DNA SEGREGATION DURING THE CELL CYCLE OF *ESCHERICHIA COLI*

An analysis of intracellular positions of fluorescently labelled DNA regions

ACADEMISCH PROEFSCHRIFT

Ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. J. J. M. Franse
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der Universiteit
op donderdag 5 juli 2001, te 11.00 uur

door Marco Roos
geboren te Amsterdam

PROMOTIECOMMISSIE

Promotor:	Prof. dr. N. Nanninga
Overige leden:	Prof. dr. K. J. Hellingwerf
	Prof. dr. C. J. F. van Noorden
	Prof. dr. H. V. Westerhoff
	Dr. J. Aten
	Dr. C.L. Woldringh
	Dr. R. Lingeman

FACULTEIT
DER NATUURWETENSCHAPPEN, WISKUNDE EN INFORMATICA

The research presented in this thesis was conducted at the Section of Molecular Cytology of the Swammerdam Institute for Life Sciences, part of the Faculty of Science of the University of Amsterdam, Kruislaan 316, Amsterdam, The Netherlands (postal address: P.O. Box 94062, 1090 GB Amsterdam, The Netherlands; homepage: http://wwwmc.bio.uva.nl).

Printed by Dione Boekproducties BV, Erasmusweg 9, 4104 AK Culemborg, tel.: 0345-533203, e-mail: dione@wxs.nl

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

1.1 Organisation of DNA in the cell

1.1.1 Introduction

In this thesis we address the dynamics of DNA organisation of the prokaryote (and eubacterium) *Escherichia coli* in relation to the cell cycle. During the cell cycle all DNA is duplicated and segregated (partitioned) to each nascent daughter cell. Of course, DNA provides a cell with an information database, which the cell uses to make its constituents. The complete genome sequence of *E. coli* has been available since 1997 (Blattner *et al.*, 1997). Nevertheless, however important DNA may be, living cells exist by virtue of a combination of DNA and many other constituents, such as RNA, phospholipids, sugars, proteins and water. Therefore, our aim as molecular cytologists is to find out how these constituents work together in their natural context: inside the cell. With the experiments described in this thesis we have tried to find a fitting description of how DNA segregates in *E. coli* during the cell cycle (chapters 2, 3, and 4). Our results pertain to global DNA-organisation;

underlying factors that determine this organisation were not assessed directly. Therefore, I use this chapter to introduce factors that play a part in determining global DNA-organisation. In chapter 6 I discuss how these factors might be involved in DNA segregation.

Some aspects of DNA organisation are equal among most species, e.g. the basic structure of the DNA-molecule itself. Many other aspects of DNA organisation differ substantially between species, e.g. *E. coli* has its genes on one circular chromosome whereas humans have their genes on 46 linear chromosomes. Treating all that is known about DNA organisation is beyond the scope of this thesis, but because prokaryotes and eukaryotes are often compared in the literature, I will also introduce some aspects of eukaryotic DNA organisation. Species are classified as eukaryotes if their DNA is contained in a membrane-bound organelle, called the nucleus. Prokaryotes lack a nucleus; instead, their protein-poor DNA is separated from protein-rich cytoplasm by physical forces (Odijk, 1998). This gives rise to the nucleoid as a distinct structure, a structure that can be visualised by phase-contrast microscopy (Mason and Powelson, 1956). I should note that conclusions drawn from a comparison between a prokaryotes and eukaryotes with respect to DNA organisation, might not be applicable to every prokaryote-eukaryote combination. For instance, as a rule prokaryotic DNA is circular and eukaryotic DNA is linear, but some bacterial species contain linear DNA (e.g. *Borrelia burgdorferie*; see Jumas-Bilak *et al.*, 1998 and references therein) and under some circumstances a eukaryote, such as fission yeast, may contain circular forms of DNA (Nakamura *et al.*, 1998).

1.1.2 The basic shape of DNA

DNA can adopt different shapes, but it is generally believed that DNA exists in cells as B-DNA: a double helix of two DNA strands, held together by H-bonds between complementary nucleotides (Bates and Maxwell, 1993a). This shape fits well with the fact that the ideal shape of long molecules is helical (Maritan *et al.*, 2000), but in the case of DNA it imposes an interesting biological problem: because the two strands of the

double helix are twisted around each other they cannot be separated without breaking the double helix, they are linked (fig. 1.1). Theoretically, this 'linkage problem' occurs only when the ends of the double helix are joined to form a circle. When the two strands of a linear DNA molecule are twisted around each other *Tw* times, and the ends of this molecule are joined, this 'twist' is constrained within the molecule. Chromosomes and plasmids in most bacteria are circular, but eukaryotic chromosomes are long linear DNA-molecules. However, if DNA is very long, or if the ends of a region of DNA are fixed in some way (by DNA-binding proteins, for instance), the two strands can also be considered 'linked'. Linkage problems are solved by topo-isomerases, which are enzymes that can break and rejoin DNA (see section 1.1.5.7; for an extensive review about topo-isomerases see Wang, 1996; for a more general description see Bates and Maxwell, 1993c).

1.1.3 DNA geometry and DNA topology

Because of the linkage problem, introduced in the previous section, the geometry of a closed-circular DNA-molecule, or a bound region of DNA, can change. Twist (*Tw*) can be converted into 'writhe' (*Wr*), which is the twisting of the double helix around its central axis. This changes the geometry of the DNA molecule, such that it forms higher-order 'supercoils'. Twist and writhe are geometrical parameters. The 'linkage number' *Lk* is a topological parameter. It represents the number of links in a closed circular DNA molecule. The linkage number is equal to the sum of twist and writhe (Bates and Maxwell, 1993b):

$$Lk = Tw + Wr$$

Thus, whilst the topology of a DNA molecule remains constant, its geometry may change (fig. 1.2).

Supercoiling can have an effect on how ‘active’ DNA is. On average, the two strands of linear DNA are twisted around each other 360° every 10.5 base pairs (Watson and Crick, 1953). This represents the 'relaxed' state of

DNA. Increasing or decreasing *Lk* may introduce torsional stress into a DNA molecule. Because this stress cannot be released from circular or otherwise bound DNA, it represents a certain amount of free energy (Bates and Maxwell, 1993b). DNA-binding proteins, such as topo-isomerases, may use this energy. Proteins that increase torsional stress do this at the expense of energy (e.g. DNA gyrase in bacteria uses ATP to increase negative supercoiling; see Hsieh *et al.*, 1991; van Workum *et al.*, 1996; Westerhoff *et al.*, 1988).

Without DNA-binding proteins, DNA forms plectonemic supercoils (fig. 1.2b); with specific DNA-binding proteins, DNA can form solenoidal supercoils (fig. 1.2c; section 1.1.5.6). Plectonemic supercoiling brings about limited compaction of a DNA molecule (~2.5 fold; Boles *et al.*, 1990). It is the predominant type of supercoiling in bacteria. Solenoidal supercoiling brings about a higher level of compaction. It is the predominant type of supercoiling in eukaryotes (nucleosomes).

In most cells DNA is negatively supercoiled, i.e. under-twisted (see for instance Higgins, 1999). In general, negative supercoiling increases the meltability of DNA. This makes sense, because DNA-related processes often require melting of the double helix (see for instance section 1.1.5.2). Environmental conditions may change the average state of supercoiling (e.g. Hsieh *et al.*, 1991), and, therewith, global gene expression.

Despite general believe, it is possible to construct double-stranded DNA in such a way that it seems supercoiled without being linked. Due to steric hindrance, a single-stranded DNA-molecule cannot rotate completely freely around its axis (Wu and Wu, 1996). Therefore, if a single-stranded DNA-molecule is twisted on itself and its ends are joined to form a single-stranded circle, it may not resume an open circular form but may remain twisted (Wu and Wu, 1996). If two complementary strands are twisted in this way and rejoined by H-bonding, then the resulting molecule will look very similar to a linked supercoiled double-helix. However, the strands are not linked and can be easily separated. There is some experimental data that supports this model of supercoiling

Figure 1.1 - The 'linkage problem' of DNA.

When opposing strands are pulled apart at one position (arrows in left image), the strands 'lock up' at two other positions (arrowheads in right image). Note that this is only a 'problem' when the helices in this figure are part of larger circular molecules or if the ends are otherwise bound; if the ends are free the 'problem' can be solved by rotating the strands around the axis of the helix (unwinding).

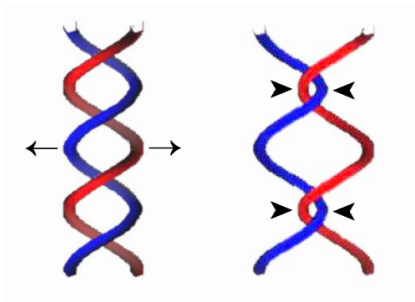
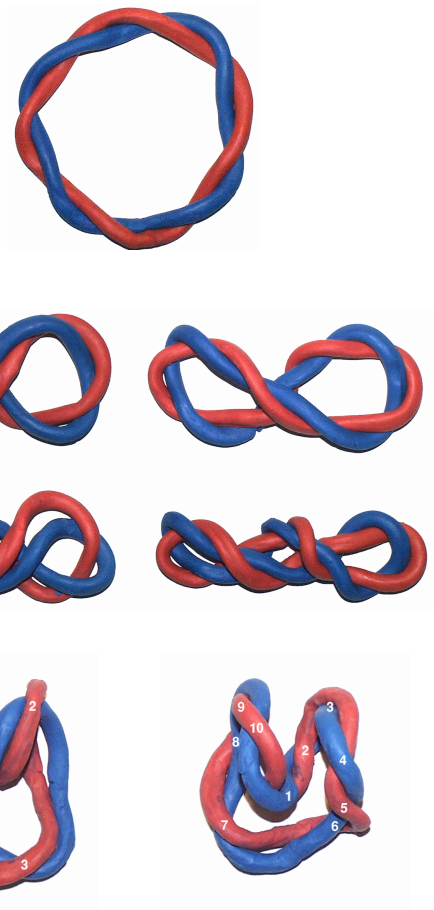


Figure 1.2 - Clay models of supercoiled circular double-stranded DNA

a: A (negatively) twisted circular molecule without higher-order supercoils (*Lk*=-4, *Tw*=-4, *Wr*=0). b: The same molecule, but with plectonemic supercoils (*Lk*=-4, top-left: *Tw*=-3, *Wr*=-1; bottom-left: *Tw*=-2, *Wr*=-2; top-right: *Tw*=-5, *Wr*=+1; bottom-right: *Tw*=-6, *Wr*=+2). c: The same molecule, but with solenoidal supercoiling (*Lk*=-4, left: *Tw*=-3, *Wr*=-1; right: *Tw*=-5, *Wr*=+1). Note that upon increased negative supercoiling (more negative writhe; left images), the molecule appears more 'loose' (less negative twist). Because this is difficult to see in the images of the solenoidial supercoils, the cross points of the two strands are numbered (*Tw*=-crosspoints/2).



(Wu and Wu, 1996), but the model requires a reinterpretation of the crystallisation data of Watson and Crick and a re-evaluation on a vast amount of literature in which a linked topology of double-stranded DNA was assumed. This includes studies of topo-isomerases, of which the main function is presumed to be to solve linking problems of the DNA double-helix (see section 1.1.5.7). Perhaps the data of Wu and Wu pertain to a special case, or an effect of the experimental procedure was overlooked (e.g. an unanticipated nick in either strand of the studied plasmid would make $Lk = 0$).

1.1.4 From sequence to structure

The structure of a DNA molecule is not uniform over its entire length. This may influence how DNA is organised within the cell. The underlying non-random sequence of nucleotides influences DNA structure, defined by parameters such as curvature, flexibility and helix stability (Bates and Maxwell, 1993e; Pedersen *et al.*, 2000; Ussery, 2000). Based on computational analysis and statistical methods Pedersen and co-workers were able to define a structural atlas of the whole *E. coli* genome (Pedersen *et al.*, 2000). They found that the regions in front of genes, which contain promoters and other regulators of gene expression, have non-average structural properties: higher levels of curvature, lower levels of flexibility, and lower degrees of helix stability. These properties suggest that the regulatory regions are more rigid and open up ('melt') more easily than average DNA. This links DNA function to DNA structure, because initiation of DNA transcription by RNA polymerase coincides with the melting of the regulatory region in front of a gene.

Pederson and co-workers also found non-average structural properties at a scale much larger than that of regulatory regions. Especially two large regions on opposite sides of the *E. coli* chromosome deviated from the average structure of the whole chromosome (fig. 1.3). The largest and most distinct one encompasses the terminus of replication: on average, it is more curved, less flexible and melts more easily than average DNA. The other, less distinct, region was found near the origin of replication.

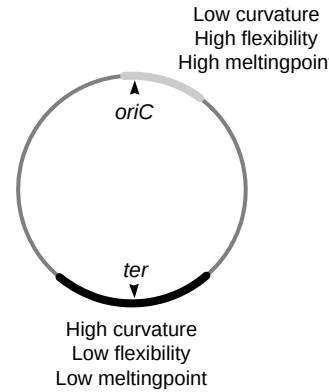


Figure 1.3 - Simplified structural map of *E. coli*

See http://www.cbs.dtu.dk/services/GenomeAtlas and Pedersen *et al.*, 2000 for a more detailed map. Two large regions of which the average structural properties deviate substantially from the genome's average have been indicated.

The structural properties of this 'origin' region were different from the terminus region: less curved, more flexible, and with a higher melting point. In addition to these large regions, some 20 regions of extreme structural properties were found. Pedersen speculated that these regions are involved in defining the transient domain-boundaries thought to exist in the *E. coli* chromosome *in vivo*. These boundaries might prevent supercoiling to diffuse throughout the chromosome (Higgins, 1999). Although these domains are an interesting aspect of DNA organisation in *E. coli*, the relationship with the position of DNA regions in a cell is not yet clear. Hence, the relationship between the non-uniform distribution of DNA structure and the organisation of DNA *with respect to the cell* is still unknown.

1.1.5 Proteins and DNA structure

Proteins that bind to DNA change the local structure of the DNA they bind to. Moreover, because DNA is a string, the effect may be felt over long distances. In addition, the sum of all DNA-binding that takes place in a cell probably contributes greatly to global DNA-organisation. Therefore, it is difficult to single out individual proteins or classes of proteins as being primary DNA-structuring elements, except perhaps for histones in eukaryotes and archaea-bacteria (see section 1.1.5.6). In the following I will introduce some well-known examples of DNA-binding proteins that have an influence on DNA structure in *E. coli.*

1.1.5.1 AraC

The AraC protein is an example of a DNA-binding protein that affects the local structure of DNA and therewith determines whether one particular gene region, that of the *araBAD* operon, is expressed (Dunn *et al.*, 1984; Martin *et al.*, 1986; Seabold and Schleif, 1998). AraC proteins form dimers via their C-terminal domains, their N-terminal domains bind to DNA. Two AraC binding sites lie just upstream of the operon. A third binding site lays 280 base pairs upstream (Dunn *et al.*, 1984). In the absence of arabinose AraC dimers create a loop between one of the proximal binding-sites and the distant binding-site, whereas AraC-dimers connect two proximal sites in the presence of arabinose (fig. 1.4). Only in the latter conformation can transcription of the *araBAD* operon be induced.

1.1.5.2 RNA polymerase

RNA polymerase is an example of a protein that changes several structural properties of DNA and - as a consequence - transcription of one gene may influence expression of other genes several hundreds of base pairs away. In addition, the enzyme may be partly responsible for both global DNA organisation (see below) and DNA segregation in *E. coli* (see section 6.1.2.5).

Upon initiation of transcription, RNA polymerase binds to the promoter region upstream of a gene (from base -35 to -10 in *E. coli*), after which the region between base -10 and -1 is melted and the polymerase can begin transcribing the gene. The structure of DNA can influence these initial stages of transcription (see Bates and Maxwell, 1993e; St. Jean, 1999 and references therein). In general, negative supercoiling increases transcription (e.g. Drew *et al.*, 1985), presumably because negative supercoiling decreases the melting point of a promoter region (see section 1.1.3).

RNA polymerase itself may induce long-range effects by altering the

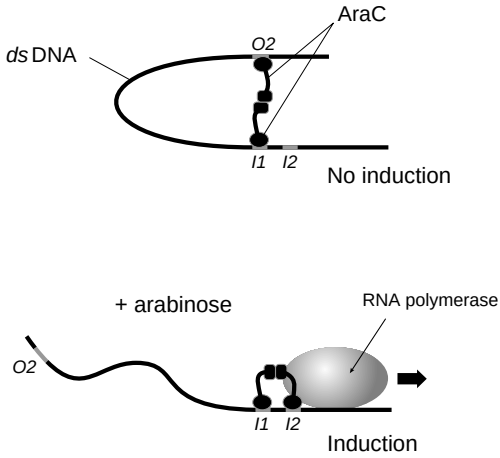


Figure 1.4 - Simplified schematic representation of the function of AraC

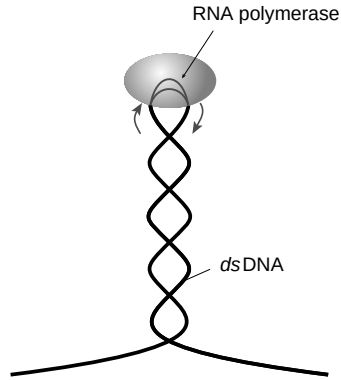
Top: two AraC proteins associate and bind to two regions (I1 and O2) 210 bp apart in front of the *araBAD* promoter, thereby bending the DNA and preventing initiation of transcription by RNA polymerase. **Bottom**: in the presence of arabinose, one of the AraC proteins binds to an adjacent binding site (I2); DNA is not bended anymore and transcription can initiate.

state of supercoiling. RNA-polymerase induces positive supercoiling in front of it and negative supercoiling behind it. Topo-isomerases counteract this, but the effect of transcription may be felt several hundreds of base pairs away from the RNA polymerase. This effect was demonstrated by experiments with the *leu-500* promoter in *E. coli* (see Bates and Maxwell, 1993e; St. Jean, 1999 and references therein). A point mutation in this promoter inactivates the promoter. Re-activation could be obtained by mutating *topA*, the gene for topo-isomerase I. This indicates that supercoiling is important for the activity of the promoter. Consequently, if RNA polymerase has an effect on supercoiling, transcription might also re-activate the mutated *leu-500* promoter. Indeed, re-activation was obtained by transcription of genes several hundreds of base pairs upstream of the promoter, now placed on a plasmid (Chen *et al.*, 1992). Activation was further enhanced by transcription of a downstream gene (Chen *et al.*, 1993).

During transcription, RNA polymerase either 'walks' along the DNA molecule or RNA polymerase is stationary and DNA moves through it. Although the latter model requires DNA to be very dynamic, it is the most likely model for transcription *in vivo*, for instance because the RNA-

Figure 1.5 - Transcription by RNA polymerase at the end of a plectonemic supercoil

DNA is 'spooled' through the stationary RNA polymerase. Note that double helices 'slither' past each other during the process.



polymerase complex is probably too large to move along the DNA (Cook, 1989; 1999; Liu and Wang, 1987). The complex probably operates at the terminal end of a plectonemic supercoil (fig. 1.5). The fact that regulatory regions, because of their intrinsic structure (section 1.1.4), favour a position at such terminal ends supports this idea (ten Heggeler-Bordier *et al.*, 1992).

In bacteria, translation of a transcript starts before transcription finishes. In fact, several ribosomes may start producing protein and nascent proteins may already start finding their destination in the cell. Consequently, DNA can become anchored to these destinations, and this may have an effect on DNA structure. For instance, it has been suggested that in bacteria transcription and translation of membrane proteins plays a role in global DNA organisation and in DNA segregation (Woldringh *et al.*, 1995).

1.1.5.3 DNA polymerase

As DNA polymerase is the enzyme that replicates DNA, its importance for global DNA-organisation in proliferating cells is obvious. The enzyme is part of a large replication complex that contains various other enzymes involved in replication (Kornberg and Baker, 1992b and references therein). In many respects replication by DNA polymerase is

similar to transcription by RNA polymerase. One important difference is, of course, that transcription may transiently change DNA structure, whilst replication irreversibly changes it.

Initiation of replication takes place at specific sites on a chromosome, the origins of replication. Prokaryotes have one such origin per chromosome; in *E. coli* the minimal origin is a sequence of 248 base pairs, called *oriC*. Eukaryotes have many origins per chromosome, but possibly the process of replication per origin is comparable (but see section 6.2.3). In most cases, the binding of initiation factors to a replication origin precedes initiation of replication. For example, binding of DnaA to DNA near *oriC* is a prerequisite for replication in *E. coli* (Kornberg and Baker, 1992a and references therein). Similar to transcription, negative supercoiling probably helps initiation of replication.

As was the case for RNA polymerase (see above), DNA polymerase either 'walks' through DNA or it is stationary and the DNA moves through the polymerase. There has been much debate about this issue in the case of eukaryotes, where it is difficult to find prove for either model because of the large number of replication origins. However, recent experimental evidence, obtained in bacteria, strongly supports the stationary model of DNA replication (Koppes *et al.*, 1999; Lemon and Grossman, 1998; 2000). I see no reason why the process of replication in eukaryotes would be fundamentally different at this point.
DNA polymerase induces positive supercoiling in front of itself, similar to RNA polymerase, but it does not induce negative supercoiling behind itself. The latter is simply because template DNA does not rejoin after passing the replication complex. Not surprisingly, topo-isomerases are active during replication to control supercoiling. The activity of specific topo-isomerases is especially important near termination of replication, when replicated DNA becomes catenated because of the linkage of the DNA double-helix (sections 1.1.2 and 1.1.3; Wasserman and Cozzarelli, 1986).

1.1.5.4 'Structural-maintenance-of-chromosomes'

Members of the family of structural-*m*aintenance-of-*c*hromosomes (SMC) proteins have gained increasing attention for their role in structuring DNA. Indeed, they are involved in a broad spectrum of DNA-related functions, such as sister chromatid cohesion, DNA condensation, DNA recombination, and DNA repair (for a review see Jessberger *et al.*, 1998). An SMC protein can be associated with various accessory proteins, which may explain the involvement of SMC proteins in so many functions. SMC proteins may have an effect on DNA structure by affecting supercoiling (Holmes and Cozzarelli, 2000; Kimura *et al.*, 1999). Such an effect was shown of 13S-condensin in yeast *in vitro* (Kimura and Hirano, 1997).

The SMC equivalent in *E. coli* is MukB, which forms a complex with MukE and MukF (Yamazoe *et al.*, 1999). MukB does not cluster with other SMC's in a phylogenetic tree, and it is not homologous to most eukaryotic SMC proteins (Melby *et al.*, 1998). Nevertheless, I call MukB an SMC protein, because it has the distinctive structural features of an SMC protein (see below; Melby *et al.*, 1998). Hence, I use a structural definition.

MukB was originally found by genetic screening for partitioning mutants (Niki *et al.*, 1991). The typical phenotype of *muk*- mutants is the production of 5-15% anucleate cells (*muk*aku means 'anucleate' in Japanese), but other consequences are a change in the average size of cells and a loss of position of *oriC* from the central and quarter-length regions of the cell (Weitao *et al.*, 2000a). The *muk*- phenotype is partly suppressed by mutations in *topA* (Sawitzke and Austin, 2000; Weitao *et al.*, 2000a; Weitao *et al.*, 2000b). *TopA* encodes topo-isomerase I, which relaxes DNA by removing negative supercoils; these *topA* mutants have more negatively supercoiled DNA (Weitao *et al.*, 2000b). Thus, apparently the loss of MukB is compensated by an increase in negative supercoiling. Assuming supercoiling has a substantial effect on the compaction of DNA, the most likely function of MukB is that it is involved in compacting DNA. This interpretation has also led to the idea

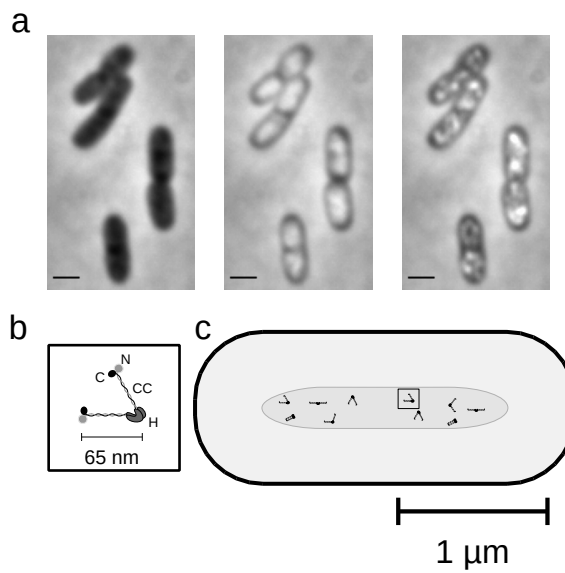


Figure 1.6 - MukB localisation (a), structure (b), and size (c)

a: phase contrast image (cells), DAPI image (nucleoid) plus phase contrast, and fluorescent immunolabel image (MukB) plus phase contrast. b: structure of MukB homodimer (C: C-terminal domain, N: N-terminal domain, CC: coiled coil, H: hinge). c: schematic depiction of cell with 10 MukB proteins, drawn more or less on scale. Note that, on average, about 150 MukB proteins are present in a cell (Kido *et al.*, 1996), and that the angle between the arms of the homo-dimer probably varies between 0 and180 degrees. Digital images courtesy of Tanneke den Blaauwen.

that compaction of replicated DNA, after leaving the replisome and mediated by MukB, could lead to DNA segregation (Dasgupta *et al.*, 2000; Graumann, 2001; Holmes and Cozzarelli, 2000 and references therein; see also section 6.1.2.4).

Immuno-labelling of MukB provided direct evidence for a role in organising DNA in *E. coli* (Den Blaauwen, submitted). An estimated number of 150 MukB molecules per cell (Kido *et al.*, 1996) produced distinct foci that co-localised with nucleoids during all of the cell cycle (fig. 1.6a). A role in organising DNA is further supported by the structure and size of MukB (fig. 1.6b). It has the distinctive five-domain structure of an SMC protein: an N-terminal globular domain containing an ATP-binding site, a long coiled-coil, a hinge, another long coiled-coil, and a C-terminal globular domain containing a DNA binding site (Melby *et al.*, 1998; Niki *et al.*, 1992). Most eukaryotic SMC proteins are present as hetero-dimers, probably with their coiled coils in anti-parallel conformation. MukB and most bacterial SMC proteins form anti-parallel

homo-dimers (Melby *et al.*, 1998). An anti-parallel dimer should be able to bind DNA at both ends. As all SMC proteins, MukB is huge (170 kDa; Niki *et al.*, 1992). The coils are around 65 nm long and the hinge is flexible: the angle between two arms may range from 1° to 180° and possibly beyond (Melby *et al.*, 1998). MukB should be able to operate on two strands of DNA 100 nm apart. This is a large distance compared to the size of *E. coli* (fig. 1.6c).

1.1.5.5 Other nucleoid-associated proteins

Although SMC proteins are now considered to play an important role in structuring the bacterial genome, about ten other DNA-binding proteins are traditionally associated with nucleoid structure (Azam and Ishihama, 1999). I will introduce the most abundant ones briefly (for a review see Pettijohn, 1996).

The five most abundant nucleoid-associated proteins in *E. coli* growing exponentially in rich medium are (in order of abundance): Fis, Hfq, HU, and the closely homologous proteins StpA and H-NS (Ali Azam *et al.*, 1999). Their relative quantities change with growth phase (Ali Azam *et al.*, 1999). In stationary phase, the level of Fis drops below detectable levels, whilst the proteins Dps and IHF, present at low levels in exponential phase, become the first and second most abundant proteins. The order of DNA binding affinity was determined to be: HU (K_d = 25 nM), Hfq, Lrp, CpbB, Fis, H-NS, StpA, CpbA, IciA, and Hfq/Dps (K_d = 250 nM) (Ali Azam *et al.*, 1999).

Fis is a small DNA-binding protein that probably influences DNA topology by regulation of topo-isomerase activity and by modulating super helical density (Schneider *et al.*, 1999; Schneider *et al.*, 1997). Hfq (*h*ost *f*actor for integration of phage Q) binds both DNA and RNA (Kajitani *et al.*, 1994). It binds preferentially to curved DNA, independent of the underlying nucleotide sequence (Ali Azam *et al.*, 1999). HU was initially called 'histone-like', because SV40 DNA together with HU and a 'nicking-closing' enzyme produced histone-like beads under the electron

microscope (Rouviere-Yaniv *et al.*, 1979). However, unlike true histones, the beads were unstable. HU is probably more analogous to the eukaryotic HMG-1 and HMG-2 proteins (Oberto *et al.*, 1994). HMG stands for high mobility group proteins, a group of minor groove-binding proteins capable of bending DNA (see Bewley *et al.*, 1998 and references therein). HU bends DNA and binds preferentially to bent or 'kinked' DNA also independent of the underlying sequence. The saturation level is 1 HU per 9 bp, but this level is never reached *in vivo*. Its function is thought to be in facilitating protein-DNA interactions, for instance of proteins that regulate expression of genes by binding to the promoter region (see Pettijohn, 1996 and references therein). H-NS, another 'histone-like' protein, has little sequence specificity, but preferentially binds to bent DNA. It affects the transcription of over 35 genes or operons (see Atlung and Ingmer, 1997). IHF (*i*ntegration *h*ost *f*actor) has 30% amino-acid-sequence identity with HU, but the arm of the protein that binds DNA is similar in sequence to the eukaryotic TATA-binding factor; HU is probably not sequence-independent (Nash and Granston, 1991; for information about the TATA binding factor see for instance Bewley *et al.*, 1998). There is evidence that IHF indirectly activates transcription by optimising the bending angle between the two arms of bent DNA (Engelhorn and Geiselmann, 1998). HU, IHF, and H-NS probably overlap in function; mutations in any of the proteins produced non-lethal phenotypes, but a triple mutant could not be produced (Yasuzawa *et al.*, 1992).

1.1.5.6 Histone-like proteins and true histones

Nucleoid-associated proteins in prokaryotes are often coined 'histone-like' proteins (previous section and Schmid, 1990), because they are basic, small, abundant, and bind DNA in a sequence-independent manner. However, proteins capable of organising DNA as stable solenoids (fig. 1.2c), a defining feature of true histones, have only been found in eukaryotes and archaea-bacteria, but not in eubacteria (Li *et al.*, 1999; Sandman and Reeve, 2000).

In eukaryotes, DNA is wrapped around complexes of eight histones, forming beads called nucleosomes. Nucleosomes constrain negative supercoiling, which protects DNA from ligands. When DNA is released from the histone-complex the negative supercoiling promotes unwinding (Bates and Maxwell, 1993d and references therein), which makes the DNA more susceptible to ligands. A fibre consisting of nucleosomes has a diameter of 10 nm. This '10 nm fibre' is further condensed, first in a 30 nm fibre, then into 'chromonema', which are fibres with a diameter of around 100 nm; in metaphase chromosomes DNA is further packed into 200 nm and 400 nm fibres (Visser, 1999 and references therein). During interphase, DNA is present in various degrees of compaction. The most compact form of DNA folding is reached in metaphase, although the average difference between the degree of compaction of DNA between interphase and metaphase is probably limited (see for instance Manders *et al.*, 1999). In HeLa cells, the amount of de-compaction of chromatin from anaphase to G1 was quantified from 3-D time-lapse images of GFP-tagged histones. All chromatin de-condensed by at least a factor of 3 (the average factor was 5), and no substantial re-arrangement of chromatin domains took place (Manders *et al.*, in preparation).

1.1.5.7 Topo-isomerases

Topo-isomerases have already been mentioned in the previous text, because of their role in DNA supercoiling. They change the linking number of DNA, hence, its topology, by transiently cutting DNA (Bates and Maxwell, 1993c; Wang, 1996 and references therein). Two types of topo-isomerases have been defined: type I topo-isomerases cut one strand of the double-helix to allow one strand to pass the other; type II topo-isomerases make a double-strand break to allow one helix to pass another (e.g. Brown and Cozzarelli, 1979).

Most topo-isomerases use the free energy contained in supercoiled DNA. An exception is DNA-gyrase, a type II topo-isomerase found only in prokaryotes. Although the enzyme is capable of relaxing negatively supercoiled DNA by increasing the linking number, it can also introduce

negative supercoiling in relaxed DNA by reducing the linking number of DNA (Gellert *et al.*, 1976). Because the latter is energetically unfavourable, it does this at the expense of ATP. Thus, in prokaryotes DNA-gyrase keeps DNA under negative torsional stress. In eukaryotes, no such topo-isomerases have been found. They may be obsolete in eukaryotes, because the combined action of DNA-organising proteins, such as histones and relaxing topo-isomerases, can also produce negative supercoiling upon protein dissociation. Chromatin remodelling in eukaryotes is achieved by ATP-dependent remodelling-enzymes, which are part of multi-subunit complexes (Peterson, 2000). Remarkably, bacterial DNA is wrapped, be it transiently, around DNA-gyrase reminiscent of the way DNA is wrapped around the histone octamere in eukaryotic nucleosomes (Bates and Maxwell, 1993c).

Even though DNA gyrase is an essential enzyme, its 'control' over supercoiling in *E. coli* is limited. Jensen and co-workers were able to modulate the expression level of DNA gyrase and quantified its amount of control as the percentage change in the amount of supercoiling in response to a 1% change in the level of DNA-gyrase expression (Jensen *et al.*, 1999). They found that the amount of control was only 0.2 % when expression of DNA gyrase was modulated around wild-type levels (see Jensen *et al.*, 1999 for details). Apparently, DNA supercoiling is subtly controlled in *E. coli*.

1.2 DNA organisation and the cell cycle

The organisation of DNA in a living cell is not static in time. Many processes influence DNA organisation at different times during the cell cycle. One of the more prominent ones is DNA replication. Thus, to understand intracellular DNA-organisation in *E. coli,* we need to know the timing of DNA-replication within the cell cycle.

1.2.1 Timing of replication within the cell cycle of E. coli

The relationship between the replication cycle and the cell-division cycle

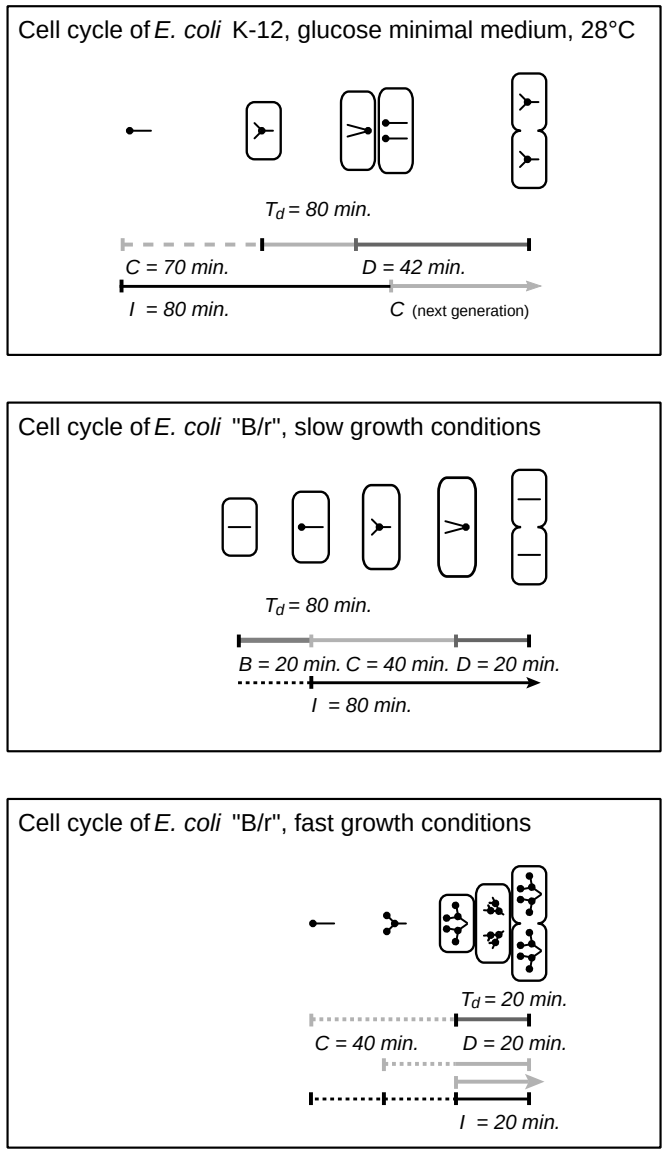


Figure 1.7 - Cell cycles in different strains of *E. coli* and under different growth conditions

T_d : cell doubling time; *C*: time to replicate one chromosome equivalent; *D*: time from termination of replication to cell division; *I* : inter-initiation time. Every *I*+*C*+*D* minutes a new *I*+*C*+*D* cycle starts (note that $I=T_d$ under steady state conditions). Top: cell cycle of *E. coli* strain K-12, derivative MC4100, grown steady state in minimal glucose medium at 28°C (conform the experiments described in this thesis; Huls *et al.*, 1999); Middle: "classic" cell cycle parameters in which *C* and *D* are more or less constant over a range of growth rates (measured in B/r strains); Bottom: multi-fork replication at a fast growth rate in the same B/r strain.

of *E. coli* can be described using the '*I*+*C*+*D* rule' (fig. 1.7; Helmstetter, 1996). *I* is the time required to achieve the capacity for initiation of replication, which is defined by the initiation mass of a cell; *C* is the time required to replicate one chromosome; *D* is the time in between termination of replication and cell division. Thus, when the initiation mass is reached, in *I* minutes, it takes *C*+*D* minutes from initiation of replication, via termination of replication, to cell division. Under steady-state conditions growth conditions are constant by definition. Consequently, *I*, *C*, and *D* remain constant: an *I*+*C*+*D* cycle starts every *I* minutes, and *I* equals the doubling time T_d. When $C+D \geq T_d$ then initiation of replication precedes cell birth. Thus, a cell is born with at least two origins, because the replication cycle is already underway. When $T_d < C$, then a new round of replication initiates before the previous one has terminated (multifork replication; fig. 1.7c). When $C+D \geq T_d$, then a single round of replication initiates $T_d - (C+D) = B$ minutes after cell birth, in which *B* is the time between cell birth and initiation of replication.

Cell-cycle parameters depend on growth conditions and on the type of strain (Bipatnath *et al.*, 1998; Helmstetter, 1996). To illustrate, figure 1.7 schematically depicts three cases: the first represents the K-12 strain used in our experiments, growing with a doubling time of T_d=80 min. (cf. chapters 2 and 3; Huls *et al.*, 1999); the second represents a different strain, growing with the same doubling time (T_d=80 min.), but with different values for *C* and *D* (cf. Cooper and Helmstetter, 1968); the third also represents this strain, but now growing with a different doubling time (T_d=20 min.) whilst *C* and *D* remain constant (cf. Cooper and Helmstetter, 1968). The third case exemplifies multi-fork replication. Helmstetter summarised cell-cycle parameters of three *E. coli* B/r strains and several *E. coli* K-12 strains grown at different growth rates (Table 1 in Helmstetter, 1996). In this list the cell cycle parameters of K-12 strains show a lot of variability, probably because the various K-12 strains are genetically and physiologically different. In the list of Helmstetter, B/r strains appear to have a more or less constant *C* and *D* period at growth

rates below 60 minutes, but more recently Bipatnath and co-workers have shown that *C* may decrease gradually from *C*=70 minutes at a doubling time of 100 minutes to *C*=33 minutes at a doubling time of 20 minutes in both a K-12 and a B/r strain (Bipatnath *et al.*, 1998). Because of these uncertainties, we chose our strain and growth conditions such that we could use experimentally determined values for *C*, *D*, and T_d (chapters 2 and 3).

1.2.2 Differences in the timing of DNA replication and DNA segregation between prokaryotes and eukaryotes

With respect to the timing of replication within the cell cycle, prokaryotes and eukaryotes are distinctly different. In *E. coli*, for instance, DNA replication coincides with DNA segregation, and sister copies are segregated completely, whilst in eukaryotes, replication coincides with the alignment and cohesion of sister chromatids. The chromatids are separated later in the cell cycle, during mitosis (fig. 1.8). Moreover, bacteria have the ability to initiate a new round of replication before previous rounds of replication have finished. In fast growing bacteria, many 'generations' of replication cycles may take place simultaneously. For instance, a fast-growing bacterium could contain 16 or more copies of the origin of replication, even though there is only one origin per chromosome. Eukaryotes replicate all their DNA once and only once during S-phase. Despite the differences, some researchers propose that the mechanism behind DNA segregation in bacteria may be compared to the mechanism underlying mitosis in eukaryotes (e.g. Begg and Donachie, 1991; Lin *et al.*, 1997; Moller-Jensen *et al.*, 2000; Niki and Hiraga, 1998; Sharpe and Errington, 1999; Webb *et al.*, 1997; Wheeler and Shapiro, 1997). I feel, however, that it may be more useful to compare the initial separation of nascent DNA in prokaryotes to that in eukaryotes, thus during S-phase (see section 6.2.3).

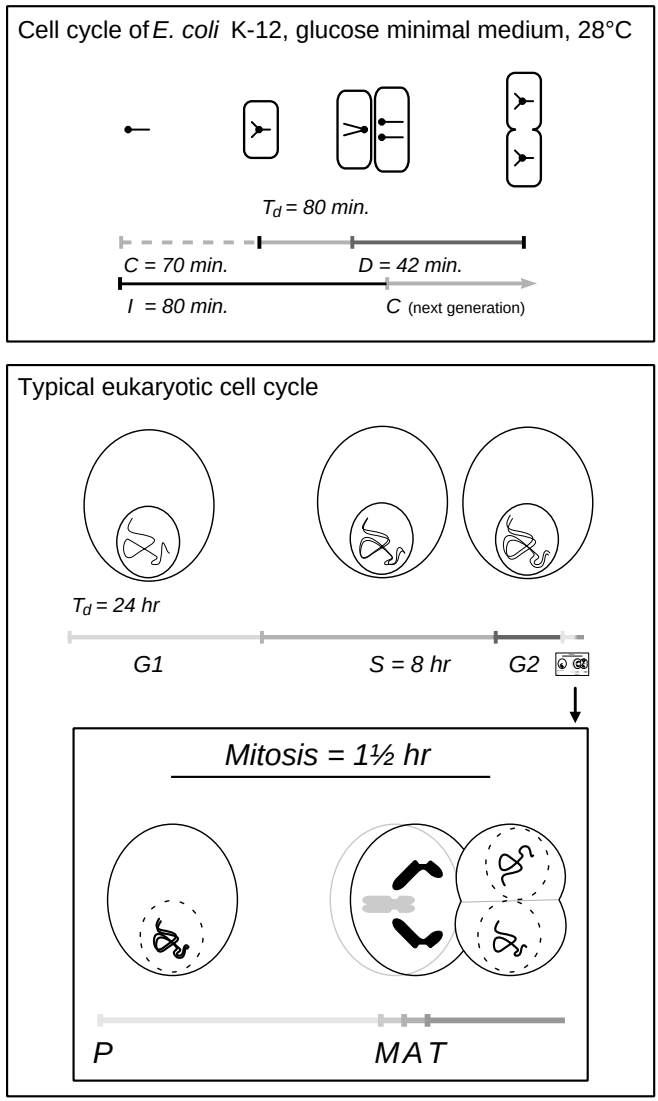


Figure 1.8 - Cell cycles of *E. .coli* K-12 and a typical eukaryote

Top: cell cycle of *E. coli* strain K-12 grown in minimal glucose medium at 28°C (see figure 1.7). Bottom: 'typical' eukaryotic cell cycle. For simplicity only one chromosome is depicted. *G1*: (gap) phase following cell division, chromosomes consist of one chromatid; *S*: phase in which all DNA is replicated (synthesised) exactly once; *G2*: (gap) phase preceding mitosis, chromosomes consist of two chromatids held together by cohesins. Mitosis is divided into four phases: *P*: during prophase chromosomes condense maximally; *M*: during metaphase chromosomes align in the metaphase plate in the middle of the cell by pulling and pushing of microtubules; *A*: during anaphase chromatids are separated and segregated to opposite cell poles by microtubules; *T*: during telophase chromosomes decondense before cell division. Note the difference in time scale between pro- and eukaryotes. Also note that in prokaryotes DNA replication and DNA segregation are not separated in time as in eukaryotes (*S*-phase and mitosis respectively).

1.3 Studying the organisation of DNA in *E. coli*

The size of many of the structures introduced in this chapter, are below the resolution of a light microscope (~250 nm; see chapter 3). Nucleoids appear as large, somewhat lobular, structures that segregate gradually in line with cell elongation (van Helvoort and Woldringh, 1994). Two methods are currently used that enable the study of sub-nucleoid DNA regions using fluorescence light-microscopy.

1.3.1 FISH

Fluorescent *in situ* hybridisation (FISH) relies on hybridising a DNA-region of interest with a complementary DNA probe. The DNA of the probe is conjugated with a fluorescent dye or with a molecule that can be detected by fluorescently labelled antibodies. FISH is not a gentle method, it requires cells to be fixed and made permeable to the DNA probe, and the samples need to be heated to about 80 °C in the presence of formamide or incubated in 0.1 M NaOH in order to 'melt' the DNA, which is needed for hybridisation with probe DNA. On the other hand, cells can be cultured under steady-state conditions, which is a prerequisite for reconstructing DNA-organisation as a function of the cell cycle.

1.3.2 GFP

The second method uses green-*f*luorescent *p*rotein (GFP) to tag a DNA region (Margolin, 2000; Robinett *et al.*, 1996). GFP is fused to a DNA-binding protein that binds specifically to a DNA region inserted at a specific position in the genome. For instance, we describe in chapter 4 preliminary results of a time-lapse experiment with GFP fused to LacI, which binds to a cassette of *lacO* sequences inserted near the origin of replication of the *E. coli* chromosome. The same system had been used earlier in *B. subtilis* (Webb *et al.*, 1998) and in *E. coli* (Gordon *et al.*, 1997). The big advantage of this method is that it can be used to monitor the dynamics of DNA in living cells. However, for time-lapse microscopy cells need to be on a slide under the microscope. These are non-ideal

circumstances and certainly not 'steady-state'. Moreover, the amount of GFP-fusion protein is difficult to control. The system that was used in these cases can lead to heterogeneity with respect to the amount of GFP per cell, and possibly to artificial clustering of an excess of fusion protein (forming inclusion bodies; see section 5.2.3). In summary, FISH is the method of choice if information about average movement of DNA as a function of the cell cycle is required. GFP studies are better suited for studies on DNA dynamics, although one should be cautious about extrapolating to a (steady-state) population.

1.4 Outline of this thesis

1.4.1 Outline in brief

In this thesis, an analysis of DNA segregation in *E. coli* as a function of the cell cycle is described. We have studied the average intracellular position of three regions of the *E. coli* chromosome by FISH (chapters 2 and 3), and used computer simulation to test if three particular models explain our FISH data sufficiently. Because we found that these models did not fully explain our data, I re-examine the assumptions on which they were based and propose a new model, based on diffusion of DNA (chapter 6). Finally, after having introduced factors that have an influence on DNA structure in this chapter, I discuss in chapter 6 how some of these factors may be involved in DNA segregation, and take a brief look at how we can make a model incorporating such factors. In the same chapter I give my view on comparisons that are made between DNA segregation in bacteria and in eukaryotes.

1.4.2 Chapter 2

In chapter 2 the position of the origin of replication as a function of cell length is described. Regression analysis was performed on sub-sets of data (distances from a cell pole) to find average positions of *oriC* as a function of cell length. Previously measured values of *C* and *D* were used

to select these sub-sets from the total data set. We concluded that the average distance from a cell pole to *oriC* remains constant, which is consistent with a model for gradual separation along with nucleoid segregation.

1.4.3 Chapter 3

In chapter 3 we extended the analysis of chapter 2 with two more regions on the *E. coli* chromosome. In this case our regression analysis incorporated all data-points. We further emphasised that when doing FISH we should always consider the possibility that not all DNA regions in a cell are labelled. This led us to propose a model that describes *average* positions of DNA regions as a function of cell length.

1.4.4 Chapter 4

In chapter 4 models based on interpretations of FISH data and GFP data were evaluated by computer simulation. We simulated FISH data by computer and compared the results with the experimental data. The results of the 'random model' came closest to the experimental results, probably because the amount of variation in the experimental data was large; a large amount of variation is an intrinsic property of the random model. We argued that the three studied DNA regions (*oriC*, *ftsQAZ*, and *minB*) were not positioned completely random, because the measured distributions of these regions were different. The random model would not predict this. We concluded that movement of a DNA-region might be the result of various factors. We proposed that one of these factors could be diffusion of DNA within a confined region (cf. Marshall *et al.*, 1997).

1.4.5 Chapter 5

In chapter 5 I summarise the results of our FISH experiments, and carefully (re-) examine assumptions on which the models that we tested in chapter 4 were based. Incorrectness of these assumptions may explain some of the variation in our FISH data, and possibly why the distribution

of the simulated data-points differed from the measured distribution.

1.4.6 Chapter 6

In chapter 6 I take another look at some of the factors that influence DNA structure in *E. coli*. In particular, I discuss how they might be involved in a mechanism that segregates replicated DNA. Because such a mechanism is probably based on non-linear interactions between many factors, it is difficult to define a proper model, but it can be even more difficult to predict what kind of data this model would produce. Nevertheless, I discuss some methods that could get us closer to finding a model that produces data similar to our FISH data. I also discuss how a comparison between prokaryotes and eukaryotes could help us to find the mechanism behind DNA segregation in bacteria. Unfortunately, I have to conclude that the mechanisms involved are probably fundamentally different. Finally, I conclude that a new model of DNA segregation in *E. coli* could be based on diffusion of DNA.

1.5 References

Ali Azam, T., A. Iwata, A. Nishimura, S. Ueda, and A. Ishihama. 1999. Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol* 181: 6361-6370.

Atlung, T., and H. Ingmer. 1997. H-NS: a modulator of environmentally regulated gene expression. *Mol Microbiol* 24: 7-17.

Azam, T. A., and A. Ishihama. 1999. Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J Biol Chem* 274: 33105-33113.

Bates, A. D., and A. Maxwell. 1993a. Chapter 1: DNA structure. Pages 1-16. *DNA topology*. IRL Press at Oxford University Press, Oxford ; New York.

...1993b. Chapter 2: DNA supercoiling. Pages 17-46. *DNA topology*. IRL Press at Oxford University Press, Oxford ; New York.

...1993c. Chapter 5: DNA topoisomerases. Pages 73-82. *DNA topology*. IRL Press at Oxford University Press, Oxford ; New York.

...1993d. Chapter 6: Biological consequences of DNA topology. Pages 83-106. *DNA topology*. IRL Press at Oxford University Press, Oxford ; New York.

...1993e. *DNA topology*. IRL Press at Oxford University Press, Oxford ; New York.

Begg, K. J., and W. D. Donachie. 1991. Experiments on chromosome separation and positioning in *Escherichia coli*. *New Biol* 3: 475-486.

Bewley, C. A., A. M. Gronenborn, and G. M. Clore. 1998. Minor groove-binding architectural proteins: structure, function, and DNA recognition. *Annu Rev Biophys Biomol Struct* 27: 105-131.

Bipatnath, M., P. P. Dennis, and H. Bremer. 1998. Initiation and velocity of chromosome replication in *Escherichia coli* B/r and K-12. *J Bacteriol* 180: 265-273.

Blattner, F. R., G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* 277: 1453-1474.

Boles, T. C., J. H. White, and N. R. Cozzarelli. 1990. Structure of plectonemically supercoiled DNA. *J Mol Biol* 213: 931-951.

Brown, P. O., and N. R. Cozzarelli. 1979. A sign inversion mechanism for enzymatic supercoiling of DNA. *Science* 206: 1081-1083.

Chen, D., R. Bowater, C. J. Dorman, and D. M. Lilley. 1992. Activity of a plasmid-borne leu-500 promoter depends on the transcription and translation of an adjacent gene. *Proc Natl Acad Sci USA* 89: 8784-8788.

Chen, D., R. P. Bowater, and D. M. Lilley. 1993. Activation of the leu-500 promoter: a topological domain generated by divergent transcription in a plasmid. *Biochemistry* 32: 13162-13170.

Cook, P. R. 1989. The nucleoskeleton and the topology of transcription. *Eur J Biochem* 185: 487-501.

...1999. The organization of replication and transcription. *Science* 284: 1790-1795.

Cooper, S., and C. E. Helmstetter. 1968. Chromosome replication and the division cycle of *Escherichia coli* B/r. *J Mol Biol* 31: 519-540.

Dasgupta, S., S. Maisnier-Patin, and K. Nordstrom. 2000. New genes with old modus operandi. The connection between supercoiling and partitioning of DNA in *Escherichia coli*. *EMBO Rep* 1: 323-327.

Den Blaauwen, T. submitted. *Mol Microbiol*.

Drew, H. R., J. R. Weeks, and A. A. Travers. 1985. Negative supercoiling induces spontaneous unwinding of a bacterial promoter. *Embo J* 4: 1025-1032.

Dunn, T. M., S. Hahn, S. Ogden, and R. F. Schleif. 1984. An operator at -280 base pairs that is required for repression of *araBAD* operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. *Proc Natl Acad Sci U S A* 81: 5017-5020.

Engelhorn, M., and J. Geiselmann. 1998. Maximal transcriptional activation by the IHF protein of *Escherichia coli* depends on optimal DNA bending by the activator. *Mol Microbiol* 30: 431-441.

Gellert, M., K. Mizuuchi, M. H. O'Dea, and H. A. Nash. 1976. DNA gyrase: an enzyme that introduces superhelical turns into DNA. *Proc Natl Acad Sci U S A* 73: 3872-3876.

Gordon, G. S., D. Sitnikov, C. D. Webb, A. Teleman, A. Straight, R. Losick, A. W.

Murray, and A. Wright. 1997. Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. *Cell* **90**: 1113-1121.

Graumann, P. L. 2001. SMC proteins in bacteria: Condensation motors for chromosome segregation? *Biochimie* **83**: 53-59.

Helmstetter, C. E. 1996. Timing of synthetic activities in the cell cycle. Pages 1627-1639 in Neidhardt, F. C. and Curtiss, R., eds. *Escherichia coli and Salmonella : cellular and molecular biology*. ASM Press, Washington, D.C.

Higgins, N. P. 1999. DNA supercoiling and its consequences for chromosome structure and function in Charlebois, R. L., ed. *Organization of the prokaryotic genome*. ASM Press, Washington, D.C.

Holmes, V. F., and N. R. Cozzarelli. 2000. Closing the ring: links between SMC proteins and chromosome partitioning, condensation, and supercoiling. *Proc Natl Acad Sci U S A* **97**: 1322-1324.

Hsieh, L. S., J. Rouviere-Yaniv, and K. Drlica. 1991. Bacterial DNA supercoiling and [ATP]/[ADP] ratio: changes associated with salt shock. *J Bacteriol* 173: 3914-3917.

Huls, P. G., N. O. Vischer, and C. L. Woldringh. 1999. Delayed nucleoid segregation in *Escherichia coli*. *Mol Microbiol* 33: 959-970.

Jensen, P. R., C. C. Van Der Weijden, L. B. Jensen, H. V. Westerhoff, and J. L. Snoep. 1999. Extensive regulation compromises the extent to which DNA gyrase controls DNA supercoiling and growth rate of *Escherichia coli*. *Eur J Biochem* 266: 865-877.

Jessberger, R., C. Frei, and S. M. Gasser. 1998. Chromosome dynamics: the SMC protein family. *Curr Opin Genet Dev* **8**: 254-259.

Jumas-Bilak, E., S. Michaux-Charachon, G. Bourg, M. Ramuz, and A. Allardet-Servent. 1998. Unconventional genomic organization in the alpha subgroup of the Proteobacteria. *J Bacteriol* 180: 2749-2755.

Kajitani, M., A. Kato, A. Wada, Y. Inokuchi, and A. Ishihama. 1994. Regulation of the *Escherichia coli hfq* gene encoding the host factor for phage Q beta. *J Bacteriol* 176: 531-534.

Kido, M., K. Yamanaka, T. Mitani, H. Niki, T. Ogura, and S. Hiraga. 1996. RNase E polypeptides lacking a carboxyl-terminal half suppress a *mukB* mutation in *Escherichia coli*. *J Bacteriol* 178: 3917-3925.

Kimura, K., and T. Hirano. 1997. ATP-dependent positive supercoiling of DNA by 13S condensin: a biochemical implication for chromosome condensation. *Cell* **90**: 625-634.

Kimura, K., V. V. Rybenkov, N. J. Crisona, T. Hirano, and N. R. Cozzarelli. 1999. 13S condensin actively reconfigures DNA by introducing global positive writhe: implications for chromosome condensation. *Cell* **98**: 239-248.

Koppes, L. J., C. L. Woldringh, and N. Nanninga. 1999. *Escherichia coli* contains a DNA replication compartment in the cell center. *Biochimie* **81**: 803-810.

Kornberg, A., and T. A. Baker. 1992a. Genome origins. Pages 511-552. *DNA replication*. W.H. Freeman, New York.

...1992b. Replication mechanisms and operations. Pages 471-510. *DNA replication*. W.H. Freeman, New York.

Lemon, K. P., and A. D. Grossman. 1998. Localization of bacterial DNA polymerase: evidence for a factory model of replication. *Science* 282: 1516-1519.

. 2000. Movement of replicating DNA through a stationary replisome. *Mol Cell* 6: 1321-1330.

Li, J. Y., B. Arnold-Schulz-Gahmen, and E. Kellenberger. 1999. Histones and histone-like DNA-binding proteins: correlations between structural differences, properties and functions. *Microbiology* 145: 1-2.

Lin, D. C., P. A. Levin, and A. D. Grossman. 1997. Bipolar localization of a chromosome partition protein in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* 94: 4721-4726.

Liu, L. F., and J. C. Wang. 1987. Supercoiling of the DNA template during transcription. *Proc Natl Acad Sci U S A* 84: 7024-7027.

Manders, E. M., H. Kimura, and P. R. Cook. 1999. Direct imaging of DNA in living cells reveals the dynamics of chromosome formation. *J Cell Biol* 144: 813-821.

Manders, E. M., A. E. Visser, A. Koppen, A. d. Leeuw, R. v. Liere, G. J. Brakenhoff, and R. v. Driel. in preparation. Chromatin dynamics during the formation of the interphase nucleus.

Margolin, W. 2000. Green fluorescent protein as a reporter for macromolecular localization in bacterial cells. *Methods* 20: 62-72.

Maritan, A., C. Micheletti, A. Trovato, and J. R. Banavar. 2000. Optimal shapes of compact strings. *Nature* 406: 287-290.

Marshall, W. F., A. Straight, J. F. Marko, J. Swedlow, A. Dernburg, A. Belmont, A. W. Murray, D. A. Agard, and J. W. Sedat. 1997. Interphase chromosomes undergo constrained diffusional motion in living cells. *Curr Biol* 7: 930-939.

Martin, K., L. Huo, and R. F. Schleif. 1986. The DNA loop model for ara repression: AraC protein occupies the proposed loop sites in vivo and repression-negative mutations lie in these same sites. *Proc Natl Acad Sci U S A* 83: 3654-3658.

Mason, J. M., and D. M. Powelson. 1956. Nuclear division as observed in live bacteria by a new technique. *J Bacteriol* 71: 474-479.

Melby, T. E., C. N. Ciampaglio, G. Briscoe, and H. P. Erickson. 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.

Moller-Jensen, J., R. B. Jensen, and K. Gerdes. 2000. Plasmid and chromosome segregation in prokaryotes. *Trends Microbiol* 8: 313-320.

Nakamura, T. M., J. P. Cooper, and T. R. Cech. 1998. Two modes of survival of fission yeast without telomerase. *Science* 282: 493-496.

Nash, H. A., and A. E. Granston. 1991. Similarity between the DNA-binding domains of IHF protein and TFIID protein. *Cell* 67: 1037-1038.

Niki, H., and S. Hiraga. 1998. Polar localization of the replication origin and terminus in *Escherichia coli* nucleoids during chromosome partitioning. *Genes Dev* 12: 1036-1045.

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

5101-5109.

Niki, H., A. Jaffe, R. Imamura, T. Ogura, and S. Hiraga. 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.

Oberto, J., K. Drlica, and J. Rouviere-Yaniv. 1994. Histones, HMG, HU, IHF: Meme combat. *Biochimie* 76: 901-908.

Odijk, T. 1998. Osmotic compaction of supercoiled DNA into a bacterial nucleoid. *Biophys Chem* 73: 23-29.

Pedersen, A. G., L. J. Jensen, S. Brunak, H. H. Staerfeldt, and D. W. Ussery. 2000. A DNA structural atlas for *Escherichia coli*. *J Mol Biol* 299: 907-930.

Peterson, C. L. 2000. ATP-dependent chromatin remodeling: going mobile. *FEBS Lett* 476: 68-72.

Pettijohn, D. E. 1996. The Nucleoid. Pages 158-166 in Neidhardt, F. C. and Curtiss, R., eds. *Escherichia coli and Salmonella : cellular and molecular biology*. ASM Press, Washington, D.C.

Robinett, C. C., A. Straight, G. Li, C. Willhelm, G. Sudlow, A. Murray, and A. S. Belmont. 1996. In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J Cell Biol* 135: 1685-1700.

Rouviere-Yaniv, J., M. Yaniv, and J. E. Germond. 1979. *E. coli* DNA binding protein HU forms nucleosomelike structure with circular double-stranded DNA. *Cell* 17: 265-274.

Sandman, K., and J. N. Reeve. 2000. Structure and functional relationships of archaeal and eukaryal histones and nucleosomes. *Arch Microbiol* 173: 165-169.

Sawitzke, J. A., and S. Austin. 2000. Suppression of chromosome segregation defects of *Escherichia coli muk* mutants by mutations in topoisomerase I. *Proc Natl Acad Sci U S A* 97: 1671-1676.

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

Schneider, R., A. Travers, T. Kutateladze, and G. Muskhelishvili. 1999. A DNA architectural protein couples cellular physiology and DNA topology in *Escherichia coli*. *Mol Microbiol* 34: 953-964.

Schneider, R., A. Travers, and G. Muskhelishvili. 1997. FIS modulates growth phase-dependent topological transitions of DNA in *Escherichia coli*. *Mol Microbiol* 26: 519-530.

Seabold, R. R., and R. F. Schleif. 1998. Apo-AraC actively seeks to loop. *J Mol Biol* 278: 529-538.

Sharpe, M. E., and J. Errington. 1999. Upheaval in the bacterial nucleoid. An active chromosome segregation mechanism. *Trends Genet* 15: 70-74.

St. Jean, A. 1999. Local genetic context, supercoiling, and gene expression in Charlebois, R. L., ed. *Organization of the prokaryotic genome*. ASM Press, Washington, D.C.

ten Heggeler-Bordier, B., W. Wahli, M. Adrian, A. Stasiak, and J. Dubochet. 1992. The apical localization of transcribing RNA polymerases on supercoiled DNA prevents their rotation around the template. *Embo J* 11: 667-672.

Ussery, D. 2000. What is a "DNA structural atlas"?

http://www.cbs.dtu.dk/services/GenomeAtlas/background.html

van Helvoort, J. M. L. M., and C. L. Woldringh. 1994. Nucleoid partitioning in *Escherichia coli* during steady-state growth and upon recovery from chloramphenicol treatment. *Mol Microbiol* 13: 577-583.

van Workum, M., S. J. van Dooren, N. Oldenburg, D. Molenaar, P. R. Jensen, J. L. Snoep, and H. V. Westerhoff. 1996. DNA supercoiling depends on the phosphorylation potential in *Escherichia coli*. *Mol Microbiol* 20: 351-360.

Visser, A. E. 1999. Organization of chromosomes in the interphase cell nucleus - general discussion. Pages 79-90. *Organization of chromosomes in the interphase cell nucleus*. University of Amsterdam, Centrum voor microscopisch onderzoek (CMO).

Wang, J. C. 1996. DNA topoisomerases. *Annu Rev Biochem* 65: 635-692.

Wasserman, S. A., and N. R. Cozzarelli. 1986. Biochemical topology: applications to DNA recombination and replication. *Science* 232: 951-960.

Watson, J. D., and F. H. C. Crick. 1953. Molecular structure of nucleic acids. *Nature* 171: 373.

Webb, C. D., P. L. Graumann, J. A. Kahana, A. A. Teleman, P. A. Silver, and R. Losick. 1998. Use of time-lapse microscopy to visualize rapid movement of the replication origin region of the chromosome during the cell cycle in *Bacillus subtilis*. *Mol Microbiol* 28: 883-892.

Webb, C. D., A. Teleman, S. Gordon, A. Straight, A. Belmont, D. C. Lin, A. D. Grossman, A. Wright, and R. Losick. 1997. Bipolar localization of the replication origin regions of chromosomes in vegetative and sporulating cells of *B. subtilis*. *Cell* 88: 667-674.

Weitao, T., S. Dasgupta, and K. Nordstrom. 2000a. Role of the *mukB* gene in chromosome and plasmid partition in *Escherichia coli*. *Mol Microbiol* 38: 392-400.

Weitao, T., K. Nordstrom, and S. Dasgupta. 2000b. *Escherichia coli* cell cycle control genes affect chromosome superhelicity. *EMBO Rep* 1: 494-499.

Westerhoff, H. V., M. H. O'Dea, A. Maxwell, and M. Gellert. 1988. DNA supercoiling by DNA gyrase. A static head analysis. *Cell Biophys* 12: 157-181.

Wheeler, R. T., and L. Shapiro. 1997. Bacterial chromosome segregation: is there a mitotic apparatus? *Cell* 88: 577-579.

Woldringh, C. L., P. R. Jensen, and H. V. Westerhoff. 1995. Structure and partitioning of bacterial DNA: determined by a balance of compaction and expansion forces? *FEMS Microbiol Lett* 131: 235-242.

Wu, R., and T. Wu. 1996. A novel intact circular dsDNA supercoil. *Bull Math Biol* 58: 1171-1185.

Yamazoe, M., T. Onogi, Y. Sunako, H. Niki, K. Yamanaka, T. Ichimura, and S. Hiraga. 1999. Complex formation of MukB, MukE and MukF proteins involved in chromosome partitioning in *Escherichia coli*. *Embo J* 18: 5873-5884.

Yasuzawa, K., N. Hayashi, N. Goshima, K. Kohno, F. Imamoto, and Y. Kano. 1992. Histone-like proteins are required for cell growth and constraint of supercoils in DNA. *Gene* 122: 9-15.

Chapter 2

CELLULAR LOCALISATION OF *oriC* DURING THE CELL CYCLE OF *ESCHERICHIA COLI* AS ANALYSED BY FLUORESCENT *IN SITU* HYBRIDISATION

Marco Roos, Anton B. M. van Geel, Mirjam E. G. Aarsman, Jacques T. M. Veuskens, Conrad L. Woldringh and Nanne Nanninga

Institute for Molecular Cell Biology, BioCentrum Amsterdam, University of Amsterdam, The Netherlands

Published in Biochimie 81: 797–802 (1999)

2.1 Abstract

The origin of replication of *Escherichia coli*, *oriC*, has been labelled by fluorescent *in situ* hybridisation (FISH). The *E. coli* K12 strain was grown under steady state conditions with a doubling time of 79 min at 28 °C. Under these growth conditions DNA replication starts in the previous cell cycle at -33 min. At birth cells possess two origins, which are visible as two separated foci in fully labelled cells. The number of foci increased with cell length. The distance of foci from the nearest cell pole has been

measured in various length classes. The data suggest (*i*) that the two most outwardly located foci keep a constant distance to the cell pole and they therefore move apart gradually in line with cell elongation and (*ii*) that at the initiation of DNA replication the labelled origins occur near the centre of prospective daughter cells.

2.2 Introduction

The *Escherichia coli* chromosome has a circumference of about 1.6 mm, whereas the width of the cell is in the order of 1 μm. This indicates that considerable compaction is required to fold the bacterial chromosome into the so-called nucleoid (reviews: Pettijohn, 1996; Woldringh and Odijk, 1999). The extreme smallness of the nucleoid has to a large extent precluded the elucidation of its organisation by microscopic techniques, even by high-resolution transmission confocal-microscopy (Valkenburg *et al.*, 1985). Recent advances in fluorescent labelling techniques, however, allow the visualisation of sub-nucleoid regions during DNA replication and cell elongation. Microscopic studies have led to the idea that the nucleoid is a polarised structure and, dependent on the replication status of the bacterial chromosome, sub-nucleoid regions occupy defined positions within the nucleoid and within the cell (Glaser *et al.*, 1997; Gordon *et al.*, 1997; Lewis and Errington, 1997; Lin *et al.*, 1997; Mohl and Gober, 1997; Niki and Hiraga, 1998; Webb *et al.*, 1997). Sub-nucleoid regions appear to move independently of cell elongation (Gordon *et al.*, 1997; Webb *et al.*, 1998), which has been interpreted to mean that motor-like proteins are involved in partitioning of the bacterial chromosome (Hiraga, 1992; Lewis and Errington, 1997; Niki and Hiraga, 1998).

The nucleoid as a whole is confined to a specific cellular position. For instance, during cell elongation of *E. coli*, the distance between the nucleoid border and the nearest cell pole remains constant (van Helvoort and Woldringh, 1994). Similarly, during a large part of the DNA replication cycle, there is a fixed distance between the polarly located origin of replication and the cell pole (Niki and Hiraga, 1998 and this paper).

Another important aspect relevant to the study of nucleoid organisation is that, in contrast to eukaryotic cells, the DNA replication-cycle is not always coincident with the cell division cycle. For instance, initiation of DNA replication can start in a previous cell division cycle. A special situation arises at fast growth, when the mass doubling time is smaller than the duration of DNA replication. In such a case, a new round of DNA replication is started on an already replicating chromosome (multifork replication, for a review see Helmstetter, 1996). Thus, to simplify the interpretation of sub-nucleoid microscopical data, it is important to use slowly growing cells and cells of which the temporal relationship between the DNA replication cycle and the cell division cycle is known.

Using such a defined culture we have carried out fluorescent *in situ* hybridisation (FISH) to label the origin of replication in an *E. coli* K12 strain growing with a doubling time of 79 min. We have addressed the following questions and we have compared the results with existing literature. (*i*) What is the temporal relationship between the DNA replication cycle and the cell division cycle, (*ii*) where are the replicating origins located in the newborn cell, (*iii*) where is the origin of replication located at initiation of replication, and finally, (*iv*) do the origins move within the nucleoid independently of cell elongation.

2.3 Materials and methods

2.3.1 Bacterial strains and growth conditions

Cells of *E. coli* strain LMC500 (MC4100 (*F-, araD139, (argF-lac) U169, deoC1, flbB5301, ptsF25, rbsRrelA1, rpsl150) lysA*) (Taschner *et al.*, 1988) were grown in steady-state in glucose minimal medium containing 6.33 g of $K_2HPO_4.3H_2O$, 2.95 g of KH_2PO_4, 1.05 g of $(NH_4)2SO_4$, 0.10 g of $MgSO_4.7H_2O$), 0.10 g of $MgSO_4.7H_2O$), 0.28 mg of $FeSO_4.7\ H_2O$, 7.1 mg of $Ca(NO_3)_2.4H_2O$, 4 mg of thiamine, 4 g of glucose and 50 µg lysine per liter (pH7.0) at 28 °C with a doubling time (T_d) of 79 min. Cells were

harvested at OD_{450} = 0.2 and fixed in 1ml 0.1% OsO_4 in TY-medium (1% tryptone, 0.5% yeast extract, 3mM NaOH, 0.5% NaCl) and stored at 4 °C.

2.3.2 Probe and probe labelling

We used an *oriC* probe obtained from plasmid pOC162 (kindly provided by dr. W. Messer, Max Planck Institut für Molekulare Genetik, Berlin, Germany). This plasmid contains the 248bp *oriC* region with 3kbp on each side of *oriC*. Because pOC162 contains no plasmid sequences homologues to chromosomal sequences, the whole construct was labelled and used as a probe for FISH. The plasmid was labelled with digoxiginin-11-dUTP (DIG, Boehringer, Mannheim, Germany) by nick translation as described by Volkers (Volkers *et al.*, 1988). DIG incorporation was checked by spot blotting and detection with mouse-anti-DIG alkaline phosphatase-conjugated antibodies and 5-bromo-4-chloro-3-indolylphate/nitroblue tetrazolium (BCIP/NBT) as substrate. The specificity of the probe was checked by Southern hybridisation on digested genomic DNA of strain MC4100 (Sambrook *et al.*, 1989).

2.3.3 Preparation of cells for fluorescence microscopy

Cells were washed in PBS (140mM NaCl, 27mM KCl, 10mM Na_2HPO_4 and 2mM KH_2PO_4, pH 7.2) and post-fixed in 0.5% formaldehyde and 0.04% glutaraldehyde in TBS (10mM Tris and 0.9% NaCl) for 15 min at room temperature. Fixed cells were centrifuged at 8,000 x g for 5 min and washed three times in PBS and subsequently incubated in 0.1% Triton X-100 in PBS for 45 min at room temperature. Cells were washed three times in PBS and incubated in PBS containing 100 mg/ml lysozyme and 5 mM EDTA (ethylene-diamine-tetra-acetate.$2H_2O$) for 45 min at 37 °C. After washing three times in PBS and once in 2xPBS/SSC (sodium salt citrate; 0.15M NaCl, 0.015M sodium citrate) the cells were incubated with 100 µl of 100 mg/ml RNase A (Boehringer, Mannheim, Germany) in 2xPBS/SSC for 60 min at 37 °C. After washing three times in PBS, 10 µl of cell suspension was applied to slides coated with 0.01% poly-L-lysine, covered with a coverslip (15 mm diameter) and left for 20 min at room temperature. Then the slides were washed in a coplin jar with 2xSSC and

put on top of coverslips with 10 µl of probe solution. The preparations were denatured for 5 min at 80 °C. Hybridisation occurred overnight in a moist chamber with 2xSSC and 50% formamide at 37 °C. Unbound probe was washed three times for 10 min at 40 °C in a coplin jar with 1xSSC and 50% formamide (pH7). Slides were subsequently washed in 2xSSC and TN buffer (100 mM Tris-HCl, pH 7.5, 150 mM NaCl).

Prior to immunofluorescence staining non-specific binding sites were blocked by incubating the cells in 0.5% (w/v) blocking reagens (Boehringer) in PBS for 60 min. at 37 °C. Digoxigenin-labelled probe detection was carried out in TNB with mouse-anti-digoxigenin (Boehringer 1333062), 1:500 in TNB, rabbit-anti-mouse conjugated Cy3 (Jackson ImmunoResearch Laboratories, Inc., West Grove, USA), 1:400 in TNB and goat-anti-rabbit conjugated Alexa 546 (Molecular Probes Inc., Eugene, USA), 1:600 in TNB. Slides with antibodies were incubated in a humid chamber at 37 °C for 60 min. Between incubations with antibodies, the slides were washed three times 5 min in TN with 0.05% (v/v) Tween 20. After immunodetection slides were rinsed in PBS/TNB and taken up in TE buffer (10 mM TrisHCl and 1mM EDTA). The preparations were mounted in 5 µl PBS with DAPI (500 ng/ml 2,4,diamidinophenyl-indole).

2.3.4 Microscopy and image analysis

Preparations were photographed with a cooled Princeton CCD camera mounted on an Olympus fluorescence microscope (BH2-RFC; Olympus, Tokyo, Japan) equipped with a 100 W mercury lamp. Images were made using the program IPlab spectrum 3.0 (Signal Analytics Co., Vienna, USA). Cells were first photographed in phase contrast mode, then with an Alexa filter (illuminated at 510-550 nm with an emission filter of 590 nm) and finally with a DAPI fluorescence filter (illuminated at 300-400 nm with an emission filter of 420 nm). Length, positions of foci and nucleoid length of each cell were determined interactively. Interactive measurements were performed as 'structured point collection' on a Macintosh 7100 computer using the public domain program Object-Image

(version 1.62n; Vischer, http://simon.bio.uva.nl/object-image.html; Vischer *et al.*, 1994) which is based on the NIH Image software (W. Rasband, http://rsp.info.nih.gov/nih-image).

2.3.5 Shift correction

The position of a single object imaged using different filter combinations can be slightly different. To correct for this shift in position, we determined the centre of gravity of spheres (1.01 μm diameter) that are visible in phase contrast, Alexa and DAPI images, respectively. Preparations of these spheres were made and from a single position on the slide ten times three images were made: a phase contrast image, an Alexa-image and a DAPI-image. As correction factors we used the average distance between the centre of gravity of corresponding spheres in a phase contrast image and an Alexa image and the average distance between corresponding spheres in a phase contrast image and a DAPI image.

2.4 Results

2.4.1 Temporal relationship between DNA replication cycle and cell division cycle

Under certain conditions *E. coli* initiates replication before cell division resulting in multiple copies of *oriC* during a cell cycle (Helmstetter, 1996). Therefore, we estimated the number of *oriC* copies per cell length using data from Huls *et al.* (Huls *et al.*, 1999). These authors determined the *C* and *D* period by measuring the DNA content per cell by integrated-density cytometry (Vischer *et al.*, 1999). We used the same strain growing under the same conditions as Huls *et al.* did (Huls *et al.*, 1999). Of this steady state culture with a T_d of 79 minutes, the *C* period was approximately 70 minutes and the *D* period was approximately 43 minutes (Huls *et al.*, 1999). A value for the *C* period of 70 min for cells grown at 28 °C is comparable to the period of 55 min for 37 °C-culture of *E. coli* cells as determined by Bipatnath *et al.* (Bipatnath *et al.*, 1998). By subtracting the sum of *C* and *D* from T_d, the time of initiation of DNA replication was calculated to be –33 minutes, indicating that replication

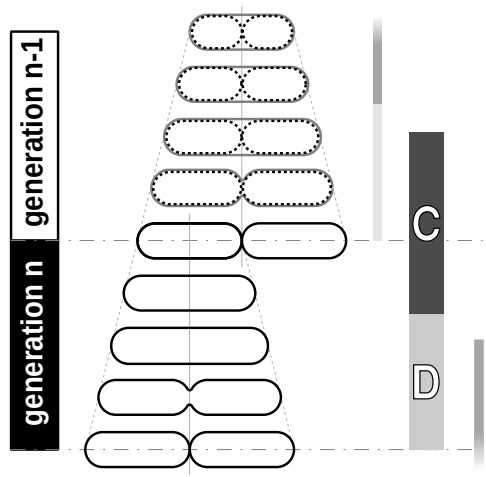


Figure 2.1 - Schematic representation of the cell cycle of *E. coli* MC4100 strain growing in steady state at 28°C

The bars on the left with 'generation n' and 'generation n-1' represent consecutive cell cycles. In the cells of generation n-1 dashed lines indicate prospective daughter cells. The bars on the right represent consecutive replication cycles: the leftmost bar corresponds to generation n-1, the thick middle bar corresponds to generation n and the rightmost bar corresponds to generation n+1. The cell doubling time was 79 min, the *C* period (dark grey) was approximately 70 minutes and the *D* period (light grey) was approximately 43 minutes (Huls *et al.*, 1999). Cell cycle times were converted to length using the formula: $l = e^{(t/T_d).\ln 2}$ (Helmstetter, 1996).

started well before cell birth. The cell cycle parameters are schematically depicted in figure 2.1.

2.4.2 Position of oriC in relation to cell length

To study the position of the origin of replication in relation to cell length we fixed cells from the aforementioned culture of strain MC4100 and used a 6kbp sequence containing *oriC* for a FISH experiment. Of 871 cells analysed, 692 (79%) contained fluorescent foci. We saw one focus in 312 cells (36%), two foci in 261 cells (30%), three foci in 77 cells (9%), four foci in 381 cells (4%), and five foci in 4 cells (<1%). In cells with fluorescent foci we measured the distance from each focus to the nearest cell pole (fig. 2.2). We found foci close to a cell pole and foci closer to mid-cell. Of foci close to a cell pole, the average distance between focus and pole was 0.5 μm (CV = 47%). As can be judged from the regression lines through the foci positions with respect to the two cell poles (fig. 2.3) the focus-to-pole distance remains largely constant. This indicates a gradual and continuous separation of the most outwardly positioned replicated origins.

2.4.3 Position of oriC in newborn cells and at initiation of DNA replication

The average length of a newborn cell was estimated using the cell length distribution of constricting cells (cf. Koppes, L. H. *et al.*, 1978). This value is indicated as an interrupted horizontal line in figure 2.3. At this length newborn cells have two copies of *oriC* (see above). The location of the two regression lines in the newborn cell indicates that the origins have separated and that their relative positions are at approximately $^1/_3$ and $^2/_3$ of cell length. Where is *oriC* located at initiation of DNA replication? From the calculated *C*-period, its temporal position in the cell cycle (fig. 2.1) and from the known relationship between cell age and cell length ($l = e^{t/T_d \cdot \ln 2}$, Helmstetter, 1996), initiation of DNA replication is expected to start at an average cell length of 1.1 μm. (Note that this is the length of the prospective daughter cell in the previous cell cycle; see fig. 2.1). One approach to assess *oriC* position at this cell length is to extrapolate the regression lines to the point where they cross (fig. 2.3). This occurs at a cell length of 0.82 μm. The estimated cell length at which replication initiated was within the 95% confidence interval around this length (i.e., between 0.5 and 1.2 μm). Though *oriC* thus seems to be located at the centre of the prospective daughter cell upon initiation of DNA replication, determination of its precise position is hampered by the limitations of the cytological technique.

2.5 Discussion

2.5.1 Replication starts before cell division and newborn cells have two origins

The *C*+*D* period of the *E. coli* strain we used was 33 minutes longer than the doubling time (Huls *et al.*, 1999), which means that a large part of the chromosome is replicated before cell birth. Hence, we should expect two origins of replication in newborn cells and four origins of replication in cells before division. Indeed, we found newborn cells with two foci and we found that the number of foci per cell increased with cell length.

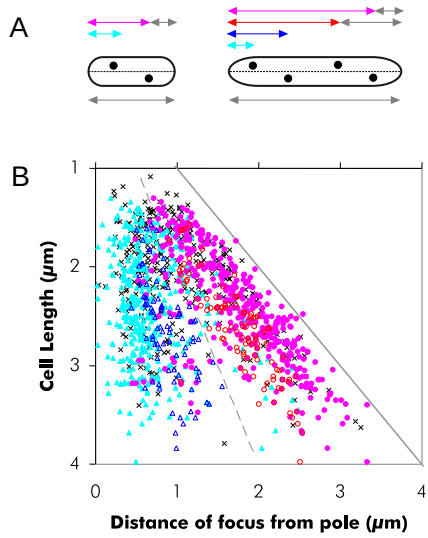


Figure 2.2 - Analysis of *oriC* localisation in *E. coli* using FISH

Cell length and distances of foci from a cell pole were measured as shown in A: one of the poles was arbitrarily chosen as reference pole. Distances of foci closest to this pole were measured directly (cyan and blue arrows). Distances of foci closer to the other pole were derived by subtracting the distance to the nearest pole (grey arrows above cells) from cell length (grey arrows below cells). B: plot of polar distances, measured as shown in A, against cell length of 692 cells. Black crosses represent polar distances of cells with a single focus, filled cyan triangles represent distances of the foci closest to the reference pole, filled magenta circles represent distances of the foci furthest from the reference pole, open blue triangles and open red circles represent distances of the remaining foci in cells with more than 2 foci.

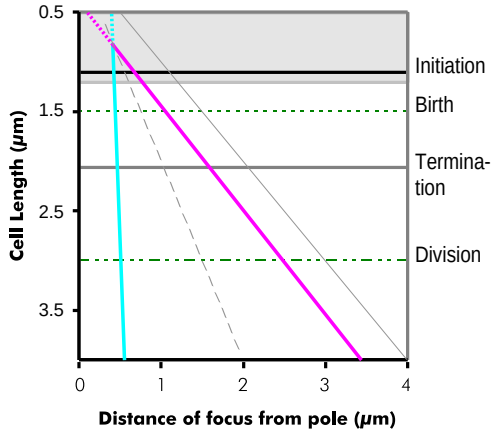


Figure 2.3 - Regression analysis of *oriC* localisation in *E. coli* during steady state growth

Data points from figure 2.2 of foci close to a cell pole were used for regression analysis. The regression equation of the cyan line is: $d1(l) = 0.37 + 2.5*10^{-3} l$ µm, where $d1(l)$ is the distance of a focus from the nearest pole and l is the cell length. The slope was statistically not significant ($p>0.05$, standard error: $1.3*10^{-3}$). The equation of the magenta regression line is: $d2(l) = l - d1(l)$. The extrapolated cross-point of the two regression lines represents the putative separation point of two replicated origins. The grey area represents the 95% confidence interval of this cross point. Black and grey horizontal lines indicate the cell lengths at which replication initiates and terminates, respectively. The interrupted horizontal lines represent the cell lengths at birth (1.5 µm, CV=12%) and division, respectively.

However, the number of foci did not always correspond to the number of origins expected. We presume this to be due partly to a less than 100% labelling efficiency of the FISH technique and partly because cells could have been impermeable to probe DNA and/or fluorescent antibodies.

The resolution of our method is limited by the optical resolving power and by biological variation. Part of this variation can be explained by variation in cell length (Koppes, L. H. *et al.*, 1978), but variation in the position of *oriC* itself cannot be excluded. To understand this variation we could speculate that two sources of *oriC* movement exist. The first source of movement is related to chromosome separation, which we believe is gradual and continuous. The second source might be Brownian motion, which would result in movement within a confined region of the nucleoid, similar to what has been found in eukaryotic cells (Marshall *et al.*, 1997). Our regression lines correspond to the first source of movement; part of the variation in *oriC* position will correspond to the second source of movement.

2.5.2 Localisation of oriC at initiation of DNA replication

The position of *oriC* when DNA replication starts could, in principle, be assessed by estimating the cell length at which four foci first appear, or by attempting to draw regression lines through the innermost pairs of *oriC* foci and estimate where they cross the regression lines through the outward ones. However, such an analysis appeared not sufficiently reliable because of less than hundred percent FISH efficiency, variation in cell length and in *oriC* position as well as the limited resolution of light microscopy. Nevertheless, our findings indicated that on average the distance between *oriC* foci and cell pole is almost constant for all cell lengths (fig. 2.2). This would mean that initiation of DNA replication starts (by definition) in the centre of a prospective daughter cell. This is in line with electron microscopic auto radiography data on slow growing *E. coli* B/r K showing that DNA replication takes place in the cell centre (Koppes, L. J. *et al.*, 1999).

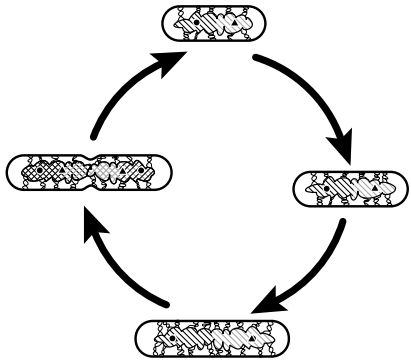


Figure 2.4 - Position of *oriC* during the cell cycle in *E. coli* MC4100 grown at 28 °C with a T_d of 79 min

The nucleoid is indicated as a hatched area (shade and direction of hatches indicate replication state). DNA loops are depicted between nucleoid and cell membrane. At birth (top) the cell contains two separated origins (circles: orgins near an old pole; triangles: origins near a new pole). During cell elongation the distance between *oriC* and nearest pole remains constant. DNA replication starts well before cell division (bottom).

2.5.3 Separation of origins

During the cell cycle the origins move apart gradually (fig. 2.4) In other words, the interpretation of our data does not require the assumption of a relatively fast movement of a just-duplicated *oriC* copy to the opposite pole (cf. Niki and Hiraga, 1998).

Gordon and coworkers visualised the vicinity of the origin region by binding of GFP-LacI fusion proteins to multiple lac operator sequence insertions (Gordon *et al.*, 1997). How do our data compare to theirs? These authors used a rich medium (L broth) and a growth temperature of 20 °C. The DNA replication cycle under those conditions was not determined. In both cases (Gordon *et al.*, 1997 and see above) two fluorescent foci occurred in newborn cells and four in older cells and the four foci population seemed to arise gradually from the two-foci ones.

A different picture emerged when applying time-lapse microphotography (Gordon *et al.*, 1997). In this case a duplicated focus appeared to move at a speed exceeding that of cell elongation. However, under time-lapse conditions, cell growth became severely impaired. It is unclear whether this affects *oriC* partitioning and therefore we have no adequate explanation for the discrepancy with respect to the speed of *oriC* movement between our results and those of Gordon *et al.* (Gordon *et*

al., 1997).

2.6 References

Bipatnath, M., P. P. Dennis, and H. Bremer. 1998. Initiation and velocity of chromosome replication in *Escherichia coli* B/r and K-12. *J Bacteriol* 180: 265-273.

Glaser, P., M. E. Sharpe, B. Raether, M. Perego, K. Ohlsen, and J. Errington. 1997. Dynamic, mitotic-like behaviour of a bacterial protein required for accurate chromosome partitioning. *Genes Dev* 11: 1160-1168.

Gordon, G. S., D. Sitnikov, C. D. Webb, A. Teleman, A. Straight, R. Losick, A. W. Murray, and A. Wright. 1997. Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. *Cell* 90: 1113-1121.

Helmstetter, C. E. 1996. Timing of synthetic activities in the cell cycle. Pages 1627-1639 in Neidhardt, F. C. and Curtiss, R., eds. *Escherichia coli and Salmonella : cellular and molecular biology*. ASM Press, Washington, D.C.

Hiraga, S. 1992. Chromosome and plasmid partition in *Escherichia coli*. *Annu Rev Biochem* 61: 283-306.

Huls, P. G., N. O. Vischer, and C. L. Woldringh. 1999. Delayed nucleoid segregation in *Escherichia coli*. *Mol Microbiol* 33: 959-970.

Koppes, L. H., C. L. Woldringh, and N. Nanninga. 1978. Size variations and correlation of different cell cycle events in slow- growing *Escherichia coli*. *J Bacteriol* 134: 423-433.

Koppes, L. J., C. L. Woldringh, and N. Nanninga. 1999. *Escherichia coli* contains a DNA replication compartment in the cell center. *Biochimie* 81: 803-810.

Lewis, P. J., and J. Errington. 1997. Direct evidence for active segregation of *oriC* regions of the Bacillus subtilis chromosome and co-localization with the SpoOJ partitioning protein. *Mol Microbiol* 25: 945-954.

Lin, D. C., P. A. Levin, and A. D. Grossman. 1997. Bipolar localization of a chromosome partition protein in Bacillus subtilis. *Proc Natl Acad Sci USA* 94: 4721-4726.

Marshall, W. F., A. Straight, J. F. Marko, J. Swedlow, A. Dernburg, A. Belmont, A. W. Murray, D. A. Agard, and J. W. Sedat. 1997. Interphase chromosomes undergo constrained diffusional motion in living cells. *Curr Biol* 7: 930-939.

Mohl, D. A., and J. W. Gober. 1997. Cell cycle-dependent polar localization of chromosome partitioning proteins in *Caulobacter crescentus*. *Cell* 88: 675-684.

Niki, H., and S. Hiraga. 1998. Polar localization of the replication origin and terminus in *Escherichia coli* nucleoids during chromosome partitioning. *Genes Dev* 12: 1036-1045.

Pettijohn, D. E. 1996. The Nucleoid. Pages 158-166 in Neidhardt, F. C. and Curtiss, R., eds. *Escherichia coli and Salmonella : cellular and molecular biology*. ASM Press, Washington, D.C.

Rasband, W. S. NIH Image Home Page. http://rsb.info.nih.gov/nih-image

Sambrook, J., T. Maniatis, and E. F. Fritsch. 1989. Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Taschner, P. E. M., P. G. Huls, E. Pas, and C. L. Woldringh. 1988. Division behaviour

and shape changes in isogenic *ftsZ*, *ftsQ*, *ftsA*, *pbpB*, and *ftsE* cell division mutants of *Escherichia coli* during temperature shift experiments. *J Bacteriol* 170: 1533-1540.

Valkenburg, J. A., C. L. Woldringh, G. J. Brakenhoff, H. T. van der Voort, and N. Nanninga. 1985. Confocal scanning light microscopy of the *Escherichia coli* nucleoid: comparison with phase-contrast and electron microscope images. *J Bacteriol* 161: 478-483.

van Helvoort, J. M. L. M., and C. L. Woldringh. 1994. Nucleoid partitioning in *Escherichia coli* during steady-state growth and upon recovery from chloramphenicol treatment. *Mol Microbiol* 13: 577-583.

Vischer, N. O.E., P. G. Huls, R. I. Ghauharali, G. J. Brakenhoff, N. Nanninga, and C. L. Woldringh. 1999. Image cytometric method for quantifying the relative amount of DNA in bacterial nucleoids using *Escherichia coli*. *J Microsc* 196: 61-68.

Vischer, N. O. E. 1994. Object-Image, for non-destructive marking and analyzing of 2D and 3D images. http://simon.bio.uva.nl/object-image.html

Vischer, N. O. E., P.G.Huls, and C. L. Wolderingh. 1994. Object-Image :An Interactive Image Analysis program Using structured Point Collection. *Binary* 6: 160-166.

Volkers, H. H., P. v. Amstel, F. M. v. d. Berg, M. M. Polak, R. Rook, and J. M. M. Walboomers. 1988. Nicktranslatie: Een DNA labelings techniek ten behoeve van *in situ* hybridisatie. *Histotechniek* 7: 2-9.

Webb, C. D., P. L. Graumann, J. A. Kahana, A. A. Teleman, P. A. Silver, and R. Losick. 1998. Use of time-lapse microscopy to visualize rapid movement of the replication origin region of the chromosome during the cell cycle in *Bacillus subtilis*. *Mol Microbiol* 28: 883-892.

Webb, C. D., A. Teleman, S. Gordon, A. Straight, A. Belmont, D. C. Lin, A. D. Grossman, A. Wright, and R. Losick. 1997. Bipolar localization of the replication origin regions of chromosomes in vegetative and sporulating cells of *B. subtilis*. *Cell* 88: 667-674.

Woldringh, C. L., and T. Odijk. 1999. Structure of DNA within the bacterial cell: physics and physiology. Pages 77-90 in Charlebois, R. L., ed. *Organization of the prokaryotic genome.* ASM Press, Washington, D.C.

Chapter 3

THE REPLICATED *FTSQAZ* AND *MINB* CHROMOSOMAL REGIONS OF *ESCHERICHIA COLI* SEGREGATE ON AVERAGE IN LINE WITH NUCLEOID MOVEMENT

M. Roos, A. B. M. van Geel, M. E. G. Aarsman, J. T. M. Veuskens, C. L. Woldringh and N. Nanninga

Published in Molecular Microbiology 39(3): 633-640 (2001)

3.1 Abstract

The average intracellular positions of the *ftsQAZ* region (2 min) and the *minB* region (26.5 min) during the cell cycle was determined by fluorescent *in situ* hybridisation using the position of *oriC* as a reference point. At the steady-state growth conditions used, newborn cells had replicated about 50% of the chromosome. By measuring the distances of the labelled *oriC*s with respect to mid-cell, we found two well-separated average *oriC* positions in cells of newborn length. These average *oriC* positions moved further apart along with cell elongation. The cellular

position of the *ftsQAZ* gene region resembled the position of *oriC*, although its average position was closer to mid-cell. In contrast, a single *minB* focus was observed at cell birth. Separated *minB* foci appeared towards the end of DNA replication. The average positions of *oriC*, *ftsQAZ* and *minB* relative to each other fitted a model in which DNA replication takes place in the cell centre and subsequent gene regions pass sequentially through this centre. We have interpreted the polarised orientation of the studied gene regions as a consequence of the mode of DNA segregation.

3.2 Introduction

Extensive cytometric analyses have demonstrated that *Escherichia coli* nucleoids move gradually, along with DNA replication and cell elongation (Van Helvoort and Woldringh, 1994). However, recent studies using light microscopic labelling techniques that visualise subnucleoid regions as fluorescent foci (Gordon *et al*., 1997; Lin *et al*., 1997; Webb *et al*., 1997; Niki and Hiraga, 1998; Niki *et al*., 2000) have indicated that such regions do not always segregate in line with the nucleoid. For instance, Gordon *et al*. (1997) found by time-lapse microscopy that, after duplication, one of the replicated origins visualised by the GFP-LacI/*lacO* system moved quickly away from the other. Movement of origins of replication has also been reconstructed from studies on fixed cells by fluorescent *in situ* hybridisation (FISH; Niki and Hiraga, 1998; Roos *et al*., 1999; Niki *et al*., 2000). The conclusions drawn by these authors diverge. Niki and Hiraga (1998) concluded that sister origins remain together near a cell pole for an appreciable time after duplication. Next, one of them moves quickly to the other pole. In contrast, we argue that such a conclusion cannot be drawn on the basis of FISH data if the position of the DNA replication cycle within the division cycle and the efficiency of FISH labelling are taken into account (Roos *et al*., 1999; this paper). By performing FISH with DNA probes that hybridise with regions around the *E. coli* chromosome, Niki *et al*. (2000) obtained evidence that gene regions tend to cluster near the origin and the terminus of replication. They suggested a circular arrangement of the

replicating nucleoid: rotation of the chromosomal circle would explain the cellular positions of gene regions relative to cell length. However, they based their interpretation on an incorrect assumption about the relationship between cell age and DNA replication (under the growth conditions used, *E. coli* strain K-12 does not contain a so-called B-period, i.e. a period in which no DNA replication takes place after cell birth; for a review, see Helmstetter, 1996).

For a correct interpretation of the movement of subnucleoid DNA regions, a thorough understanding of the relationship between cell growth and DNA replication is essential. This relationship can be determined experimentally (Bipatnath *et al.*, 1998; Huls *et al.*, 1999). To use it, however, growth conditions should be constant, i.e. the culture should be in steady state (Campbell, 1957; Maaløe and Kjeldgaard, 1966; Fishov *et al.*, 1995; Huls *et al.*, 1999).

We wished to determine the position of two subnucleoid DNA regions in between origin and terminus of DNA replication during the cell cycle. One probe recognises the *ftsQ*, *ftsA* and *ftsZ* genes in the 2 min region or *dcw* cluster, where most of the cell division genes are encoded. The other probe is specific for the *minB* region at 25.6 min, which contains the genes *minC*, *minD* and *minE* (De Boer *et al.*, 1988). As a positional reference, we used FISH-labelled *oriC* (cf. Roos *et al.*, 1999).

We found that, on average, the *oriC*, *ftsQAZ* and *minB* regions segregated in the order in which they became replicated. Their average positions followed the separation of the nucleoids. In newborn cells, the duplicated origins were already symmetrically positioned in the cell, whereas the non-replicated *minB* region was asymmetrically located. The *minB* region moved from an off-centre position to a central position upon duplication. The data fitted a model in which DNA replication takes place in the cell centre and gene regions move apart after duplication in line with nucleoid separation (Dingman, 1974; Lemon and Grossman, 1998; Koppes *et al.*, 1999).

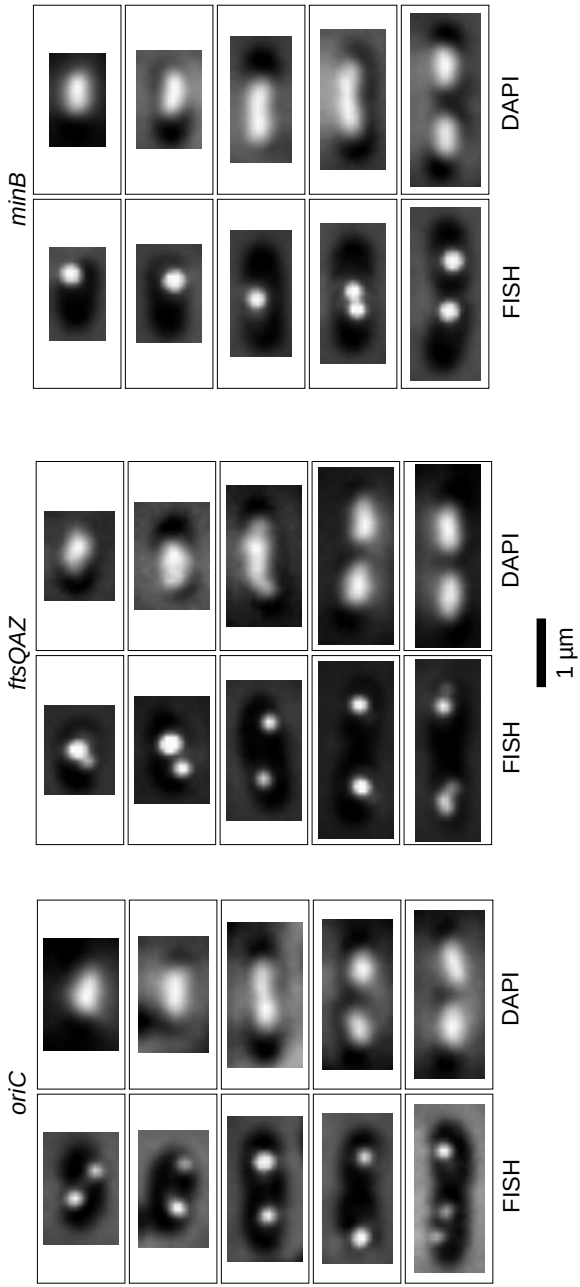


Figure 3.1 - hybridisation foci of *oriC*, *ftsQAZ* and *minB* and DAPI- stained nucleoids in fixed cells of increasing length

Cells and foci were photographed in phase contrast and with an Alexa filter (left); cells and nucleoids were photographed in phase contrast and with a DAPI filter (right). Cells hybridised with *oriC* probe (top); cells hybridised with *ftsQAZ* probe (middle); cells hybridised with *minB* probe (bottom). The bar represents 1 µm.

3.3 Results

The duration of major cell cycle events such as DNA replication and cell division are only reproducible if cultures are grown under steady-state conditions (Campbell, 1957; Maaløe and Kjeldgaard, 1966; Fishov *et al.*, 1995; Huls *et al.*, 1999). The steady-state culture of *E. coli* strain MC4100 that we grew had a doubling time of 79 min, and the periods *C* (DNA replication time) and *D* (time between termination of DNA replication and cell division; Helmstetter, 1996) were approximately 70 min and 43 min respectively (Huls *et al.*, 1999; Vischer *et al.*, 1999). By subtracting the sum of *C* and *D* from the doubling time, it was derived that DNA replication initiates about 34 min before cell birth (fig. 1.7, top). Thus, in the strain used, the *C* period is distributed over two subsequent cell cycles. Because cell length can be used as a measure of time in a steady-state culture, we could calculate the expected number of *oriC*, *ftsQAZ* and *minB* DNA regions per cell length from the determined *C* and *D* periods (Nanninga *et al.*, 1982 and references therein). Assuming that replication proceeds at a constant rate, we expected two *oriC* regions, two *ftsQAZ* regions and one *minB* region in newborn cells; these numbers double before the next cell division. The cellular positions of *oriC*, *ftsQAZ* and *minB* regions were compared by three DNA hybridisations *in situ* on cells from the same steady-state culture.

3.3.1 Efficiency of FISH labelling

Micrographs of fluorescent foci in various length classes are shown in figure 3.1. In the case of *oriC* and *ftsQAZ*, the images resemble each other and they suggest two foci in small cells. These foci duplicate at longer cell length, whereas in the case of *minB*, one focus is present in small cells, which becomes duplicated towards cell division. Not all cells were labelled or contained the number of expected labelled targets. Only 55% of the cells were labelled in the case of *oriC*, whereas this was 94% for *ftsQAZ*. This indicates that the *ftsQAZ* probe is more effective in labelling its target than the *oriC* probe. To assess this point further, we have listed the expected number of targets per cell for the three gene regions in Table1. Out of 496 cells successfully hybridised with *oriC* probe, we

observed four foci in only four out of the predicted 291 cells with four targets. We observed two foci in 71 cells out of the remaining 205 cells with two targets. Out of 498 cells successfully hybridised with *ftsQAZ* probe, we observed four foci in one out of 19 cells with four targets, two foci in 183 out of the predicted 465 cells with two targets and one focus in six out of 14 cells with one target.

Table 1. Predicted and observed number of cells from a sample of cells with foci, with calculated number of DNA targets and corresponding number of foci.

	oriC		ftsQAZ			minB	
Targets per cell	2	4	1	2	4	1	2
Predicted	205	289	14	461	19	161	826
Observed	71	4	6	183	1	121	214
Percentage of predicted	35	1	43	40	5	75	26
Percentage or predicted with fewer foci	61	99	NA	42	95	NA	72
Percentage of predicted with more foci	4	0	57	19	0	25	2

Out of 999 cells successfully hybridised with the *minB* probe, we observed two foci in 214 out of the predicted 835 cells with two targets. We observed one focus in 121 out of 164 cells with one target. Clearly, the *minB* probe also did not label all potential targets.

There are a number of possible reasons why the observed number of foci cell [1] is reduced. Foci might be too close together to be resolved by microscopy, foci might be on top of each other and/or targets may not be labelled because DNA regions are not accessible to DNA probes (cf. Roos *et al.*, 1999). These considerations, in our view, underpin the importance of establishing the DNA replication cycle at a particular growth rate to calculate the number of potential FISH targets.

3.3.2 Average position of *oriC*, *ftsQAZ* and *minB* DNA foci

To investigate the intracellular distribution of fluorescent foci with respect to cell length, we measured the distance from the centre of each focus to mid-cell and plotted this against cell length (fig. 3.2A). The scatter patterns of *oriC* foci and *ftsQAZ* foci suggest gradual separation of

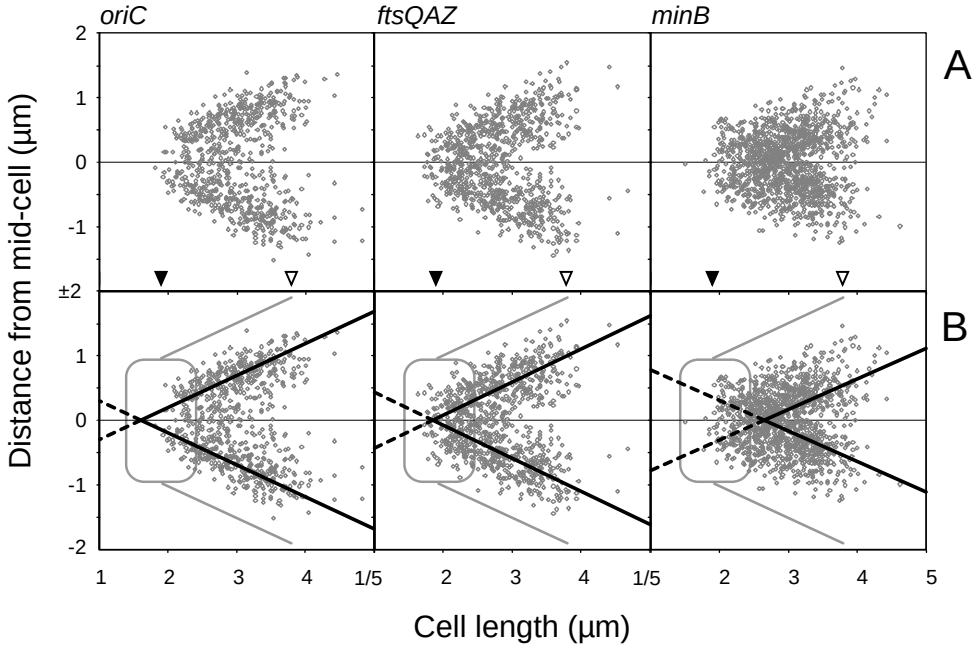


Figure 3.2 - Distances of foci from mid-cell in relation to cell length

Black arrowheads indicate the estimated average size of newborn cells (1.9 µm); open arrowheads indicate the estimated average size of dividing cells (3.8 µm). Negative values refer to the arbitrary left side of a cell; positive values refer to the arbitrary right side of a cell. Black lines represent regression lines (B). Regression coefficients were found using a maximum likelihood procedure (see Appendix). Grey lines represent the cell poles (a newborn cell is drawn to scale in each graph). Regression coefficients: *oriC*, slope = 0.50 ± 0.02 µm, y-intercept = 0.79 ± 0.09 µm; *ftsQAZ*, slope = 0.51 ± 0.11 µm, y-intercept = 0.96 ± 0.34 µm; *minB*, slope = 0.46 ± 0.09 µm, y-intercept = 1.20 ± 0.29 µm.

these regions as cell length increases. The scatter pattern of *minB* foci is more compact, but still suggests gradual separation in longer cells.

Because gradual separation seemed appropriate at a first approximation, we assumed a linear relationship between focusmid-cell distance and cell length, which enabled us to use a linear regression procedure to estimate the average cellular position as a function of cell length (fig. 3.2B).

Furthermore, we used our predictions about the number of DNA targets (Table1), assuming that the reduced number of foci was a result of missing label instead of targets clustering together. Note that the slope of the regression line of the distance between a cell pole and mid-cell as a function of cell length is 0.5. The slopes of regression lines through scatter points of *oriC*, *ftsQAZ* as well as *minB* (fig. 3.2B) were not significantly different from 0.5 ($P>0.05$), suggesting that, on average, separation of DNA regions took place in line with cell elongation. The point where regression lines cross, corresponding to the *x*-intercept, was denoted as the 'separation point', the point at which duplicated DNA regions start to separate. The separation point of *oriC* was at a cell length of about 1.6 μm (the length of a prospective daughter cell in the previous cycle), which indicates that separation occurred before cell birth (the estimated length of a newborn cell was 1.9 μm). The separation point of *ftsQAZ* foci was at a cell length of about 1.9 μm, and the separation point of *minB* was at a cell length of about 2.6 μm.

As can be seen in figure 3.1, a single *minB* focus is located asymmetrically in the smallest cell, whereas in longer cells, the focus occurs in the cell centre. To assess this further, we made histograms of the *minB* positions in various length classes (fig. 3.3). In the shortest length classes, the average position of a *minB* focus is off centre. Near the *minB* separation point (length class 2.7 μm), the average positions of the two foci are closely apposed, and they separate further when cells become longer (fig. 3.3).

Thus, on average, *oriC* regions separated first, followed by *ftsQAZ* regions and *minB* regions. The asymmetrically located *minB* gene region seems to migrate to the cell centre before duplication.

3.4 Discussion

3.4.1 The interpretation of FISH data

We have analysed the location of subnucleoid DNA regions in *E. coli* relative to cell cycle events, such as cell birth and initiation of DNA

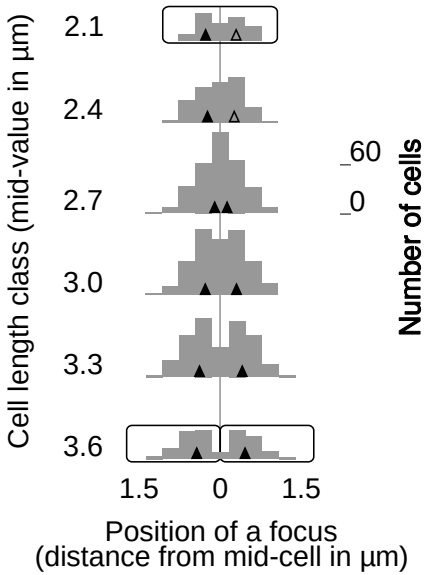


Figure 3.3 - Histograms of the cellular position of minB foci in various cell length classes measured as distance from mid-cell

Foci were randomly placed on the left side of a cell or on the right side of a cell. Arrowheads indicate the average position of foci in a cell length class. Because of randomisation, two average positions are found in cells with one predicted DNA target (cells shorter than 2.4 m m); the second position is indicated by an open arrowhead.

replication. We concluded that, on average, replicated DNA regions segregate in line with cell growth and nucleoid segregation. However, examination of the number of foci found in cells did not directly support this conclusion. The number of predicted DNA targets exceeded the number of visible foci in most cells. For instance, although we expected two origin regions in short cells, many cells contained one focus. An explanation may be that two (or more) DNA regions are too close to be resolved by light microscopy, which would suggest that DNA regions stick together for some time after duplication. Although we do not exclude this possibility, we favour the explanation that not all DNA targets became labelled to produce a visible signal. This is illustrated by the high number of cells with foci of the *ftsQAZ* probe (94%) and the low number of cells with foci of the *oriC* probe (55%). The higher efficiency of the *ftsQAZ* probe shows that probe DNA can penetrate cells; we have no reason to believe that this is not also true for the *oriC* probe. Our conclusion is that the *oriC* probe did penetrate the cells but, in many cases, did not label its target. For our interpretation of FISH data, we

therefore used the expected number of DNA targets per cell length and not the number of visible foci.

In contrast, the assumption that each DNA target is labelled may lead to a different interpretation of FISH data (cf. Niki and Hiraga, 1998). Under growth conditions (T_d = 55 min at 37 °C in M9 medium supplemented with glucose, proline and thiamine) that would also produce two origins at birth (Hiraga *et al.*, 1998; our unpublished results), Niki and Hiraga (1998) detected short cells with one focus and short cells with two well-separated foci. Although the labelling patterns correspond to our findings, they concluded that one of the newly duplicated origins migrates quickly to the other side of the cell. However, we believe that, for the interpretation of FISH data, the efficiency of the technique should be taken into account. If a single focus is present, one has to choose between the possibility that a duplicated gene region cannot be resolved or whether a focus is lacking because a target has not been labelled. According to these considerations, we interpret our *oriC* data to mean that, in a newborn cell, when 50% of the genome has been replicated, the two *oriC* targets are on average well separated (this paper and cf. Roos *et al.*, 1999).

3.4.2 Separation of DNA regions

We have looked at the separation behaviour of two regions at different locations on the *E. coli* chromosome [*ftsQAZ* (2 min) and *minB* (26.5 min)] using the position of *oriC* as a reference (fig. 3.2). As for *oriC*, our data suggest that, on average, these regions move apart in line with cell elongation as soon as they are detectable after duplication and that separation follows the order of their replication. Thus, the average position of *ftsQAZ* was in between the origin and *minB*. In *Bacillus subtilis*, the cellular positions of two opposite gene regions in between origin and terminus have been compared (Teleman *et al.*, 1998). It was found that these regions not only occupied comparable positions in the cell, they also occurred in between origin and terminus. Our *E. coli* data therefore closely resemble those obtained with *B. subtilis*.

Note, however, that the above analysis refers to average positions of the foci and also note that considerable scatter occurs around the separation points (fig. 3.2). This makes it difficult to assess whether initial separation after duplication of a gene region occurs in line with cell elongation. In individual living cells of *E. coli* and *B. subtilis*, DNA regions close to the origin have been shown to separate non-gradually by time-lapse microscopy (Gordon *et al.*, 1997; Webb *et al.*, 1998 respectively). We can reconcile these results with our data, if we assume that movement of DNA regions is the sum of an overall movement along with nucleoid segregation and of local short-distance movements unrelated to nucleoid movement. In our analysis, local short-distance movements would be averaged out. They would instead constitute an important part of the observed variation in distances of foci from mid-cell. This implies that the expected result of our method would be gradual movement, whereas this would not necessarily be so in the case of an individual cell.

3.4.3 DNA segregation and the replisome

Recent evidence in *B. subtilis* (Lemon and Grossman, 1998) and in *E. coli* (Koppes *et al.*, 1999) indicates that DNA replication occurs in the cell centre. These observations are compatible with the replisome model of Bingman (1974), in which bi-directional DNA replication takes place at a fixed cellular site (the replisome) and the DNA strands to be replicated move through the replisome. In this model, duplicated origins move apart during DNA replication. How do our data fit into this model? In a newborn cell, the two origins have already separated, because they were duplicated in the previous cell cycle (fig. 1.7, top). In contrast, the *minB* region has not yet duplicated, and it is located asymmetrically in the cell near the youngest pole (fig. 3.1). Presumably, this asymmetry has arisen because the *minB* regions were closer to mid-cell during the preceding division. When the newborn cells become longer, the *minB* region goes to mid-cell towards the replisome, whereas the origins move further apart while maintaining a constant distance to the respective nearest poles (Roos *et al.*, 1999). Obviously, *minB* and one of the *oriC*s

have to pass each other. How can these opposing DNA movements be visualised? The original Dingman figure with the terminus perpendicular to the fixed replisome has been redrawn in figure 3.4A. Bending of the terminus towards mid-cell brings *minB* into an asymmetrical position on the length axis of the cell (fig. 3.3 and fig. 3.4B). During DNA replication, *minB* approaches the replisome at mid-cell (fig. 3.3) and passes one of the replicated *oriC*s that moves in the opposite direction (fig. 3.4C). Thus, two different movements are responsible for the passing of *oriC* and *minB* along each other. On the one hand, duplicated origins separate, initially independently and later on in line with cell elongation. On the other hand, the asymmetrically located *minB* region migrates to the cell centre, possibly pulled by the active replisome.

It has been suggested (Niki *et al.*, 2000) that, in order to start replication, *oriC* would move from a polar position to mid-cell, and this rearrangement has been supposed to take place during the *B* period. The *B* period is, by definition, the time between cell birth and initiation of DNA replication, provided that the *C* period is contained within the same cell cycle, i.e. when *C*+*D* T_d (Helmstetter, 1996). This has been observed in *E. coli* B/r strains, in particular at slow growth rates (T_d > 60 min; for instance, see Koppes *et al.*, 1999). So far, a *B* period has not been observed in K-12 strains with T_ds of 19230 min (Helmstetter, 1996). In our view, these considerations argue against the existence of a *B* period in *E. coli* K-12 strains for the repositioning and activation of *oriC*.

3.4.4 Concluding remarks

We emphasise that our data refer to average movement of DNA regions during the cell cycle. Many of the forces behind the presumed movements of DNA regions have not been found yet. A mitotic-like mechanism that actively segregates DNA has been suggested (for a review, see Glaser *et al.*, 1997; Shapiro and Losick, 1997; Sharpe *et al.*, 1998), and it has also been suggested that this mechanism exists alongside a mechanism for gradual movement of bulk DNA (Sharpe *et al.*, 1998). We would like to add the possibility that DNA regions exhibit

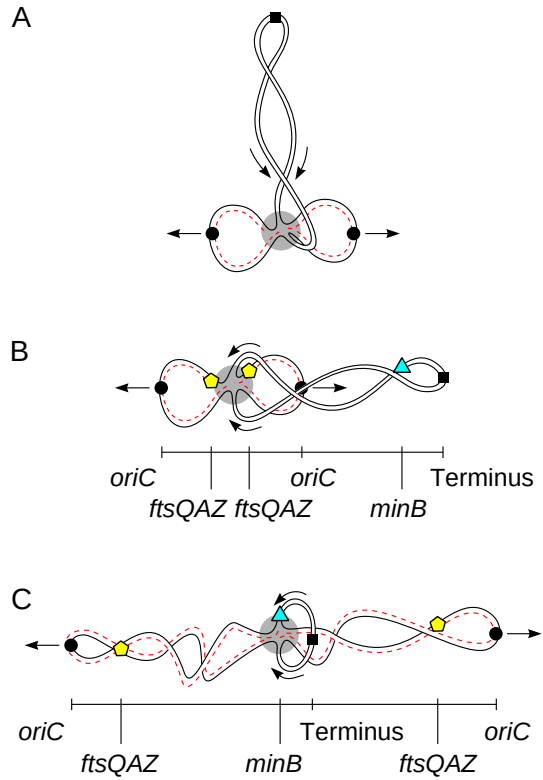


Figure 3.4 - Model for sequential separation of DNA subregions on the *E. coli* chromosome

A. Redrawing of the original Bingman figure with the terminus perpendicular to the fixed replisome, shortly after initiation of DNA replication (Bingman, 1974). B. Adaptation of (A) with the terminus-containing loop in the length axis of the cell. C. Organisation of DNA subregions before termination of DNA replication. Arrows indicate the direction of movement of DNA (towards the replisome or away from the replisome in the length axis of the cell). Black circles, *oriC* region; black squares, terminus region; yellow pentagons, *ftsQAZ* region; blue triangles, *minB* region; red dashed line, newly synthesised DNA. Note that, as replication progresses from (B) to (C), *minB* , moving towards the replisome, passes one of the *oriC*s, which is moving away from the replisome.

constrained diffusional motion, similar to that found in yeast and *Drosophila* (Marshall *et al.*, 1997).

3.5 Experimental procedures

3.5.1 Bacterial strains and growth conditions

Cells of *E. coli* strain LMC500 [MC4100 (*F*$^-$, *araD139*, (*argF-lac*) *U169*, *deoC1*, *flbB5301*, *ptsF25*, *rbsR relA1*, *rpsl150*) *lysA*] (Taschner *et al.*, 1988) were grown at 28 °C in steady state with a doubling time (T_d) of 79 min in glucose minimal medium containing 6.33 g/l $K_2HPO_4.3H_2O$, 2.95g/l KH_2PO_4, 1.05g/l $(NH_4)2SO_4$, 0.10g/l $MgSO_4.7H_2O$,

0.10g/l $MgSO_4.7H_2O$, 0.28mg/l $FeSO_4.7H_2O$, 7.1mg/l $Ca(NO_3)_2.4H_2O$, 4mg/l thiamine, 4g/l glucose and 50mg/l lysine (pH7.0). We harvested 50ml of cell culture at an OD_{450} of 0.2 and fixed cells in 1ml of 1% OsO_4 in TY medium (1% tryptone, 0.5% yeast extract, 0.5% NaCl, 3 mM NaOH, pH7.0), which was stored at 4 °C.

3.5.2 Probes

We used an *oriC* probe obtained from plasmid pOC162 (kindly provided by Dr W. Messer, Max Planck Institut für Molekulare Genetik, Berlin, Germany). This plasmid contains the 248bp *oriC* region with 3kbp on each side of *oriC*. Because pOC162 contains no sequence homologies to chromosomal sequences, the whole construct was labelled and used as a probe for FISH. The *ftsQAZ* probe was obtained from plasmid pZAQ, which contains the *ftsZAQ* operon (kindly provided by R. d'Ari, Institut Jacques Monod, CNRS, Université de Paris, France). To obtain a probe without any other chromosomal sequences from the plasmid, a 4.4kbp *Eco*RI*Cla*I fragment containing *ftsZAQ* was excised and labelled.
We obtained the *minB* probe from plasmid pDB103, a plasmid that contains the complete *minB* operon of *E. coli* (De Boer *et al.*, 1988). To obtain our probe without other chromosomal sequences from the plasmid, a 1kbp segment containing part of the *minD* gene and part of the *minE* gene was amplified by polymerase chain reaction (PCR) and labelled.

3.5.3 Probe labelling

Plasmids were labelled with digoxigenin-11-dUTP (DIG; Hoffmann-La Roche) by nick translation. DIG incorporation was checked by spot blotting on Hybond nylon filter (Hoffmann-La Roche) and detection with mouse anti-DIG alkaline phosphatase-conjugated antibodies (Hoffmann-La Roche) and 5-bromo-4-chloro-3-indolylphosphate/nitroblue tetrazolium (BCIP/NBT; Hoffmann-La Roche) as substrate. The specificity of the probe was checked by Southern hybridisation on digested genomic DNA of strain MC4100 (Sambrook *et al.*, 1989).

3.5.4 Preparation of cells for fluorescence microscopy

Cells were washed in PBS (140 mM NaCl, 27 mM KCl, 10 mM Na_2HPO_4 and 2 mM KH_2PO_4, pH7.2) and post-fixed in 0.5% formaldehyde and 0.04% glutaraldehyde in TBS (10 mM Tris and 0.9% NaCl, pH7.5) for 15 min at room temperature. Fixed cells were centrifuged at 8000*g* for 5 min, washed three times in PBS and subsequently incubated in 0.1% Triton X-100 in PBS for 45 min at room temperature. Cells were washed three times in PBS and incubated in PBS containing 100µg/ml lysozyme (Sigma-Aldrich) and 5 mM EDTA for 45 min at 37 °C. After washing three times in PBS, the cells were incubated with 100µg/ml RNase A (Hoffmann-La Roche) in PBS for 60 min at 37 °C. After washing three times in PBS, 10µl of cell suspension was applied to slides coated with 0.01% poly L-lysine, covered with a coverslip (15 mm diameter) and left for 20 min at room temperature. Then, the slides were washed in a Coplin jar with 2× SSC (sodium salt citrate; 0.15M NaCl, 150 mM sodium citrate, pH7) and put on top of coverslips with 10µl of probe solution [1ng/µ probe in hybridisation mix, containing 2×SSC with 300 µg/ml herring testes DNA, 5% polystyrenesodium sulphonate 100000 (Acos), 1× Denhardt's reagent (Sambrook *et al.*, 1989) and 50% formamide]. The preparations were denatured for 5 min at 80 °C. Hybridisation occurred overnight in a moist chamber at 37 °C. Unbound probe was washed three times for 10 min at 40 °C in a Coplin jar with 1× SSC and 50% formamide (pH7). Slides were subsequently washed in 2×SSC and TN buffer (100 mM Tris-HCl, pH7.5, 150 mM NaCl).

Before immunofluorescence staining, non-specific binding sites were blocked by incubating the cells in TNB [0.5% (w/v) blocking reagents (Hoffmann-La Roche) in TN] for 30 min at 37 °C. DIG-labelled probe detection was carried out with mouse anti-DIG (Hoffmann-La Roche), 1:500 in TNB, rabbit anti-mouse-conjugated Cy3 (Jackson ImmunoResearch Laboratories), 1:400 in TNB and goat anti-rabbit-conjugated Alexa 546 (Molecular Probes), 1:600 in TNB. Slides with antibodies were incubated in a moist chamber at 37 °C for 60 min. Between incubations with antibodies, the slides were washed three times

for 5 min in TNT [TN with 0.05% (v/v) Tween 20]. After immunodetection, slides were rinsed three times in TNT, once in TN and once in PBS. The preparations were mounted in 5µl of PBS with DAPI (500ng/m 2,4,diamidino-phenylindole).

3.5.5 Microscopy and image analysis

Preparations were photographed with a cooled Princeton CCD camera mounted on an Olympus fluorescence microscope (BH2-RFC) equipped with a 100W mercury lamp. Images were made using the program IPLAB spectrum 3.0 (Signal Analytics). Cells and fluorescent foci were photographed using a combination of phase contrast and a filter combination for Alexa 546 (illumination 510-550nm; emission 590nm). To check nucleoid morphology, cells and DAPI-stained DNA were photographed using a combination of phase contrast and a filter combination for DAPI (illumination 300-400nm; emission 420nm). Length, position of foci and nucleoid length of each cell were determined interactively. Interactive measurements were performed as 'structured point collection' on an Apple computer (Power Macintosh 7100) using the public domain program ObjectImage 1.62n (Vischer *et al.*, 1994; http://simon.bio.uva.nl/object-image.html), which is based on the NIH Image software (W. Rasband; http://rsp.info.nih.gov/nih-image).

3.6 Acknowledgements

We thank Michiel Müller and Grace Sickmann for suggestions and discussion, Rob Lingeman for advice on statistical issues, and Tanneke den Blaauwen and Arieh Zaritsky for critical reading of the manuscript.

3.7 References

Bipatnath, M., Dennis, P.P., and Bremer, H. (1998) Initiation and velocity of chromosome replication in *Escherichia coli* B/r and K-12. *J Bacteriol* 180: 265273.

Campbell, A. (1957) Synchronization of cell division. *Bacteriol Rev* 21: 263272.

De Boer, P.A.J., Crossley, R.A., and Rothfield, L.I. (1988) Isolation and properties of *minB*, a complex genetic locus involved in correct placement of the division site in *Escherichia coli*. *J Bacteriol* 170: 21062112.

Bingman, C.W. (1974) Bidirectional chromosome replication: some topological considerations. *J Theor Biol* 43: 187195.

Fishov, I., Zaritsky, A., and Grover, N.B. (1995) On microbial states of growth. *Mol Microbiol* 15: 789794.

Glaser, P., Sharpe, M.E., Raether, B., Perego, M., Ohlsen, K., and Errington, J. (1997) Dynamic, mitotic-like behavior of a bacterial protein required for accurate chromosome partitioning. *Genes Dev* 11: 1160 1168.

Gordon, G.S., Sitnikov, D., Webb, C.D., Teleman, A., Straight, A., Losick, R., Murray, A. W., and Wright, A. (1997) Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. *Cell* 90: 11131121.

Helmstetter, C.E. (1996) Timing of synthetic activities in the cell cycle. In Escherichia coli and Salmonella: Cellular andMolecular Biology. Neidhardt, F.C. (ed. in chief). Washington, DC: American Society for Microbiology Press, pp. 16271639.

Hiraga, S., Ichinose, C., Niki, H., and Yamazoe, M. (1998) Cell cycle-dependent duplication and bidirectional migration of SeqA-associated DNA-protein complexes in *E. coli*. *Mol Cell* 1: 381387.

Huls, P., Vischer, N.O.E., and Woldringh, C.L. (1999) Delayed nucleoid segregation in *Escherichia coli*. *Mol Microbiol* 33: 959970.

Koppes, L.J.H., Woldringh, C.L., and Nanninga, N. (1999) *Escherichia coli* contains a DNA replication compartment in the cell center. *Biochimie* 81: 803810.

Lemon, K.P., and Grossman, A.D. (1998) Localization of bacterial DNA polymerase: evidence for a factory model of replication. *Science* 282: 15161519.

Lin, D.C.-H., Levin, P.A., and Grossman, A.D. (1997) Bipolar localization of a chromosome partitioning protein in *Bacillus subtilis*. *Proc Natl Acad Sci USA* 94: 47214726.

Maaløe, O., and Kjeldgaar, N.O. (1966) Control of Macromolecular Synthesis. A Study of DNA, RNA and Protein Synthesis in Bacteria. New York: Benjamin.

Marshall, W.F., Straight, A., Marko, J.F., Swedlow, J., Dernburg, A., Belmont, A., Murray, A. W., Agard, D. A., and Sedat, J. W. (1997) Interphase chromosomes undergo constrained diffusional motion in living cells. *Curr Biol* 7: 930939.

Nanninga, N., Woldringh, C.L., and Koppes, L.J.H. (1982) Growth and division of *Escherichia coli*. In *Cell Growth*. Nicolini, C. (ed.). New York: Plenum, pp. 225270.

Niki, H., and Hiraga, S. (1998) Polar localization of the replication origin and terminus in *Escherichia coli* nucleoids during chromosome partitioning. *Genes Dev* 12: 10361045.

Niki, H., Yamaichi, Y., and Hiraga, S. (2000) Dynamic organization of chromosomal DNA in *Escherichia coli*. *Genes Dev* 14: 212223.

Roos, M., van Geel, A.B.M., Aarsman, M.E.G., Veuskens, J.T.M., Woldringh, C.L., and Nanninga, N. (1999) Cellular localization of *oriC* during the cell cycle of *Escherichia coli* as analyzed by fluorescent *in situ* hybridization. *Biochimie* 81: 797802.

Sambrook, J., Fritsch, F., and Maniatis, T. (1989) Analysis of genomic DNA by southern hybridization. In Molecular Cloning: a Laboratory Manual, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 9.319.57.

Shapiro, L., and Losick, R. (1997) Protein localization and cell fate in bacteria. *Science* 276: 712718.

Sharpe, M.E., Hauser, P.M., Sharpe, R.G., and Errington, J. (1998) *Bacillus subtilis* cell cycle as studied by fluorescence microscopy: constancy of cell length at initiation of DNA replication and evidence for active nucleoid partitioning. *J Bacteriol* 180: 547555.

Taschner, P.E., Huls, P.G., Pas, E., and Woldringh, C.L. (1988) Division behaviour and shape changes in isogenic *ftsZ, ftsQ, ftsA, pbpB,* and *ftsE* cell division mutants of *Escherichia coli* during temperature shift experiments. *J Bacteriol* 170: 15331540.

Teleman, A.A., Graumann, P.L., Lin, D.C.-H., Grossman, A.D., and Losick, R. (1998) Chromosome arrangement within a bacterium. *Curr Biol* 8: 11021109.

Van Helvoort, J.M.L.M., and Woldringh, C.L. (1994) Nucleoid partitioning in *Escherichia coli* during steady state growth and upon recovery from chloramphenicol treatment. *Mol Microbiol* 13: 577583.

Vischer, N.O.E., Huls, P.G., and Woldringh, C.L. (1994) Object-image: an interactive image analysis program using structured point collection. *Binary* 6: 160166.

Vischer, N.O.E., Huls, P.G., Ghauharali, R.I., Brakenhoff, G.J., Nanninga, N., and Woldringh, C.L. (1999) Image cytometric method for quantifying the relative amount of DNA in bacterial nucleoids using *Escherichia coli*. *J Microsc* 196: 6168.

Webb, C.D., Teleman, A., Gordon, S., Straight, A., Belmont, A., Lin, D.C.-H., Grossman, A. D., Wright, A., and Losick, R. (1997) Bipolar localization of the replication origin regions of chromosomes in vegetative and sporulating cells of *B. subtilis*. *Cell* 88: 667674.

Webb, C.D., Graumann, P.L., Kahana, J.A., Teleman, A., Silver, P.A., and Losick, R. (1998) Use of time-lapse microscopy to visualize rapid movement of the replication origin region of the chromosome during the cell cycle in *Bacillus subtilis*. *Mol Microbiol* 28: 883892.

3.8 Appendix

Regression lines were calculated using a linear regression model:

$$d(l) = A + B \cdot l$$

in which $d(l)$ is the distance from mid-cell at cell length l, A is the y-intercept of the regression line, and B is the slope of the regression line.

Regression coefficients A and B were found using a maximum likelihood procedure. We assumed that distances were distributed normally around a mean distance at each cell length, and we took into account that foci can be positioned on the left side of a cell as well as on the right side of a

cell (negative and positive distance values respectively). We used the following likelihood function:

$$L_{l,s}(md) = [N_{-d(l),s}(md) + N_{+d(l),s}(md)] / 2$$

in which $L_{l,s}$ represents the likelihood that measured distance md is found at cell length l, and N is a Gaussian likelihood function with mean $d(l)$ or $d(l)$ and standard deviation s. For s, we assumed a value of 0.2 µm. We used a bootstrap method to find those A and B for which the sum of likelihoods of all measured distances was maximal.

Chapter 4

EXPERIMENTS ON MOVEMENT OF DNA REGIONS IN *ESCHERICHIA COLI* EVALUATED BY COMPUTER SIMULATION

Marco Roos, Rob Lingeman and Nanne Nanninga

Swammerdam Institute for Life Sciences, BioCentrum Amsterdam, University of Amsterdam, The Netherlands

4.1 Summary

During the cell cycle of *Escherichia coli,* DNA is replicated and segregated over two prospective daughter cells. Nucleoids as a whole segregate gradually, in line with cell elongation, but sub-nucleoid DNA regions may behave differently, segregating non-gradually. We tested the ability of three models to predict the outcome of a fluorescent *in situ* hybridisation (FISH) experiment. We did this by comparing computer-simulated data with experimental data. The first model predicts gradual segregation in line with cell elongation. The second model predicts that origins stick together for some time after duplication before one copy jumps to the other side of the cell (non-gradual segregation). The simulated data of these models were very similar, indicating that FISH is not a suitable

method to distinguish between these two models. The third model predicts that origins may be anywhere within the nucleoid(s). We found that simulated data using the third model resemble the experimental data most. However, DNA regions are not randomly localised in the cell, although their localisation is fuzzy. We propose that movement of DNA regions is the result of a combination of factors. Nucleoid segregation (or the forces behind it) dictates the overall direction of movement. Other factors, of which we show that diffusion could be an important one, move DNA in other directions giving rise to non-gradual movement in individual cells and contributing to variation in intracellular position per cell length in a population of cells.

4.2 Introduction

Cell growth, transcription, replication and DNA segregation occur simultaneously and with high efficiency in bacteria. It is likely that the (dynamic) organisation of DNA in the cell is essential to reach this efficiency. Therefore, movement of replicated DNA to the prospective daughter cells is probably the result of one or more mechanisms (for a recent review see Gordon and Wright, 2000). Nucleoids as a whole have been shown to segregate gradually in line with cell elongation in *E. coli* (van Helvoort and Woldringh, 1994) and in *Bacillus subtilis* (Sharpe *et al.*, 1998, but how do individual DNA regions move?

Our studies on fixed cells in which DNA regions were fluorescently labelled *in situ* via hybridisation with a DNA probe (FISH), suggested that, on average, sub-nucleoid DNA regions move gradually in line with nucleoid segregation (Roos *et al.*, 1999; 2001). However, the FISH data showed a lot of variation in intracellular position when plotted as a function of cell length; a typical example is the origin of replication of *E. coli*, *oriC* (fig. 4.1). Additionally, movement of sister DNA regions directly after duplication could not be addressed by FISH, because DNA regions are too close to each other to be resolved by light microscopy. Thus, statistical treatment of FISH data can only give information about average movement of DNA regions as a function of cell length of a

(steady-state) population (Roos *et al.*, 2001).

Movement of DNA regions in individual living cells can be studied by tagging DNA regions by green fluorescent protein (GFP; see Margolin, 2000 and references therein). GFP tagged regions close to *oriC* or containing *oriC* have been shown to segregate much quicker than cell elongation in *E. coli* (Gordon *et al.*, 1997) and in *B. subtilis* (Sharpe and Errington, 1998; Webb *et al.*, 1998). This supports models which predict that duplicated DNA regions segregate quickly, not in line with cell elongation, and possibly in an active, mitotic-like process (see for instance Sharpe and Errington, 1999 and references therein). Consequently, Niki and Hiraga explained their FISH data with a model in which one copy of *oriC* 'jumps' to the other side of the cell after duplication (Niki and Hiraga, 1998).

Even if nucleoids segregate in a different way than sub-nucleoid DNA regions, a model of movement of DNA regions in bacteria as function of cell length should explain average movement of DNA in a (steady-state) population of cells as well as movement of DNA in individual living cells. However, it can be difficult to see if a particular model explains the observed data sufficiently. An objective method to test the validity of a model is to compare computer simulated data based on such a model with experimental data for a population of cells. We have done this using three models of DNA movement in *E. coli.* Model I predicts gradual segregation of DNA regions in line with cell elongation; model II predicts that DNA regions segregate non-gradually after replication; model III predicts that DNA regions have no preferred position within the nucleoid. We present here the results for the origin of replication in *E. coli*, *oriC,* because this is probably the best studied DNA region. Our results indicate that DNA organisation in *E. coli* is not random, although it is fuzzy, i.e. labelled DNA regions are not precisely located in the cell during the cell cycle. We discuss a model in which DNA movement is the result of a number of factors. Nucleoid segregation, or the force(s) behind it, is one of these factors and 'diffusion in a confined region' (cf. Marshall *et al.*, 1997) is one additional factor.

4.3 Materials and methods

4.3.1 Simulation of FISH data

For each cell from a population, theoretical positions were calculated for each copy of *oriC* present in the cell according to three models (see appendix, section 4.8). Because no distinction could be made between the two sides of a cell, we randomised cell orientation by multiplying every distance in a cell by +1 or -1 (numbers obtained from Microsoft Excel's Bernouilli random number generator with $p = 0.5$). To each theoretical position a random value drawn from a normal distribution with mean 0 and a standard deviation of 0.2 μm was added. This value for the standard deviation was derived from previously published data (Roos *et al.*, 2001). FISH labelling efficiency was incorporated by randomly selecting DNA regions in the data set with an experimentally derived probability $p = 0.3$ (see appendix, section 4.8). If DNA regions were closer to each other than the experimentally derived value of 250 nm (see below), then they were scored as one focus with its centre in between the DNA regions. Finally, a measure of error was incorporated by adding a random value drawn from a normal distribution with mean 0 and a standard deviation of 15 nm (see below).

4.3.2 Measurement error

To determine the measurement error of manually indicating the centre of a FISH focus in an image, we created a thousand images with a simulated FISH focus at various positions. An experimenter scored the centre of the focus in each image manually. The standard deviation of the average difference between the positions of the simulated FISH foci and the positions indicated by the experimenter was used as measurement error (0.015 μm). FISH foci were simulated by a 2-D normal intensity distribution with a standard deviation similar to that of a FISH focus (approximately 0.14 μm).

4.3.3 Resolution of fluorescence microscopy

We estimated the resolution of fluorescence microscopy by scoring if two foci could be resolved visually in a thousand images with two simulated FISH foci at various distances from each other. The average distance at which two foci could be resolved was used as a measure of resolution (250 nm).

4.3.4 Diffusion in a confined region

To simulate positions that are the result of diffusion in a confined region we calculated the position that a putative DNA region reaches after a random walk of 10^6 equidistant steps of 0.075 μm in random directions within a sphere with a radius of 0.3 μm (cf. Marshall *et al.*, 1997), and we added that position to the intracellular position of the confined regions as determined by the theoretical positions of model I and model II, respectively.

4.4 Results

In order to evaluate models of movement of *oriC* in slow growing *E. coli* K-12 cells, we first calculated the theoretical position of *oriC* as a function of cell length according to each of the three models (fig. 4.2A). Model I predicts that segregation of origins proceeds at exactly the same speed as cell elongation. Consequently, origins remain at a fixed distance from the nearest (prospective) cell pole. Model II is equal to model I except that after duplication the two duplicated origins first stick together, jointly moving away gradually from the cell centre on one side of the cell. When half of the cell length between initiation of DNA replication and termination of DNA replication is reached, one of the origins 'jumps' to the other side of the (prospective daughter) cell. In model III DNA has no preferred position: a DNA region has an equal probability to be anywhere within the nucleoid. For each model, basic assumptions about cell size, growth rate, and replication rate were chosen conform our previously published data (Huls *et al.*, 1999; Roos *et al.*, 2001). Thus, because replication initiated in a previous cell cycle, there are two copies

of *oriC* at cell birth in all models. Except for model III, the position of the origins at initiation is at ¼ and ¾ of the cell, because we assumed that replication takes place in the centre of each prospective daughter cell when replication initiates before cell birth (cf. Koppes *et al.*, 1999; Lemon and Grossman, 1998).

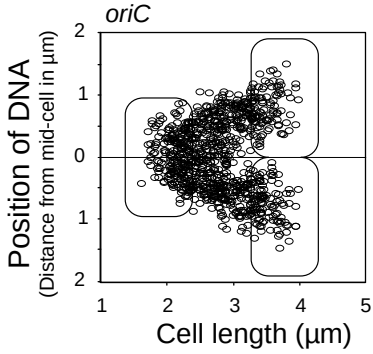


Figure 4.1 - Distances of *oriC* foci from mid-cell as a function of cell length.

Reprinted from figure 3.2.

We simulated FISH experiments by calculating the theoretical intracellular positions of *oriC* for a population of cells and incorporating the following factors: measurement error, resolution of fluorescence microscopy, the probability that an *oriC* region is labelled and a normally distributed variation in the intracellular position of *oriC* regions (fig. 4.2B). If we compare the simulated data of model I and model II, we find that, although the bases of these models differ, the difference in simulated data is small. If we compare the simulated data of these models with the simulated data of model III, we find that models I and II lack data points in the cell centre when compared to model III. If we compare our simulated data (fig. 4.2B) with the experimental data of figure 4.1, then model I and model II both lack data points in the cell centre. So far, model III seems to resemble the experimental data most. To investigate these results quantitatively we calculated the 'likelihood' of the experimental data of figure 4.1 to be the result of each of the three models (fig. 4.2). In agreement with our visual comparison we found that the experimental data are less likely the result of model I or model II than the result of model III. Note that the data plots and the likelihood-values of model I and model II are very similar, indicating that FISH is not a suitable method for discriminating between these two models.

In the simulations above we assumed a value of 0.2 µm for the standard

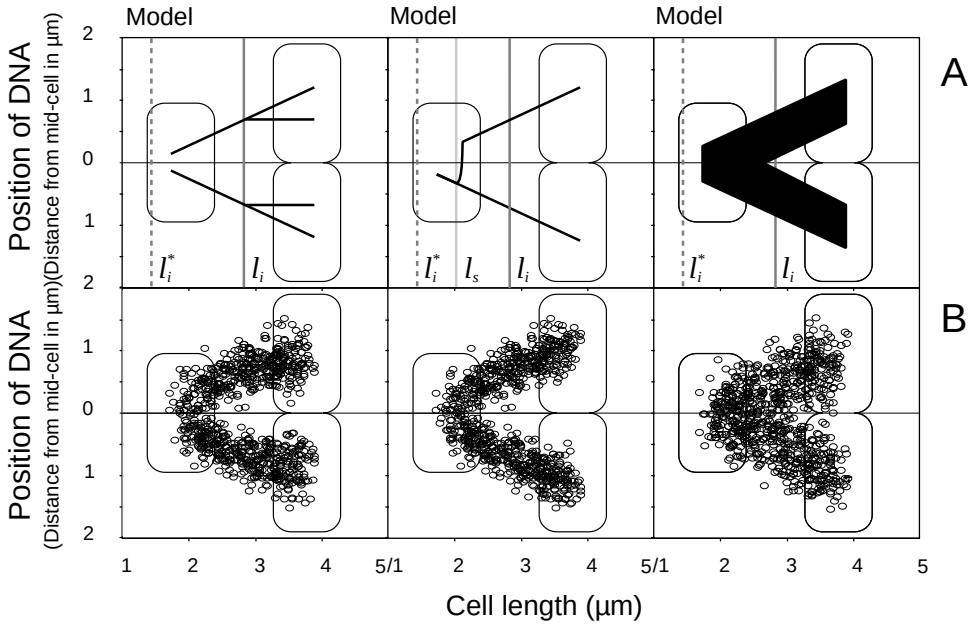


Figure 4.2 - Theoretical (A) and simulated (B) distances of *oriC* regions from mid-cell as a function of cell length as predicted by three different models

Model I predicts gradual separation of duplicated *oriC* regions in line with cell elongation; Model II predicts that duplicated *oriC* regions stick together until a cells reach a cell length halfway between initation of DNA replication and termination of DNA replication, after which one of the copies 'jumps' to the other side of the cell. Model III predicts that DNA regions are distributed uniformly over the nucleoid. FISH data was simulated by incorporating measurement error (standard deviation = 15nm), resolution of fluorescent microscopy (250 nm), FISH labelling efficiency (p = 0.3), and normally distributed variation in distance between foci and mid-cell (standard deviation = 0.2 µm estimated from experimental FISH data; see Roos *et al.*, 2001). Likelihood values of experimental data of figure 4.1 to be the result of model I, II and III were 821, 824, and 372, respectively. l_i^*, virtual cell length at initiation of DNA replication before cell birth; l_i, cell length at initiation of DNA replication; l_s, cell length when sister copies of *oriC* separate.

deviation of distance from mid-cell, because this was the estimated value of our experimental data (Roos *et al.*, 2001), but what is the source of this variation? The measurement error was 15 nm, which is much smaller than the observed variation. Indeed, a plot of simulated data as function

of cell length using only measurement error as a source of variation looked very similar to the theoretical lines of figure 4.2A (data not shown). Incorporating resolution and FISH labelling efficiency had an effect on the average number of foci cell, but no noticeable effect on the amount of variation. This means that other sources of variation exist.

To test if diffusion might be a source of variation, we simulated data using a model in which a DNA region moves by diffusion within a spherical region with a radius of 0.3 μm (cf. Marshall *et al.*, 1997). We produced two data sets using the theoretical positions of model I and model II to define the positions of the confined regions (fig. 4.3). The distributions of simulated data points per cell length of the two confined diffusion models resemble those of model I and model II (compare fig. 4.2B with fig. 4.3), although the distribution of the confined diffusion models are best described by a quadratic function instead of a normal distribution function. These results indicate that diffusion could be an important source of variation. However, diffusion in a spherical confined region with a radius of 0.3 μm did not explain the number of foci found in the cell centre of small to intermediate sized cells. This implies that additional factors are required to explain the distribution of *oriC* regions per cell length.

4.5 Discussion

In this study we show by comparing computer-simulated data with experimental data that DNA regions in *E. coli* may not be precisely localised; of the three models tested the model that predicts random distribution of the origins of replication within the nucleoid, most closely resembled experimental data. However, a 'random model' predicts equal distributions for every DNA region, whereas previous work showed that this is not the case; the average position of two regions in between origin and terminus, *ftsQAZ* and *minB*, analysed by FISH, differed from the average position of the origin (Roos *et al.*, 2001). Moreover, although the positions of regions in a cluster around the origin were similar and the positions of regions in a cluster around the terminus were also similar,

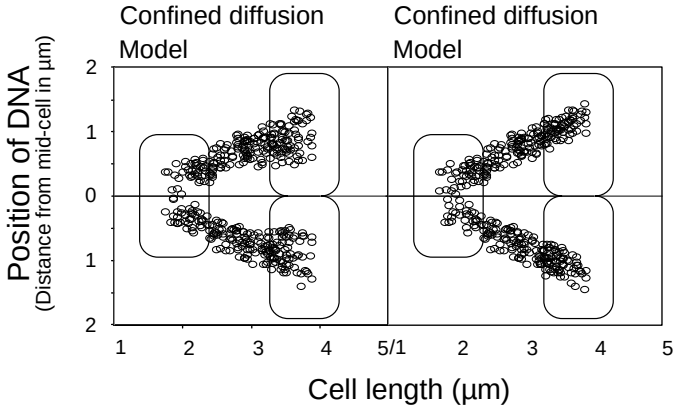


Figure 4.3 - Distance from mid-cell as a function of cell length as predicted by 'diffusion in a confined region'

The theoretical positions of model I and model II (figure 4.2A) were used to determine the central position of the confined regions. The confined region was a sphere with a radius of 0.3 µm.

the positions of regions in between these clusters and the positions of the clusters themselves were different (Niki *et al.*, 2000). This means that the organisation of sub-nucleoid DNA regions in growing *E. coli* may be fuzzy, but it is not random.

The first model that we tested predicts gradual segregation at a speed exactly equal to the speed of cell elongation (fig. 4.2A, model I). Our previous studies showed that the slope of the linear regression lines through the data points of three regions on the *E. coli* chromosome, *oriC*, *ftsQAZ*, and *minB*, did not significantly differ from 0.5, which corresponds to gradual segregation at the speed of cell elongation (Roos *et al.*, 2001). We should stress, however, that we used linear regression to find an estimate of the average position of foci within cells of increasing length. Non-gradual motion of DNA regions during the cell cycle cannot be ruled out. For instance, whether separation is gradual directly after duplication of a DNA region cannot be visualised by methods based on light-microscopy, because of the limited resolution of these methods. However, if initial separation would be gradual, then the cell length at which the regression lines cross should correspond to the cell length at which DNA regions start segregating as calculated from the *C* and *D* periods. We calculated cross-points and cell lengths at initiation of replication from previously published data (Roos *et al.*, 2001) and indeed found that these values correspond: for *oriC* 1.6 µm and 1.4 µm

respectively, for *ftsQAZ* 1.9 μm and 1.8 μm respectively, and for *minB* 2.6 μm and 2.4 μm respectively. These properties make the gradual segregation model an attractive model, although the simulated data based on this model did not sufficiently explain the data points observed in the cell centre of small to intermediate length cells (fig. 4.1).

The second model that we tested was based on GFP studies that indicated quick non-gradual movement of duplicated DNA regions after replication (fig. 4.2B; Gordon *et al.*, 1997; Sharpe and Errington, 1998). Non-gradual movement of the origin has also been used to explain FISH data (Niki and Hiraga, 1998), implicitly assuming that in cells with fluorescent signal all DNA regions were labelled. We have argued that this is probably not the case for FISH (Roos *et al.*, 2001). Therefore, we assumed here that inefficient labelling of DNA regions reduces the number of observed foci per cell (model I), or that, in addition, sister origins remain together until halfway between the length at initiation of replication and the length at termination of replication (model II). The 'jump' that follows in model II, should occur after cell birth, because otherwise it would offer no explanation for the occurrence of large numbers of small cells with a single focus as is observed in FISH experiments (Niki and Hiraga, 1998; Roos *et al.*, 2001). This implies that the origin jumps after 50% of the DNA is replicated, because according to our calculations nearly 50% of the *E. coli* chromosome has been replicated at birth (Huls *et al.*, 1999; Roos *et al.*, 2001). In any case, the simulated data of the first and the second model were very similar. In our view, this means that FISH is not a suitable method to distinguish between a gradual model and a model in which DNA regions stick together for some time before quickly segregating. The small difference we found between the likelihood of the experimental data to be the result of these two models supports this conclusion. Consequently, our results neither support nor rule out non-gradual, 'mitotic-like' segregation of DNA regions.

However, we do show here that both the model for gradual segregation and the model for non-gradual segregation are insufficient to explain the FISH data completely.

The 'fuzzy' nature of DNA localisation is reflected by the large amount of variation in experimental data. A standard deviation of 0.2 μm of the average distance of foci from mid-cell is substantial for an organism with a width of about 1 μm and a cell length of a newly born cell of about 1.9 μm. Because FISH is not a gentle method, it could enhance variation. One would expect that different labelling protocols would lead to different amounts of variation. However, FISH data obtained using a protocol slightly different from ours (Niki and Hiraga, 1998), and data obtained using a fundamentally different protocol using GFP (Gordon *et al.*, 1997) show similar amounts of variation to our data. Unfortunately, we could not compare variation quantitatively, because measures of variation were either incompatible or absent. We have tried to quantify some experimental artefacts, such as measurement error, resolution of fluorescent microscopy, and FISH labelling efficiency (the probability that a DNA target is indeed labelled by probe), but the effects on overall variation were marginal (data not shown). Therefore, it seems likely that a substantial amount of variation is caused by other, biologically relevant, factors.

One of the factors that may contribute to variation in DNA localisation is diffusion, because it moves DNA in directions other than those belonging to nucleoid segregation. In yeast and *Drosophila* it has been shown that DNA regions move as a function of diffusion, but that diffusion is limited to a confined region (Marshall *et al.*, 1997). We have tested this idea for *E. coli* by computer simulation and found that diffusion in a confined region could be responsible for a substantial amount of movement of DNA in bacteria (fig. 4.3). Irregular motion of GFP-tagged DNA regions has been reported earlier (Gordon *et al.*, 1997), but to the best of our knowledge, diffusion of individual DNA regions in bacteria has not been addressed directly. Our preliminary results of a time-lapse experiment in *E. coli* with GFP-tagged DNA regions in cells in which all active processes were blocked by NaN_3 (effectively ruling out any movement except diffusion) indicated that DNA regions indeed show substantial diffusional motion, which would allow a DNA region

to cross the width of a cell easily within 20 seconds. Marshall (Marshall *et al.*, 1997) argued that diffusion may be sufficient for the motion required for most of the nuclear processes in eukaryotic nuclei. We suspect that diffusion is at least as important in bacteria, because bacteria are so small.

Based on the analysis described here, we propose that several independent forces determine the position of a DNA region in a growing cell (fig. 4.4). In our view, nucleoid segregation, or the forces behind it, determines the direction of the average movement of DNA regions, which is gradual in line with cell elongation (van Helvoort and Woldringh, 1994). Other factors are responsible for the movement of DNA in various other directions giving rise to non-gradual movements in individual cells, but contributing only to variation when a population is studied. If diffusion is important, it is interesting to see how and to what extent diffusion is constrained: we theoretically tested a spherical confinement region with a radius of 0.3 μm (cf. Marshall *et al.*, 1997), but this was not sufficient to explain the relatively high number of foci found in the centre of small to intermediate sized cells. Overall movement of DNA regions is probably not solely connected to cell elongation; nucleoids in fast growing cells appear triangular in shape, suggesting that origins are not always oriented in the length axis of the cell.

4.6 Acknowledgements

The authors would like to thank Conrad Woldringh for support and discussion.

4.7 References

Gordon, G. S., D. Sitnikov, C. D. Webb, A. Teleman, A. Straight, R. Losick, A. W. Murray, and A. Wright. 1997. Chromosome and low copy plasmid segregation in *E. coli:* visual evidence for distinct mechanisms. *Cell* **90**: 1113-1121.

Gordon, G. S., and A. Wright. 2000. DNA segregation in bacteria. Annu. *Rev Microbiol* 54: 681-708.

Huls, P. G., N. O. Vischer, and C. L. Woldringh. 1999. Delayed nucleoid segregation in *Escherichia coli*. *Mol Microbiol* 33: 959-970.

Koppes, L. J., N. Overbeeke, and N. Nanninga. 1978. DNA replication pattern and cell wall growth in *Escherichia coli* PAT 84. *J Bacteriol* 133: 1053-1061.

Koppes, L. J., C. L. Woldringh, and N. Nanninga. 1999. *Escherichia coli* contains a DNA replication compartment in the cell center. *Biochimie* 81: 803-810.

Lemon, K. P., and A. D. Grossman. 1998. Localization of bacterial DNA polymerase: evidence for a factory model of replication. *Science* 282: 1516-1519.

Margolin, W. 2000. Green fluorescent protein as a reporter for macromolecular localization in bacterial cells. *Methods* 20: 62-72.

Marshall, W. F., A. Straight, J. F. Marko, J. Swedlow, A. Dernburg, A. Belmont, A. W. Murray, D. A. Agard, and J. W. Sedat. 1997. Interphase chromosomes undergo constrained diffusional motion in living cells. *Curr Biol* 7: 930-939.

Niki, H., and S. Hiraga. 1998. Polar localization of the replication origin and terminus in *Escherichia coli* nucleoids during chromosome partitioning. *Genes Dev* 12: 1036-1045.

Niki, H., Y. Yamaichi, and S. Hiraga. 2000. Dynamic organization of chromosomal DNA in *Escherichia coli*. *Genes Dev* 14: 212-223.

Roos, M., A. B. M. van Geel, M. E. G. Aarsman, J. T. M. Veuskens, C. L. Woldringh, and N. Nanninga. 1999. Cellular localization of *oriC* during the cell cycle of *Escherichia coli* as analyzed by fluorescent in situ hybridization. *Biochimie* 81: 797-802.

.....2001. The replicated *ftsQAZ* and *minB* chromosomal regions of *Escherichia coli* segregate on average in line with nucleoid movement. *Mol Microbiol* 39: 633-640.

Sharpe, M. E., and J. Errington. 1998. A fixed distance for separation of newly replicated copies of *oriC* in *Bacillus subtilis:* implications for co-ordination of chromosome segregation and cell division. *Mol Microbiol* 28: 981-990.

.....1999. Upheaval in the bacterial nucleoid. An active chromosome segregation mechanism. *Trends Genet* 15: 70-74.

Sharpe, M. E., P. M. Hauser, R. G. Sharpe, and J. Errington. 1998. *Bacillus subtilis* cell cycle as studied by fluorescence microscopy: constancy of cell length at initiation of DNA replication and evidence for active nucleoid partitioning. *J Bacteriol* 180: 547-555.

Van Helvoort, J. M. L. M., and C. L. Woldringh. 1994. Nucleoid partitioning in *Escherichia coli* during steady-state growth and upon recovery from chloramphenicol treatment. *Mol Microbiol* 13: 577-583.

Webb, C. D., P. L. Graumann, J. A. Kahana, A. A. Teleman, P. A. Silver, and R. Losick. 1998. Use of time-lapse microscopy to visualize rapid movement of the replication origin region of the chromosome during the cell cycle in *Bacillus subtilis*. *Mol Microbiol* 28: 883-892.

4.8 Appendix

4.8.1 Determination of the cell length at initiation and termination of DNA replication

We assume that within the time period 0 t T_d cell length growth in a

steady-state population is exponential (Koppes *et al.*, 1978), thus

$$l(t) = 2^{\frac{t}{T_d}} \cdot l_0 \qquad (1)$$

in which T_d = the moment of cell division, and l_0 = the length of a new born cell.

Consequently

$$ln\{l(t)\} = \frac{ln\,2}{T_d} t + ln(l_0) \qquad (1')$$

The moment of initiation of DNA replication, t_i, is derived from

$$t_i = T_d - (C + D)$$

in which C = duration of one round of replication, D = period from termination of replication to cell division. Consequently, the length at the moment of initiation of DNA replication, l_i^*, is found by:

$$l_i^* = l(t_i) = 2^{\frac{t_i}{T_d}} l_0$$

or

$$l_i^* = 2^{\frac{T_d - (C + D)}{T_d}} l_0$$

Similarly the length at termination (or end) of DNA replication, l_e, was derived by substituting t in (1) by the moment of termination of DNA replication

$$t_e = T_d - D$$

Experimental values of the model-parameters are T_d = 79 min., C = 70 min., and D = 42 min. (Huls et al., 1999). So t_i = -33 min.,

l_i^* 0.75 l_0, t_e = 37 min, and l_e 1.4 l_0. Thus replication initiates before cell birth in a virtual cell with length , corresponding to a prospective daughter cell before cell division (Roos *et al.*, 2001).

4.8.2 Modelling the distance of *oriC* from mid-cell

The theoretical position of the origin of replication of *E. coli*, *oriC*, was represented as the distance of that DNA region from mid-cell as a function of cell length. Definitions:

- d = d_r^n = distance from cell centre of copy number r out of n copies of *oriC* present in the cell
- l = cell length
- l_0 = length at birth of the cell
- l_i = virtual length at initiation of DNA replication before cell birth
- l_s = length when sister copies of *oriC* separate
- l_i = length at initiation of DNA replication
- l_e = length at termination of DNA replication
- l_d = length of a dividing cell
- L = length of a nucleoid

4.8.2.1 Model I (fig. 4.4A)

For model I we assume that each of the two or four copies of *oriC* in the cell increases its distance from mid-cell (or from the centre of a prospective daughter cell) according to the model

$$d = b \cdot l + a \tag{2}$$

in which the slope b = ½ by definition, because in a steady-state population the distance between a DNA region and the centre of a prospective daughter cell at cell division is equal to the distance between a DNA region and the cell centre at cell birth. By assuming that DNA replication takes place in the cell centre we find that initiation of DNA replication takes place in the cell centre, such that upon initiation of DNA replication

$$d(l_i^*) = 0$$

Consequently, the y-intercept of equation (2)

$$a = -\frac{1}{2} l_i^*$$

Substituting *b* and *a* in equation (2) gives

$$d = \frac{1}{2} l - \frac{1}{2} l_i^*$$

For each copy of *oriC* we find

If $l_0 \quad l \quad l_i$ then $d_1^2 = -d$, and $d_2^2 = d$

If $l_i < l \quad l_d$ then $d_1^4 = -d$, $d_2^4 = -d(l_i) + \frac{1}{2}(l - l_i)$,

$d_3^4 = d(l_i) - \frac{1}{2}(l - l_i)$, and $d_4^4 = +d$

4.8.2.2 Model II (fig. 4.4B)

For model II we assumed that sister copies of *oriC* move away from the cell centre together until reaching a cell length halfway between initiation and termination of DNA replication:

$$l_s = l_i^* + \frac{1}{2}(l_e - l_i^*)$$

After reaching ls one of the copies jumps to the other side of the cell, where it continues to move away from the cell centre. Thus for each copy of *oriC*:

If $l_0 \quad l \quad l_s$ then $d_1^2 = d_2^2 = -d$

If $l_s < l \quad l$ then $d_1^2(l) = -d$, and $d_2^2 = +d$

If $l_i < l \quad l_d$ then $d_1^4(l) = d_2^4 = -d$, and $d_3^4 = d_4^4 = d$

4.8.2.3 Model III (fig. 4.4C)

In this model each copy of *oriC* is somewhere within the boundaries of the nucleoid(s), hence their theoretical position depends on the length of

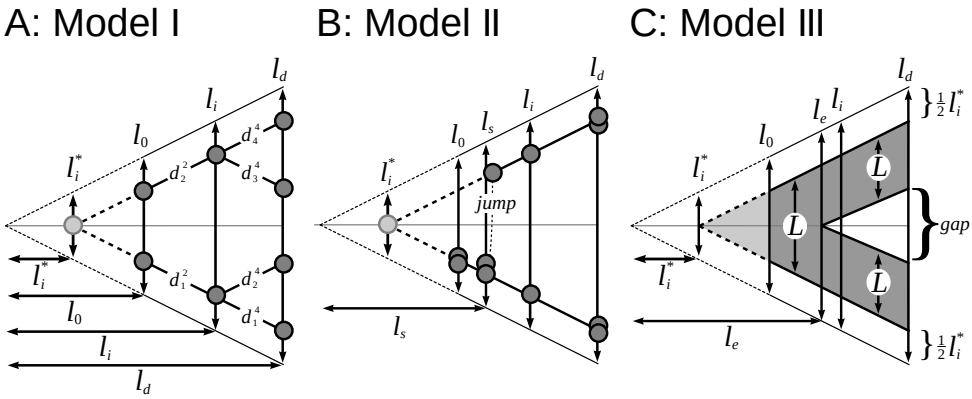


Figure 4.4 - Schematic representation of the parameters used to derive model I (A), model II (B), and model III (C)

See the appendix for descriptions of these parameters. Dotted lines represent the length of and distances in the virtual cell before cell birth.

the nucleoid(s). In the virtual cell at initiation of DNA replication, the length L of the nucleoid

$$L(l_i^*) = 0$$

which means that l_i^* represents the total length of the DNA free zones at the poles of the cell. Assuming that the length of these zones remains constant during the cell cycle we find that

$$\text{If } l_0 \quad l \quad l_e \text{ then } L = l - l_i^*$$

We assume that after termination of DNA replication two separate nucleoids appear. The length of the 'gap' between the two nucleoids in the cell centre increases until it equals l_i^* upon cell division, thus

$$gap(l) = l_i^* \frac{l - l_e}{l_d - l_e}$$

from which we derive the length of each separate nucleoid:

$$\text{If } l_e < l \quad l_d \text{ then } L = \frac{1}{2}\{(l - l_i^*) - gap(l)\}$$

For each copy of *oriC* in a cell we obtain a distance value by taking a random item L_r^n out of a uniform distribution:

0 $\overline{L_r^n}$ L :

If $\overline{l_0}$ l l_e then $d_1^2 = \overline{L_1^2} - \frac{1}{2}L$, and $d_2^2 = \overline{L_2^2} - \frac{1}{2} - \frac{1}{2}L$

If $l_e < l$ l_d then $d_1^4 = \overline{L_1^4} - \left\{L + \frac{1}{2}gap(l)\right\}$, $d_2^4 = \overline{L_2^4} - \left\{L + \frac{1}{2}gap(l)\right\}$,,

$d_3^4 = \overline{L_3^4} + \frac{1}{2}gap(l)$, and $d_4^4 = \overline{L_4^4} + \frac{1}{2}gap(l)$

The second term of each d_r^n places the cell centre at position 0 (the origin), assuming that nucleoids are symmetrically positioned in the cell.

4.8.3 FISH labelling efficiency

The probability of a cell to have r out of n copies of *oriC* labelled is

$$_r^n P = \binom{n}{r} \cdot p^r \cdot (1 - p)^{r - n} \tag{3}$$

in which p is the probability of a single DNA region to be labelled. The probability to have at least one copy of *oriC* labelled in a cell is defined by

$$_{r>0}^{n} P = {}^2P \cdot \left({}_1^2P + {}_2^2P\right) + {}^4P \cdot \left({}_1^4P + {}_2^4P + {}_3^4P + {}_4^4P\right) \tag{4}$$

which can be estimated by scoring of the fraction of cells with at least one focus out of 100 randomly selected cells. ^2P, and ^4P are the probabilities of a cell to contain two or four copies of *oriC* respectively. We assume that the number of cells in a steady-state population is described by

$$N(l) = N_0 2^{-\frac{l - l_0}{l_0}} \tag{5}$$

from which we derive ^2P by integrating (5) between l_0 and l_i.
Consequently $^4P = 1 - {}^2P$. By substituting in equation (4) each $_r^n P$ by

equation (3), and the estimates of $_{r>1}^{n}P$, ^{2}P, and ^{4}P we can derive p. The experimental values of the model parameters for *oriC* were $_{r>1}^{n}P$ 0.5 (Roos *et al.*, 2001), ^{2}P 0.8 , and ^{4}P 0.2 ; consequently p 0.3 .

Chapter 5

ASSUMPTIONS REVISITED

5.1 Introduction

To reconstruct how replicated DNA regions might move as a function of cell growth, we used FISH to reveal the positions of the copies of specific DNA regions within *E. coli* cells and we applied what we know about the relationship between replication and the cell cycle (chapters 2 and 3).

In our view, the most straightforward interpretation of our data is that, *on average*, replicated copies of specific DNA regions segregate gradually, in line with cell elongation. Their respective positions in the cell correspond to those on the chromosome map (fig. 5.1; chapter 3). However, it should be noted that this interpretation is based on average positions as determined by linear regression, and it is therefore based on the assumption that the intracellular position of a DNA region is linearly related to cell length.

An alternative interpretation is that copies of DNA regions stay together for some time after replication and then separate abruptly. This interpretation is based on time-lapse studies in bacteria (Gordon *et al.*, 1997; Teleman *et al.*, 1998), which revealed that separation of replicated GFP-

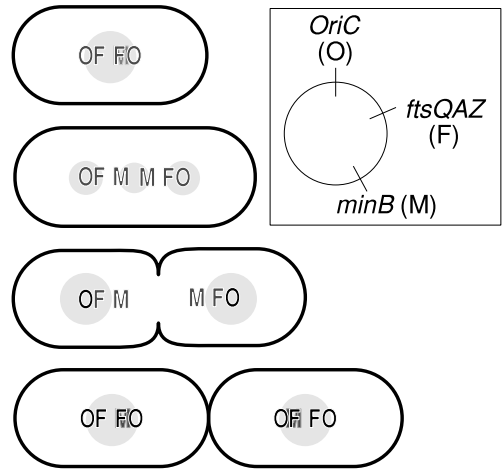


Figure 5.1 - Average order of *oriC* (●), *ftsQAZ* (F), and *minB* (M) regions within slow growing *E. coli* cells

See chapter 3 for details about the positions of the regions. Grey circles indicates the positions of the replisomes (smaller circles indicate replisomes that are being broken down or that are being build-up). The inset shows the relative positions of the DNA regions on the *E. coli* chromosome.

tagged DNA-regions close to or encompassing *oriC* may be abrupt, thus not in line with cell elongation. This has led authors to propose that replicated DNA in bacteria is pulled apart by unknown factors in an active process, reminiscent of chromosome segregation during mitosis in eukaryotes. However, the observed movements of GFP foci do not necessarily correspond to average movement of DNA regions in a population of cells. Instead, they could reflect movements that are unrelated to DNA partitioning, such as Brownian motion of DNA (Roos, M. *et al.*, 2001a; sections 5.2.3 and 6.3). Indeed, variation observed in the intracellular position of GFP-tagged DNA regions (e.g. Webb *et al.*, 1997) indicated that the arrangement of the chromosome is dynamic (Teleman *et al.*, 1998).

Before speculating about the biological factors underlying either of these interpretations (gradual or abrupt segregation), we decided to evaluate the models by computer simulation. We simulated a FISH experiment for cells in which a particular model applies, and then compared the simulated data with our observed data (fig. 5.2). Computer simulation forces us to make every implicit assumption in a model explicit. One example is the assumed amount of variation in the position of a DNA region. The amount of variation in our experimental data was more than we could account for theoretically (see chapter 4 and section 5.2.3).

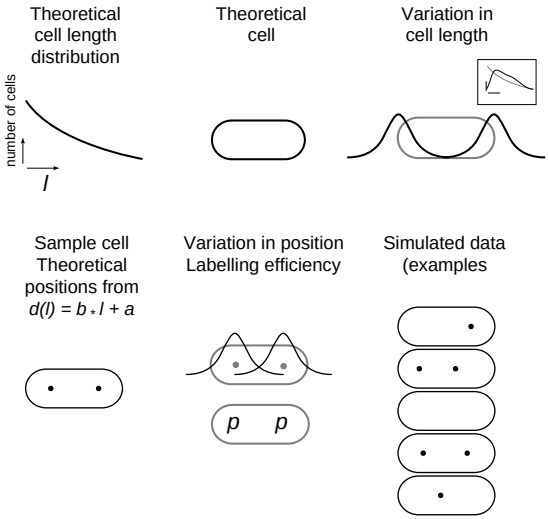


Figure 5.2 - Scheme of simulation method

Top-left: Theoretical cells were selected from an exponential cell length distribution (e.g. top-middle). Top-right: to each cell a value was added from a Gaussian frequency distribution (s= 0.1 μm; inset: typical cell length distribution after applying this step to many cells). Bottom-left: from the cell length of each cell the positions of DNA regions were calculated. Bottom-middle: to each position a value was added from a normal distribution (s=0.2 μm) or from a 'diffusion distribution' (see chapter 4 for details). Furthermore, positions were discarded using the probability p that a DNA region is labelled.

Therefore, we introduced Brownian motion of DNA as a source of variation and as an important aspect of DNA organisation in *E. coli*. A second example is the time period that replicated-copies of *oriC* stick together before they separate in the 'mitotic-like' model. Niki and co-workers proposed that in newborn cells a single copy of *oriC* is on one side of the cell (Niki and Hiraga, 1998). In fast growing cells (T_d=55 min.), following duplication, one copy jumps to the other side of the cell. Niki also proposed that in slow growing cells (e.g. T_d=80 min.), origin and terminus rotate from polar positions to the middle of the cell during the B period (Niki *et al.*, 2000). However, in our case a B-period was not observed (Huls *et al.*, 1999; fig. 1.7a), even though the doubling time of the cells in our culture was the same as Niki's (~80 min.; the carbon source was glucose). Consequently, in order to have a replicated copy of *oriC* jump after cell birth, we assumed that replicated copies of *oriC* stick together until at least 50% of the genome is replicated (cf. Hiraga, 2000). Therefore, we called this model the 'sticky-jump' model (chapter 4).

Our simulations showed that we could not accept the gradual segregation model or the sticky-jump model without further evidence.

Neither of them produced data sufficiently similar to our experimental data (see chapter 4). Thus, even though we have experimental data that reflect DNA organisation as a function of the cell cycle in *E. coli*, we do not have a proper model to describe them yet. How should we proceed? In the next section I take a critical look at the assumptions on which the models and simulations treated in this thesis were based. This may help us to define a new model, which I suggest could be a three-dimensional model in which Brownian motion of DNA is an important factor (see section 6.3).

5.2 Assumptions

In this section I treat the following assumptions: *The intracellular position of a DNA region is dependent on cell length* (section 5.2.1, page 88); *DNA moves along the length axis of the cell only* (section 5.2.2, page 90); *variation in position is normally distributed (standard deviation 0.2 μm) or defined by diffusional motion within a confined region* (section 5.2.3, page 90); *DNA organisation is intact after FISH* (section 5.2.4, page 92); *DNA regions are not always labelled* (section 5.2.5, page 92); *a FISH focus corresponds to no more than one DNA region* (section 5.2.6, page 94); *the centre of a FISH focus corresponds to the intracellular position of a hybridised DNA region* (section 5.2.7, page 95); *there is no visible distinction between the two poles of E. coli cells* (section 5.2.8, page 96); *cell cycle parameters at 28 °C in minimal glucose medium: C = 70 minutes, D = 42 minutes, and T_d = 79 minutes* (section 5.2.9, page 100); *replication takes place in the middle of prospective-daughter cells* (section 5.2.10, page 101). These assumptions were used for the analysis of FISH data in chapters 2 and 3, or for the simulation of FISH data in chapter 4. In most cases the assumption applies to both the experimental and the simulated data.

5.2.1 The intracellular position of a DNA region is dependent on cell length

The primary assumption on which the analyses of chapters 2, 3 and 4 are based is that, statistically, the position of a DNA region in a cell is

dependent on cell length. Under this assumption, there may be variation in position, but not in length (the independent variable). This is a prerequisite of linear regression of type I (Sokal and Rohlf, 1995).

Because cell length is a measured quantity, length as a variable will show some variation. However, this variation appeared negligible with respect to other sources of variation (see chapter 4). Nevertheless, is the intracellular position of a DNA region really so strictly dependent on cell length?

GFP studies have shown that replicated copies of a DNA region may move apart at a different rate than cell elongation (Glaser *et al.*, 1997; Gordon *et al.*, 1997; Sharpe and Errington, 1998). However, part of the observed movement may have been the result of Brownian motion, hence, diffusion of DNA (see section 5.2.3). Because Brownian motion is random, its net contribution to overall DNA movement is zero, and the assumption that the intracellular position of a DNA region depends on cell length could still hold.

Variation in cell length could be in conflict with the assumption that the intracellular position of a DNA region is dependent on cell length. For instance, the average cell length of newborn cells has a coefficient of variation of about 10%. This is largely due to variation between individual cells (see for instance Koppes, 1980), but this could imply variation in parameters, such as cell shape and DNA organisation. For instance, compared to our slow-growing cells (28 °C, minimal glucose medium, T_d=80 min.), the same cells grown at 37 °C in rich medium (T_d=20 min.) are thicker on average and show multi-fork replication (e.g. fig. 1.7c). In response to multi-fork replication, internal DNA organisation is likely to change, e.g. the shape of the nucleoids changes, possibly because the segregation axes are tilted with respect to each other (Woldringh *et al.*, 1994). Clearly, in these cases it would be too much of a simplification to assume that cell length is the only determinant of DNA position.

5.2.2 DNA moves along the length axis of the cell only

Although a bacterial cell is a three-dimensional structure, we analysed the positions of DNA regions in one dimension: the length axis of the cell. We did score the positions of FISH foci in two dimensions (x and y coordinates of the computer image), but calculated the projection of each focus on the length axis of the cell. This was done for simplicity and because the size of a bacterium barely allows for a two-dimensional or three-dimensional analysis of nucleoid structure. The width of a slow growing *E. coli* cell is less than 1 ∝m, the width of the nucleoid is even less. We estimated that a typical nucleoid is about 1.7 times wider than a FISH focus (data not shown, but compare FISH and DAPI images of fig. 2.3). A FISH focus may serve as a representation of the point-spread function (PSF) of a microscope, assuming that the physical size of a labelled DNA-region is well below the optical resolution (see section 5.2.7). A PSF is the microscope-image of an infinitely small point source; the PSF of our microscope was measured to be about 270 nm wide at half its height (data not shown).

5.2.3 Variation in position is normally distributed (standard deviation 0.2 μm) or defined by diffusional motion within a confined region

The amount of variation in our experimental data was substantial with respect to the size of *E. coli*, but the sources of this variation are largely unknown (see chapter 4). One possible source of variation is Brownian motion, hence diffusion, of DNA regions. The variation in the position of GFP foci (Gordon *et al.*, 1997; Teleman *et al.*, 1998; Webb *et al.*, 1998) may be an indication of Brownian motion (see section 5.2.1; chapter 4), and computer simulations confirmed that a certain amount of variation in our FISH data might be explained in this way (Roos, M. *et al.*, 2001a). In yeast and *Drosophila* nuclei, Brownian motion of DNA was shown by tagging either with GFP (yeast) or by fluorescently labelled topo-isomerase II (*Drosophila*). The motion was confined to a spherical region; radii of the confinement-regions were 0.3 μm in yeast, and 0.9 μm in

Drosophila (Marshall *et al.*, 1997). Does DNA in *E. coli* move in a similar way?

Preliminary data that we obtained using a protocol resembling that of Marshall indicated that diffusion is important in *E. coli* (chapter 4). We used a GFP-LacI fusion-protein to indicate the position of a cassette of *lacO* regions inserted near *oriC*; expression of the fusion-protein was under the control of the *araBAD* promoter (cf. Gordon *et al.*, 1997). From 20 minutes after induction, some cells were uniformly filled with fluorescent signal, some had no signal at all, some had distinct foci in the nucleoid, and some had extremely bright foci. The heterogeneity might be related to the use of the *araBAD* promoter, which may produce a mixture of induced and uninduced cells (Siegele and Hu, 1997). We suspect that the extremely bright foci represent inclusion bodies. Inclusion bodies are a means to store excess material and are often formed by recombinant proteins (see for instance Swartz, 1996, or Carrio *et al.*, 1998; 1999; Carrio *et al.*, 2000). The putative GFP-inclusion-bodies are most prominent when cells are grown at 37°C in rich medium. GFP-foci are then clearly visible in phase-contrast images as black dots, sometimes with a white core. Roughly estimated, such dots could contain about 40% more material than the surrounding cytoplasm (van Munster, personal communication). At lower temperatures this effect is supposed to be less of a problem (e.g. Gordon *et al.*, 1997), but we feel that we cannot rule out the possibility that the accumulation of GFP-fusion protein is less visible, but not absent. It is thinkable that the observed movement of GFP foci may actually be related to the formation and movement of inclusion bodies. In our experience inclusion bodies are often found at cell poles (cf. Carrio *et al.*, 1998) and in between nucleoids. Possibly, GFP-inclusion-bodies form near *oriC* first, but then move towards a cell pole or in between nucleoids. Because of these uncertainties, we were unable to quantify to what extent diffusion in *E. coli* was confined. Therefore, if we want to study DNA-movement by time-lapse microscopy and GFP-tagging, it has to be ascertained that the GFP-foci reflect the natural positions of the DNA region under study.

5.2.4 DNA organisation is intact after FISH

When studying DNA organisation as a function of the cell cycle it is important that a labelling experiment does not interfere with DNA organisation. FISH is not a gentle method and it may introduce severe artefacts (see for instance Trask *et al.*, 1993), but it is not entirely clear in what specific way(s) DNA organisation is affected. Our methods to test for artefacts caused by FISH are limited. Basically, we evaluated cell morphology by phase contrast microscopy and nucleoid morphology by fluorescence light microscopy, using DAPI to stain DNA. We tested various intermediate steps of the FISH protocol in this way. Although we found no clear indications of substantial morphological re-arrangements, we cannot exclude that DNA is re-organised by FISH. For instance, osmium-tetroxide (OsO_4), which we used to fix cells, has been reported to cause a contraction of the nucleoid, possibly as a result of protein dissociation (Valkenburg and Woldringh, 1984). In the case of FISH, this release of proteins is quite useful as native proteins could make DNA less accessible to a DNA probe.

5.2.5 DNA regions are not always labelled

We have assumed that FISH may not label all possible DNA regions in a cell. This is obvious in cases where no FISH signal was present: cells with DNA should have at least one copy of all DNA regions. For some DNA probes, the number of cells without any detectable signal was substantial (e.g. *oriC* and *minB*, chapter 3). In general, the number of foci was below the expected number of DNA regions (see table 1, chapter 3). There are two possible explanations: (*i*), probe DNA was not able to enter the cell, and (*ii*), probe DNA did enter the cell, but was not able to hybridise with the target region. We assumed that the second reason would be the most important, because one of our probes seemed to be able to produce a signal in almost all cells (*ftsQAZ*, see chapter 3). In our view, this shows that the pieces of DNA of this probe were capable of entering the cells after these were made permeable. As is common for FISH, we used pieces of DNA of around 300 bp to hybridise with a region of a few kbp.

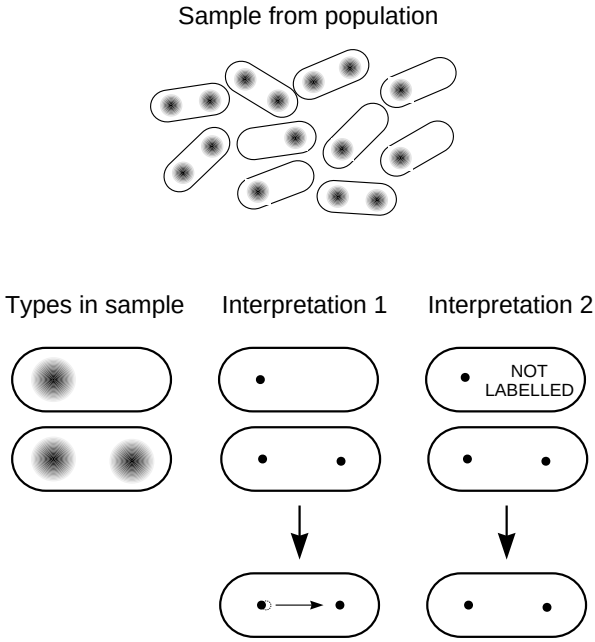


Figure 5.3 - Different interpretations of FISH data

Top: a sample from a hypothetical population of newborn cells with one or two foci (bottom left). Interpretation 1: after replication of a DNA region, one copy moves quickly ('jumps') to the other side of the cell. This interpretation is based on the assumption that all DNA regions are labelled, and follows from the fact that no intermediate focus positions are found in the sample. Interpretation 2: two DNA regions are positioned at opposite poles in the cell. This interpretation is based on the assumption that in cells with one focus, the second DNA region was not labelled.

Because the pieces of DNA of the other probes were also around 300 bp, we could see no reason why they would not be able to enter the cells. Therefore, we had to assume that it is possible that some DNA regions in a cell are labelled, whilst others are not. We estimated the probability of labelling a DNA region from the ratios of the average number of foci observed and the number of DNA regions expected per cell (see appendix of chapter 4).

Making a different assumption about FISH efficiency could lead to a different interpretation of FISH data. Suppose we have a population of newborn cells with either one focus near a pole, or two foci near opposite poles (fig. 5.3). If we assume all DNA regions were labelled we might conclude that upon duplication one DNA region jumps to the other side of the cell, because we did not find intermediate configurations (e.g. a cell with one polar and one centrally located focus). However, if we

assume that not all DNA regions were labelled, we would simply conclude that in cells with one focus, the DNA region on the other side of the cell was not labelled. Because we chose the second interpretation, we did not plot data of cells with one focus separate from data of cells with two (or more) foci (see also section 5.2.8).

5.2.6 A FISH focus corresponds to no more than one DNA region (chapters 2 and 3)

For simplicity, we assumed that a FISH focus was the result of the labelling of one DNA region. However, if two DNA regions were less than 250 nm apart, we would see only one FISH focus, because of the limited resolution of light microscopy. This situation worsens, when a specimen is out of focus. Furthermore, we scored the majority of foci that appeared extended in shape as single foci, whilst they could have represented two (or more) neighbouring DNA-regions at a distance slightly over 250 nm (for an example see the bottom left image of *ftsQAZ* foci in fig. 3.3). In all these cases, we may not only have misjudged the number of DNA regions, but we may also have misjudged the position of either of the two regions.

If we would know exactly how many DNA-regions were present in a focus, we might be able to find their positions using image-processing, even if their FISH signals were inseparable by eye (Noordmans and Smeulders, 1998). Unfortunately, we could not tell how many DNA regions were present in an individual focus. The intensity of a focus could not be used, because the variation in intensity between individual foci was too large; some foci were barely visible, others were very bright. In theory, we could use the *C*, *D*, and T_d to calculate the number of DNA regions per cell length (chapters 2 and 3; Huls *et al.*, 1999), but it would be statistically wrong to use this *average* number of DNA regions as the number of DNA regions in an *individual* cell (see also section 5.2.5). Therefore, we have not tried to resolve this issue when scoring the positions of foci. Consequently, when the two replicated DNA regions

started to separate, we first scored single foci. When they had separated far enough (at least 250 nm), we scored two foci. Hence, we did not find pairs of foci with less than 250 nm between them. Naturally, we did not interpret this as a sudden jump of DNA regions from 0 to 250 nm.

5.2.7 The centre of a FISH focus corresponds to the intracellular position of a hybridised DNA region

We assume that a FISH focus emanates from a complex formed by hybridisation of our DNA probe with a single DNA region, and that the diameter of this complex is below 250 nm, i.e. below the resolution of our microscope. We may then consider the complex as a point source and take the centre of the FISH focus as the exact position of the DNA region. However, is the diameter of the complex of labelled DNA smaller than 250 nm?

The success of FISH is related to the size of the target region. Therefore, the DNA probe is often directed against a larger region than the actual target of interest. For instance, the minimal *oriC* region is 248 bp, but the probe we used was constructed from a 6-kbp region containing *oriC* (see chapter 2). We may consider the sizes of the fluorescing antibodies negligible, but a 6 kbp stretch of DNA could, when completely linear, stretch along the full length of a new-born *E. coli* cell (~2 μm). Hence, we could expect extended FISH foci or even multiple foci, because a DNA probe is a collection of pieces of DNA around 300 bp in length. However, the length of DNA must be reduced about a 1000-fold in *E. coli*, and although the level of the required compaction may vary locally, a 10-fold degree of compaction already brings a 6 kbp region below optical resolution limits.

To be more certain about what kind of FISH signal one stretch of DNA would produce, we could evaluate the signal resulting from two-colour FISH on adjacent DNA-regions. However, we would need to know exactly how many DNA regions are in a focus. Consequently, we would need to be sure about the number of DNA regions present in each

individual cell. Unfortunately, we are not (see section 5.2.5).

In eukaryotic nuclei the number of FISH-targets is better defined, because the amount of DNA does not change except during S-phase, when all DNA is copied exactly once. To the best of my knowledge, when FISH is performed on regions of a few thousand base pairs in eukaryotic nuclei, the number of foci is never higher than the number of targets (apart from background signal). This shows that, at least in eukaryotes, it is unlikely that a probe targeted at a single region of a few thousand base pairs, produces multiple foci or a large labelled region that does not emanate from a point source. Hence, if we find a FISH signal that is different from what we expect from a point-source, we may assume that more than one neighbouring DNA-region has been labelled.

5.2.8 There is no visible distinction between the two poles of *E. coli* cells

When the intracellular positions of FISH foci are plotted, one of the poles is used as reference. How this 'reference-pole' is chosen can have important consequences on the interpretation of the data. Even though one of the poles of an *E. coli* cell is a new pole, having been created during the preceding cell division, it was not possible to discriminate between the two poles in our images. Therefore, we chose a random pole as reference. This prevents artificial bias in our plots, but may have introduced extra variation in intracellular position.

In a number of publications that pole is chosen as reference which is nearest to any one of the foci (e.g.Niki and Hiraga, 1998; Niki *et al.*, 2000), in other cases the nearest pole is chosen for each focus in a cell (e.g. Sharpe and Errington, 1998). In a number of cases, I could not find which of these methods was used, although I suspect the former; e.g. Gordon *et al.*, 1997; Weitao *et al.*, 2000. This 'sorting' of data can lead to biased interpretations. Suppose we have cells of the same length and with two labelled DNA-regions. On average, the two regions are symmetrically

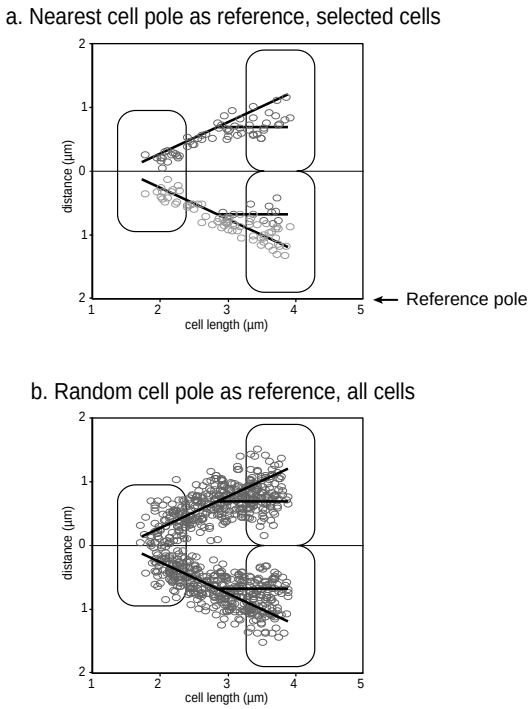


Figure 5.4 - Two plot methods

a: simulated distances of foci from mid-cell based on Model 1 (black line, see chapter 4 for details)were classified using the 'nearest cell pole' as reference (indicated at the bottom of the chart; see text).. Cells were selected on the basis of the number of visually separable foci; data from cells with two foci are depicted here. Note that most data points are below the black lines; they are shifted towards the reference pole.
b: distances from the same simulation, but now using a randomly chosen cell pole as reference without selecting cells on the basis of visually seperable foci. Note that the data points are evenly distributed around the black lines.

positioned on either side of the cell, but because there is variation in the position of these regions, the exact position of a DNA region in an individual cell deviates from the average position. Let us assume that the shape of the distribution of these positions is Gaussian (a 'normal' distribution), and because we have no further information, we assume that the positions of the two DNA regions in a cell are independent of each other. Consequently, which cell pole is nearest to a DNA region is a matter of chance. To show that using the 'nearest cell pole' as reference can lead to biased results, we re-plotted the simulated data based on model 1 in chapter 4, using cells with two foci (fig. 5.4a). Comparing the data points with the line that represents model 1, we see that all data points are shifted towards the reference pole (they are below the black lines in fig. 5.4a). In contrast, no shift of data is observed when a randomly chosen cell pole is used as reference (fig. 5.4b).

Similarly, in cells with *one* visible focus, using the nearest cell pole as reference may lead to biased results (fig. 5.5). To illustrate this: suppose we have a hypothetical class of cells, all 2 μm long with one DNA region that is normally distributed around 0.25 μm from the cell centre (standard deviation s = 0.2 μm; dark grey curve in fig. 5.5). One consequence is that some positions scored on one side of the cell belong to the distribution of which the average is on the other side of the cell (cross-hatched section in fig. 5.5). If we use the nearest cell pole as reference, we would assign all positions to a distribution on one side of the cell (black curve in fig. 5.5). Because this distribution is not normal, it is more difficult to estimate the average position of the DNA regions correctly, e.g. the peak of the distribution does not correspond to this position. In contrast, if we randomly choose a reference pole, we find two peaks on either side of the cell (dashed curve in fig. 5.5). The distribution is the sum of two normal distributions (light grey curves in fig. 5.5). Consequently, it is relatively easy to find the original average position. Indeed, we used this kind of reasoning to find average positions per cell length in chapter 3 (Roos *et al.*, 2001b).

In some cases, data of cells with one focus have been plotted and analysed separately from cells with two foci (e.g. Gordon *et al.*, 1997; Niki and Hiraga, 1998; Niki *et al.*, 2000; Sharpe and Errington, 1998; Weitao *et al.*, 2000). I already mentioned that we did not do this, because we felt we could not be certain that the number of foci corresponded to the number of DNA regions in a cell (section 5.2.5). However, if FISH data are separated according to the number of foci observed, and distances are plotted using the nearest cell pole as reference, one might get a wrong impression about the movement of DNA regions, because the average position of the DNA regions in cells with one focus is determined differently from that in cells with two foci. We chose the reference pole randomly and did not find a clear difference in the distribution of foci in cells with one focus, and the distribution of foci in cells with two foci (fig. 5.6).

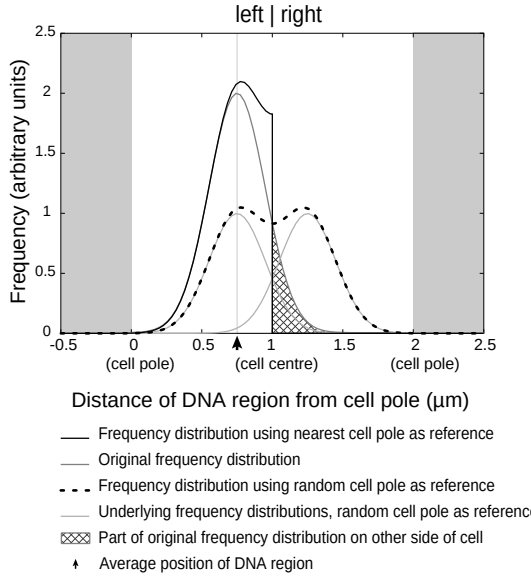


Figure 5.5. - Bias in focus positions in cells with one focus

The theoretical average position of the focus is on the left side of the cell (arrow and grey line). Dark grey curve: Gaussian frequency distribution of the position of the focus (using the left pole as reference). Note that part of this distribution is on the right side of the cell (hatched region). Black curve: frequency distribution when the nearest cell pole (placed on the left) is chosen as reference. Note that the shape of this curve is not Gaussian, and that the position of the maximum does not coincide with the theoretical average position. Dotted curve: frequency distribution of positions when a random pole is chosen as reference. This distribution is the sum of two Gaussian distributions of which the absolute average position coincides with the theoretical average position.

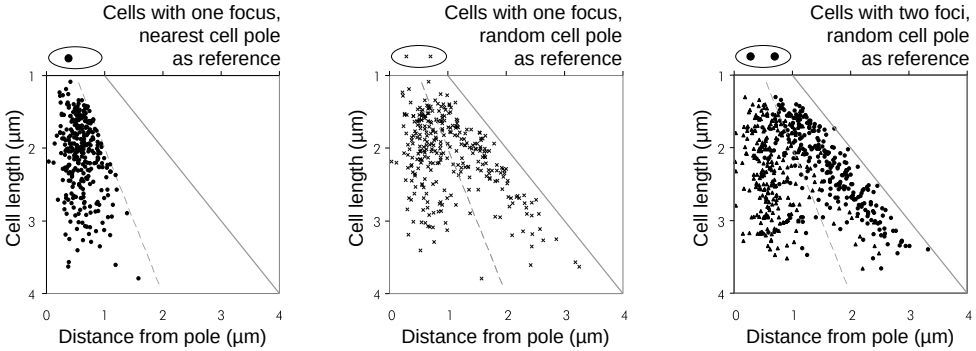


Figure 5.6 - Different ways of plotting *oriC* FISH data

Left: *oriC* positions of cells with one focus using the 'nearest cell pole' (placed on the left) as reference. Middle: the same *oriC* positions, but using a randomly chosen cell pole as reference. Right: *oriC* positions of cells with two foci, using a randomly chosen cell pole as reference. Note that the general patterns of data points in the middle chart and in the right chart appear similar. See chapter 2 for experimental details.

5.2.9 Cell cycle parameters at 28 °C in minimal glucose medium: C = 70 minutes, D = 42 minutes, and T_d = 79 minutes

To assist in the interpretation of intracellular positions of (copies of) DNA regions visualised by FISH (or GFP), it is important to know how many of those DNA regions are present in a cell on average. This number might be estimated from the number of FISH foci observed, but this is unreliable because of the limited resolution of light microscopy, the possibility that not all DNA regions are labelled, and the possibility that foci lie on top of each other (Roos, Marco *et al.*, 2001b). Therefore, we chose to estimate the average number of DNA regions per cell length from measured cell cycle parameters (chapters 2 and 3; Huls *et al.*, 1999).

The cell cycle parameters C (the time of one round of replication), and D (the time from termination of replication to cell division) were obtained from a run-out experiment with rifampicin performed by Huls and co-workers (Huls *et al.*, 1999). Rifampicin prevents initiation of replication, but allows current replication cycles to finish. Consequently, the amount of DNA per cell continually increases due to replication, but the increase diminishes until all replication forks terminate. Eventually, the amount of DNA per cell reaches a plateau and all cells contain whole numbers of genome equivalents. The time from rifampicin addition to reaching this plateau is used as an estimate for the C period. From C and T_d, the D period can be calculated (Bipatnath *et al.*, 1998; Huls *et al.*, 1999). T_d can readily be obtained from cell density measurements. Note that in order to estimate the average number of DNA regions per cell length, we have to make two further assumptions. First, we assume that replication proceeds linearly in time. Secondly, we assume that cell length increases exponentially in time (equation (1) in chapter 4; e.g. Koppes, 1980; Trueba and Koppes, 1998).

We thus estimated the copy-number of specific DNA-regions per cell length. However, there is a certain amount of variation in measurements of cell cycle parameters, which we did not take into account. Moreover, cell-cycle parameters vary between individual cells, even in a steady-

state population. Consequently, the relationship between cell length and the timing of replication is probably not as strong as we had assumed. Nevertheless, we could assume that replication initiated before cell division, because it had been shown that the cell cycle does not contain a B period in case of *E. coli* strain K-12, grown on minimal glucose medium at 28 °C (Huls *et al.*, 1999).

5.2.10 Replication takes place in the middle of prospective-daughter cells

We assumed that replication takes place in the cell centre, or at the ¼ and ¾ positions in the case of two replicating chromosomes. The latter positions represent the cell centres of the two prospective-daughter cells. This means that, at initiation of DNA replication, the distance of an origin of replication from the centre of the corresponding prospective-daughter cell is 0. This enabled us to determine the y-intercept of the equation for distance of a focus from the cell centre (equation (2) in chapter 4).

There is strong experimental evidence that replication takes place in the centre of (prospective-daughter) cells in bacteria. In a slow growing strain of *E. coli* (T_d = 210 min.) it was shown by electron-microscopic auto-radiography that newly synthesised DNA is localised in the cell centre (Koppes *et al.*, 1999), and in *B. subtilis* it was shown by GFP-tagging that DNA polymerase is localised in the cell centre or in the centres of prospective-daughter cells (Lemon and Grossman, 1998; 2000). A DNA polymerase that is localised in a fixed position within the cell implies that the replisome is stationary with respect to the DNA it is replicating. Support for a stationary replication-machine in eukaryotic nuclei is also increasing (see Cook, 1989; 1999 and references therein). Nevertheless, the resolution of auto-radiography and light microscopy are limited, leaving room for movement of the replication machinery within a region with a diameter of at least 250 nm (Koppes *et al.*, 1999). In addition, in the data of Lemon and Grossman (Lemon and Grossman,

1998; 2000) a certain amount of variation can be observed in the position of GFP-tagged DNA-polymerase. Possibly, part of the variation that we observed in our FISH data is due to movement of the replisome as such.

For our statistical analysis of FISH data (chapters 2 and 3), we did not need to assume that replication takes place in the centre of prospective-daughter cells, because it is simply a consequence of linear regression: any two lines, mirrored in the x-axis, will meet somewhere on the x-axis, unless their slope is exactly 1. Remarkably, in the case of the regression lines through the data-points of *oriC*, *ftsQAZ*, and *minB*, the length at which the regression lines intercepted the x-axis, corresponded roughly to the length at which these regions would be replicated as calculated from C, D, and T_d (chapter 4).

5.3 References

Bipatnath, M., P. P. Dennis, and H. Bremer. 1998. Initiation and velocity of chromosome replication in *Escherichia coli* B/r and K-12. *J Bacteriol* 180: 265-273.

Carrio, M. M., J. L. Corchero, and A. Villaverde. 1998. Dynamics of in vivo protein aggregation: building inclusion bodies in recombinant bacteria. *FEMS Microbiol Lett* 169: 9-15.

...1999. Proteolytic digestion of bacterial inclusion body proteins during dynamic transition between soluble and insoluble forms. *Biochim Biophys Acta* 1434: 170-176.

Carrio, M. M., R. Cubarsi, and A. Villaverde. 2000. Fine architecture of bacterial inclusion bodies. *FEBS Lett* 471: 7-11.

Cook, P. R. 1989. The nucleoskeleton and the topology of transcription. *Eur J Biochem* 185: 487-501.

...1999. The organization of replication and transcription. *Science* 284: 1790-1795.

Glaser, P., M. E. Sharpe, B. Raether, M. Perego, K. Ohlsen, and J. Errington. 1997. Dynamic, mitotic-like behaviour of a bacterial protein required for accurate chromosome partitioning. *Genes Devel* 11: 1160-1168.

Gordon, G. S., D. Sitnikov, C. D. Webb, A. Teleman, A. Straight, R. Losick, A. W. Murray, and A. Wright. 1997. Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. *Cell* 90: 1113-1121.

Hiraga, S. 2000. Dynamic localization of bacterial and plasmid chromosomes. *Annu Rev Genet* 34: 21-59.

Huls, P. G., N. O. Vischer, and C. L. Woldringh. 1999. Delayed nucleoid segregation in *Escherichia coli*. *Mol Microbiol* 33: 959-970.

Koppes, L. J. 1980. *Duplication of Escherichia coli: Variation of division cycle parameters as*

studied by electron microscopy. University of Amsterdam, Amsterdam.

Koppes, L. J., C. L. Woldringh, and N. Nanninga. 1999. *Escherichia coli* contains a DNA replication compartment in the cell center. *Biochimie* 81: 803-810.

Lemon, K. P., and A. D. Grossman. 1998. Localization of bacterial DNA polymerase: evidence for a factory model of replication. *Science* 282: 1516-1519.

...2000. Movement of replicating DNA through a stationary replisome. *Mol Cell* 6: 1321-1330.

Marshall, W. F., A. Straight, J. F. Marko, J. Swedlow, A. Dernburg, A. Belmont, A. W. Murray, D. A. Agard, and J. W. Sedat. 1997. Interphase chromosomes undergo constrained diffusional motion in living cells. *Curr Biol* 7: 930-939.

Niki, H., and S. Hiraga. 1998. Polar localization of the replication origin and terminus in *Escherichia coli* nucleoids during chromosome partitioning. *Genes Dev* 12: 1036-1045.

Niki, H., Y. Yamaichi, and S. Hiraga. 2000. Dynamic organization of chromosomal DNA in *Escherichia coli*. *Genes Dev* 14: 212-223.

Noordmans, H. J., and A. W. M. Smeulders. 1998. Detection and characterization of isolated and overlapping spots. *Computer vision and image understanding* 70: 23-35.

Roos, M., R. Lingeman, C. L. Woldringh, and N. Nanninga. 2001a. Experiments on movement of DNA regions in Escherichia colievaluated by computer simulation. *Biochimie* 83: 67-74.

Roos, M., A. B. M. van Geel, M. E. G. Aarsman, J. T. M. Veuskens, C. L. Woldringh, and N. Nanninga. 2001b. The replicated *ftsQAZ* and *minB* chromosomal regions of *Escherichia coli* segregate on average in line with nucleoid movement. *Mol Microbiol* 39: 633-640.

Sharpe, M. E., and J. Errington. 1998. A fixed distance for separation of newly replicated copies of *oriC* in *Bacillus subtilis*: implications for co-ordination of chromosome segregation and cell division. *Mol Microbiol* 28: 981-990.

Siegele, D. A., and J. C. Hu. 1997. Gene expression from plasmids containing the *araBAD* promoter at subsaturating inducer concentrations represents mixed populations. *Proc Natl Acad Sci U S A* 94: 8168-8172.

Sokal, R. R., and F. J. Rohlf. 1995. *Biometry : the principles and practice of statistics in biological research*. W.H. Freeman, New York.

Swartz, J. R. 1996. *Escherichia coli* recombinant DNA technology. Pages 1693-1711 in Neidhardt, F. C. and Curtiss, R., eds. *Escherichia coli and Salmonella: cellular and molecular biology*. ASM Press, Washington, D.C.

Teleman, A. A., P. L. Graumann, D. C. H. Lin, A. D. Grossman, and R. Losick. 1998. Chromosome arrangement within a bacterium. *Curr Biol* 8: 1102-1109.

Trask, B. J., S. Allen, H. Massa, A. Fertitta, R. Sachs, G. van den Engh, and M. Wu. 1993. Studies of metaphase and interphase chromosomes using fluorescence in situ hybridization. *Cold Spring Harb Symp Quant Biol* 58: 767-775.

Trueba, F. J., and L. J. H. Koppes. 1998. Exponential growth of *Escherichia coli* B/r during its division cycle is demonstrated by the size distribution in liquid culture [In Process Citation]. *Arch Microbiol* 169: 491-496.

Valkenburg, J. A., and C. L. Woldringh. 1984. Phase separation between nucleoid and cytoplasm in *Escherichia coli* as defined by immersive refractometry. *J Bacteriol* 160: 1151-1157.

Webb, C. D., P. L. Graumann, J. A. Kahana, A. A. Teleman, P. A. Silver, and R. Losick. 1998. Use of time-lapse microscopy to visualize rapid movement of the replication origin region of the chromosome during the cell cycle in *Bacillus subtilis*. *Mol Microbiol* 28: 883-892.

Webb, C. D., A. Teleman, S. Gordon, A. Straight, A. Belmont, D. C. Lin, A. D. Grossman, A. Wright, and R. Losick. 1997. Bipolar localization of the replication origin regions of chromosomes in vegetative and sporulating cells of *B. subtilis*. *Cell* 88: 667-674.

Weitao, T., S. Dasgupta, and K. Nordstrom. 2000. Role of the *mukB* gene in chromosome and plasmid partition in *Escherichia coli*. *Mol Microbiol* 38: 392-400.

Woldringh, C. L., A. Zaritsky, and N. B. Grover. 1994. Nucleoid partitioning and the division plane in *Escherichia coli*. *J Bacteriol* 176: 6030-6038.

Chapter 6

DNA organisation and DNA segregation in *Escherichia coli*

Both the gradual-segregation model and the 'sticky-jump' model (chapter 4) intend to give a description of observed FISH data. However, they do not tell us what the mechanism of DNA segregation is. In this chapter I have re-examined some of the factors involved in organising DNA (chapter 1), but now more explicitly with respect to their possible role in DNA segregation (section 6.1). Subsequently, I give my view on how and whether knowledge of DNA segregation in eukaryotes could help us to understand DNA segregation in E. coli (section 6.2).

6.1 Factors involved in DNA segregation

6.1.1 DNA

The double-stranded circular DNA-molecule itself has intrinsic structural properties. For instance, it forms supercoils. Supercoiling promotes compaction of DNA to some degree (Sottas *et al.* 1999), and during replication it promotes compaction of daughter DNA's upon themselves, rather than upon each other or upon unreplicated DNA (Alexandrov *et*

al. 1999). Thus, DNA segregation could be the result, at least in part, of DNA compaction by supercoiling after replication.

Also, the distribution of genes around the chromosome of *E. coli* is non-random (Blattner *et al.* 1997; Pedersen *et al.* 2000; Ussery *et al.* 2000). This suggests that transcriptional activity is non-randomly distributed over the chromosome. In addition to genes, structural properties of DNA, such as DNA curvature and DNA bending are also distributed non-randomly over the chromosome (see section 1.1.4; Pedersen *et al.* 2000). For instance, in *E. coli* the intrinsic structure of a region near *oriC* is clearly different from that of a large region of DNA encompassing the terminus. How this may affect DNA organisation or DNA segregation is not yet clear. Nevertheless, the replisome and the segregation mechanism have to deal with DNA with variable structural properties during replication.

6.1.2 Proteins

6.1.1.1 'Mitotic-like' proteins

The most direct involvement of proteins in DNA segregation is predicted by the 'mitotic-like'-segregation model. This model proposes that specific proteins actively pull replicated DNA apart towards opposite cell poles, in a process reminiscent of chromatid segregation during mitosis in eukaryotes. However, to the best of my knowledge these proteins have not yet been found in bacteria (see also section 6.2.2).

6.1.1.2 DNA polymerase

Is the enzyme that replicates DNA also involved in DNA segregation? If replisomes are specifically localised in the centres of (prospective daughter) cells, and if they are stationary with respect to the DNA they are replicating (section 1.1.5.3), then the replisome positions replicating DNA within the cell during the cell cycle. More specifically, DNA polymerase may be an important factor in the movement of unreplicated

DNA towards the replisome, and, possibly, in giving an initial direction to replicated DNA away from the replisome. It probably does not exert a direct pushing force on DNA over long distances, because when replication is inhibited, but not transcription, DNA is still being moved apart (Woldringh *et al.* 1994).

6.1.1.3 Topo-isomerases

Topo-isomerases are involved in maintaining DNA in a negatively supercoiled state during replication, and in decatenating replicated DNA-molecules during termination of replication (see for instance Kornberg and Baker 1992b). Some topo-isomerases are part of the replication machine (Kornberg and Baker 1992a, table 15-4), and replication cannot proceed without any topo-isomerases. Thus, topo-isomerases are likely to have an effect on DNA segregation via their effect on DNA supercoiling (see section 1.1.5.7).

6.1.1.4 Nucleoid-associated proteins

Through dynamic binding to DNA, nucleoid-associated proteins may influence DNA segregation. This influence could be direct, e.g. by actively pulling or pushing DNA, or indirect, e.g. by compacting DNA after replication. MukB, the SMC equivalent of *E. coli*, co-localises with the nucleoid (fig. 1.6; Den Blaauwen submitted), and may thus serve as an example. It has been suggested that MukB actively moves DNA via a presumed motor function (Lockhart and Kendrick-Jones 1998). However, more recent findings suggest that MukB is involved in DNA segregation because of a role in compacting DNA after replication (Sawitzke and Austin 2000; Weitao *et al.* 2000a; Weitao *et al.* 2000b, or see section 1.1.5.4 or, for a review, see Dasgupta *et al.* 2000; Graumann 2001; Holmes and Cozzarelli 2000).

6.1.1.5 RNA polymerase

Are RNA polymerases involved in DNA segregation? Just as DNA polymerase, RNA polymerase is probably stationary with respect to the

DNA it is transcribing (see section 1.1.5.2). RNA polymerase is a very strong motor protein (Yin *et al.* 1995), 5-6 times stronger than kinesin (force of procession between 25 and 30 pN; Wang *et al.* 1998). However, how could this strength be used to segregate DNA?

In bacteria, whilst RNA polymerase is transcribing a gene, ribosomes already translate the transcript, and nascent translation-products may start moving to their destination. This chain of gene
RNA polymerase transcript ribosomes translation-products may anchor DNA to, for instance, the membrane (Norris 1995; Woldringh *et al.* 1995). Through these anchors, RNA polymerase may be able to move DNA within the cell (Woldringh personal communication). In fact, because cytoplasm is viscous, anchoring to the cytoplasm may have a similar effect. Considering that many RNA polymerases are active in a cell, transcription may be an important driving force behind DNA segregation. However, for segregation DNA has to move in specific directions: replicated DNA away from the replisome, unreplicated DNA towards the replisome. It is not yet clear how directional movement of DNA is accomplished by transcription combined with anchoring of DNA to the membrane or the cytoplasm.

6.1.1.6 Transcription factors and other DNA-binding proteins

Transcription factors and DNA-binding proteins that I did not mention before may transiently alter the structure of DNA regions (e.g. AraC in section 1.1.5.1). They may thus be responsible for a certain amount of DNA movement, especially because they are so plentiful. However, we can probably assume that the net contribution to DNA segregation is negligible, because of a lack in directionality. Nevertheless, it is possible that these movements explain part of the observed variation in FISH data.

6.1.2 Membranes

Membranes may not have been given the proper amount of attention in

this thesis. As any other cellular constituent they influence DNA organisation, at least indirectly. For instance, DNA partitioning is related to cell growth, hence, related to membrane growth. In fact, in combination with co-transcriptional-translation-and-insertion of membrane proteins, membrane-growth may even play an essential part in DNA segregation (see previous section). Another example of an important role of the membrane might be the supposed transient sequestering of DNA near the origin, following initiation of replication. It is assumed that sequestering prevents premature re-initiation of replication (Ogden *et al.* 1988, for a review see Crooke 1995), but it would also link a specific DNA region to a growing membrane. However, to the best of my knowledge, sequestering of *oriC* has not been shown by microscopy.

6.1.3 Towards a model of the mechanism underlying DNA segregation in *E. coli*

In the foregoing, a number of factors (DNA, various proteins, membranes) were presented that might contribute to DNA segregation. If we are to make a model of the mechanism of DNA segregation in *E. coli*, we may assume that it should be based on a combination of these factors. So, what is the relative importance of each of these factors and how do they interact to give rise to DNA segregation?

6.1.3.1 The relative importance of specific factors

Even if the mechanism behind a process is unknown, it is possible to determine the relative importance of a specific factor involved. For instance, using a biochemical method (Groen *et al.* 1982; Westerhoff *et al.* 1998), the 'inherent control' of DNA gyrase on DNA supercoiling was measured to be only 0.2 % in *E. coli* (see section 1.1.5.7 and Jensen *et al.* 1999 for details); homeostatic regulation of DNA supercoiling appears to dominate (Jensen *et al.* 1999). Unfortunately, as far as I know, not many factors involved in DNA segregation have been analysed using these kinds of methods.

6.1.3.2 Interactions between factors

To find out which interactions determine DNA segregation, we may define a putative model and predict what kind of data the model would produce. However, if some of the factors in our model interact non-linearly, this becomes a non-trivial task. In contrast, if a model is based on a few linear equations, it is relatively easy to simulate data and compare the predicted data with, for instance, FISH data (chapter 4 gives examples). However, most biological systems are based on non-linear interactions (Westerhoff *et al.* 1998); we have to assume DNA segregation is no exception.

In some cases, the consequences of a 'non-linear' model can be predicted. For instance, by applying a set of biophysical calculations, Odijk showed that DNA supercoiling and depletion forces in *E. coli* can give rise to the nucleoid as a distinct structure (Odijk 1998). Computer simulation can also be used to predict the consequences of a non-linear model. For instance, Brownian-dynamics simulations have been used to predict both the structure and the dynamics of small DNA molecules, such as plasmids (e.g. Klenin, K. *et al.* 1998; Klenin, K. V. and Langowski 2001; Langowski *et al.* 1999). To the best of my knowledge, these methods have not yet been used specifically to study DNA segregation in bacteria. (A further discussion on the use of computer simulation methods is beyond the scope of this thesis.)

6.1.4 Conclusion

The purpose of this section was to review, briefly, some of the factors that may contribute to DNA segregation in *E. coli*. We may assume that none of these factors is much more important than the others; a combination of factors is probably responsible for DNA segregation. Ideally, we would define a model that incorporates all known factors, and then evaluate it by comparing computer-simulated data with experimental data (cf. chapter 4). However, the foregoing makes clear

that both defining this type of model, as well as predicting what kind of data it produces is difficult. Nevertheless, without such an evaluation, models remain non-committal. Some aspects of DNA segregation may be examined using biophysical or biochemical methods, but I think that, in addition, computer-simulation methods should be used to examine possible mechanisms underlying DNA segregation in *E. coli* (see section 5.3).

6.2 DNA segregation in *E. coli* versus DNA segregation in eukaryotes

6.2.1 Introduction

Several aspects of prokaryotes and eukaryotes may be compared. For instance, bipolarity is a common concept for cell division and DNA segregation in both classes (see Nanninga 2001). At the molecular level, many bacterial components that are presumed to be involved in organising DNA are supposed to be similar to eukaryotic counterparts. For example, some nucleoid-associated proteins, such as HU, are called 'histone-like' (e.g. Jaffe *et al.* 1997; Rouviere-Yaniv *et al.* 1979; Shellman and Pettijohn 1991; Yasuzawa *et al.* 1992; see also section 1.1.5.5), others, such as MukB, are called 'SMC-like' (e.g. Graumann 2001; Melby *et al.* 1998), nucleotide sequences that contain *oriC* or are close to *oriC* are called 'centromere-like' (e.g. Gordon and Wright 2000; Marston and Errington 1999; Sharpe and Errington 1999), the terminus region has been called 'telomere-like' (e.g. Gordon and Wright 2000), and the whole nucleoid has even been called 'nucleosome-like' (e.g. Azam *et al.* 2000). Finally, the process of DNA segregation in bacteria has been called 'mitotic-like' (e.g. Begg and Donachie 1991; Lin *et al.* 1997; Moller-Jensen *et al.* 2000; Niki and Hiraga 1998; Sharpe and Errington 1999; Webb *et al.* 1997; Wheeler and Shapiro 1997). Unfortunately, the use of the additive '-like' does not seem to be bound by any definition. For instance, in the case of proteins and sequences, it is unclear if '-like' implies sequence homology, structural similarity, functional similarity, or a combination of these. I suggest restricting the use of '-like' to those cases where there is

distinct sequence-homology. Perhaps we should make an exception for classes of proteins that have a structural definition, such as the class of SMC proteins (section 1.1.5.4).

6.2.2 Size

I will base my comparison of DNA segregation in prokaryotes and eukaryotes on size, and I will do so briefly, as an in-depth treatment is beyond the scope of this thesis (for a more general comparison between DNA organisation in prokaryotes and eukaryotes see for instance Nanninga 2001; Woldringh and van Driel 1999). On the basis of size, DNA segregation in bacteria and mitosis in eukaryotes are of a different order (fig. 6.1). Although the function of both processes is to partition replicated DNA over two prospective daughter-cells, bacteria move their DNA over less than one micrometer, whilst eukaryotes may move their DNA over more than 5 micrometers. Consequently, the cellular structures and forces required to move the DNA are different. Whereas diffusion may be enough to accomplish DNA segregation in bacteria (chapter 4 and next section), in eukaryotes DNA segregation is achieved through a specialised system consisting of large regions of heterochromatin (centromeres*), large protein complexes attached to centromeres (kinetochores), specific motor proteins and microtubules. The microtubules attach to the centromere via the kinetochore, and with the aid of motor-proteins the chromatids are segregated towards opposite cell poles. The centromere may be several thousands base pairs long, and the size of the centromere/kinetochore structure is about half a micrometer (the structure in figure 6.1 is more or less on scale). This is enormous compared to the size of a bacterium and bacterial DNA.

* I use centromere here to denote the region of heterochromatin that gives rise to the visible constriction in metaphase chromosomes, which is involved in chromatid segregation. The underlying nucleotide sequences differ from species to species or even between chromosomes of one species. They contain a variety of heterochromatic sequences (see for instance Vig. 1994. *Mutat Res* 309: 1-10.).

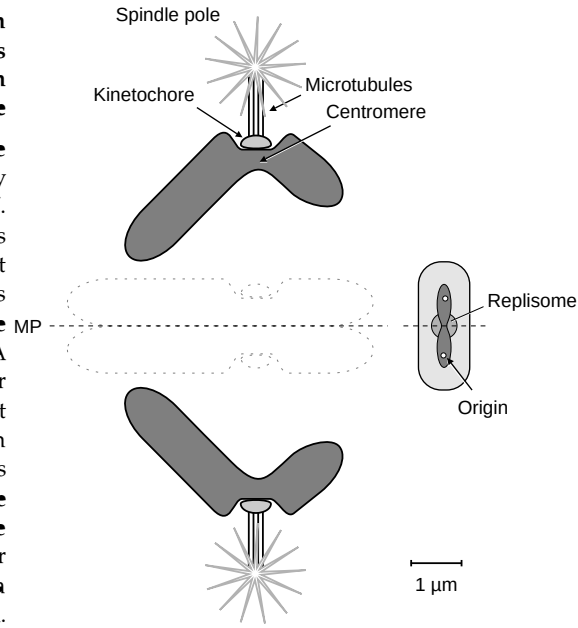


Figure 6.1 - Comparison between DNA segregation in eukaryotes (mitosis) and DNA segregation in *E. coli*, based on size

The sizes are more or less on scale. The eukaryotic chromosome is loosely based on human chromosome 7. During the anaphase stage of mitosis the two chromatids are pulled apart over a distance of about 5 µm towards spindle poles on opposite sides of the cell. The chromatids are copies of DNA molecules that were replicated earlier during *S*-phase and subsequently kept together until mitosis (see text). In *E. coli*, DNA segregates whilst DNA is replicating. Thus, there is no stage during the cell cycle of *E. coli* where chromosomes are held together over their full lengths. Segregation is over a distance less than 1 µm.

Hence, based on size, we cannot expect to find any 'centromere-like' sequences or structures in bacteria. Indeed, typical components of the eukaryotic mitotic-system, such as centromeric DNA-sequences, kinetochores, or microtubules, have not been found in eubacteria.

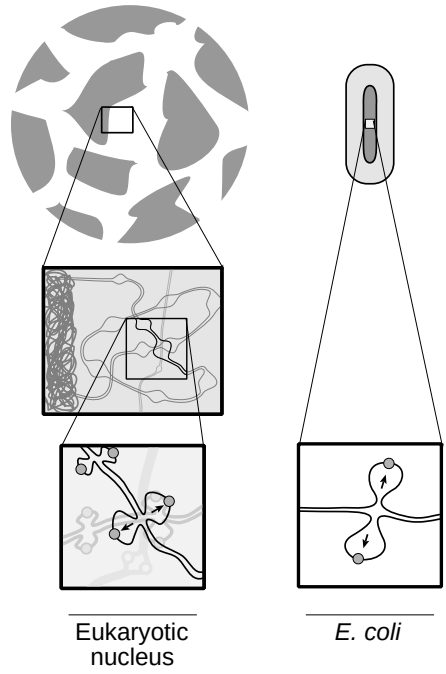
6.2.3 Initial DNA segregation during replication

On the basis of size, a more appropriate comparison may be that between DNA-segregation in bacteria and initial DNA-segregation in eukaryotes during S-phase (fig. 6.2). In both classes of species a complex containing DNA polymerase replicates DNA. Replication initiates from an origin of replication and a replication bubble is created (one per chromosome in bacteria, many per chromosome in eukaryotes). Replicated DNA may initially segregate in opposite directions using a similar mechanism in both cases. In slow-growing bacteria segregation is towards opposite cell

Figure 6.2 -

Comparison of initial segregation of replicated DNA in eukaryotes and *E. coli*

Left: initial segregation of replicated DNA during *S*-phase in eukaryotes; right: segregation of replicated DNA during the *C*-period in *E. coli*. For simplicity, the cartoon depicts a B/r strain with one replisome replicating in the cell centre (cf. Koppes *et al.*, 1999). In both species a complex containing DNA polymerase replicates DNA, thereby forming a replication bubble. Eukaryotes have multiple bubbles per chromosome, because they have multiple origins per chromosome. *E. coli* has one bubble per replicating chromosome, because it has only one origin per chromosome.



poles; in fast-growing bacteria, undergoing multi-fork replication, sister copies do not move towards opposite cell poles, but they may still segregate in opposite directions.

If the above comparisons make sense, the idea that diffusion of DNA is a driving force behind DNA segregation in bacteria may also be applicable to initial DNA segregation in eukaryotes during S-phase. That is, in obtaining a defined distance between sister chromatids. Similarly, the idea that compaction of DNA after replication segregates DNA in bacteria (see section 6.1.2.4; Dasgupta *et al.* 2000; Graumann 2001; Holmes and Cozzarelli 2000 and references therein), may be applicable as well.

Nevertheless, because of important differences at later stages of DNA segregation we should be careful with these comparisons. For instance,

in bacteria sister copies have fully segregated at termination of replication, whilst in eukaryotes, sister copies are cohered together after replication until mitosis (Hirano 2000 and references therein). Cohesion may take place almost immediately after replication. Recent evidence suggests that also in eukaryotes the position of replicating DNA is related to compaction (see section 6.1.2.4 and above). Electron-microscopic data from nuclei from V97 Chinese hamster cells indicated that nascent DNA is positioned in less-dense areas of chromatin, whereas DNA that is not replicating is confined to more-condensed areas of chromatin (Jaunin *et al.* 2000). However, the type of compaction is different in bacteria and eukaryotes. In bacteria, compaction is accomplished mainly through *plectonemic* supercoiling in combination with depletion forces and, presumably, SMC-proteins such as MukB (e.g Holmes and Cozzarelli 2000). In eukaryotes, a much higher level of compaction is achieved through *solenoidal* supercoiling. Histones define this type of structure, although SMC proteins play an important role as part of a protein complex, condensin, that appears to play a central role in chromosome compaction (Hirano 2000 and references therein). Some histones are released during replication, but the majority of histones remain attached (although some biochemical properties of these histones change).

In summary, differences in size, but also differences in some basic properties of DNA-segregation during replication, suggest that prokaryotes and eukaryotes require fundamentally different mechanisms to segregate their DNA.

6.3 Concluding remarks

In chapter 4 we show that two current models, designed to explain FISH data, do not explain all of the variation in our data. From re-examining the assumptions on which these models were based (chapter 5), we concluded that this is so because a number of these assumptions are not met in practice. For instance, the models assume that the position of

DNA is dependent solely on cell length ('a 1-D model'), and that variation between cells, regarding parameters that might influence DNA segregation, is negligible. We might be able to explain more of the variation in our FISH data, if we could define a 3-D model that would allow us to incorporate more variation in these parameters. However, deriving such a model analytically is not a trivial task, and there is little information available about 3-D organisation of DNA in *E. coli*.

Another way to explain DNA segregation is by a model based on factors that organise DNA in the cell (previous sections). An important factor in determining DNA organisation in *E. coli*, put forward in this thesis, is Brownian motion (diffusion) of DNA regions. Our simulations and preliminary data have shown that Brownian motion (i.e. diffusion) within a confined region (cf. Marshall *et al.* 1997) may contribute greatly to the total of observed variation (Roos *et al.* 2001). However, it is not yet clear what its precise function is in positioning DNA during the cell cycle. This is mainly because we were unable to quantify the position and size of the putative confined-regions properly (section 5.2.3).

As a next step, I suggest to test the possibility that diffusion, replication, and cell growth determine how replicated DNA is segregated in *E. coli* (fig. 6.3). I will refer to this model as the 'diffusional-drift' model. Diffusion is the motive force in this model, replication gives initial direction to replicated DNA, and cell growth produces the necessary space for the increasing amount of DNA. Whilst template DNA moves towards the replisome, replicated DNA 'drifts' towards the cell poles as the cell is growing in the length axis of the cell. In addition to these three factors, repulsion between DNA molecules may help in separating the stretches of replicated DNA and the template DNA.

Because it is not immediately clear what mathematical function would represent this model, the method of adding variation to mathematically determined positions, such as described in chapter 4 of this thesis, cannot be applied. An alternative method is to use a Brownian-dynamics procedure. This procedure has already been used successfully to model

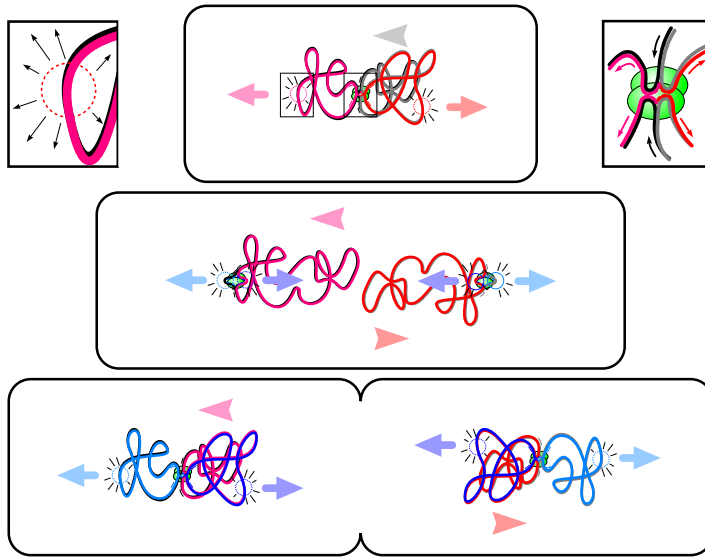


Figure 6.3 - Simplified model of DNA segregation in E. coli based on diffusion and cell growth. The dashed circles indicate the position of an oriC region

Three stages of the cell cycle are depicted. All DNA (the curves) displays Brownian motion (shown for *oriC* by black arrows, see the enlarged image on the left). On average, each *oriC* region moves in the direction of a cell pole (old or new), as the cell is growing (large coloured arrows). In this way segregation of replicated DNA is achieved. Parental DNA moves towards the replisome (large arrowheads). The replisome gives initial direction to daughter strands and pulls parental DNA (arrows in enlarged image on the right). Note that DNA is not drawn on scale and is depicted without twist.

the structure of super helical DNA as a function of salt concentration (in combination with Monte Carlo simulations; Langowski *et al.* 1999), and to model a diffusion-controlled reaction between two DNA regions 470 bp apart on a 2.5 kbp-long supercoiled plasmid (Klenin and Langowski 2001). In these studies, the DNA molecule is represented as a worm-like polyelectrolyte chain. A limited set of parameters is used to define properties, such as bending and twisting elasticity, hydrodynamic diameter, and electrostatic interactions between parts of the molecule (Klenin *et al.* 1998). Shape and dynamics of DNA are further determined by the surrounding fluid, which was represented as a continuous viscous fluid. The simulations in these studies are of DNA up to a few kbp in

length. The whole *E. coli* chromosome is $4.6 * 10^6$ bp long. So, if we are to simulate DNA organisation as a function of replication in *E. coli* within a realistic time period, then we probably need to make some simplifications. For instance, we could use the data of Pedersen to define large-scale structural properties of the DNA molecule (Pedersen *et al.* 2000; section 1.1.4), instead of using the properties of each individual nucleotide.

The process of segregation is a result of the interaction of many factors. One of these is certainly diffusion of DNA regions, as outlined in chapter 4. Presumably, it will not be possible to predict, beforehand, how such a model would result in segregation of replicated DNA, because many (non-linear) interactions are involved. Nevertheless, using sophisticated computer simulation we may be able to gain insight into the process of DNA segregation. For instance, we may ask if the combination of factors that determines DNA segregation in slow growing cells can also explain the more complex DNA movements in fast growing cells. In addition, we may learn about the relationship between DNA segregation, DNA replication and cell division. For instance, we may ask if there is a common principle that positions both replication and cell division at the cell centre, be it at different times. In any case, I think it is important to evaluate such models by comparing simulated data to experimentally obtained data as we did in chapter 4. I see this approach as an important methodological aspect of this thesis. It led us to propose the diffusional-drift model.

6.4 References

Alexandrov, A. I., N. R. Cozzarelli, V. F. Holmes, A. B. Khodursky, B. J. Peter, L. Postow, V. Rybenkov, and A. V. Vologodskii. 1999. Mechanisms of separation of the complementary strands of DNA during replication. *Genetica* 106: 131-140.

Azam, T. A., S. Hiraga, and A. Ishihama. 2000. Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes Cells* 5: 613-626.

Begg, K. J., and W. D. Donachie. 1991. Experiments on chromosome separation and positioning in *Escherichia coli*. *New Biol* 3: 475-486.

Blattner, F. R., G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A.

Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* 277: 1453-1474.

Crooke, E. 1995. Regulation of chromosomal replication in *E. coli*: sequestration and beyond. *Cell* 82: 877-880.

Dasgupta, S., S. Maisnier-Patin, and K. Nordstrom. 2000. New genes with old modus operandi. The connection between supercoiling and partitioning of DNA in *Escherichia coli*. *EMBO Rep* 1: 323-327.

Den Blaauwen, T. submitted. *Mol Microbiol*.

Gordon, G. S., and A. Wright. 2000. DNA segregation in bacteria. *Annu Rev Microbiol* 54: 681-708.

Graumann, P. L. 2001. SMC proteins in bacteria: Condensation motors for chromosome segregation? *Biochimie* 83: 53-59.

Groen, A. K., R. J. Wanders, H. V. Westerhoff, R. van der Meer, and J. M. Tager. 1982. Quantification of the contribution of various steps to the control of mitochondrial respiration. *J Biol Chem* 257: 2754-2757.

Hirano, T. 2000. Chromosome cohesion, condensation, and separation. *Annu Rev Biochem* 69: 115-144.

Holmes, V. F., and N. R. Cozzarelli. 2000. Closing the ring: links between SMC proteins and chromosome partitioning, condensation, and supercoiling. *Proc Natl Acad Sci U S A* 97: 1322-1324.

Huls, P. G., N. O. Vischer, and C. L. Woldringh. 1999. Delayed nucleoid segregation in *Escherichia coli*. *Mol Microbiol* 33: 959-970.

Jaffe, A., D. Vinella, and R. D'Ari. 1997. The *Escherichia coli* histone-like protein HU affects DNA initiation, chromosome partitioning via MukB, and cell division via MinCDE. *J Bacteriol* 179: 3494-3499.

Jaunin, F., A. E. Visser, D. Cmarko, J. A. Aten, and S. Fakan. 2000. Fine structural in situ analysis of nascent DNA movement following DNA replication. *Exp Cell Res* 260: 313-323.

Jensen, P. R., C. C. Van Der Weijden, L. B. Jensen, H. V. Westerhoff, and J. L. Snoep. 1999. Extensive regulation compromises the extent to which DNA gyrase controls DNA supercoiling and growth rate of *Escherichia coli*. *Eur J Biochem* 266: 865-877.

Kido, M., K. Yamanaka, T. Mitani, H. Niki, T. Ogura, and S. Hiraga. 1996. RNase E polypeptides lacking a carboxyl-terminal half suppress a *mukB* mutation in *Escherichia coli*. *J Bacteriol* 178: 3917-3925.

Klenin, K., H. Merlitz, and J. Langowski. 1998. A Brownian dynamics program for the simulation of linear and circular DNA and other wormlike chain polyelectrolytes. *Biophys J* 74: 780-788.

Klenin, K. V., and J. Langowski. 2001. Diffusion-controlled intrachain reactions of supercoiled DNA: Brownian Dynamics simulations. *Biophys J* 80: 69-74.

Kornberg, A., and T. A. Baker. 1992a. Replication mechanisms and operations. Pages 471-510. *DNA replication*. W.H. Freeman, New York.

...1992b. Topoisomerases. Pages 379-402. *DNA replication*. W.H. Freeman, New York.

Langowski, J., M. Hammermann, K. Klenin, R. May, and K. Toth. 1999. Superhelical DNA studied by solution scattering and computer models. *Genetica* 106: 49-55.

Lin, D. C., P. A. Levin, and A. D. Grossman. 1997. Bipolar localization of a chromosome partition protein in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* 94: 4721-4726.

Lockhart, A., and J. Kendrick-Jones. 1998. Nucleotide-dependent interaction of the N-terminal domain of MukB with microtubules. *J Struct Biol* 124: 303-310.

Marshall, W. F., A. Straight, J. F. Marko, J. Swedlow, A. Dernburg, A. Belmont, A. W. Murray, D. A. Agard, and J. W. Sedat. 1997. Interphase chromosomes undergo constrained diffusional motion in living cells. *Curr Biol* 7: 930-939.

Marston, A. L., and J. Errington. 1999. Dynamic movement of the ParA-like Soj protein of *B. subtilis* and its dual role in nucleoid organization and developmental regulation. *Mol Cell* 4: 673-682.

Melby, T. E., C. N. Ciampaglio, G. Briscoe, and H. P. Erickson. 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.

Moller-Jensen, J., R. B. Jensen, and K. Gerdes. 2000. Plasmid and chromosome segregation in prokaryotes. *Trends Microbiol* 8: 313-320.

Nanninga, N. 2001. Cytokinesis in pro- and eukaryotes: common principles and different solutions. *Microbiol Mol Biol Rev*: in press.

Niki, H., and S. Hiraga. 1998. Polar localization of the replication origin and terminus in *Escherichia coli* nucleoids during chromosome partitioning. *Genes Dev.* 12: 1036-1045.

Norris, V. 1995. Hypothesis: chromosome separation in *Escherichia coli* involves autocatalytic gene expression, transertion and membrane-domain formation. *Mol Microbiol* 16: 1051-1057.

Odijk, T. 1998. Osmotic compaction of supercoiled DNA into a bacterial nucleoid. *Biophys Chem* 73: 23-29.

Ogden, G. B., M. J. Pratt, and M. Schaechter. 1988. The replicative origin of the *E. coli* chromosome binds to cell membranes only when hemimethylated. *Cell* 54: 127-135.

Pedersen, A. G., L. J. Jensen, S. Brunak, H. H. Staerfeldt, and D. W. Ussery. 2000. A DNA structural atlas for *Escherichia coli*. *J Mol Biol* 299: 907-930.

Roos, M., R. Lingeman, C. L. Woldringh, and N. Nanninga. 2001. Experiments on movement of DNA regions in Escherichia colievaluated by computer simulation. *Biochimie* 83: 67-74.

Rouviere-Yaniv, J., M. Yaniv, and J. E. Germond. 1979. *E. coli* DNA binding protein HU forms nucleosomelike structure with circular double-stranded DNA. *Cell* 17: 265-274.

Sawitzke, J. A., and S. Austin. 2000. Suppression of chromosome segregation defects of *Escherichia coli muk* mutants by mutations in topoisomerase I. *Proc Natl Acad Sci U S A* 97: 1671-1676.

Sharpe, M. E., and J. Errington. 1999. Upheaval in the bacterial nucleoid. An active chromosome segregation mechanism. *Trends Genet.* 15: 70-74.

Shellman, V. L., and D. E. Pettijohn. 1991. Introduction of proteins into living bacterial cells: distribution of labeled HU protein in *Escherichia coli*. *J Bacteriol* 173: 3047-3059.

Sottas, P. E., E. Larquet, A. Stasiak, and J. Dubochet. 1999. Brownian dynamics simulation of DNA condensation. *Biophys J* 77: 1858-1870.

Ussery, D., H. H. Stærfeldt, and L. J. Jensen. 2000. DNA structural analysis of sequenced prokaryotic genomes. http://www.cbs.dtu.dk/services/GenomeAtlas/

Vig, B. K. 1994. Do specific nucleotide bases constitute the centromere? *Mutat Res* 309: 1-10.

Wang, M. D., M. J. Schnitzer, H. Yin, R. Landick, J. Gelles, and S. M. Block. 1998. Force and velocity measured for single molecules of RNA polymerase. *Science* 282: 902-907.

Webb, C. D., A. Teleman, S. Gordon, A. Straight, A. Belmont, D. C. Lin, A. D. Grossman, A. Wright, and R. Losick. 1997. Bipolar localization of the replication origin regions of chromosomes in vegetative and sporulating cells of *B. subtilis*. *Cell* 88: 667-674.

Weitao, T., S. Dasgupta, and K. Nordstrom. 2000a. Role of the *mukB* gene in chromosome and plasmid partition in *Escherichia coli*. *Mol Microbiol* 38: 392-400.

Weitao, T., K. Nordstrom, and S. Dasgupta. 2000b. *Escherichia coli* cell cycle control genes affect chromosome superhelicity. *EMBO Rep* 1: 494-499.

Westerhoff, H. V., P. R. Jensen, J. L. Snoep, and B. N. Kholodenko. 1998. Thermodynamics of complexity: The live cell. *Thermochimica Acta* 309: 111-120.

Wheeler, R. T., and L. Shapiro. 1997. Bacterial chromosome segregation: is there a mitotic apparatus? *Cell* 88: 577-579.

Woldringh, C. L., P. R. Jensen, and H. V. Westerhoff. 1995. Structure and partitioning of bacterial DNA: determined by a balance of compaction and expansion forces? *FEMS Microbiol Lett* 131: 235-242.

Woldringh, C. L., and R. van Driel. 1999. The eukaryotic perspective: Similarities and distinctions between pro- and eukaryotes. Pages 77-90 in Charlebois, R. L., ed. *Organization of the prokaryotic genome*. ASM Press, Washington, D.C.

Woldringh, C. L., A. Zaritsky, and N. B. Grover. 1994. Nucleoid partitioning and the division plane in *Escherichia coli*. *J Bacteriol* 176: 6030-6038.

Yasuzawa, K., N. Hayashi, N. Goshima, K. Kohno, F. Imamoto, and Y. Kano. 1992. Histone-like proteins are required for cell growth and constraint of supercoils in DNA. *Gene* 122: 9-15.

Yin, H., M. D. Wang, K. Svoboda, R. Landick, S. M. Block, and J. Gelles. 1995. Transcription against an applied force. *Science* 270: 1653-1657.

Summary

Background

DNA, the carrier of genetic code, becomes biologically meaningful in the context of the cell. The way it is organised is important for its function. In general, DNA is much longer than a cell, e.g. the chromosome of *Escherichia coli* is a thousand times longer than a newborn cell. Thus, DNA is highly compacted, of course without losing functionality. In exponentially growing *E. coli* cells, DNA is not statically organised. During the cell cycle DNA is replicated and segregated towards the future daughter cells (replicated DNA is 'partitioned'). In *E. coli* replication, segregation, and cell growth take place concurrently. The manner in which DNA segregates in *E. coli* is not yet clear. Nucleoids as a whole segregate gradually, in line with cell elongation, but GFP studies in living cells suggest that specific DNA regions segregate non-linearly.

Aim

Our aim was to describe how specific DNA regions segregate, *on average*, during the cell cycle of *E. coli*, using cells from a steady state culture (chapters 2 through 4). Ultimately, we aim to define a model that explains our observations and those obtained with GFP experiments (chapters 1, 5 and 6).

Methods

The experimental method that we used to study DNA segregation in *E. coli* was fluorescent *in situ* hybridisation (FISH). With this method it is possible to visualise specific DNA regions in fixed cells from a steady state culture. Measuring the intracellular positions of these regions allowed us to construct a model of DNA segregation as a function of the cell cycle. Our models predict the position of all copies of a DNA region per cell length (see below). The number of DNA regions is dependent of the timing of replication within the cell cycle. Because FISH may not label all possible DNA regions in a cell (chapter 3 and section 5.2.5), we estimated the number of DNA regions per cell length using measured values of *C* (the time to fully replicate the bacterial chromosome), and *D* (the time from termination of replication to cell division).

In addition to FISH, we embarked on using a GFP-fusion protein to measure the speed and extent of diffusion of a 'GFP-tagged' DNA region in *E. coli* (chapter 4). Unfortunately, uncertainty about possible artefacts prevented us from doing quantitative measurements (see section 5.2.3).

Our analyses were based on linear regression (chapters 2 and 3), and on a comparison of computer simulated data with experimentally obtained data (chapter 4). Linear regression was used to estimate the average position of a DNA region as a function of cell length. The estimated numbers of DNA regions per cell were used to select sub-sets of data (chapter 2), or for the number of regression lines to be fitted (chapter 3). Computer simulation was used to produce data on the basis of three different models, two of which were based on previously proposed models, which we adapted to conform to measured *C* and *D* periods. The first model predicted gradual segregation of replicated DNA regions, completely in line with cell growth (conform nucleoid segregation). The second model predicted that replicated DNA regions stick together for some before one of them 'jumps' abruptly to the other side of the cell (conform models based on results from GFP experiments). The third model was a control-model, which predicted that a DNA region may be

anywhere in a nucleoid.

Results

The results of FISH with DNA probes targeted at the DNA region containing the origin of replication of *E. coli* (*oriC*) suggested that the *oriC* region nearest to a cell pole keeps a constant distance from the cell pole (chapter 2). This corresponds to gradual segregation in line with cell elongation. In chapter 3 we estimated the average positions of two additional DNA regions (*ftsQAZ*, and *minB*), now using all available data points and the estimated number of DNA regions per cell length. The slope of the regression lines of all three regions suggested that the average speed of segregation of replicated DNA regions corresponded to the speed of cell elongation. The cell length at which regression lines crossed, corresponded roughly to the cell length at which the DNA regions would replicate as calculated from C, D, and T_d. This is in support of a model in which replicated DNA starts to segregate immediately and in a gradual manner. To investigate this further we simulated data based on such a linear model and on a model that included a jump of replicated DNA regions. The distribution of data points of either model did not fully comply with experimental data. Both lacked data points in the cell centre of intermediate sized cells. More importantly, the amount of variation was much larger than we could account for by, for instance, measurement errors. Therefore, we tested the possibility that diffusion of DNA within a confined region is an important aspect of the localisation of DNA regions. Indeed, we included diffusion in our models and found that a substantial amount of variation may be explained by diffusion of DNA.

Conclusions

Initially, we considered the gradual segregation model the most straightforward model to describe our data. However, we concluded from our computer simulation studies that when using FISH it is not possible to discriminate between this model and a model that

incorporates a jump of replicated DNA. Moreover, we concluded that neither of the two models explained our data sufficiently. Therefore, I reconsidered some of the assumptions on which these models were based in chapter 5. Possibly, less strict assumptions would make simulated data resemble our data more. However, the implications of 'loosening' the assumptions are not always clear. For instance, we do not know how variation in the timing of replication within the cell cycle would affect the intracellular organisation of DNA.

In any case, the models that we tested were not based on some underlying mechanism that incorporates factors present in a cell. In chapter 1 I introduced a number of factors that organise DNA within the cell; in chapter 6 I review some of them for their possible role in DNA segregation. In this chapter I also compare DNA segregation in *E. coli* with that in eukaryotes. Although there may be some similarities (e.g. the Brownian motion of DNA regions), the differences in size are so large that we do not consider the comparison useful at this time.

Factors that organise DNA in *E. coli* include DNA itself, because it has intrinsic structure, macromolecular crowding, which helps to keep supercoiled DNA compact, and DNA binding proteins, which may (transiently) influence the structure and/or position of DNA. Because DNA segregation is controlled by a number of factors that interact non-linearly, it is difficult to define a straightforward model. I suggested to start with a model based on diffusion of DNA, DNA-DNA repulsion, and cell growth. Using sophisticated computer simulation it should be possible to simulate data to compare to experimental data obtained by, for instance, FISH or GFP experiments.

Samenvatting voor de leek

DNA IN BEWEGING

DNA is vooral bekend als de drager van erfelijke informatie, maar hoe is DNA georganiseerd in de cel en hoe verandert deze organisatie in de tijd, oftewel: hoe beweegt het? Met die vragen heb ik me de afgelopen jaren beziggehouden. In het eerste deel van deze samenvatting leg ik uit wat erfelijke informatie is en waarom bovenstaande vragen belangrijk zijn. In het tweede deel vat ik de inhoud van dit proefschrift samen, aan de hand van zijn titel.

1. Wat is erfelijke informatie en hoe functioneert het?

Erfelijke informatie

Organismen zoals de mens bestaan uit heel veel cellen. Er zijn levercellen, hersencellen, huidcellen, noem maar op. Al die cellen bij elkaar bepalen hoe wij eruit zien, hoe we bewegen, hoe we denken, enzovoorts.

Hoewel cellen in uiterlijk en in functie behoorlijk kunnen verschillen, komt toch elk van onze cellen voort uit één enkele cel: bij de bevruchting versmelten een eicel van de moeder en een zaadcel van de vader tot één cel. Vanuit deze cel ontstaan twee, iets verschillende, cellen. Deze cellen verdubbelen zich ook weer, de vier nieuwe cellen doen dat ook weer, enzovoorts, enzovoorts. Uiteindelijk ontstaat weer een nieuwe mens, bestaande uit vele miljarden cellen.

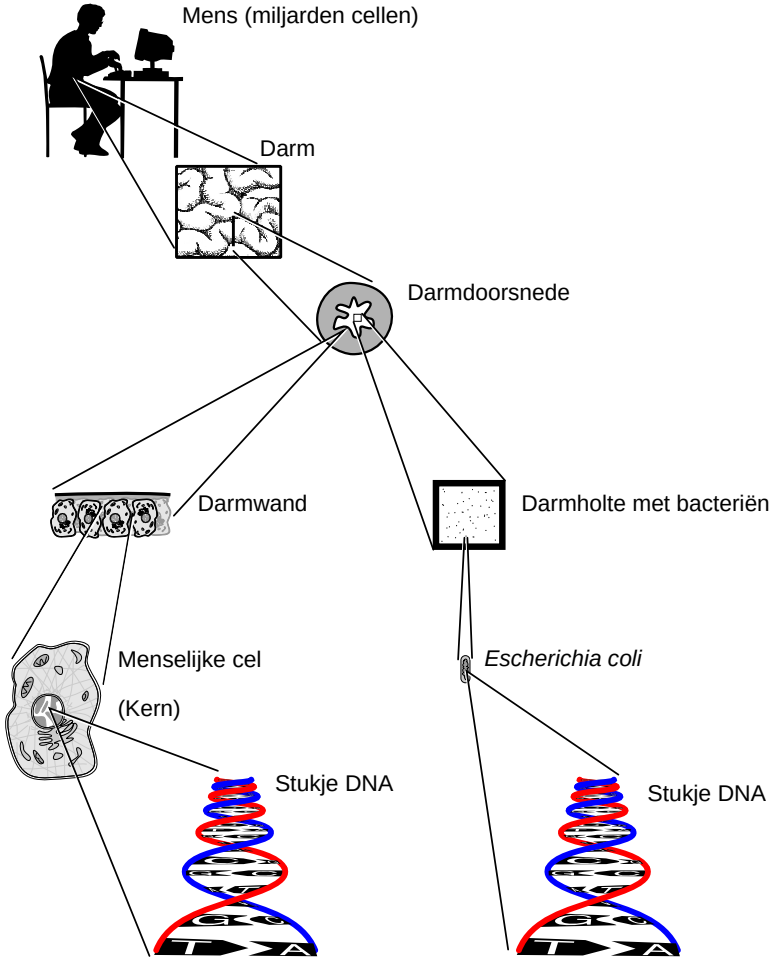
Hoeveel lijken de cellen van ons lichaam nog op die eerste cel? Als we binnen in een cel kijken zal het opvallen dat bijvoorbeeld een hersencel nog veel lijkt op een spiercel. Sterker nog, in theorie kan je met elk van onze cellen weer een mens maken. Dit komt omdat in elke cel een bibliotheek met 'receptenboeken' aanwezig is. Deze receptenboeken bevatten alle informatie om een nieuwe mens te maken. Omdat deze informatie wordt overgedragen van generatie op generatie noemen we het *erfelijke* of *genetische informatie*.

Bij verdubbeling van een cel krijgt elke dochtercel een volledige set erfelijke informatie. Daarom wordt, voordat een cel deelt, eerst de erfelijke informatie verdubbeld. Vervolgens wordt deze informatie verdeelt over de twee toekomstige dochtercellen. Bij een bevruchting wordt een combinatie gemaakt van de informatie van moeder en van vader, zodat we eigenschappen van beide ouders *erven*.

De erfelijke informatie is essentieel voor het functioneren van een cel. Neem bijvoorbeeld een bouwsteen van de celwand. Als daar behoefte aan is, dan wordt eerst het recept voor deze bouwsteen 'gelezen' in onze erfelijke informatie. Maar wat wordt dan gelezen? Hoe ziet erfelijke informatie eruit?

DNA

Erfelijke informatie is aanwezig in een cel als een code die terug te vinden is in DNA moleculen (chemische naam: *desoxyribonucleinezuur*, of



Figuur 1 - DNA van de mens en van de bacterie *Escherichia coli*

De mens bestaat uit miljarden cellen. Links zijn cellen uit de celwand van de dunne darm schematisch afgebeeld. Elk van onze cellen bevat een kern, waarin DNA is opgeslagen. In de darmholte leven bacteriën. Rechts is een *Escherichia coli* (*E. coli*) cel afgebeeld. Alle organismen hebben DNA, dus ook *E. coli*. Menselijke cellen zijn 0.01-0.03 mm lang; ze bevatten circa twee meter DNA opgedeeld in 46 stukken (chromosomen); in totaal is het DNA opgebouwd uit 6 miljard bouwstenen (A's, C's, T's, en G's); *E. coli* cellen zijn 0.002-0.004 mm lang; het ene chromosoom van *E. coli* is 1.5 mm lang; het bestaat uit 4.6 miljoen bouwstenen. De erfelijke code van de mens en *E. coli* verschillen aanzienlijk.

in het Engels: *deoxyribonucleic acid*). Je kunt je een DNA-molecuul voorstellen als twee in elkaar gevlochten lange linten (figuur 1). De linten bestaan uit een aaneenschakeling van vier bouwstenen, aangeduid met de letters A, C, G, en T *. Met deze letters worden de recepten voor de componenten van de cel geschreven. Een recept voor één component wordt meestal een *gen* genoemd. Zo is er een gen voor het enzym** dat suiker afbreekt en voor het pigment dat onze ogen een kleur geeft.

De ontcijfering van de genetische code heeft veel discussie losgemaakt. Vaak wordt gesuggereerd dat als je de code weet, je precies weet hoe cellen funtioneren. Belangrijker nog, je zou in staat zijn dat functioneren gericht te veranderen door 'genetisch te manipuleren'. Het is echter goed te bedenken dat het ontcijferen van, bijvoorbeeld, de circa 30 duizend genen van de mens, niet meer opgelevert dan een woordenboek van een nog onbekende taal. Om de taal te leren spreken moeten we de verbanden tussen de woorden vinden, en moeten we leren wanneer welke woorden worden gebruikt. Hoe cellen functioneren is nog grotendeels onbekend. Dat is niet zo verwonderlijk als je bedenkt hoeveel interacties 32 duizend componenten kunnen aangaan.

Hoe is DNA georganiseerd in een cel?

Ons doel is om de werking van een cel te ontrafelen door de componenten van de cel te onderzoeken in hun natuurlijke context, dus in een cel. Met betrekking tot DNA vragen we ons het volgende af: 1. 'Hoe is DNA georganiseerd in de cel' en 2. 'Hoe verandert deze organisatie in de loop van de tijd?' Als we ons DNA moleculen weer voorstellen als receptenboeken dan vragen we eigenlijk: 1. 'Welke boeken staan in de schappen en welke liggen open op de leestafel?', en 2.

* A, C, G, en T zijn de eerste letters van de chemische bouwstenen (*nucleotiden*), waaruit DNA is opgebouwd: Adenine, Cytosine, Guanine, en Thymine.

** Een 'enzym' is een molecuul dat helpt om een stof A in een stof B om te zetten, het is een catalysator; enzymen spelen een rol bij bijna alle chemische reacties in een cel.

'Wanneer en hoe worden de boeken gewisseld tussen schap en leestafel?'. Het antwoord op vraag 1 zegt iets over het functioneren van de cel *op een bepaald moment*. Het antwoord op vraag 2 zegt iets over het functioneren van de cel *in de loop van de tijd*.

In de cel van bijvoorbeeld de mens zit een enorme hoeveelheid DNA. Als je het DNA van één cel achter elkaar zou leggen kom je op zo'n twee meter uit (circa 6 miljard A's, C's, G's. en T's). Dit moet passen in een ruimte met een diameter die niet groter is dan 10 μm (0.01 mm). Dit is vergelijkbaar met circa 6 km draad in een ping-pong balletje!

DNA in de cel is daarom sterk opgevouwen, maar het mag hierdoor niet ontoegankelijk worden. Ten eerste moeten genen worden gelezen; ten tweede moet, als de cel zich wil verdubbelen, het DNA eerst verdubbeld worden. Ik heb me vooral beziggehouden met de beweging van zich verdubbelend DNA, maar ik deed mijn onderzoek niet met menselijke cellen.

Er zijn vele soorten cellen, maar omdat de cellen van alle organismen DNA bevatten, kunnen we voor fundamentele vragen uitwijken naar eenvoudiger cellen dan die van de mens. Ik ben begonnen met cellen van het Indiase blafhert, omdat die minder, maar grotere, DNA moleculen heeft. Uiteindelijk heb ik me toegelegd op *Escherichia coli*, omdat deze darmbacterie nog veel eenvoudiger is. Zo heeft een mens, maar ook een blafhert, zo'n duizend maal meer DNA per cel dan *E. coli*.

2. Titelverklaring

De titel van mijn proefschrift vertaald in het Nederlands luidt:

> ***DNA segregatie gedurende de celcyclus van Escherichia coli***
> *Een analyse van intracellulaire posities van fluorescent gemarkeerde gengebieden*

Hieronder vat ik de inhoud van mijn proefschrift samen aan de hand van de termen uit deze titel.

DNA segregatie gedurende de celcyclus van Escherichia coli

'Onze' bacterie *E. coli* heeft één cirkelvormig DNA-molecuul opgebouwd uit 4.6 miljoen bouwstenen (de A's, C's, G's en T's), samen 1.5 mm lang. Ter vergelijking: een menselijke cel bevat 46 lineare DNA-moleculen opgebouwd uit 6 miljard bouwstenen, in totaal twee meter lang . Een bacterie heeft veel minder DNA per cel dan wij, maar de cel is ook navenant kleiner.

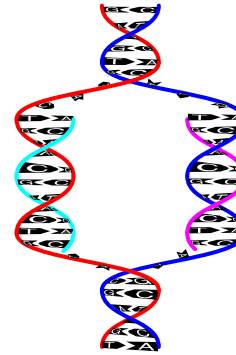
DNA segregatie gedurende de celcyclus van Escherichia coli

DNA segregatie is het uit elkaar bewegen van verdubbeld DNA. Als een cel wil verdubbelen, wordt eerst het DNA verdubbeld. Dat gaat als volgt. De twee in elkaar gevlochten linten die een DNA molecuul vormen, worden van elkaar gehaald (zie figuur 2). In elk van de oude linten wordt een nieuw lint gevlochten. Zo ontstaan uit één DNA molecuul twee nieuwe DNA moleculen, elk bestaand uit een oud en een nieuw lint. Het van elkaar halen begint op één plek van het DNA. Van daaruit wordt steeds meer van elkaar gehaald. Met invlechten van nieuw DNA wordt meteen begonnen; er wordt niet gewacht tot alles van elkaar is.

Nieuw DNA moet verdeeld of *gesegregeerd* worden over twee

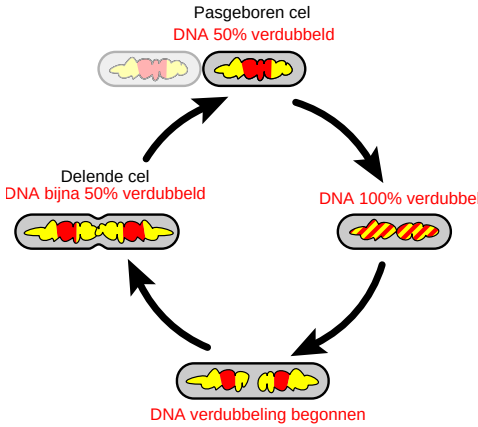
Figuur 2 - DNA verdubbeling

De twee 'linten' van DNA (rood en blauw) gaan uit elkaar; nieuwe linten (magenta en cyaan) worden in de oude gevlochten.



***Figuur 3 - E. coli* celcyclus**

Van bovenaf: een pasgeboren cel groeit tot ze twee keer zo groot is, dan deelt ze en beginnen we weer van voren af aan. Gedurende de celcyclus wordt het DNA van *E. coli* verdubbelt, maar het begin van verdubbeling valt niet samen met celdeling. In ons geval begon het verdubbelen van DNA halverwege de celcyclus.



toekomstige dochtercellen. Over hoe dit precies verloopt is het laatste woord nog niet gezegd. Dat geldt voor bacteriën, maar nog meer voor menselijke cellen. Waarschijnlijk begint segregatie zodra het eerste stukje nieuw DNA is ingevlochten. Bij bacteriën gaat het segregeren verder waarschijnlijk gelijk op met het groeien van de cel. Mijn onderzoek heeft ons nieuwe inzichten verschaft over het mechanisme achter segregatie (zie onder), maar we hebben het probleem nog niet geheel opgelost. Bij mensen gaat het segregeren in stappen; eerst een klein stukje (misschien op dezelfde manier als bij *E. coli*), pas later helemaal.

DNA segregatie gedurende de celcyclus van Escherichia coli

Zich vermeerderende cellen doorlopen een cyclus: de celcyclus. Een pasgeboren cel groeit totdat ze twee keer zo groot is, groot genoeg om weer te delen (figuur 3). De hoeveelheid voedsel en de temperatuur bepalen hoe snel deze cyclus verloopt (en de vorm van de cellen; hoe dik ze zijn bijvoorbeeld). Onze bacteriën doorliepen hun celcyclus in 79 minuten. Dat noemen we langzame groei (een cyclus van 20 minuten noemen we snelle groei). De celcyclus van menselijke cellen is veel langzamer, bijvoorbeeld 24 uur (snelle groei, van bijvoorbeeld kankercellen, duurt toch nog een paar uur, en is vaak niet foutloos).

Tijdens het groeien worden alle componenten van de moedercel verdubbeld en verdeeld over de toekomstige dochtercellen, zo ook DNA. Het moment waarop de verdubbeling van DNA wordt ingezet hangt af van de groei-omstandigheden. Voor ons onderzoek hebben we zulke groei-omstandigheden gekozen dat we precies wisten wanneer DNA-verdubbeling begon en hoelang het duurde. Dit was nodig voor een juiste interpretatie van onze data.

De celcyclus van bacteriën kan heel complex zijn. Een bacterie wil zich soms sneller vermeerderen dan DNA kan verdubbelen; het verdubbelen van DNA moet dan maar eerder beginnen, desnoods in een voorgaande celcyclus. Om uit te leggen hoe dit tot snellere groei leidt, maak ik de volgende vergelijking: stel dat mensen zich sneller zouden willen voortplanten dan eens per 9 maanden, terwijl de dracht-tijd 9 maanden blijft. De 'bacterie-oplossing' zou zijn om reeds zwangere kinderen te baren. Dan wordt binnen 9 maanden de volgende generatie geboren. (Gelukkig is dit onmogelijk.)

DNA segregatie gedurende de celcyclus van Escherichia coli

Escherichia coli (*E. coli*) is een darmbacterie die bij ons allemaal voorkomt (figuur 1). We leven ermee samen; zonder bacteriën zou ons eten niet verteren. *E. coli* is een ééncellig organisme (de mens is een meercellig

organisme). De *E. coli* cellen die wij kweekten waren ongeveer 0.0015 tot 0.004 mm lang en 0.001 mm breed. Het DNA van *E. coli* (1.5 mm) is dus zo'n duizend maal langer dan een pasgeboren cel. Menselijke cellen zijn minimaal 10 keer groter dan een bacterie.

DNA segregatie gedurende de celcyclus van Escherichia coli

Analyse van intracellulaire posities van fluorescent gemarkeerde DNA gebieden

Het is niet mogelijk direct de bewegingen van DNA in een cel te zien, zelfs niet met krachtige lichtmicroscopen. Het kleinste deeltje dat onze lichtmicroscopen zichtbaar maken is 0.0002 mm in doorsnee; een pasgeboren *E. coli* cel is maar tien keer groter. We kunnen DNA kleuren, maar het totaal aan DNA is dan zichtbaar als één structuur, zonder enig detail (zie figuur 3.2). Vergelijk het met het bekijken van een bos op grote afstand: de contouren zijn zichtbaar, de bomen niet. We zijn niet overgestapt op vormen van microscopie die meer detail zichtbaar maken, omdat die een behandeling vereisen die tot vervorming van de cellen leidt.

Om toch de beweging van DNA te bestuderen, hebben we een techniek gebruikt waarbij fluorescerende moleculen hechten aan een klein stukje DNA. Zodra licht op deze moleculen valt, gaan ze licht uitzenden (vergelijkbaar met een fietsreflector) en 'verraden' daarmee de positie van het stukje DNA. Door de microscoop zien we een cel met daarin één of meer lichtvlekjes (figuur 3.2). Meer vlekjes betekent dat het stukje DNA is verdubbelt. Door in een groot aantal cellen van verschillende lengte de positie van de vlekjes te meten, konden we de beweging van DNA in de tijd reconstrueren.

Wij vonden dat in de gemeten posities van drie specifieke stukjes DNA veel variatie zat. We hebben daaruit de conclusie getrokken dat DNA continu kleine willekeurige bewegingen maakt: het 'wiebelt'. Gemiddeld genomen beweegt verdubbeld DNA geleidelijk uit elkaar, maar dat doet

het al wiebelend.

Omdat het DNA zo wiebelt, kan je in individuele cellen grillige bewegingen van DNA waarnemen. Men heeft wel gedacht dat zulke bewegingen erop duiden dat er specifieke moleculen zijn die aan DNA trekken of het duwen. Wiebelen is echter een normaal fenomeen voor moleculen ('Brownse' bewegingen of diffusie). Wij hebben daarom voorgesteld dat het wiebelen van DNA een belangrijke bron van beweging is voor het uit elkaar bewegen van verdubbeld DNA. Ons model is dat DNA verdubbelt en dan 'wiebelend' uit elkaar 'drijft', terwijl de cel groeit.

Het zou goed kunnen dat de bewegingen die menselijk DNA maakt, lijken op wat er bij bacteriën gebeurt. De vergelijking moet echter niet te ver doorgevoerd worden. Zo moet het verdubbeld DNA van een mens uiteindelijk over een veel grotere afstand verplaatst worden, en daarvoor bestaan speciale mechanismen (figuur 6.2). Dat neemt niet weg dat het 'wiebelen' van DNA altijd een rol speelt, zowel bij bacteriën als bij mensen.

DANKWOORD

Acknowledgements

Eindelijk, het mag nu: ongegeneerd mensen bedanken. Ik begin niet op het lab, maar thuis. Tenslotte hebben mijn collega's mijn aanwezigheid aan de mensen daar te danken. Pa, ma: jullie hebben mogelijk gemaakt dat ik alles wat ik wilde doen, ook kon doen. Jullie vertrouwen heeft mij altijd veel zelfvertrouwen gegeven. Mede dankzij jullie ben ik nu geen doorsnee bioloog, maar heb ik een eigen plekje gevonden, bovendien een die toekomst lijkt te hebben. Wat het kopen van die eerste BBC computer al niet tot gevolg heeft gehad! Roland, bedankt voor het zijn van 'mijn grote broer' (ook al ben ik langer), voor je computer-hulp, en natuurlijk voor je muziek. Je hebt een trots 'klein' broertje!

Bedankt ook Mirjam, Jaap, Barbara, de ouders en zus van mijn vriendin Simone. Zij zijn anders dan mijn familie (zij komen bijvoorbeeld wel op tijd), maar op één punt zijn ze hetzelfde: ook zij hebben alles voor elkaar over, en nu ook voor mij. Recent voorbeeld: dit boekwerk, een cadeau van deze familie; ik sta nog steeds met mijn mond vol tanden.

Simone, met jou ben ik nog steeds heel gelukkig, maar ik heb het jou de laatste tijd niet gemakkelijk gemaakt om ook gelukkig met mij te zijn.

Het schrijven van dit proefschrift was niet alleen voor mij een beproeving, mijn lange dagen en stress hebben ook veel van jouw energie en nachtrust gevraagd. Maar je komt nog elke dag thuis, dus ik geloof dat we het doorstaan hebben. Bedankt voor je hulp bij mijn werk (zowel geestelijk als inhoudelijk), voor Simone zijn en voor bij mij zijn.

Zonder mensen bij naam te noemen bedank ik familie, vrienden en de Snippen. Zij zijn voor mij van wezenlijk belang om geestelijk gezond te blijven.

Dan nu de collega's. Mijn eerste stage bij Michel Haring was leerzaam, maar niet in alle opzichten leuk. Adriaan Houtsmuller, een unieke onderzoeker, was mijn tweede stagebegeleider. Ik zag dat de combinatie computer en biologie iets voor mij was, daar dank ik Adriaan en Moleculaire Cytologie voor. Het is niet toevallig dat ik bij dezelfde groep ben gaan promoveren. Na mijn stage heb ik vervangende dienst op het AMC gedaan (plaatjes maken van rundersperma bij de groep Radiobiologie, later Centrum voor Microscopisch Onderzoek). Dat ik daar kwam dank ik aan mijn goede vriendin Astrid. Astrid, bedankt, want het is een hele fijne tijd geweest. Dat is niet in het minst te danken aan de collega's daar: Jan, Carel, Ron, Jacob, Erik, Klaas, Chris, Greet, Henny, Jacob, Gertrude, Bart Kipp, Bart Maes, Marieke, en Lieve (ben ik nu niemand vergeten?).

Na mijn vervangende dienst begon mijn promotie bij Moleculaire Cytologie op de biologiefaculteit Anna's Hoeve. Eerst werkte ik aan de chromosomen van de Muntjac. Cellen kon ik niet op Anna's Hoeve kweken, waardoor ik kon blijven samenwerken met de mensen op het AMC. Jan heeft me daar het kweken geleerd, maar hij heeft me vooral vertrouwen op het lab gegeven. Wat betreft de Muntjac-tijd, wil ik Jacky en Oof speciaal bedanken. Jacky heeft een enorme inzet, ook voor anderen zoals voor mij destijds. Samen met Oof heb ik regelmatig onderwijs verzorgd, wat heel leerzaam voor mij was. Oof heeft zich toegelegd op onderwijs, wat ergens jammer is, want ik kon Oof zijn originele manier van wetenschap bedrijven altijd zeer waarderen.

I would also like to thank Harry Scherthan. Harry, before I switched to E. coli, you were my 'remote supervisor'; via e-mail you helped me tremendously in working with Muntjac cells. Your attitude towards working with colleagues from other laboritories should be an example to all scientists.

Onze plannen met de chromosomen van de Muntjac bleken niet uitvoerbaar (in ieder geval niet in de beschikbare tijd). Daarom ben ik overgestapt naar *E. coli*. Proeven heb ik haast niet meer gedaan; ik kon me volledig toeleggen op de analyse. Daarom ben ik zeer veel dank verschuldigd aan Jacky, Anton en Mirjam. Toen ik nog stoeide met de Muntjac, hebben zij de in dit proefschrift beschreven FISH methode voor bacteriën ontwikkelt. Vooral Mirjam wil ik speciaal bedanken, want zonder haar zou ik niet promoveren. Mirjam, je doet je werk niet alleen ontzettend goed, maar het is ook zeer prettig met jou samen te werken. Je hebt me wel eens gezegd dat je het liefst langer met één AiO zou meewerken; dat gun ik jou, maar vooral die AiO.

Het is niet de traditie, maar ik ben zeer te spreken over de begeleiding die ik tijdens mijn promotie heb gehad, eerst van Jacky en Oof, en in de laatste jaren van mijn promotor, prof. dr. Nanne Nanninga. Bij deze wil ik u hartelijk danken voor een - in mijn ogen - heel vruchtbare samenwerking. Uw planning en daardoor overzicht benijd ik zeer.

Ik heb het erg naar mijn zin gehad bij Moleculaire Cytologie. Dat komt natuurlijk door de fijne collega's. Naast de al genoemden bedank ik daarom Peter, Raquel (en Oscar), Martin, Jan, Ellen, Evelien, Thessa, Sonia, Erhan, Conrad (ook voor het 'brainstormen'), Norbert, Cecile-Marie ('nog tijd voor thee?'), Tanneke, Wijnand, Fred, Michiel, Mark, Erik Manders, Gerard, Erik van Munster (en Annemarie en Erik speciaal voor het me op het spoor zetten van het simuleren), Thijs, Julion, Nicoletta, en de mensen die inmiddels elders werken zoals Jos, Herke-Jan, Geert, Rick, Philipine, Hans, Jitze, Bert, Bob, Rogier, Joop, en Nienke. Alle studenten die ons hebben aangedaan, zoals recentelijk Luísa en Maarten, hebben natuurlijk ook bijgedragen aan de goede sfeer. Grace, 'mijn' stagiaire, wil ik speciaal bedanken omdat ze een byzondere en

gezellige studente was, en een goede bovendien.

Voor de hulp bij het schrijven van dit proefschrift wil ik mijn promotor en Conrad Woldringh nog speciaal bedanken. Jullie schat aan kennis heeft mij een groot voordeel gegeven. Voor wiskundige zaken ben ik vaak bij Rob Lingeman te rade gegaan. Rob, dat was niet alleen nuttig, maar ook altijd een groot plezier voor mij. In de laatste fase heeft Jaap Kaandorp met me meegedacht over hoe DNA bewegingen in bacteriën gemodelleerd en gesimuleerd zouden moeten worden. Er is niet zo heel veel van in dit proefschrift terecht gekomen, maar ik heb er persoonlijk veel aan gehad; bedankt daarvoor.

Tenslotte; ik weet dat velen deze tekst als eerste (en laatste) lezen, en zich dan stiekem afvragen: 'word ik genoemd'? Mezelf kennende, vergeet ik mensen die ik niet had willen vergeten. Dat je dit nu leest, maakt de kans groot dat je één van die mensen bent. Daarom, bij deze, voor jou: bedankt!