restrict the DNA, effectively dividing the chromosome into a series of 50-kb circular plasmids.

Now, if all 30 to 200 supercoiled chromosome domains could be gathered together at their bases, the overall size of the chromosome would be reduced. The precise reduction factor depends on how the anchor points are arranged. However, a rough, but reasonable, estimate for the size of the chromosome is based simply on the random-walk size of one of the loop domains (each containing a length L/n of DNA), or $\sqrt{AL/n}$. Using L and A as above, and assuming there are $n = 100$ domains, the *E. coli* chromosome radius reduces to 1 µm, nearly small enough to fit into the cell. In fact, this estimate closely matches the size of carefully isolated, compact *E. coli* chromosomes (Fig. 1B).

Does supercoiling reduce the size of the loop domains? Theoretically, supercoiling should slightly reduce the size of large circular DNAs, but by a factor that increases only very slowly with the overall DNA length. For a 50-kb domain, for example, supercoiling is expected to reduce its overall radius by less than a factor of 2. The precise factor depends on a balance of size reduction, due to the branched structure of a supercoil, and swelling, due to self-avoidance; no one has yet directly measured this effect.

Supercoiling very effectively brings DNA near to itself, thereby accomplishing two things. First, DNA-DNA interactions are potentially enhanced simply because distant DNA sequences are juxtaposed. Second, supercoiling changes the global shape of a large circular DNA from that of something determined by a random walk to a more ordered, branched tree-like structure (Fig. 2C). Thus, two supercoiled DNAs are relatively unlikely to become entangled, an event that is more costly in terms of free energy than are entanglements between relaxed circular DNAs.

Supercoiling may therefore play a role in promoting late-stage chromosome separation, a notion that is supported by the observation that mutants defective in topoisomerase II (gyrase) and topo IV are impaired in partitioning chromosomes between daughter cells during cell division.

Prokaryotes and Eukaryotes Condense Chromosomes Differently

Eukaryotic chromosomes are subject to at least three levels of condensation. The best characterized of these involves the wrapping of 140 bp of double-stranded DNA around histone proteins to form nucleosomes. Once wound into nucleosomes, eukaryotic DNA is wound into a fiber that is 30 nm in diameter. The detailed structure of the 30-nm fiber is not yet completely understood.

Eukaryotic DNA is subject to a third level of condensation when the 30-nm fiber is folded into large loops that are anchored at their base. Each loop is estimated to contain 50–100 kb of DNA, with the average loop size being species specific. Whether these anchor points are fixed or are dynamic remains to be determined.

In prokaryotes, despite many attempts to find one, no nucleosome-like structure has been detected. However, electron microscopists have detected a 30-nm fiber in material from *E. coli*, but details of this structure are yet to be described. Nonetheless, *E. coli* DNA does fold into 50- to 100-kb loop domains. Thus, although DNA molecules in eukaryote and prokaryote cells are condensed, the compaction methods at work in such cells appear to be distinct. In both cases, however, loop domain structures lead to substantial compaction of DNA.
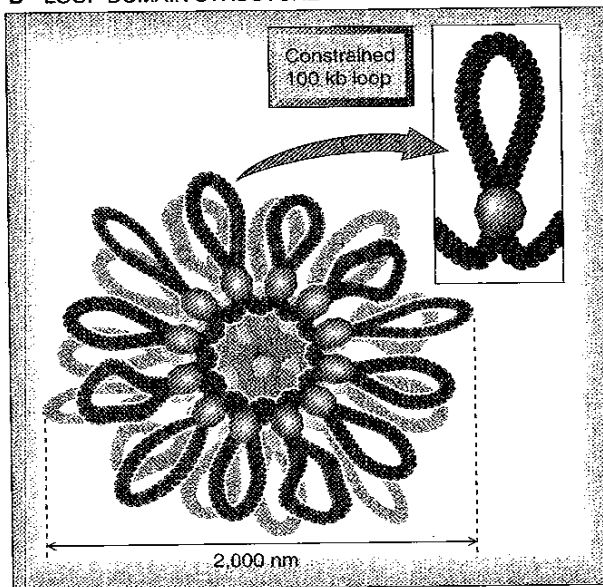
> Supercoiling enhances DNA-DNA interactions and changes the shape of a large circular DNA molecule to a more ordered structure
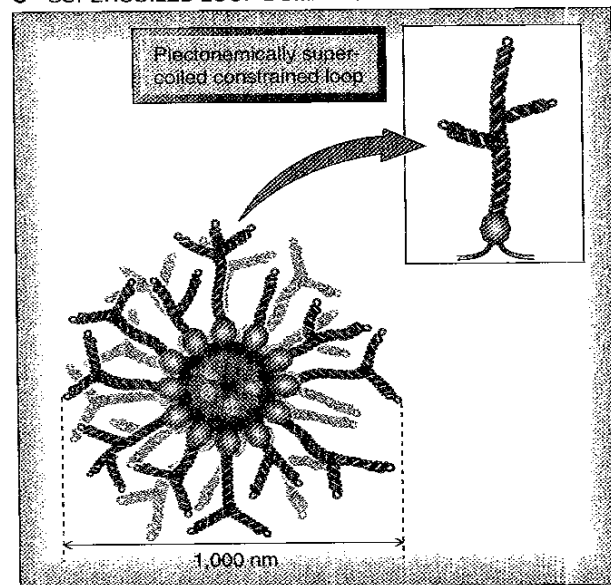
FIGURE 2

A RANDOM COIL STRUCTURE

Thermal
50 nm = A = persistence
length

10,000 nm

B LOOP DOMAIN STRUCTURE

Constrained
100 kb loop

2,000 nm
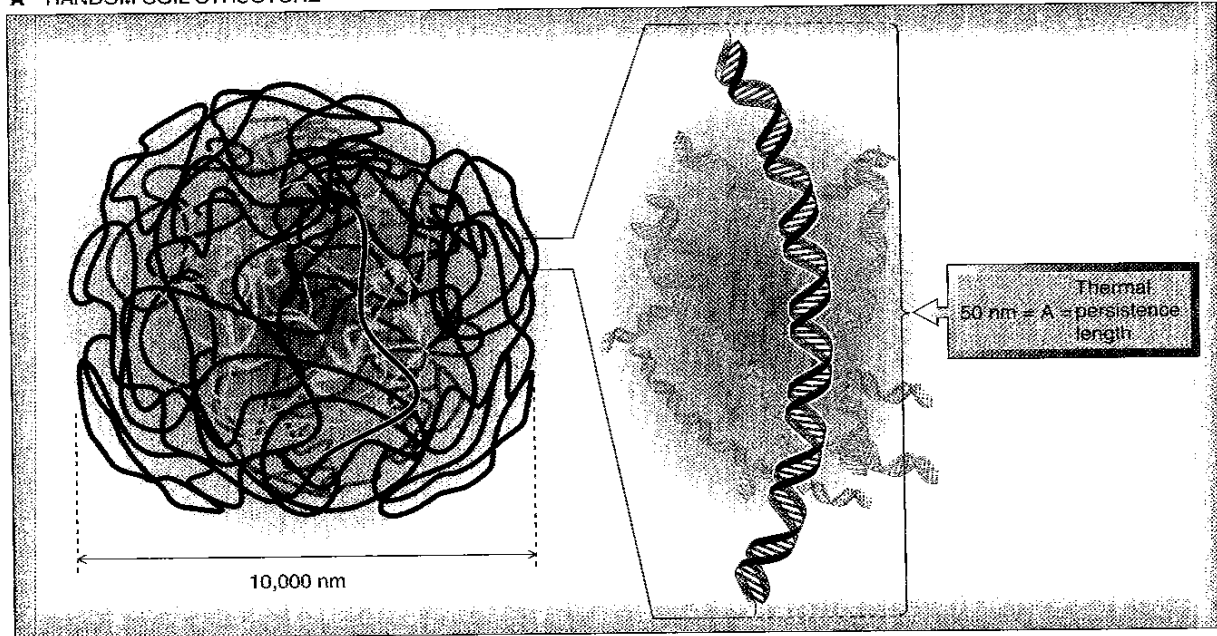
C SUPERCOILED LOOP DOMAIN STRUCTURE

Plectonemically super-
coiled constrained loop

1,000 nm

Folding schemes for chromosomes. (A) A long DNA molecule in buffer adopts a random coil conformation, the result of one random bend per persistence length of 50 nm or 150 bp. The *E. coli* chromosome would become a random coil of radius of ≈10,000 nm, or 10 μm, under such conditions. (B) The *E. coli* chromosome is attached to itself roughly once every 50 kb, to form a loop domain. These self-attachments would reduce the chromosome radius to ≈2,000 nm, nearly small enough to fit into the cell. (C) Supercoiling only slightly reduces the size of the chromosome (≈1,000 nm), but brings DNA in each loop close to itself, discouraging interloop and interchromosome entanglements, while encouraging intraloop DNA-DNA interactions.

What Is in a Bacterial Nucleoid: In Vitro Studies

Much of what is known about bacterial nucleoid structure is based on inherent physical properties of DNA plus microscope-based analysis of isolated nucleoids. Additional in vitro studies are further defining nucleoids.

In the laboratory, nucleoids may be isolated from cellular materials either as membrane-free (MF-nucleoids) or membrane-attached (MA-nucleoids). MF-nucleoids apparently represent a minimal structure containing approximately 60% DNA, which is supercoiled and organized into independent domains, 30% RNA, and 10% protein by weight. The RNA in the MF-nucleoids consists mainly of nascent mRNA, tRNA, and rRNA attached to the complex through RNA polymerase, which is the major protein within the MF-nucleoids. Meanwhile, MA-nucleoids contain DNA, RNA, RNA polymerase components, and also phospholipids. However, in MA-nucleoids, DNA makes up only about 17% of the MA-nucleoid by weight, with the remainder being RNA (19%), proteins (56%), and phospholipids (8%).

Treating both types of nucleoids with heat, DNase, RNase, or proteinase causes them to unfold. Such experiments implicate all three types of molecules—DNA, RNA, and proteins—as playing roles in maintaining the structural integrity of the compacted nucleoid.

E. coli cells contain several small, stable DNA binding and bending proteins that may also play a role in nucleoid structure. These specialized proteins include HNS, a neutral protein with a high, non-specific DNA binding affinity; HU, a non-specific DNA binding and bending protein; and two site-specific DNA binding and bending proteins, FIS and IHF. None of these proteins is found in MF-nucleoids, while all are components of MA-nucleoids. However, the 1 M salt used in preparing MF-nucleoids in vitro might displace these proteins.

When the MA-nucleoid structure is exposed to DNase I, about 13% of the total protein of the nucleoid is released. It consists of HU, FIS, HNS, RNA polymerase, and lysozyme, which had been introduced to lyse the cells. The re-

maining 87% of insoluble protein has not been characterized.

Raising salt concentrations releases HU molecules, of which there are 60,000 copies per cell, from the nucleoid as well as entrapped molecules of lysozyme. However, in their absence, the nucleoids remain compact. This procedure seems to point to several remaining proteins—HNS, of which there are 20,000 copies per cell; FIS, of which there are fewer than 100 copies per cell in stationary phase and 50,000 copies in exponential phase; and RNA polymerase, of which there are about 2,000 copies per cell—as likely mediators of chromosome compaction.

One problem with this analysis is its focus on the most abundant proteins. In principle, only about 100 proteins are needed to anchor the loop domains in the E. coli chromosome. Perhaps very rare but still not identified proteins are important for holding the nucleoid in a compact structure.

What's in a Nucleoid: In Vivo Studies

One of the first in vivo studies of nucleoid structure tested the independent supercoiled domain model. The study involved irradiating cells with X-rays to introduce a limited number of double-stranded DNA breaks.

According to this study, more than 160 double-stranded breaks are required before 95% of the E. coli chromosome loses its negative supercoiling. In other experiments, this number has ranged between 30 and 200. Pretreating cells with rifampin to disrupt RNA synthesis has no effect on the number of breaks required. Together, these data suggest that the chromosome is divided into loop domains in vivo and that the chromosome relaxes only when the majority of the loops break. RNA synthesis is not implicated in the formation or maintenance of the loops.

Inhibiting RNA synthesis dramatically changes the nucleoids, leading them to spread over the whole cytoplasm. Although the double-strand break experiments do not implicate RNA synthesis in maintaining the independent chromosomal domains, these experiments inhibiting RNA synthesis suggest it plays a role—perhaps

> A fully functional, condensed chromosome may be balancing between very condensed DNA and partially unfolded DNA

in compacting individual domains or at some other stage of DNA folding.

In contrast to inhibiting RNA synthesis, inhibiting protein synthesis with chloramphenicol makes nucleoids appear more condensed than normal. They occupy less space than in untreated cells. Fluorescent DNA dyes do not stain their centers, making them appear doughnut-shaped.

Condensing nucleoids more tightly than usual implies several possibilities to us. First, a fully functional, condensed chromosome may be balancing between very condensed DNA, which readily fits inside the cell, and partially unfolded DNA, which is more accessible for enzymes that produce mRNA and then couple transcription with translation. Second, because some proteins are exported while their cognate mRNA transcripts are still attached to the chromosome, these complexes may "fluff up" the nucleoid when bits of chromosome are being pulled toward the cell membrane. If translation indeed alters chromosomal compaction, the nucleoid structure could prove to be more highly dynamic than realized, with different regions locally responding to the translational requirements of the cell.

Inhibiting DNA initiation or synthesis does not substantially change the compactness of the nucleoid. Likewise, inhibiting cell division does not affect chromosome condensation. Even when treatments make cells filamentous, the nucleoids remain condensed and evenly spaced along the undivided cells.

Studies of mutations in genes involved in chromosome condensation suggest that these genes are also in involved in many other cellular processes. This range of gene products and genes includes HNS (*hns*), HU (*hupA* and *hupB*; HU is a heterodimer), FIS (*fis*), and IHF (*himA* and *hip*; IHF is also a heterodimer). Mutations in these genes produce varied phenotypes, with some of the phenotypes being unique to mutants of a given gene and others being shared by mutants of all of the genes.

The mutations affect cellular processes such as DNA superhelicity, illegitimate and site-specific recombination, plasmid partitioning, initiation of DNA replication from *oriC*, and the expression of a wide array of genes. Because many of these proteins have a role in gene regulation,

the caveat from the genetic studies is that any in vivo phenotype they exhibit may be an indirect effect caused by a change in the expression of other gene(s).

Nonetheless, deleting any one of these four genes produces interesting phenotypes. First and foremost, none of the four genes is essential for life. Although some of the deletions affect cell growth rates, cells can live without them, perhaps because redundancy is built into their functions. For example, if either of these gene products, HU and FIS or HU and IHF, is deleted, cells remain viable with little detriment. However, when all three gene products—HNS, HU, and IHF—are removed by deletions, cells die. Thus, these genes appear to have overlapping functions.

Comparison of In Vitro and In Vivo Results

Both in vitro and in vivo studies indicate that RNA plays an important role in DNA compaction. According to in vitro experiments, nascent RNA copurifies with the compacted nucleoid, and digesting that RNA disrupts the nucleoid structure. Although mRNA is not required for maintaining independently supercoiled loops in vivo, disrupting its synthesis disrupts nucleoid structure. Precisely what role mRNA plays in maintaining nucleoid structure is not yet known.

Proteins, particularly RNA polymerase, are also important for maintaining nucleoid structure. Further experiments are needed to determine whether RNA polymerase plays a central role in DNA compaction, or if it is simply there adventitiously because it tightly binds DNA.

The small DNA-bending and -binding proteins associated with nucleoids represent an unusual case. For instance, in vivo data implicate HU as a major player, whereas in vitro data indicate that HU can be removed from the nucleoids without disrupting their structure. The four proteins HNS, HU, FIS, and IHF are probably best thought of as a group, with at least some protein from the group being required for condensation. Perhaps the four proteins are significantly interchangeable but are used preferentially under different growth conditions.

> Both in vitro and in vivo studies indicate that RNA plays an important role in DNA compaction

FEATURES ★

Additional Chromosome Condensation Systems in *E. coli*

Because chromosomes must be folded before they can be moved into daughter cells, eukaryotic cells expend considerable energy condensing chromosomes during mitosis. The major family of yeast proteins implicated in this process is known as the structural maintenance of chromosomes (SMC) proteins. The amino terminus of these proteins forms a globular domain that contains a nucleotide-binding motif. The center of the proteins forms a long coiled-coil domain, with a hinge region to disrupt the coiled-coil at its center. The carboxyl terminus is also globular and contains a conserved sequence known as the DA box.

When the yeast SMC genes are mutated, chromosome segregation in such cells is defective and the chromosomes decondense incompletely. In *Xenopus* cells, SMC-like proteins are required to form and maintain mitotic chromosomes. Such findings further implicate SMCs in the folding of eukaryotic chromosomes.

Some bacteria contain proteins that are similar to the SMC proteins of yeast and other eukaryotes. In *E. coli* cells, for instance, the amino acids of the MukB protein are 20% identical and 45% similar to those of SMC1 or SMC2, with the best match-ups in the coiled-coil regions of the proteins.

According to electron microscopic studies, the structure of the MukB protein is like that proposed for the SMCs. According to genetic analysis, MukB contains an ATP-binding domain in the N-terminal region and a DNA-binding domain in the carboxy-terminal domain. Deletion of *mukB* leads to a defect in chromosome partitioning and a noticeable defect in chromosome condensation in *E. coli* cells. MukB shares the overall structure of the SMCs, as well as some of the phenotypes associated with loss-of-function *smc* yeast mutants.

Several additional *E. coli* genes, including *crcA*, *cspE*, and *crcB*, appear to be involved in another chromosome-condensing system. CrcA, a 20.5-kDa protein, appears to serve as the regulator of this system, while CspE, which was identified initially as a 7.6-kDa cold shock-like protein, is one of the nine highly homologous Csp proteins in *E. coli*.

Despite its name and unlike some of its homologs, CspE is not induced by cold shock. CspE binds to the DNA of certain promoters and also to the corresponding mRNA in vitro when it is just protruding from RNA polymerase. It does not bind to free RNA polymerase, free mRNA, or an elongated mRNA transcription complex in vitro. CrcB, an 8.9-kDa protein, has a homolog in *Haemophilus influenzae* but no known function.

The most unusual feature of this condensing system is that it seems to bind DNA and mRNA. None of the other identified condensing proteins displays mRNA binding. How the CrcAB-CspE system functions is not known, but it implicates a mRNA-binding protein as an important component for condensing bacterial DNA.

Macromolecular crowding is a third means for condensing bacterial chromosomes. The cytoplasm surrounding the chromosome contains a very high concentration of RNA and protein, about 340 mg/ml, while the DNA concentration within the nucleoid is much lower, somewhere between 50 and 100 mg/ml. The forces exerted by the cytoplasmic contents on the DNA could cause condensation by two different mechanisms: direct compaction by excluded volume effects and indirect compaction by the increased binding of HNS, HU, IHF, FIS, or other proteins under crowded conditions. In vitro experiments suggest that both of these mechanisms can compact the chromosome. The advantage of such a system is that it requires only a specific concentration of cellular components, not any specific component, and it should be relatively insensitive to changes in salt concentration or pH.

Model To Describe Overall Behavior of the *E. coli* Chromosome

In cells, chromosomes must be structured to function in multiple reactions simultaneously, including replication, transcription, translation, recombination, and physical movement. No one process appears to be accorded preference over others. Synthesizing what we know about chromosomal architecture in *E. coli*, we present the following model containing several testable hypotheses.

The loop domain structure of *E. coli* DNA (Fig. 2C) imparts order to and compacts the

★ **FEATURES**

chromosome. If one of the loops unfolds, the vast majority of the structure is maintained. Individual loops are capable of independent reactions, suggesting that a major benefit of the structure is stability. The loop-structure ordering also suggests that, with modest additional folding, the degrees of chromosomal compaction approaches those routinely observed in living cells and in carefully purified nucleoids.

Several proteins that bind or bend DNA also help to fold chromosomal DNA in *E. coli* cells. For example, the MukB protein appears to participate in *E. coli* DNA folding, possibly acting in a manner similar to members of the SMC family of condensing proteins in yeast cells. In addition, the CspE-CrcB proteins together with nascent mRNA molecules help to fold *E. coli* DNA, as does macromolecular crowding.

SUGGESTED READING

Hecht, R. M., and D. E. Pettijohn. 1976. Studies of DNA bound RNA molecules isolated from nucleoids of *Escherichia coli*. Nucleic Acids Res. 3:767–788.

Hirano, T., T. J. Mitchison, and J. R. Swedlow. 1995. The SMC family: from chromosome condensation to dosage compensation. Curr. Opinion Cell Biol. 7:329–336.

Hu, K. H., E. Liu, K. Dean, M. Gingras, W. DeGraff, and N. J. Trun. 1996. Overproduction of three genes leads to camphor resistance and chromosome condensation in *Escherichia coli*. Genetics 143:1521–1532.

Murphy, L. D., and S. B. Zimmerman. 1997. Stabilization of compacted spermidine nucleoids from *Escherichia coli* under crowded conditions: implications for in vivo nucleoid structure. J. Struct. Biol. 119:336–346.

Niki, H., R. Imamura, M. Kitaoka, K. Yamanaka, T. Oguru, and S. Hiraga. 1992. *E. coli* MukB protein involved in chromosomal partition forms a homodimer with a rod-and-hinge structure having DNA binding and ATP/GTP binding activities. EMBO J. 11:5101–5109.

Pettijohn, D. 1976. Prokaryotic DNA in nucleoid structure. CRC Critical Rev. Biochem. 4:175–202.

Schmid, M. B. 1990. More than just "histone-like" proteins. Cell 63:451–453.

Sinden, R. R., and D. E. Pettijohn. 1981. Chromosomes in living *Escherichia coli* cells are segregated into domains of supercoiling. Proc. Natl. Acad. Sci. USA 78:224–228.

Vologodskii, A. V., and N. R. Cozzarelli. 1994. Conformational and thermodynamic properties of supercoiled DNA. Annu. Rev. Biophys. Biomol. Struct. 23:609–643.

Zechiedrich, E. L., and N. R. Cozzarelli. 1995. Roles of topoisomerase IV and DNA gyrase in DNA unlinking during replication in *Escherichia coli*. Genes Dev. 9:2859–2869.