# Phase Variation of Type 1 Fimbriae:
# a Single Cell Investigation

# Phase Variation of Type 1 Fimbriae: a Single Cell Investigation

ACADEMISCH PROEFSCHRIFT

ter verkrijging
van de graad van doctor aan de Universiteit van Amsterdam,
op gezag van de Rector Magnificus prof.dr. D.C. van den Boom,
ten overstaan van een door het college voor promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op vrijdag 16 januari 2009, te 14:00 uur

door

Aileen Martinia Adiciptaningrum

geboren te Semarang, Indonesie

Promotiecommissie :

Promotor : prof. dr. D. Frenkel
Copromotor: dr. S.J Tans

Overige leden: dr. T. den Blaauwen
prof. dr. P. Borst
prof. dr. Kuipers
prof. dr. P.R. ten Wolde
dr. C.L Woldringh

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

The work described in this thesis was performed at the FOM Institute for Atomic and Molecular Physics (AMOLF) in Amsterdam, The Netherlands. This work is part of the research program of the ‘Stichting voor Fundamenteel Onderzoek der Materie (FOM)’, which is financially supported by the ‘Netherlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)’.

ISBN - 9789077209295

Cover design by Julien Husson and Aileen Adiciptaningrum. The image is an inverted phase contrast micrograph of a microcolony on agar surface. Printed in the Netherlands by Ponsen and Looijen BV graphical company, Wageningen.

Variability is a hallmark of biological systems.

*Brehm-Stecher & Johnson*
*(Microbiol. Mol. Biol. Rev., 2004)*

## Contents

**Chapter 1 Introduction** 9

1.1 The single cell approach 10
1.2 Phenotypic diversity and phase variation 11
1.3 The *fim* system of *E.coli* 12
1.4 Cell cycle and DNA replication of *E.coli* 15
  1.4.1 The cell cycle 16
  1.4.2 Models of DNA replication initiation 17
1.5 Methods of C and D period determination 19
  1.5.1 Methods for measuring the C period 20
  1.5.2 Methods for measuring the D period 21
1.6 DNA replication initiation 24
1.7 DNA sequestering protein SeqA 25
1.8 This thesis 26

**Chapter 2 Experimental Material and Methods** 29

2.1 Genetic manipulation 30
  2.1.1 Allelic exchange 30
  2.1.2 Construction of a mCherry-SeqA fusion 30
  2.1.3 Strains and plasmids 32
  2.1.4 Materials 34
2.2 Microscopy 34
  2.2.1 Sample preparation 34
  2.2.2 Microscopy 35
  2.2.3 Materials 36
  2.2.4 Data analysis 36
2.3 Determination of GFP*mut2* maturation time 38

**Chapter 3 SeqA Dynamics in Slow Growing Cells** 41

3.1 Introduction 42
3.2 Results 43
  3.2.1 SeqA foci in rapid and slow-growing cells 43
  3.2.2 Regular pattern of SeqA foci dynamics in slow growing cells 43
  3.2.3 Quantification of SeqA foci dynamics and cell cycle parameters 46
3.3 Discussion and outlook 50

**Chapter 4 Direct Observation of Type 1 Fimbrial Expression in *E.coli*** **55**

4.1 Introduction 56
4.2 Results 57
 4.2.1 *fim* phase variation in real-time at the single cell level 57
 4.2.2 Distinct *fim* expression patterns are observed near switching 59
 4.2.3 *fim* expression pattern depends on growth rate and cell cycle 61
 4.2.4 The OFF to ON *fim* switching rate is higher at beginning of cell cycle 65
4.3 Discussion and outlook 65

**Chapter 5 The Effect of Cell History on the OFF Switching Rate of Type 1 Fimbriation in *E.coli*** **69**

5.1 Introduction 70
5.2 Results 71
 5.2.1 ON and OFF switching rates of the *fim* switch 71
 5.2.2 The effect of switching history on the OFF switching frequency 74
5.3 Discussion and outlook 77

**References** **79**

**Summary** **91**

**Samenvatting** **95**

**Publications** **98**

**Acknowledgements** **99**

**Curriculum vitae** **101**

# Introduction 1

A single cell approach to study biological systems has gained popularity due to recent technological advances and a growing appreciation for non-genetic cell-to-cell diversity. One example of this heterogeneity is phase variation, a reversible and stochastic switching between two types of expression in individual cells within a clonal population. In *E.coli*, the expression of type 1 pili (fimbriae) is phase variable, using DNA inversion as the underlying molecular mechanism. In this thesis, we present the results of our investigation on the *fim* system at the single cell level together with our investigation of the cells' physiological diversity during the cell cycle. In this study the DNA replication process was monitored simultaneously with cell growth and the link between the two processes was examined.

## 1.1 The single cell approach

In the field of microbiology, studies of complex cell processes, interactions and behaviors have traditionally been done at the bulk or population level. Typically, the culture of interest is inoculated with a sample of an isogenic microbial population and is then further assumed to be a population of identical cells. Any individual cell response that may exist is therefore subjected to averaging which results in mean population outcomes.

Increasing interest in cell-to-cell heterogeneity has led to the recent development of single cell approaches. These approaches, such as microscopy, flow cytometry, single-cell microarray, single cell PCR, fluorescence *in situ* hybridization, etc., allow measurement of individual cell properties. Work on microbial [39,92] and mammalian [98] cells has demonstrated that isogenic populations show a wide variety in gene expression levels, which can be attributed to fluctuations in cellular components stemming from the stochastic nature of biochemical processes. Since then, other studies (for reviews: [3,15,73]) have demonstrated that cell-to-cell variation is a widespread phenomenon in biology and that investigation at the single cell level is necessary to reveal individual heterogeneity.

The recent increase in single cell research has benefited from technological developments in fluorescent dyes and proteins, automated microscopy, growing computer storage capacity and computing ability. Even so, the interest in investigating cell-to-cell diversity already arose several decades ago. In 1953, Benzer realized a problem in studying kinetics of β-galactosidase induction with batch cultures [9]. He considered whether all cells participate equally and simultaneously in the synthesis of an enzyme. Since a direct method to quantify protein level in individual cells did not exist, he used a bacteriophage to quantify the participation of each cell in a culture growing solely on lactose. The phage duplication *in vivo* depended on the host β-galactosidase concentration already present before the moment of infection. Upon infection, the bacteriophage blocked β-galactosidase synthesis, and uniquely prevented its own development. Hence, the time at which an individual host cell was lysed depended on the enzyme content before infection. Benzer quantified the amount of β-galactosidase released to the medium at various stages of lysis as a measure of culture heterogeneity, and found that the population response is essentially homogenous in respect to enzyme content. His innovative experiment illustrates one of the many advantages that single cell methods have to offer.

The general advantage of a single cell approach is the evasion of averaging effects that are characteristic of bulk-phase population-scale methods. Many cell properties such as viability, number of proteins per cell, and the number of structures expressed on the cell surface are discrete and intrinsic states or properties of each individual cell. Single cell techniques also allow a high spatial and temporal resolution of dynamic events *in vivo*. And in several cases, they allow connections to be made between apparent macroscopic phenomena and their microscopic cellular origins. Of course, this approach also has its caveats. Several of the techniques are low throughput, so that sufficient statistics can be difficult to obtain. In addition, it is not possible to exclude the influences of the techniques on the measured cell, since an experiment and control cannot be carried out on the same cell. Therefore, while the single cell approach can be a very powerful tool of investigation, the limitations mentioned above must be kept in mind.

## 1.2 Phenotypic diversity and phase variation

Brehm-Stecher and Johnson categorized factors contributing to cell-to-cell diversity in four major classes: genetic diversity, biochemical or metabolic diversity, physiological diversity, and behavioral diversity [15]. Genetic diversity can arise from a number of random, semi-random, or programmed DNA modification events. Biochemical diversity is characterized by individual differences in cellular macromolecular composition or activity. Physiological diversity describes morphological differences between individual cells, including cell cycle-related characteristics such as size and shape. Behavioral diversity comes from the difference in observable cellular responses as a consequence of other cell-to-cell variation. These different classes of diversities are not strictly separated and in many cases are intertwined. For example, an apparent behavioral diversity, such as the phase varying competence state of *Bacillus subtillis*, can be traced back to the biochemical diversity underlying it, which is the noise in *comK* expression level [79]. Or, an apparent physiological diversity such as the start of DNA replication might have to do with the concentration of a certain protein required for initiating the replication. Often, the diversities also influence the best suitable methods and techniques of investigation.

Phase variation in general, refers to a reversible “all-or-none” type of expression of one or more proteins in an individual cell within a clonal population. The reversible switching between the two types of expressions is considered to be stochastic in nature, and occurs at a frequency exceeding that of a random mutation. A certain heritable genetic or epigenetic mechanism is underlying the process and in some cases, environmental factors modulate the switching rates (for excellent reviews:[48,116,117]).

The classical view on the biological role of phase variation is that it helps bacteria evade a host immune system. Phase varying mechanisms are commonly found to regulate external structures exposed to the environment. Therefore, phase variation is often associated with the regulation of virulence factors in bacterial pathogens. In support of this view, it has been shown that fimbrial phase variation in *Salmonella* is a means to evade cross-immunity between serotypes [90]. A recent theoretical study has shown that when the evasion is frequency dependent, phase variation can provide a fitness advantage by generating population heterogeneity [123].

This classical view is being challenged or complemented by a "bet-hedging" view, in which the suggested biological role of phase variation is to provide bacteria with a strategy to better survive fluctuating and unknown future environments. In a mixed population, a certain subset of the population is always prepared for any sudden experimental change, increasing the chance for overall population survival [36]. Recently, it has been shown that an irreversible decision to sporulate in a subset of a *Bacillus subtilis* population results in specialization in carbon source usage and different reproductive potential under different environments [118]. Theoretical models propose that in some fluctuating environments, a dynamically heterogeneous population could have higher net growth rate than a homogenous one [115], and that if the fluctuations are infrequent [70] or cannot be sensed precisely enough [122], stochastic switching could be preferred over a sensing mechanism.

## 1.3 The *fim* system of *E.coli*

As mentioned above, phase variation is a common regulatory mechanism for external structures. In *E.coli*, this includes various pili (P, S, type1, CS18, CS31A) and the Ag43 outer membrane protein. Among those pili, one of the best-studied systems is the *fim* system, involved in the expression of type 1 pili (fimbriae) widely expressed by both pathogens and commensal strains. Fimbriae are hair-like structures made of protein expressed on the cell surface (fig. 1.1.A) with lengths of about 1 μm and diameter of 6-7 nm [20,45]. A typical fimbriated *E.coli* expresses 200-500 pili on its surface. One pilus is an arrangement of more than 1000 major structural subunits (FimA) in a right-handed helical pattern (2.4 nm pitch and 3.4 subunits per turn) assembled via an usher/chaperone pathway at the outer membrane (fig. 1.1.B) [99]. At the tip of these pili, a short thin thread (tip fibrillum) made of FimF and FimG anchors an adhesin (FimH) that binds to the host mammalian receptors [54,55].

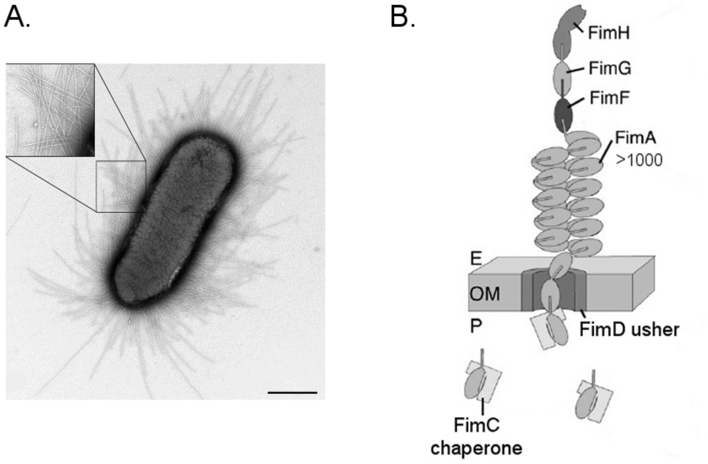

**Figure 1.1.** E.coli fimbriae.
(A) Electron microscopy image of fimbriated *E.coli* (adapted from [45]).
(B) Fimbrial subunits (FimA as major subunit) are expressed by the *fim* operon. The symbol E stands for extracellular, OM outer membrane and P periplasm (adapted from [99]).

In *E.coli,* the expression of fimbriae is thought to facilitate adhesion to host tissue at the early stage of infection, for example in Urinary Tract Infection (UTI) (fig. 1.2) (for a review see [106]). The expression of type 1 fimbriae, together with P fimbriae, is commonly found amongst clinical uropathogenic strains. In accordance with the classical view on phase variation mentioned earlier, the phase-varying feature in *fim* is also thought to represent the fine balance between pathogen and host interaction. While the nutritional advantage from infecting host cells is obvious, a high level of fimbriation evokes the host's inflammatory response, a threat to the bacterium's survival. Consequently, the level of fimbriation must be regulated such that those trade-offs are balanced.

Regulation for fimbrial expression lies on the invertible 314 bps DNA fragment *fimS* containing a promoter sequence. Upon inversion, the promoter sequence in *fimS* can obtain different orientations, acting as a reversible genetic switch. In the ON orientation, *fimS* drives the expression of multiple fimbrial structural genes, *fimA-fimH* (fig. 1.3). In the OFF orientation, the promoter is in an inverted configuration, and thus silent. The switching property of *fimS* is believed to be stochastic in nature, resulting in a mixed population of fimbriated and non-fimbriated cells in most environmental and physiological conditions. In contrast with other types of regulation such as repression, the stochastic inversion of *fimS*

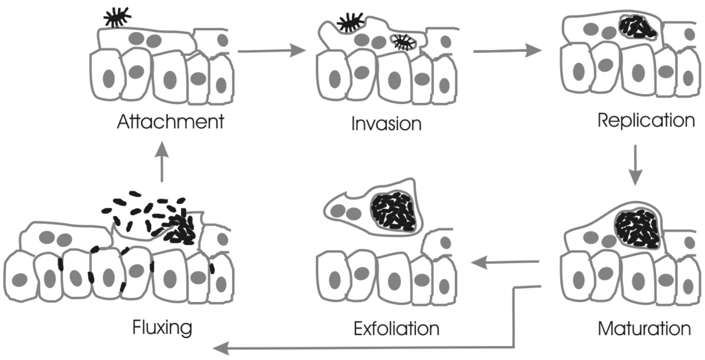

**Figure 1.2.** Schematic picture of Urinary Tractk Infection sequence of events.
The level of fimbriae expression has to be well regulated in order to balance the advantage of invasion (escaping starvation) and its disadvantage (exfoliation by the host immune system).

allows a partial, but immediate population response to sudden environmental changes. It also allows a unique all-or-nothing type of expression, a key feature for tuning the fine balance between a bacterial population and its host immune system.

The inversion of *fimS* is performed by two site-specific recombinases, *fimB* and *fimE*. While FimB is able to switch in both directions, FimE has an extreme bias to inverse the switch to the OFF orientation with a much higher rate (about 100-1000X more; up to 0.1 OFF switching events per cell per generation). FimB and FimE are homologous proteins (52% amino acid identity) and are members of the integrase family of tyrosine site-specific recombinases. Tyrosine recombinases break and rejoin single stranded DNA in pairs to form a Holliday junction intermediate [21,43]. The Holliday junction intermediate is formed by the interaction of multimeric protein complexes on the two Fim-recombinase binding sites flanking a 9 bps inverted-repeat on each side of *fimS* (IRR and IRL), thereby looping out the DNA between the sites (fig. 1.3). There are a total of four binding sites, called half sites, to be occupied for a successful inversion to occur. Temperature, environmental nutrients and host chemical responses all influence *fim* switching rates [41].

Besides regulating the expression of fimbrial structural genes, the orientation of *fimS* also regulates the expression of FimE. This phenomenon, called orientational control, is the expression of FimE only when *fimS* is in the ON orientation. Sequence differences mandate that stable mRNAs are only produced from the *fimS* ON orientation. In the opposite orientation, *fimE* transcription is prematurely terminated in a rho-dependent manner and more likely to be degraded [50,56,110].

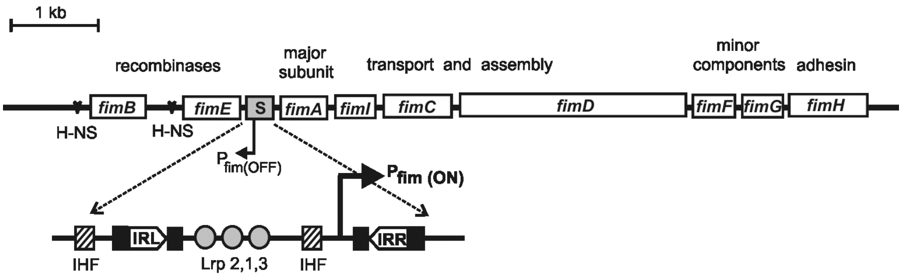

**Figure 1.3 The *fim* system of *E.coli*.**
Only when the switch, *fimS,* is in the ON orientation (as in the zoomed area), does the promoter sequence drive the expression of multiple fimbrial structural genes. The black boxes next to IRL and IRR are binding sites for the FimB and FimE recombinases, called half sites. There are also multiple binding sites for accessory proteins such as HNS, IHF and Lrp within and in close vicinity to the switch.

Previous models have proposed different biological reasoning for the existence of orientational control. One model proposed that the orientational control of *fim* acts as a timer to provide sufficient time for fimbrial development [121]. Other theoretical work has proposed that FimB and FimE act as antagonistic forces, each pushing towards a different steady state to enable rapid adjustment to fimbriation levels in response to environmental changes while keeping the overall fimbriation level low [25].

The regulation of the *fim* system is currently quite well understood. On the other hand, the dynamics of switching is still largely obscured. Previous investigations were done in bulk, and this approach poses certain constraints for fully understanding the system. For example, the correlation between switching events and cell processes as well as cell growth remains vague, and the effect of cell-state history on switching rates is elusive. We address these questions in the following chapters of this thesis as a report of our direct observation of individual events.

## 1.4 Cell cycle and DNA replication of *E.coli*

Single cell investigation also provides the possibility to correlate different factors contributing to cell-to-cell diversity. For example, in the case of the phase varying behavior of the *fim* switch of *E.coli*, it is still unknown whether a general physiological source of diversity, such as age within the cell cycle, volume and growth rate does effect the apparent behavioral diversity of individual cells. In this section, the principle behind the *E.coli* cell cycle is introduced (for a comprehensive review: [28]).

### 1.4.1 The *E. coli* cell cycle.

The division cycle not only describes the continuous birth, elongation and division of *E.coli* cells, but further includes processes during cellular growth, such as DNA replication, nucleoid partition and septum formation. Each process has to be coordinated in space and time to ensure a successful round of cell duplication. Cooper summarized the two main principles believed to govern *E.coli* cell cycle [28]. First, to ensure survival of daughter cells, a cell will not divide unless it has at least two copies of the genome. At the moment of division, the DNA content (and other materials) of a mother cell should be twice as much as that of a daughter cell so that the cell composition is stably maintained over time. Second, DNA replication will not start unless there is enough cytoplasm. Many important species in the cytoplasm are at low molecular concentrations, so that there needs to be enough proteins, enzymes, and other species for daughter cells to continue growth in the next cell cycle. Cell growth rates vary with the growth medium: the richer the medium, the shorter the cell interdivision time.

In *E.coli*, in order to double the DNA content before division, DNA replication is initiated only once in a single cell cycle. The DNA replication starts at the origin of the replication site *oriC*, and finishes at the termination site *ter*. Replication process is bidirectional, meaning that there are two replication forks proceeding in opposite directions around each half of the chromosome simultaneously. Experimental results with fast growing cells support a model for a constant period of DNA replication (about 40 minutes), called the C period. Thus, DNA replication proceeds at a constant rate of about 500 nucleotides per second. Another constant period (about 20 minutes), called the D period was observed between the end of DNA replication and cell division. Though the physiological function of the D period is unknown, it is proposed to be the time the cell needs to prepare for cell division, but the evidence for a molecular mechanism underlying it is lacking.

According to the Cooper-Helmstetter model of DNA replication [27], in a cell with an interdivision time of 60 minutes, DNA replication begins around birth time and terminates 20 minutes before division. Due to this required D period of 20 minutes, a cell that has an interdivision time of 50 minutes will have to start replicating its DNA 10 minutes before its birth, thus in the D period of its mother cell (fig. 1.4). Since it is possible for a cell to obtain a very short interdivision time (less than 40 minutes), a new round of replication will start before the previous one has ended. So, for example, in a cell that has an interdivision time of 30 minutes, DNA replication terminates 10 minutes after cell birth, and starts around the birth time of its mother cell, while the mother cell still undergoes the last 10 minutes of its own of DNA replication process. DNA replication patterns for fast growing cells with various interdivision times are shown in figure 1.4. Variation in growth rates was normally obtained by changing the type of medium and carbon sources in which the cells were growing.

Different experimental methods for determining the C and D period have led to a consensus on the DNA replication pattern for fast growing cells. There is less consensus, however, regarding the DNA replication pattern in slow growing cells (interdivision time more than 60 minutes). There are two main groups of observations involving slow growing cells. In the first group, several studies found that DNA replication starts around birth time and the length of the C and D periods increases (and rescales) to occupy 2/3 and 1/3 of the cell cycle [47,94] . In contrast to these observations, others found that when interdivision time increases significantly, DNA replication starts a significant time after birth. Consequently, there is a period devoid of DNA replication at the beginning of the cell cycle, called the B period [24,63,93]. An extreme case showed a constant C and D period of 40 and 20 minutes even for cells with an interdivision time of 50 hours grown in chemostat [66,68]. To summarize, different strains and different methods have lead to different conclusions. Therefore, it is useful to also review the common methods used to determine the C and D period. Short descriptions of those methods are summarized in section 1.5.

### 1.4.2 Models of DNA replication initiation

The pattern of DNA synthesis during division cycle at different growth rates, leads to the question of what regulates the timely firing of DNA replication initiation. Possibly, cell division triggers DNA replication initiation. But, as growth become faster and faster, DNA replication is initiated before cell division up to 2 generations earlier. So, cell division can be excluded as the sole DNA replication initiator.

The question remains then, how can a cell or a genome know when to initiate DNA replication, when a waiting time is required for one growth condition but not for others? An interesting calculation was reported by Donachie in 1968 [33]. By combining the data of the average initiation age for *E.coli* B/r [27] with measurements of the DNA/mass ratio from *Salmonella typhimurium* [104], he proposed that the average cell mass per *oriC* site at inititation (called the initiation mass) was constant and did not vary with growth rate.

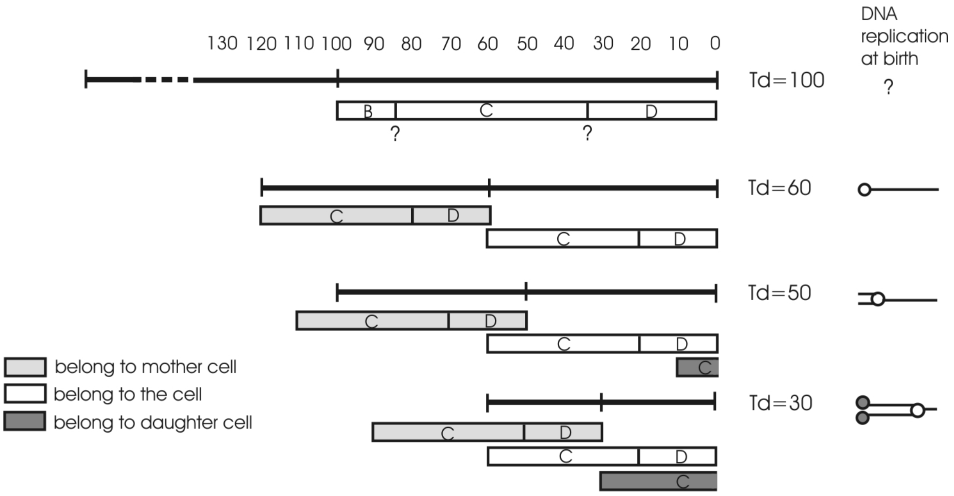

**Figure 1.4. DNA replication patterns for *E.coli* with different growth rates.**
As interdivision time (Td) decreases, the start of DNA replication occurs earlier. According to the Cooper-Helmstetter model of DNA replication, the time of DNA replication initiation is determined from the division time of the cell of interest (t=0). A D period of 20 minutes and a C period of 40 minutes are aligned from the end of each interdivision time. The start of the C period corresponds to the start of DNA replication. Each thin black bar represents two interdivision times, the first belongs to the mother cell and the second belongs to the cell of interest. The number of active replication rounds corresponds to the number of C periods. For simplicity, DNA replication at birth is depicted as unidirectional instead of bidirectional, and the chromosome as linear instead of circular. The replicated part of the chromosome appears as double lines, while the non-replicated part appears as a single line. The replication forks are depicted as circles with colors corresponding to the legend.

Experiments further validating the concept of initiation mass were performed following the introduction of the hypothesis. In one investigation, cells with abnormally high mass per origin started DNA replication simultaneously at all origins. First, DNA synthesis was inhibited in a thymine-requiring *E.coli* strain by thymine starvation. During this period of time, cell mass increased normally, resulting in a greater mass per origin than the normally required initiation mass for all origins in the culture. The starvation was then relieved and all origins were observed to initiate immediately. After approximately one interdivision time, the cell origin to mass ratio returned to the prestarvation level [34]. In another investigation, DNA replication rates were varied and the measured initiation mass was highly constant. When a thymine-requiring strain of *E.coli* was grown with less than the optimum thymine concentrations, the rate of DNA synthesis and thus the C period varied with thymine concentrations. It was shown that despite large variations in the DNA replication rate, the measured initiation mass was the same [97].

The two experiments mentioned above are among the earliest investigations. They strongly suggest that there is a parallel accumulation of the cell initiation potential with cell mass, and that it is independent of the actual replication process. This leads to the long-held belief that cell mass, or something that correlates with it, is the initiator of DNA replication. A few later experiments found that initiation mass does change with growth rate. One experiment showed that initiation mass increases with increasing growth rate [26], while others found that it decreases[120]. These results indicate that initiation mass might vary, and that is important to seek the molecular entity that determines the initiation of DNA replication.

Models proposed for regulation of initiation can be classified as based on either a positive or negative regulation. It was first proposed by Cooper and Helmstetter [27] that DNA replication initiation may depend on the amount of an initiator protein which is produced continuously during the cell cycle. When a critical amount has been reached (proportional to cell mass or size), DNA replication commences (positive regulation). In order to ensure that initiation only occurs once, the hypothetical initiator is either made periodically or consumed immediately. Alternatively, a model based on negative regulation was proposed by Pritchard, Barth and Collins [96]. They proposed that there could be a burst in synthesis of an inhibitor protein, and that DNA replication occurs when the concentration of the inhibitor decreases to sub-inhibitory level as the cell volume increases during growth. To further explain the high precision and synchrony of replication initiation, both models suggested cooperation between multiple initiators/inhibitors. A more detailed description on the molecular mechanism of replication initiation is covered in section 1.6.

## 1.5 Methods of C and D Period Determination

To measure the C period, two major types of cell cycle analyses are commonly used: integral methods, which measure accumulation of a certain compound (in this case DNA), or differential methods, which measure the synthesis rate of a particular compound. Generally, differential methods are more successful, due to higher accuracy. Upon examination of the accumulated compound, the difference between e.g. linear and exponential accumulation could be very small. While upon comparison of the synthesis rates, the difference between a constant and exponentially increasing synthesis rate is more readily interpretable.

Generally, these methods can estimate the mean length of the C and D period quite precisely, but only a few techniques provide an estimate of the variance of these periods. The first technique is flow cytometry. By fitting theoretical curves to experimental data, both the mean and the degree of variability of B, C and D periods could be estimated. The drawback of this method is that the assumptions underlying the theoretical model could influence the data interpretation considerably [83]. Another technique used to estimate the variance of the C period is autoradiography. By measuring variability in lengths of labeled DNA strands for a given pulse duration, the maximum spread in a C period was found to be about 33% [71]. However, inherent errors exist in this method, DNA tends to shrink in quite a variable way during fixation and lengths of the tracks differ by at least a factor of two in any preparation. An alternative technique to estimate the variance of the D period combines DNA synthesis rate measurements in synchronous culture and modeling [18]. However, artifacts in the membrane elution technique could affect the data interpretation. Moreover, whether or not the assumptions incorporated in the model are also valid for slow growing cells is still unclear. So far, direct observations of these periods and their variances do not exist. Below, a summary of different methods to measure the C and D periods are described.

### 1.5.1 Methods for measuring the C period

1. Increment of DNA synthesis [16,78].

This technique is based on the assumption that upon inhibition of protein or RNA synthesis, replication initiation is prevented while elongation proceeds normally. This assumption is based on an observation that when thymine-requiring bacteria are starved for thymine, cells lose their viability in a characteristic manner. DNA still increases by about 40% or more, and during this DNA synthesis period a fraction of the population becomes immune to thymine-less death. The difference in accumulated DNA before and after protein synthesis inhibition *($\Delta G_t$)* is measured and supposed to be a function of only the C period *($C$)* and cell interdivision time *($T_d$)* according to the relation:

$$\Delta G_t = \frac{G_t}{G_o} = \frac{C(\ln 2)}{T_d}(1-2^{-C/T_d})^{-1}$$

Thus the C period can be estimated.

2. Rate stimulation [128].

This technique is based on the observation that when an exponentially growing thymine-requiring culture is starved for thymine, the rate of DNA synthesis (thymine incorporation rate) upon thymine restoration is faster than before starvation. The degree of this increase is proportional to the length of the starvation time, up to a maximum value reached after one mass-doubling time. With a model that accounts for cell cycle parameters, DNA replication parameters, and culture age distribution, it was shown that the ratio of DNA synthesis rate after and before starvation *(R)* is a function of the mean interdivision time $(T_d)$ and the C period *(C)*:

$$R = (2^{C/Td+1} - 1) / (2^{C/Td} - 1)$$

Therefore, by measuring this DNA synthesis rate (for example by a [$^{14}$C]thymine incorporation technique) and the culture mean interdivision time (by optical density measurements) before and after starvation, the C period can be estimated.

3. DNA-to-mass ratios [104].

Schaechter, Maaloe and Kjeldgaard first showed that mass increases exponentially as a function of growth rate, as in the case of DNA synthesis, though with lower exponent. Thus, as the growth rate increases, the amount of DNA per unit mass (DNA concentration) decreases. Cell mass is measured by the optical density at 450 µm while DNA content is determined by colorimetric technique. Interestingly, the same data showed that cell mass at initiation is constant over various growth rates. Experimental data for different cellular growth rates fit a model of DNA concentration as a function of only the C period, cell interdivision time and initiation mass. If the initiation mass is a constant, then the DNA-to-mass ratio at one growth rate can be used to predict the DNA-to-mass ratio at other growth rates. This would confirm the constancy of C. Alternatively, if the value of C and the DNA-to-mass ratio for one growth rate is known, then the C period for other growth rates can be calculated.

4. Autoradiographic analysis of DNA chain extension [22].

The basic idea of this technique is to label DNA by exposing a cell to radioactive thymidine for a short time. The cell is then gently lysed, the DNA is spread out on a slide, and subsequently processed for autoradiography. Short sequences of radioactive segments are measured, and with the knowledge of the previous labeling time, the DNA synthesis rate of a replication fork can be obtained. The size of the entire genome can also be estimated so that an approximation of the C period can be calculated.

5. Analysis of gene frequency [24].

The frequency of a certain gene depends on the number of replication points in the chromosome and its relative position from the origin of replication. In this method, two different chromosomal markers at different positions are measured for their frequency ratio (by DNA:DNA hybridization). This ratio *(a/b)* is then investigated as a function of growth rate by examining cells grown in various media. It is found that the shorter the interdivision time, the higher the frequency difference. A mathematical relationship that describes the frequency ratio as a function of the culture's interdivision time *($T_d$)*, mean replication time *(C)*, and distance between the two genes *($\Delta d$)* are subsequently used to estimate the C period:

$$a / b = 2^{C * \Delta d / Td}$$

6. The backward membrane elution method [17].

Also called the baby-machine, exponentially grown cells are pulse-labeled (for example with [$^{14}$C]thymidine for DNA labelling), then filtered over a nitrocellulose membrane, and washed with fresh, prewarmed medium to remove excess label. Subsequently, the membrane is inverted, and fresh medium is pumped through the membrane so that bound cells can continue to grow and divide. During the elution, newborn cells are released into the medium (synchronized cells). These newborn cells are daughter cells from the oldest (closest to division) membrane bound cells. In a steady state exponential culture, the number of newborn cells is always twice as much as the number of dividing cells, hence the number of eluted cells over time resembles a saw tooth pattern.

The radioactivity patterns per cell during elution reflect the "backward" synthesis rate of the compound of interests (in this case DNA). For example, when slow growing cells with interdivision time longer than C+D are pulse labeled, the label will only be incorporated in the cells during DNA replication, or in other words in their C period. These cells will give rise to radioactive eluted daughter cells that would elute between D and D+C minutes, and thus the length of the C and D periods can be estimated. By investigating DNA synthesis patterns of cells with various interdivision times, a model of the DNA replication pattern can be developed.

7. Sucrose synchronized cultures [17].

In this method, cells are separated according to size by putting a concentrated cell sample on top of a linear sucrose gradient and then centrifugating them. Cells within a narrow size class, are selected to produce synchronized culture and the DNA synthesis pattern is subsequently analyzed. The smallest class of cells is assumed to be newborn cells, although the smallest cells can also originate from sick and abnormally growing cells.

8. Synchronization of DNA replication [74].

First, all ongoing replication in the cells are allowed to finish without allowing any increase in cell mass by amino acid starvation. Then the cells were starved of thymine in the presence of amino acids to induce synchronous DNA synthesis initiation. Samples were exposed to radioactive thymine at intervals and the replicated portions of the genome were determined by DNA-DNA hybridization.

The C period was found to be 40 minutes in rich medium and 52 minutes in glucose minimum medium.

9. Transductional analysis [46].

A different approach to determine the C period is used in this method. *E.coli* cells are transducted with a P1 bacteriophage and then samples are taken at discrete timesteps. Each sample is plated on different selective plates to follow the appearance of transductants with genetic markers positioned at various points on the chromosome. Apparently, the closer a certain marker is to the *ter* site, the earlier its corresponding transductants increase in numbers. The time interval between the appearances of two different transductants ($\Delta T$) is a simple function of their markers chromosomal positions ($d_a$ and $d_b$) and the DNA replication rate, whereby the C period ($C$) can be easily deduced:

$$\Delta T = \frac{d_a \cdot C}{50} - \frac{d_b \cdot C}{50}$$

10. DNA steptime measurement [80].

DNA steptime is the average time for adding one more nucleotide to a growing DNA chain. This technique takes advantage of DNA chain growth in the 5’ to 3’direction. When DNA is hydrolyzed, nucleotides at the end of a replicating chain stay in the form of 3’-OH while interior nucleotides yield 3’-monophosphate. By incorporating a radioactive label into the cell’s DNA (labeling time 6-18 seconds), the ratio of internal to end label in the DNA can be determined. By developing a mathematical model that considers the relationship between measured radioactivity values, labeling time and enzyme kinetics, the DNA steptime could be estimated. The result shows a chain growth rate of a predicted 500-700 nucleotides per second. This result combined with the length of the whole chromosome results in a very close agreement with other measured values of C.

11. Flow cytometry [108].

This method allows single-cell measurement of fluorescently labeled DNA content in individual bacteria in an exponentially growing culture by observing the cells pass over an optical system that measures cell fluorescence. Then, the experimental distribution of the cell’s DNA content is fit with a theoretical distribution that considers variation in the cell population’s age, interdivision time, DNA replication initiation with methodological variation. The average B, C and D periods of the population are inferred from the best theoretical fit. However, assumptions underlying the model could influence the obtained values.

### 1.5.2 Methods for measuring the D period

1. Residual cell division after inhibition of DNA synthesis [47].

This method assumes that except by lack of termination, inhibition of DNA replication does not disturb cell division. The D period is estimated from the residual division time. DNA replication in membrane bound cells is inhibited by either UV radiation, mitomycin addition, or thymine starvation. After 20 minutes of normal increase in the eluted cell number, there is a dramatic decrease in the eluted cell number. The data is interpreted as the amount of time needed for cells that have terminated their DNA replication to divide. This value of the D period was found to be constant and independent of interdivision time.

2. Residual cell division after inhibition of protein synthesis [67].

This particular method of measuring D period is quite simple and requires only the ability to count cells. Based on the proposal that the replication of the terminal part of chromosome requires the ability to synthesize protein [81], it is believed that cells that divided after the addition of chloramphenicol are cells that have finished DNA replication. Chloramphenicol is added to the medium to inhibit protein synthesis. Thus, by measuring the increase in cell number after chloramphenicol addition (the residual cell division) and growth rate, one can infer the length of the D period. The main critique on this method is the usage of chloramphenicol, which might have side effects on the cell division process in general (not D period related). The average D periods for exponentially growing cultures obtained with this method is constant, and does not vary with a vast range of growth rates.

## 1.6 DNA replication initiation

Although the model for the regulation of DNA replication initiation is still under debate, there are currently more findings supporting the positive regulation model. A protein that is specifically involved in the initiation process, DnaA, is commonly believed to be the initiator protein. Before DNA can be replicated, several discrete steps have to occur. First, the origin recognition complex (ORC) is formed at *oriC* by the intitiator DnaA protein, in either its ATP or ADP bound form. DnaA protein binds to three 9 bps binding sites within *oriC* called R1, R2 and R4. Then, the weaker recognition sites R3 and R5 are occupied in addition to the three I-boxes specific for DnaA-ATP (for reviews: [57,86,89]).

Next, accessory proteins such as IHF, HU and DiaA induce conformational changes on the DNA structure [62,101], where it becomes wrapped around the right handed DnaA-ATP filament. This complex promotes double strand separation at adjacent AT rich regions. The separation is stabilized by extra DnaA-ATP protein binding to specific 6 bps sequences in the single stranded region. Subsequently, the 'licensing' for replication starts with the recruitment of the DnaB helicase by DnaC-ATP in the form of a B6C6 complex [6] to further open the unwound DNA. DnaC is released after hydrolysis of its ATP. The whole complex is now called the pre-replicative complex (pre-RC).

DNA replication starts when two or three DNA polymerase III holoenzymes are loaded on the *oriC*, and when dNTP are available, the replication begins. Once replication begins, three processes occur to prevent extra-unwanted replication initiation. First, the conversion of DnaA–ATP into the inactive DnaA-ATP form by the complex of an Hda protein and β–clamp of the DNA polymerase III holoenzyme [60]. Second, *oriC* is sequestered by SeqA protein. Once replication has started, hemimethylated regions are produced. The hemimethylated *oriC* is a target of sequestering by the SeqA protein (for a more elaborate explanation refer to section 1.7) rendering it inaccessible to the DnaA protein. In addition, the promoter of the *dnaA* gene which is close to the *oriC* is sequestered at the same time, preventing transcription and DnaA protein synthesis [23]. Third, *in vivo* DnaA protein is titrated by the newly duplicated reservoir sites during the completion of DNA replication. There are 308 evenly distributed R-type DnaA boxes with different affinities along the *E.coli* chromosome. All three mechanisms mentioned above work together to lower the replication initiation potential of the cell, highly regulating the frequency of chromosome replication. Nonetheless, the exact timing of DNA replication initiation still varies from cell to cell. The individual DnaA protein concentration (in its ATP and ADP forms) plus other components of the complex in a cell could be responsible for this timing "noise".

## 1.7 DNA sequestering protein SeqA

Initiation of the DNA replication process in *E.coli* has to be well regulated to avoid multiple replication initiations in an untimely manner. One protein that is involved in negative regulation of DNA replication initiation is the DNA sequestering protein SeqA. In *E.coli*, when fully methylated parental DNA is replicated, the resulting daughter strand is hemimethylated until it is methylated by Dam methyltransferase. By binding to the hemimethylated, GATC-rich origin of replication, the DNA sequestering protein SeqA prevents re-initiation of DNA replication [13,58,76,82].

SeqA forms a homotetramer from two interacting dimers to bind stably to the hemimethylated DNA. One SeqA dimer binds to a pair of hemimethylated GATC sites separated by a maximum distance of about 31 bps. There are 1750 pairs of

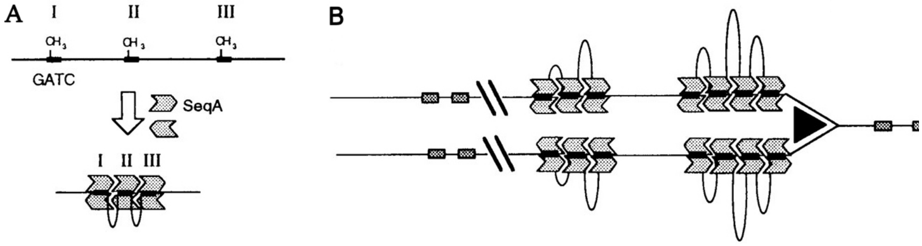

**Figure 1.5 Proposed model of SeqA cluster (adapted from [19]).**

(A) Hemimethylated GATC sites are targets for SeqA protein.

(B) DNA bound SeqA proteins in close proximity are able to form filaments and clusters.

GATC sites with spacings up to 31 bps rather evenly distributed along the length of the whole chromosome [19]. A recent study has shown that SeqA dimers are able to oligomerize into a filament, and mutations that disrupt filament formation lead to asynchronous DNA replication [44]. SeqA is also known to bind double-stranded methylated DNA, but with much less affinity. A model of DNA sequestering by SeqA was proposed where SeqA arranges itself into clusters along the DNA so that the DNA in between GATC sites forms loops (fig. 1.5).

*In vivo*, fluorescently labeled SeqA clusters are visible as discrete foci [51]. These foci are thought to be SeqA clusters on DNA plus aggregate of free SeqA proteins [59]. The number of these foci increases as the interdivision time decreases. Recently, a ChIP assay has shown that the SeqA protein binds sequentially to hemimethylated nascent DNA segments following replication fork movement in synchronized cultures of *E.coli* [125]. Therefore, the SeqA foci are assumed to represent the replication forks [85]. In this thesis the dynamics of SeqA foci *in vivo* is studied in slow growing cells.

## 1.8 This thesis

The outline of this thesis is presented below. The *E.coli fim* system, its cell cycle and their interconnection were studied at the single cell level using timelapse microscopy. The content of this thesis is organized as follows:

Chapter 1 introduces all subjects relevant for later chapters and some background information.

Chapter 2 covers experimental methods and materials involved in chapter 3, 4 and 5.

Chapter 3 contains the results of our single-cell level investigation into the *E.coli* cell cycle and DNA replication with a fluorescently labeled SeqA protein.
Chapter 4 contains the results of our investigation of cell-to-cell behavioral diversity within the *fim* system of *E.coli* and its relation to the cell cycle.
Chapter 5 contains the results of our investigation on the effect of switching history on *fim* OFF switching behavior.

# Experimental Materials and Methods

# 2

To observe *fim* switching events at the single cell level, an allelic exchange method was used to place GFP*mut2* under the control of the native *fimS*. In addition, the SeqA protein was labeled by fusing it with fluorescent protein mCherry to follow the progress of the replication process. The states of *fimS* over time were monitored by growing single cells into microcolonies on various agar media and subsequently by performing automatic timelapse microscopy. Imaging was carried out using both phase contrast and fluorescence microscopy. Movies were then analyzed with segmentation software to extract various cell parameters and to construct family trees.

## 2.1 Genetic Manipulation

All genetic manipulation was performed according to general microbiology protocols [103]. A more extended explanation for GFP insertion into the chromosome by means of allelic exchange is explained below.

### 2.1.1 Allelic Exchange

To differentiate between ON and OFF states of *fimS*, a method called allelic exchange [12] was used to insert the fast folding *GFPmut2* into the HpaI site of chromosomal *fimA* under the control of the native *E.coli* K-12 *fimS*. The main steps involved were two homologous recombinations, one of plasmid integration into the chromosome of an intermediate strain, and the other of plasmid excision and curing (fig. 2.1).

In this method, an intermediate strain derived from wild type *E.coli* MG1655, in which the destined locus was replaced with a *sacB-kan* cassette (strain Δ3'*fimE-5'fimA* ΩSacB-Kan$^R$), was transformed with a temperature-sensitive plasmid containing homologous fragments of the destined locus (*fimE–fimD*) and GFPmut2 downstream of *fimS*. The plasmid was also carrying a chloramphenicol resistance gene. By growing the transformants at the non-permissive temperature for plasmid replication (42°C) with chloramphenicol, the plasmid was selected for integration into the chromosome.

Once integrated, a second recombination for plasmid excision was selected. This second recombination could result in either a successful allelic exchange or an abortive one yielding the original strain. The excised plasmid was then cured from both strains by growing the cells in medium without antibiotics at 42°C. The cells that had successful allelic exchange were then selected by growing the cells at 28°C on medium with sucrose (without sodium chloride). Expression of *sacB* under these conditions is lethal for the original, untransformed strain. The resulting cells were then plated and re-checked for newly acquired susceptibility for chloramphenicol and kanamycin.

### 2.1.2 Construction of a mCherry-SeqA fusion

To follow progress of the replication fork *in vivo*, we fluorescently labeled the SeqA protein with mCherry (a generous gift from R. Tsien). We cloned the SeqA gene from pGAP40 (a generous gift from T. Brendler) into the multiple cloning sites in pSAV047 (kindly provided by S. Alexeeva), in between BsrGI and EcoRI

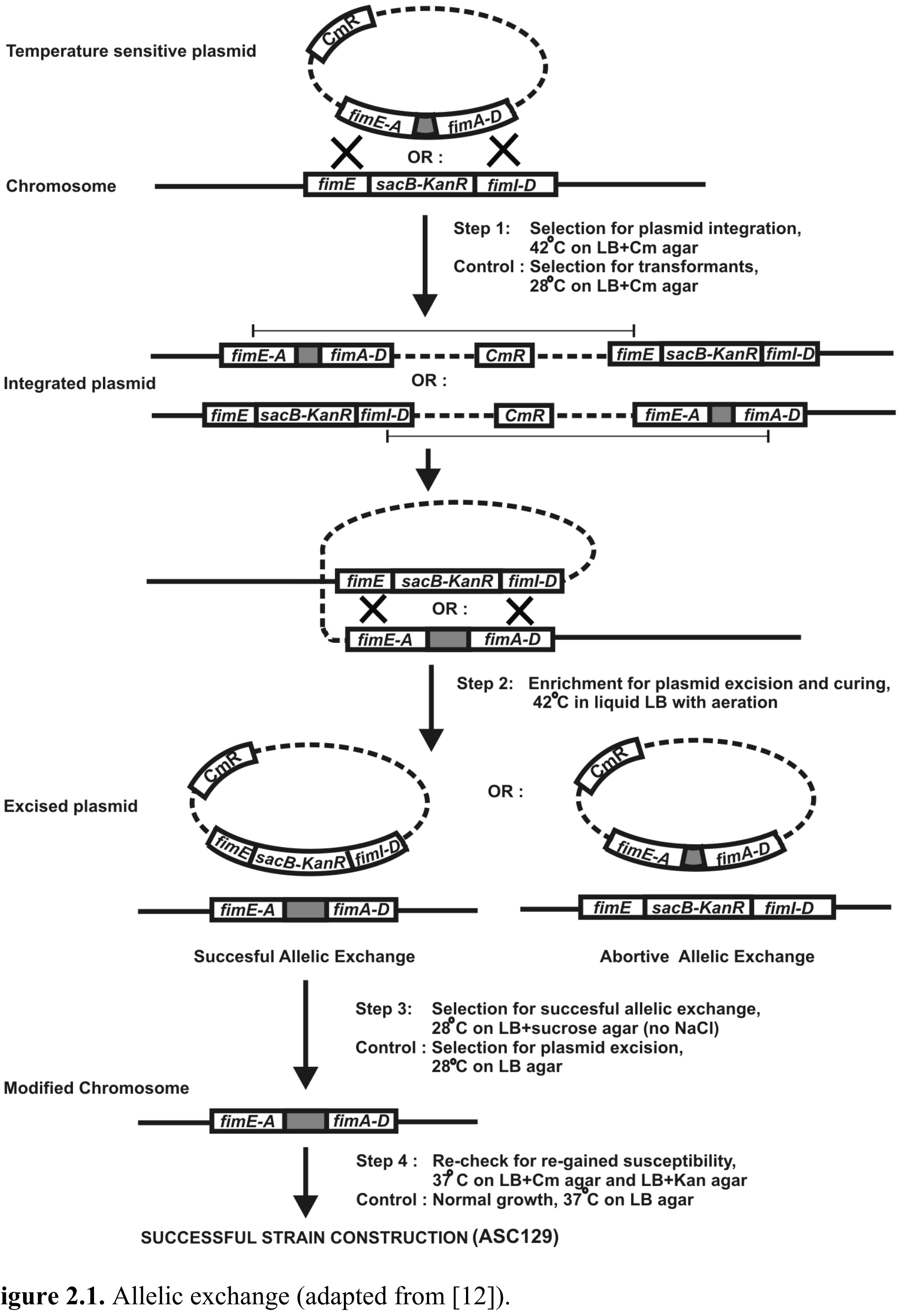
Temperature sensitive plasmid
CmR
fimE-A
fimA-D
OR :
Chromosome
fimE
sacB-KanR
fimI-D
Step 1: Selection for plasmid integration, 42°C on LB+Cm agar
Control : Selection for transformants, 28°C on LB+Cm agar
Integrated plasmid
OR :
Step 2: Enrichment for plasmid excision and curing, 42°C in liquid LB with aeration
Excised plasmid
Succesful Allelic Exchange
Abortive Allelic Exchange
Step 3: Selection for succesful allelic exchange, 28°C on LB+sucrose agar (no NaCl)
Control : Selection for plasmid excision, 28°C on LB agar
Modified Chromosome
Step 4 : Re-check for re-gained susceptibility, 37°C on LB+Cm agar and LB+Kan agar
Control : Normal growth, 37°C on LB agar
SUCCESSFUL STRAIN CONSTRUCTION (ASC129)

**Figure 2.1.** Allelic exchange (adapted from [12]).
The grey colored block represents the GFP*mut2* gene. The crosses represent homologous recombinations. The resulting strain is ASC 129.

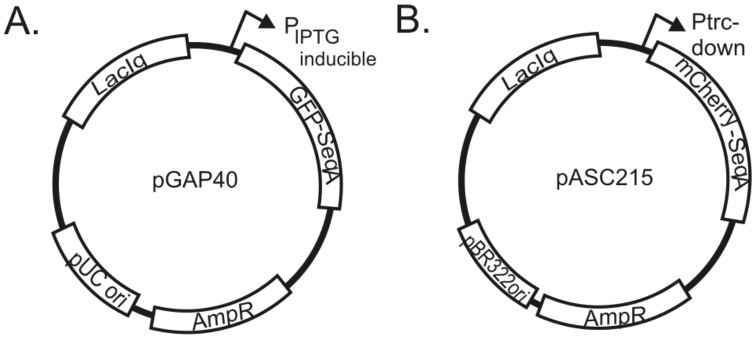

**Figure 2.2** Construction of plasmid based mCherry-SeqA fusion
(A) Plasmid pGAP40 containing the fusion of GFP and SeqA protein.
(B) Plasmid pASC215 containing the fusion of mCherry and SeqA protein.

sites, to place it under the control of the IPTG inducible promoter and fuse it with mCherry (fig. 2.2 B). Five amino acids were acting as a linker (Glu-Phe-Asn-Asn-Asn). The resulting plasmid (pASC215) was then transformed into strain ASC129 resulting in strain ASC215. Upon imaging, we were able to observe our red fluorescently labeled SeqA as discrete foci.

In order to investigate the possible adverse effects of SeqA-mCherry expression, we characterized its effect on growth rate and cell length. High overexpression of SeqA had previously been reported to result in filamentous cells [5], consistent with its role as a negative regulator for DNA replication initiation. The fusion expression was therefore kept at a minimal level, by omitting the inducer IPTG. We note that the fusion was expressed in addition to the endogenous unlabelled SeqA, which was expressed at wild-type levels in both cells. Expression leakage from the promoter appeared sufficient to produce visible SeqA foci. The interdivision time of the rapidly growing cells appeared similar with (Luria Bertani medium, Td=29.5±1 min) and without (Td=29.9±0.8 min) expression of the SeqA-mCherry fusion (fig. 2.3). Histograms of cell lengths were similar for both type of cells, as determined by a Kolmogorov-Smirnov test (D=0.068, P=0.39) (fig.2.4). For both strains, the length distributions were skewed as expected for exponentially growing culture, which is composed more of younger cells than older ones [69]. Filamentous cells were only rarely observed (less than 1 in 100). These results indicate no significant adverse effect of SeqA labeling on cell growth.

### 2.1.3 Strains and plasmids

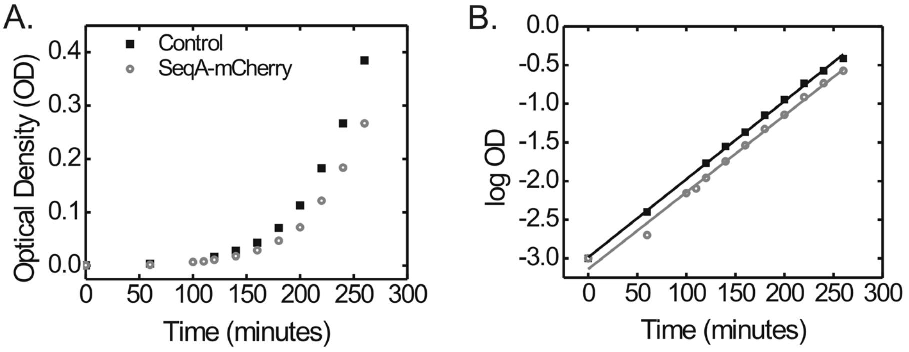

**Figure 2.3** Growth rate determinations
(A) Optical density at 600nm of control cells (ASC129, carrying no plasmid) and cells carrying expressing SeqA-mCherry (ASC215) in LB Lennox medium at 37°C.
(B) Optical density data on a log scale (A) fitted with linear regression lines. The interdivision time for ASC129 is 29.9 minutes and ASC215 is 29.5 minutes. The symbols are the same as image A.

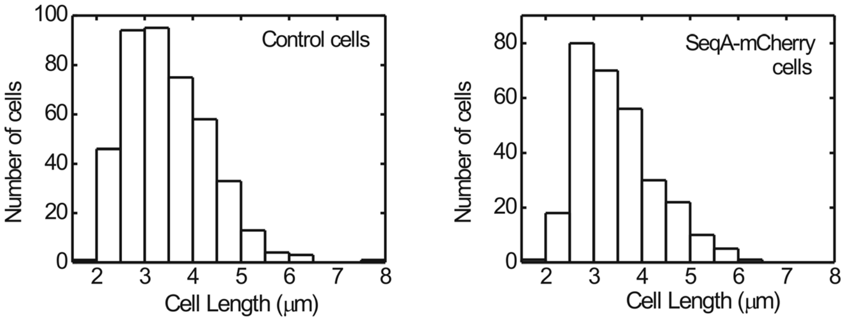

**Figure 2.4** Histogram of cell lengths.
Length data of control cells expressing no SeqA-mCherry (ASC129, left panel) and cells expressing low amount of SeqA-mCherry protein (ASC215, right panel) grown exponentially in MOPS succinate medium. The data show similar cell length distibution. A Kolmogorov-Smirnov test indicated that there was no significant difference between the two data sets (D=0.068, P=0.39).

The bacterial strains used in this study are derivatives of *E.coli* K-12. An intermediate strain AAEC112 (Δ*3'fimE-5'fimA* ΩSacB-Kan$^R$), plasmid pIB315 (Cm$^R$, temperature sensitive) containing *fimE-fimD* genes and *E.coli* strain strain MG1655 were generous gifts from I.Blomfield [11]. *GFPmut2* [30] as a reporter for the Fim switch was cloned into the HpaI site of the *fimA* gene within pIB315. By means of allelic exchange (as described in section 2.1.1) the *fimA* gene with *gfpmut2* was inserted within the original chromosomal wild-type *fimA* site, resulting in strain ASC129 (*fimA*::*gfpmut2*).

Into vector pSAV047 (Amp$^R$, pBR322 ori; *rop*-; a generous gift from S.Alexeeva), seqA gene from pGAP40 (kindly provided by T.Brendler [19]) was fused to the N-terminus of mCherry fluorescent protein under the control of a repressed trc-down promoter resulting in plasmid pASC215. Strain ASC215 is ASC129 transformed with pASC215. Strain ASC181 is ASC129 transformed with pGAP40.

### 2.1.4 Materials

The medium used for cloning and allelic exchange included LB Lennox broth (Difco), LB Lennox agar (Difco) and sucrose agar. Sucrose agar is LB agar (5 g of sodium chloride, 5 g of yeast extract and 10 g of tryptone per liter without sodium chloride and supplemented with 6% sucrose (Sigma)). Whenever necessary chloramphenicol or kanamycin, or ampicilin was supplemented to the medium to a final concentration of 30, 25 and 100µg/mL. Cells were grown at 28, 37 and 42ºC accordingly.

Restriction enzymes and polymerase were purchased from New England Biolabs, and DNA ladders were purchased from Fermentas and Promega.

## 2.2 Microscopy

### 2.2.1 Sample preparation

We modified a method previously developed by E. Stewart [113] to grow cells on agar slab. Before imaging, cells were grown overnight in the appropriate medium with the addition of antibiotics when necessary. The next day, the overnight culture was diluted with pre-warmed medium and grown for another 4 hours to be in the exponential phase (OD between 0.05 and 0.1). When a rich medium was used, the dilution was done twice, each 1000X and grown for 2 hours. When a minimum medium was used, the dilution was done once and such that the OD of the culture did not exceed 0.1 in 4 hours of growth.

Two microscope slides, one of which had a hole of size 18 mm x 55 mm, were attached to each other by use of silicon grease (Dow Corning) and heated to 80ºC. Next, 10X concentrated medium of interest was mixed with hot agaroseMP solution to its final concentration (total volume is 1 mL) and spread horizontally across the center of the microscope slides cavity (fig. 2.5). Immediately afterwards, the microscope slides were transferred to a cool flat surface, and covered with a 24

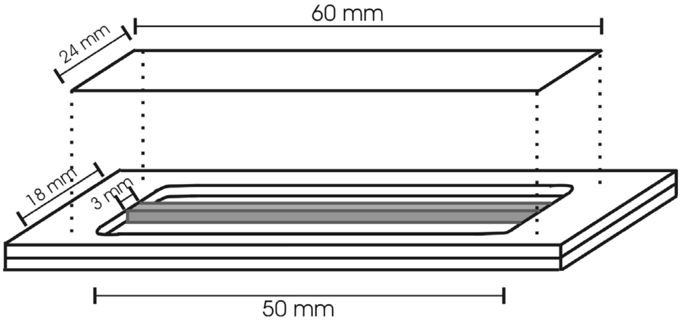

**Figure 2.5** Sample holder for timelapse microscopy**.** The agar strip is shown in grey.

mm x 60 mm silanized coverslip (a silanized coverslip was made by dipping normal coverlips into RepelSilane, and then left to dry at room temperature) by dropping it right on top of the liquid agar so that the agar surface would become flat upon cooling. The agar solidified at room temperature within 2-3 minutes.

When the color of the agar had become less transparent, the slides were incubated in the microscope's 37ºC incubation chamber to avoid any cold shock to the cells when they were deposited on the agar surface. Inside the incubation chamber, the silanized coverslip was carefully removed, revealing a thin, flat agar slab. Excess agar slab was trimmed to leave a thin strip of agar of dimension 2.5 mm x 55 mm across the center of the cavity. The rest of the cavity served as an air reservoir for cell growth. Using a pre-warmed pipette tip, 2-3 tiny droplets of exponential liquid culture were dropped on top of the agar strip. The sample was then closed immediately by dropping a pre-warmed normal coverslip on top of the cavity and the gap between the coverslip and the microscope slide was sealed with silicon grease. The cells could then spread evenly across the agar strip and become immobile. Over time, each single cell grew into microcolonies.

### 2.2.2 Microscopy

Imaging was performed with a Nikon Eclipse TE2000 inverted microscope equipped with a 37ºC incubation chamber, automated stage (Marszhauser) and CCD camera (Coolsnap, Roper Scientific). Phase contrast and green fluorescence images were taken automatically by imaging software Metamorph (RoperScientific) with a 100X magnification objective lens. For SeqA experiments, an additional 1.5X magnification was used. Light from a xenon lamp (Lambda LS) was filtered by a GFP filter (Nikon) and an HCRed filter (Chroma) for GFP and mCherry detection. mCherry images were taken every timestep for SeqA experiments and every other timestep for *fim* experiments. There were two kinds of time steps used, 4 minutes (chapter 3 and 4) and 10 minutes (chapter 5). Depending on the growth medium and the experimental purpose, typical experimental times ranged from 4 to 16 hours.

### 2.2.3 Materials

For microscopy experiments, cells were grown in EZ Defined Rich medium (Teknova) or MOPS medium (Teknova) supplemented with 0.4% sodium succinate (Sigma) or 0.4% glucose (Teknova) and 0.8 mM L-leucine (Merck). AgaroseMP was purchased from Roche. Whenever SeqA was visualized, 100µg/mL ampicilin (Sigma Aldrich) was added to the medium. For cell immobilization and microcolony growth, MOPS agar was made by mixing 10X concentrated medium (Teknova) with hot 2% agaroseMP solution to a final volume of 1 mL. Cells were grown at 37ºC. All liquid cultures were grown aerobically. To make silanized coverslips, Repel Silane was purchased from Amersham Biosciences.

### 2.2.4 Data analysis

Simple image manipulation and viewing was done with Metamorph, ImageJ and IrfanView. Cell fluorescence quantification, cell parameters extraction and lineage tree construction was performed with Schnitzcell, software written in MATLAB kindly provided by M.Elowitz [100]. First, cells from the phase contrast images were segmented automatically, in which the border of each cell was determined and dividing cells were separated accordingly (fig. 2.6). The segmentation result was then manually checked. Next, frame-by-frame, each cell was tracked through the following frames to establish familial connections. The resulting cell lineage connections were then manually checked. Finally, the SeqA foci positions were indicated manually when available and then all parameters for each cell were extracted from both the phase contrast and fluorescence images. Those included cell size, age, background corrected mean fluorescence (cell's fluorescence divided by its pixels area), relatives, etc. Family trees and growth profiles of each cell could then be plotted. An example of an analyzed image of a growing ASC129 microcolony on LB agar is shown in figure 2.5. The mean interdivision time is about29 minutes, in accordance with our previous growth measurement in liquid culture for the same strain.

Cell growth on the agar slab typically proceeded exponentially for a significant amount of time, and then it decreased over time until an abrupt decrease of growth occurred. The observed brightness of GFP expressing cells also decreased moderately at the end of each experiment. We excluded these data with decreased

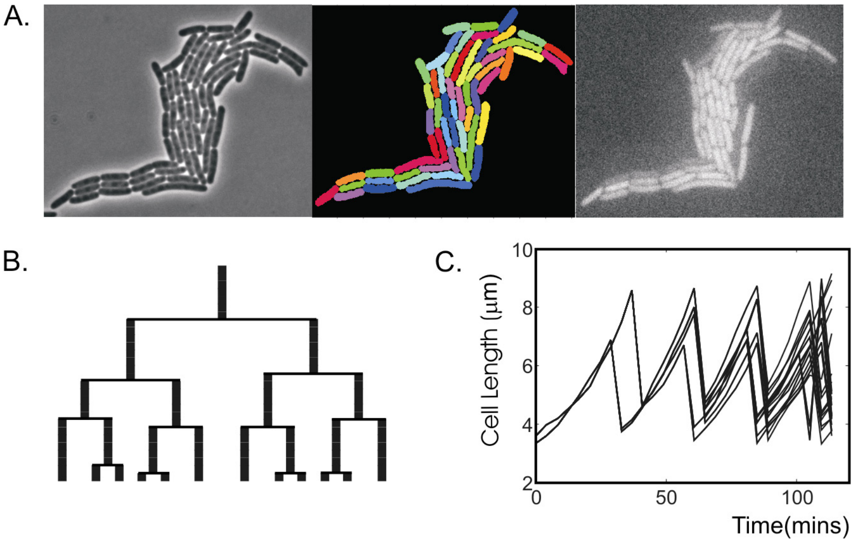

**Figure 2.6** Data analysis
(A) A phase contrast image of a microcolony growing on LB agar slab (left), image where each cell is an individual segment (middle) and its corresponding fluorescence image (right).
(B) A part of family tree of the microcolony in A.
(C) Cell length versus time for cells in figure A showing exponential growth.

growth rate obtained late in the experiments. The time period a certain microcolony stayed in the exponential phase varied from microcolony to microcolony and also varied with colony size and type of medium. Presumably, this was partly due to microenvironment nutrient or oxygen exhaustion and partly due to phototoxicity. Exposure times were chosen to be as short as possible to minimize cell damage, while fluorescence images were binned 2x2 to improve the signal-to-noise ratio and enhance visibility. Generally, a rich medium sustained longer exponential growth as well as a smaller sized microcolony.

Good quality images are beneficial for better success in automatic cell segmentation. The dryness of the agar surface played an important role in image quality. In some experiments where the agar surface was more slippery the cells were able to spread better, thus preventing multiple layered cells to occur earlier.

Upon growth of cells with a labeled *fim* switch (ASC129) in LB medium, we observed a sudden increase in background fluorescence in large microcolonies. Interestingly, stationary cells from an overnight LB culture were also very bright.

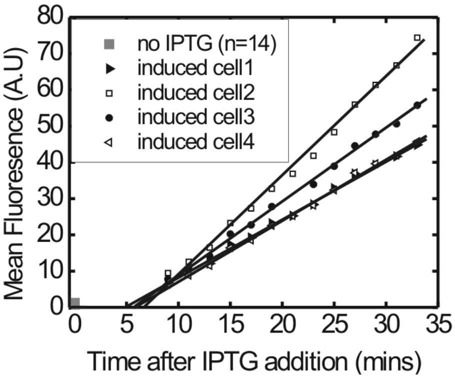

**Figure 2.7** Determination of GFP*mut2* maturation time.
The mean fluorescence levels per unit area were followed in time after IPTG addition (t=0). The gray square at t=0 represents the average background fluorescence level for non-induced cells ($9.1 \pm 1.7 \cdot 10^{-3}$ a.u). The straight lines are linear regression lines over experimental data (symbols). The average GFP maturation time in MG1655 is 5.7 ±0.7 minutes.

We propose that autofluoresence is a sign of a cell entering a stationary phase. This observation is only observed in LB medium, which has its own independent autofluoresence properties.

## 2.3 Determination of GFP*mut2* maturation time

To be able to optimally pinpoint the moment of switching, a fast folding version of GFP is used to label the state of *fimS*. GFPmut2 is a triple mutant (S65A, V68L, S72A) of the wild type *Aequorea Victoria* GFP, and is about 100 times brighter [30]. We investigated its maturation time in our strain MG1655, by performing IPTG induction experiment with a plasmid containing the GFPmut2 under the control of a LacI repressor (pTHV048, a generous gift from Tanneke den Blaauwen, UVA).

Cells containing pTHV038 plasmid (ASC401) were grown overnight in MOPS medium supplemented with 0.4% succinate at 37°C and then diluted 400 times the next morning to promote exponential growth. After about 4 hours of growth, IPTG was added to the medium (end concentration of 1 mM) and immediately spread on agar slab containing the same medium and IPTG. Microscopy images of cells were taken over time (exposure time 100ms for both fluorescence and phase contrast).

As a measure of background fluorescence level, non-induced control cells from the same culture were also imaged.

During the time necessary for microscope and sample preparations directly prior to imaging, the induced cells became much brighter than the control cells (fig. 2.7). On average, we found that increase in brightness started 5,7 minutes after IPTG was added, indicating a fast GFP maturation time.

# SeqA Foci Dynamics in Slow Growing Cells

# 3

The study of the cell cycle has a long and rich history. By combining results from separate cell cycle and DNA replication investigations through modelling efforts, classical studies have managed to greatly advance our understanding of the rules governing both the cell cycle and DNA replication processes. However, separate investigations of the cell cycle and DNA replication may miss correlations between the two processes. In this work we simultaneously follow the cell cycle and the DNA replication processes *in vivo* and in real time by fluorescently labeling the DNA sequestering SeqA protein and performing timelapse microscopy. A regular 0-1-2-1-0 pattern in the number of SeqA foci over time for slowly growing cells is observed, corresponding to an absence of replication at the beginning and end of the cell cycle and bidirectional movement of two replication forks during replication. Interestingly, our study reveals a strong anticorrelation between cell interdivision time and its elongation rate. Moreover, the period between the birth and start of DNA replication (B period) and between the end of replication and division (D period) are found to be highly variable, while the period during DNA replication (C period) is relatively constant, moderately correlated with cell elongation rate. Our results provide support for the critical mass model for DNA replication initiation.

## 3.1 Introduction

The study of the cell cycle dates back to as early as 1932, when movies of budding *Saccharomyces cereviciae* cells were recorded by Bayne-Jones and Adolph, detailing cell size and shape over time [8]. Their work was followed by later investigations of bacterial cell growth using manual timelapse microscopy, while at approximately the same time, various techniques involving radioactive labeling were performed separately to investigate the DNA replication process during the cell cycle. Results from the cell cycle and the DNA replication processes were then connected to each other by modelling to better understand the rules governing each process. One classic example is the concept of critical mass as suggested by Donachie [33]. By combining the observation of Schaehter, Maaloe, Kjeldgaard [104] on the sizes of *Salmonella* that increase exponentially with growth rate and the DNA replication pattern in *E.coli B*/*r* by Cooper and Helmstetter [27], Donachie showed that DNA replication is initiated at certain masses, an integer multiple of a particular critical mass.

By investigating the cell cycle and DNA replication separately, possible correlations between the two processes may be missed. In this work we present novel data obtained by simultaneously and directly following the cell cycle and the DNA replication processes *in vivo* and in real time. We propose that following SeqA foci dynamics over time leads to a unique opportunity to directly follow the DNA replication process, as well as to simultaneously obtain cell cycle parameters for the same cell, and to correlate these. This implies a direct and simultaneous measurement of the B period (the period between birth and the start of DNA replication), the C period (the DNA replication period), and the D period (the period after DNA replication until cell division) of the cell. Moreover, in this way we can also express the two processes mentioned above in the currency of both time and cell size, each of which is important for different cell processes.

Current advances in microbiology and imaging techniques make it possible to directly follow various intracellular processes. In *E.coli*, when fully methylated parental DNA undergoes replication, the resulting daughter strand is hemimethylated and is sequestered by DNA sequestering protein SeqA [82]. *In vivo*, this results in visible foci when SeqA is fluorescently labeled. Here, we label SeqA by fusing it with the fluorescent protein mCherry and quantify the number and position of the resulting foci *in vivo* as an indicator of the DNA replication process. We simultaneously investigate cell growth by phase contrast microscopy.

The study reveals a regular 0-1-2-1-0 pattern in the number of SeqA foci over time for slowly growing cells, which indicates an absence of replication at the start and end of the cell cycle, and bidirectional movement of two replication forks during replication. Furthermore, quite surprisingly, we find a strong anticorrelation between cell interdivision time and its elongation rate, a relationship we believe has never explicitly been shown before. We observe that the B period (in between birth and start of DNA replication) and D period (in between end of replication and division) are highly variable periods, while the C period (during DNA replication) is relatively constant and moderately anticorrelated with cell elongation rate. Our results support the previously suggested critical mass model for DNA replication initiation.

## 3.2 Results

### 3.2.1 SeqA foci in rapid and slow-growing cells

We first set out to determine how growth rate affects the number of SeqA foci. Isogenic cells are grown exponentially either in defined rich medium, resulting in a division time Td of about 24 minutes, or in MOPS minimal medium supplemented with succinate, which results in a Td of about 120 minutes. The cells were then spread onto an agarose slab and imaged with phase contrast and fluorescence microscopy. Distinct SeqA foci are observed in both cases (fig. 3.1). On average 3.5 foci per cell are observed for the fast growing cells, and 1.3 foci for slow growing cells. These observations are consistent with the Cooper-Helmstetter model of DNA replication involving multiple nested replications for fast growing cells [27].

### 3.2.2 Regular pattern of SeqA foci dynamics in slow growing cells

To follow SeqA dynamics in real time, we perform timelapse microscopy as the cells continued to grow and divide. Microcolonies of about 30 cells are formed over a period of about 8-10 hours on MOPS-succinate agar. In total about 150 single-cell trajectories are monitored and analyzed. The majority of the cells (54%) show the same pattern in the number of SeqA foci. The data of the other cells appear irregular and diverse, which may have been due to SeqA foci that cluster or are out of focus, or caused by insufficient mCherry-SeqA expression.

A typical sequence of images is shown in figure 3.2A. The white arrow indicates the cell of which the SeqA pattern is being followed. At birth, no foci are observed. At some point within the cell cycle, one SeqA focus appears. After some time, this focus elongates, and then separates into 2 nearby foci. The distance between the two foci increases over time and then decreases until a single focus is again observed, which subsequently disappeared nearing cell division. Five out of eight

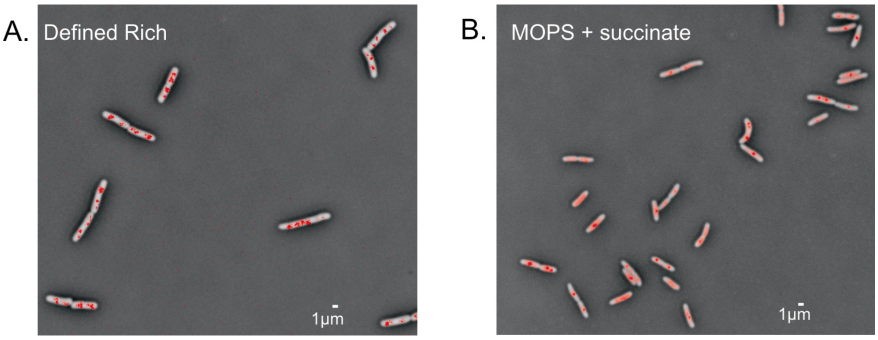

**Figure 3.1** SeqA foci *in vivo*.
Due to a time delay of the methylation process following the replication process, nascent DNA is hemimethylated. DNA sequestering protein SeqA binds to hemimethylated DNA and when fluorescently labelled, SeqA aggregate appears as discrete foci. Images show cells deposited on agar: phase contrast micrographs are in black and white, overlaid with the background corrected mCherry-SeqA fluorescence signal.
(A) Cell previously grown in defined rich medium.
(B) Cell previously grown in MOPS minimum medium supplemented with succinate. More SeqA foci are observed in faster growing cells.

cells in this particular microcolony displayed this 0-1-2-1-0 pattern. Several SeqA foci trajectories covering one cycle are shown in figure 3.2C.

The observed SeqA pattern is consistent with current models of DNA replication and SeqA function in *E.coli*. In *E.coli* DNA replication starts at *oriC*, and therefore a single SeqA foci at the beginning of the cell cycle represents the early phase of DNA replication, when the hemimethylated part of the DNA is near *oriC* (fig. 3.2B). Consequently, the two foci represent the two bidirectional forks moving away from *oriC* during chromosome replication. At the end of the cell cycle the two forks end at the terminus, resulting in only one visible SeqA focus. This model agrees very well with the results of two recent studies, the first being a ChIP assay showing that the SeqA protein binds sequentially to hemimethylated nascent DNA segments following replication fork movement in synchronized cultures of *E.coli* [125]. The second study using immunofluorescence microscopy shows colocalization of labelled SeqA and BrdU (5-bromo-2’-deoxyuridine) foci in slow growing cells [1]. Our results suggest that the ratio of SeqA foci to number of replication fork in slow growing cells is 1:1.

Interestingly, we only rarely (<1%) observe a 0-1-2-1-2-0 pattern, which would point to a separation of the two hemimethylated termini after the chromosome

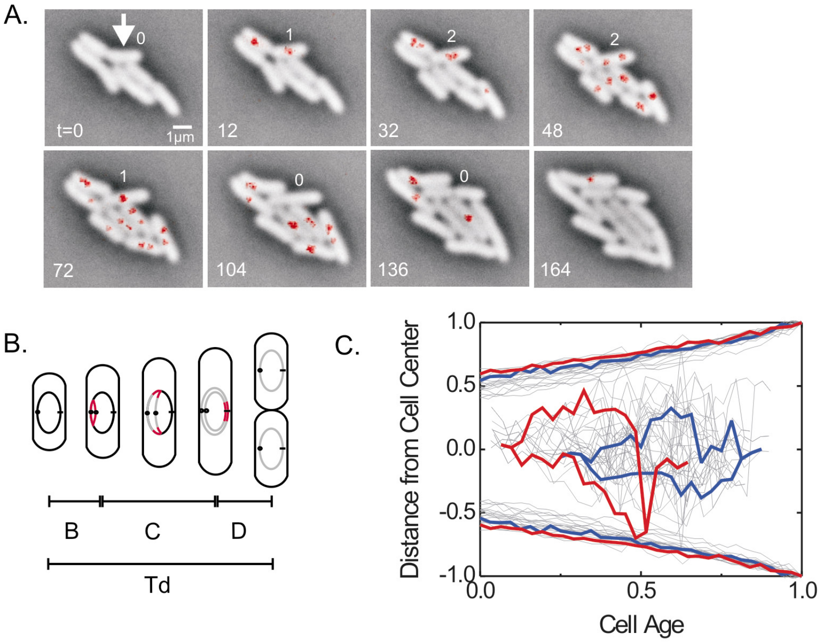

**Figure 3.2** SeqA foci dynamics in slow growing cells.
(A) SeqA foci in a cell (indicated by the white arrow) within a growing microcolony are followed over time. Distinct SeqA foci pattern according to the number 0-1-2-1-0 is observed.
(B) Schematic representation of bidirectional chromosomal DNA replication *in vivo*. The unreplicated chromosome is colored black and the replicated part, grey. Red areas indicated SeqA foci. Td is the cell interdivision time. It consists of the B period (the period between birth and the start of DNA replication), the C period (the DNA replication period), and the D period (the period after DNA replication until cell division).
(C) Typical trajectories of SeqA foci during one cell cycle normalized to the cell long axis. The maximum cell length at division is normalized to 1 and a newly born cell has a value around 0.5. Cell age is shown from birth (cell age =0) to division (cell age= 1).

replication. By using a fluorescence *in situ* hybridization (FISH) technique, Bates and Kleckner [7] probe *ter* loci of synchronized cells and observe that separated *ter* foci are only visible long after replication has ended, and they stayed next to each other until the cell divided. Because in our experiment we probe the hemimethylated DNA, and the time Dam methylase takes to fully methylate the DNA is much smaller than the time it takes for the two *ter* to separate as observed by Bates and Kleckner, we observe one focus before all foci disappear.

We limit our investigation to slowly growing cells due to the complexity of data interpretation for faster growing cells. In faster growing cells where the new replication round would already have started before cell birth, the difficulty in recognizing foci that belong to the old replication cycle and the new replication cycle can lead to misinterpretations of SeqA foci pattern. Furthermore, for fast growing cells, different patterns are expected for only a small change of interdivision time, therefore pattern subclasses are to be expected. A small percentage (about 10%) of the slowly growing cells also show this phenomenon and are excluded from further analysis, thus the variances for parameters produced by SeqA foci observation below reflect a minimum variance value.

### 3.2.3 Quantification of SeqA foci dynamics and cell cycle parameters

We have analyzed and quantified SeqA foci dynamics and cell cycle parameters for about 80 cycles. First we characterize cell growth by measuring cell interdivision time (Td) and elongation rate ($\mu$) from the phase contrast images. The elongation rate is obtained from the slope of a linear regression of logarithmic cell length over time and is the measure of the time it took for each cell to double in size. We find a mean elongation rate of 0.47 (size doubling per hour) with coefficient of variation (CV) of about 22% for both interdivision times and elongation rate among these cells. Large variation in interdivision times was observed previously [95]. We then simply plot cell elongation rate versus cell interdivision time to investigate if there is any relationship between the two parameters. Interestingly, a strong anticorrelation between cell division time and its corresponding elongation rate is observed as shown in figure 3.3A (correlation coefficient, $\rho$ =0.7). This anti-correlation suggests that cells with slower cellular growth (processes) display longer division times.

To our knowledge, no study has explicitly correlated the interdivision time and elongation rate. Early microscopy work measuring elongation rate did not quantify the extent of variation among individual cells, but qualitatively concluded that cell elongation rate is very uniform within a steady state culture [105]. In subsequent cell cycle modelling studies, a single elongation rate has been frequently assumed for all cells, indirectly implying that cells with longer Td are larger. Therefore, we next investigate whether or not a correlation exists between cell size at birth and the measured interdivision time of that cell (fig. 3.3A, lower panel). Cells are divided into three different sub-populations according to their size at birth. Interestingly, for a certain elongation rate, the smallest cells are predominantly positioned towards the longer interdivision time, while the largest cells are positioned towards the shorter interdivision time.

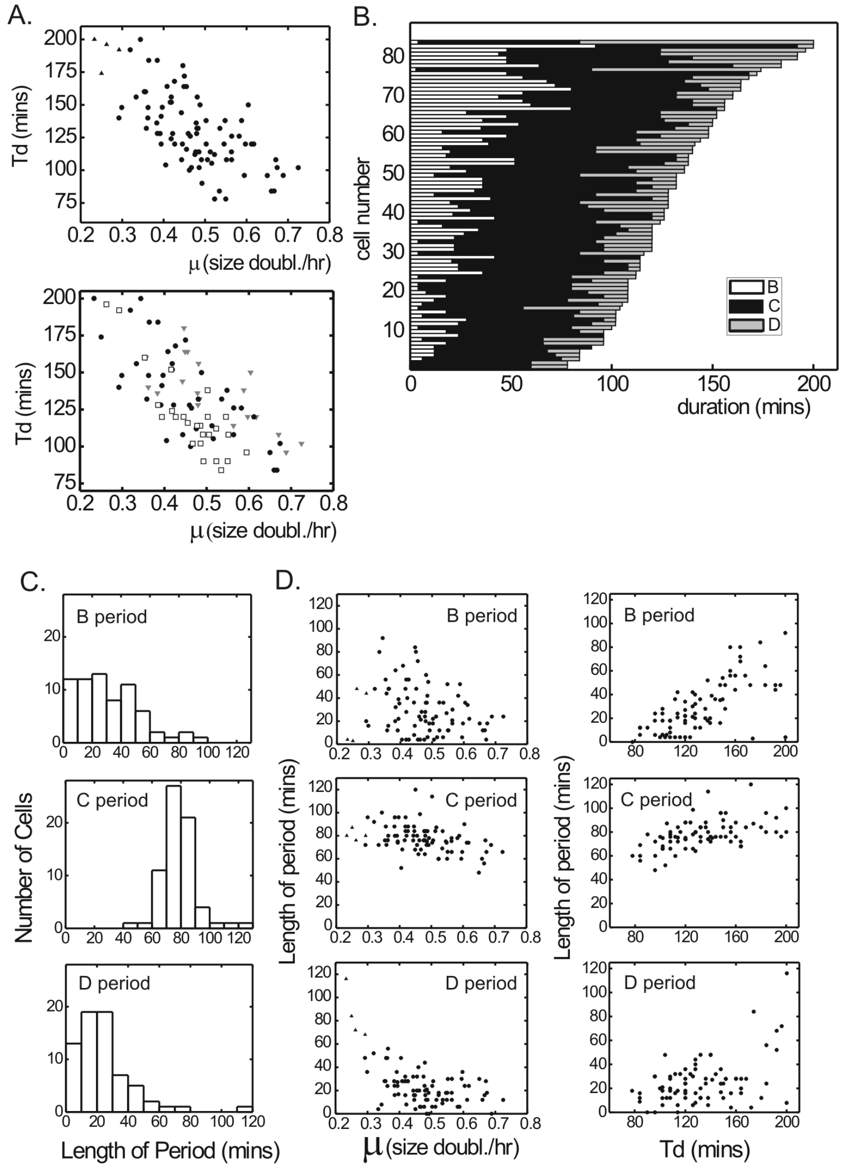

**Figure 3.3** Cell cycle and DNA replication parameters of slow growing cells.
(A) A strong anticorrelation is observed between the cell elongation rate (μ) and interdivision time (correlation coefficient, $\rho$ =0.7). Cells are then divided into three sub-populations according to their sizes at birth (lower panel). The smallest (grey triangles), intermediate (closed circles) and largest (open squares). For a certain elongation rate value, larger cells tend to have shorter interdivision time.

(B) The quantified B, C and D periods are sorted according to the cells interdivision time.
(C) Histograms of the duration of B, C and D periods (in minutes).
(D) Correlation of each period to the cell interdivision time.

Next, we investigate how the DNA replication process associates with the cell cycle. First, we quantify time periods preceding the appearance of the first focus (B period), the total time foci are observed (C period), and the time after the disappearance of the second single focus (fig. 2B). A previous investigation [23] estimates that the time chromosomal DNA stays in the hemimethylated form (except *oriC*) is about 1.5 minutes. Moreover, it has been shown that SeqA binds one to two orders of magnitude more strongly at hemimethylated DNA compared to fully methylated DNA [109]. Thus, we expect that SeqA foci do not persist for long after the nascent hemimethylated DNA become methylated. The effect of longer persisting foci will be the overestimation of the C period and the underestimation of the D period, but it should not strongly affect correlations.

Figure 3.3B shows the B, C and D periods for 80 cells, sorted according to their cell division time. The histogram of each period is shown in figure 3.3C, which shows that the B period is most widely distributed (coefficient of variation, CV=70 %), followed by the D period (CV=60 %, excluding the slowest growing cells denoted by closed triangles), and then the C period with the least spread (CV=16%). The narrower distribution of the C period compared to the spread in elongation rate (CV=22%) suggests that the DNA replication process is relatively insensitive to overall growth efficiency.

The B period shows a strong positive correlation with the interdivision time (correlation coefficient, $\rho$=0.7), while the C period shows only a moderate positive correlation ($\rho$=0.5) (fig. 3.3D). The D period shows almost no correlation with division time ($\rho$=0.2, excluding a few rare slow growing cells) and spreads around an average value of about 21 minutes. When plotted against elongation rate, all three periods are observed to have slight anti correlation with elongation rate ($\rho$ = -0.3, -0.4 and -0.4 for B,C and D period respectively). The results for the C period suggest that cells with slower metabolic processes also replicate their DNA at a slower pace. The results for the B period may be explained by the accumulation of a certain molecular entity in parallel with the accumulation of cell mass or size. Certain cells with more efficient metabolic processes may accumulate more mass and therefore elongate faster.

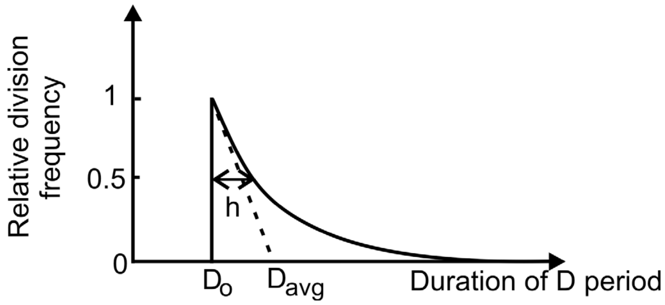

**Figure 3.4** Idealistic distribution of cell division frequency as a function of D period duration. $D_o$ is the minimum D period for each cell before division could take place. $D_{avg}$ is the mean division time, and h is the half life of the exponential decay of the curve [18].

Previous attempts to quantify variability in the B and D periods are scarce. By means of modelling and data fitting, an estimated variation coefficient for the B period in a synchronized culture was 60%, close to what we observe [64]. Previous methods investigating the variability of the D period consider the D period to have two components [18,108]. The first component is the minimum D period, $D_o$, which is a constant value. The other is the stochastic component of D, which distributes the D period beyond $D_o$ as an exponential decay (fig. 3.4). The variability in the D period is often expressed as a half-life time, h, the time it took for half of the population to divide. The proposed D period distribution has zero occurrence before $D_o$, then a sharp peak starting at $D_o$ followed by an exponentially decreasing curve characterized by h. The average D period, $D_{avg}$, is the sum $D_o$+(h/ln 2).

By means of direct observation, we measure an average D period that is very similar to the commonly reported value of about 20 minutes. However, the distribution appears rather different from that of the proposed model (fig. 3.4). We observe a significant number of cells with a very short D period. Multiple interpretations can be considered for this result. It can result from a certain sub-population of cells with longer persisting SeqA foci, and therefore giving the apparent shorter D period. Alternatively, it may indicate that the D period is more loosely regulated than previously thought. The classical interpretation of the D period is the time for a cell to prepare for cell division and that it has a rather constant value. However, if it is assumed that its high variability did not arise from persisting SeqA foci, then our data suggests that the completion of DNA replication has little influence on starting the process of cell division.

Early work with timelapse microscopy has revealed that familial relationships influence cell cycle parameters such as cell size and the correlation between the B period daughter and D period mother cell [88]. Unfortunately, only a few cells with mother daughter relationships showing clear SeqA pattern are observed, and thus prevent a reasonable interpretation from data plots.

Next, we investigate cell length at different times in the cycle as shown in figure 3.5A. We observe much less variation in cell sizes throughout the whole cell cycle as compared to the observed variation in interdivision time. Coefficient of variation is smallest for the length at division (CV=11.5 %), followed by a slightly higher coefficient of variation at the start of replication (12%) and a similar coefficient of variation for both the length at birth and at the end of replication (13.5% and 13.7%). This variation in cell size, however, is not only caused by the unequal division of sister cells. In general we observe relatively equal sized daughter cells immediately after division (CV 7% from the exact half size of mother cell).

Next, we investigate whether cell length at the beginning of the B, C, and D periods correlates with the duration of that period. Interestingly, for the B and D period a moderate anticorrelation exists ($\rho$ = -0.35 and -0.45 respectively), but not for C period (fig. 3.5B). Plots of cell size at the start of DNA replication (the end of B period) or at division (the end of cell cycle) versus cell interdivision time or elongation rate are shown in figure 3.5C. Intriguingly, a rather large variation in elongation rate (almost a factor of 3 between the slowest and the fastest) result in about only a micron spread of cell length at the time of DNA replication initiation. In contrast to earlier findings with *Salmonella typhimurium* grown in various media [104], we observe no strong correlation between cell size and either elongation rate or interdivision time in our experiment (the correlation coefficient value never exceeds 0.25). These results suggest that at least in our experimental conditions, different elongation rates have little effect on cell size.

### 3.3 Discussion and outlook

Since the first observation of SeqA as discrete foci *in vivo* by immunofluorescence staining [51], a number of unresolved issues have remained regarding the correlation between SeqA foci and the number of replication forks as well as the interpretation of SeqA dynamics. In most studies, the distribution of the number of SeqA foci per cell is determined from images of many cells taken at one time point. Using this method, it has been reported that the ratio between the number of SeqA foci and replication forks is about 1:1 [19,51], based on the observation of up to 14 foci for rapidly growing cells, which is close to the predicted number of replication

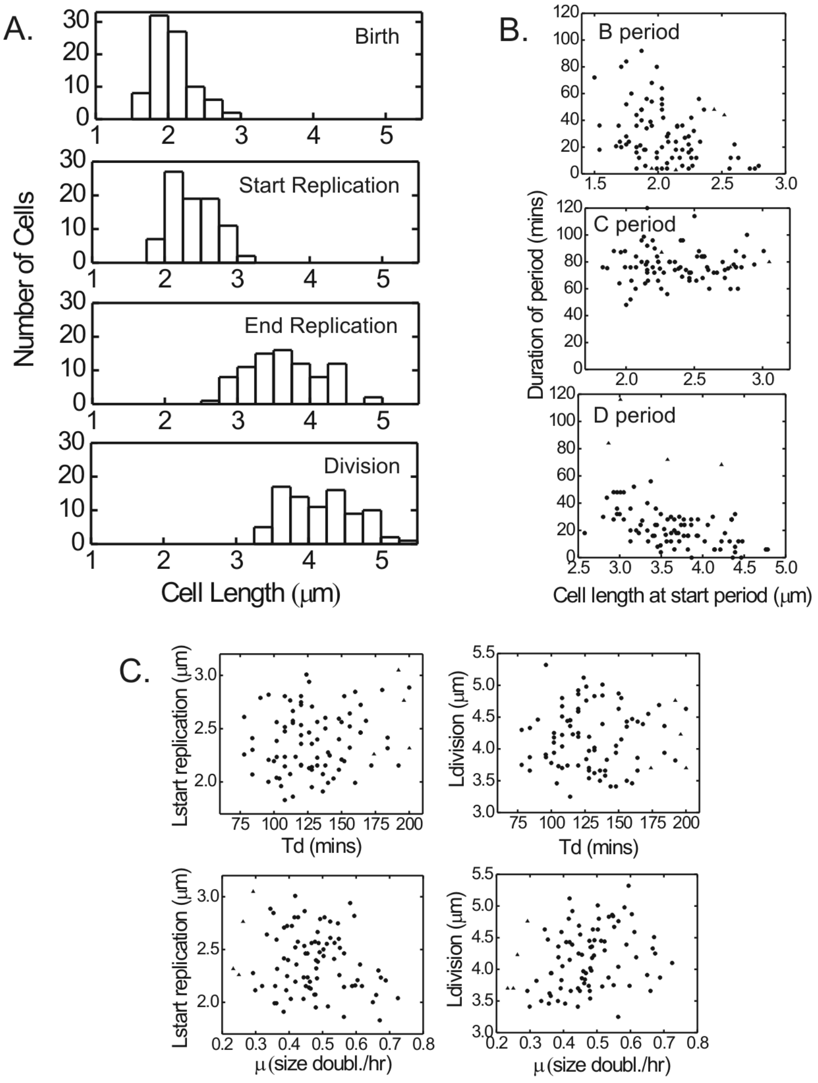

**Figure 3.5** Correlations between various cell parameters.

(A) Histograms of cell lengths at various times within the cell cycle.

(B) Correlation between the duration of the B, C, and D periods and cell length at the beginning of the corresponding period.

(C) Correlation between cell lengths at different times and the cell division times or elongation rates.

forks in the Cooper-Helmstetter model. However, a ratio of 1:2 has been suggested as well [4,32,85,91]. For cells growing at a moderate growth rate, the observation of a single SeqA focus in smaller (younger), and two foci in longer (older) cells has been observed [32]. The appearance of the second focus may indicate the start of a new replication round, thus implying that each focus represents two co-localized replication forks. However, these observations are also consistent with a 1:1 ratio, as the two foci may also represent two separate replication forks instead of a new round of replication. These issues can be resolved directly by tracing the SeqA foci in real time.

By performing timelapse imaging experiments in slow growing cells, we find a pattern of 0-1-2-1-0 for the number of SeqA foci. The results show that the second focus does not represent the initiation of a new replication round but rather a single replication fork, as evidenced by the close initial proximity of the two foci. The results are consistent with a model where two replication forks are co-localized at initiation, but then separate in a bidirectional manner. If a second round of replication were to occur, it would not be expected to initiate in close proximity to the replication machinery of the first round, so we can eliminate the possibility of a second replication round. This method has led to a unique opportunity to not only directly follow the chromosomal DNA replication and the cell-cycle process, but also to extract multiple parameters for the same cell cycle. These data can easily provide information about correlations between DNA replication and the cell cycle.

We study cell-to-cell heterogeneity in *E.coli* division and the chromosome replication cycle, for isogenic populations in a constant environment. The resulting spontaneously occurring variability both contrasts and complements traditional studies of the cell cycle for different steady state growth rates in different media. We observe a strong correlation between the cell interdivision time and its elongation rate, showing that elongation rate is an important contribution to the noise in interdivision times. The data also shows that the B period is the main contributor to this trend: slower cells have longer B periods, while faster cells have shorter B periods. Such a trend is expected if a certain critical mass (or certain molecular entities accumulate in parallel with the mass) would need to accumulate before DNA replication initiation can proceed.

Since it was first proposed, the concept of a critical mass at the time of DNA replication initiation has been challenged many times (for reviews: [14,29,49]). Several studies using various growth media have found a growth dependent initiation mass [26,120]. However, their results contradict one another and large variations in growth rate (as much as 8 times) result in a maximum factor of two difference in initiation mass. Alternatively, the time elapsed after a certain event can be viewed as the determining factor for DNA replication initiation. Recently, a study with synchronized cells obtained by a membrane elution method proposed that cell division itself may license the chromosome(s) for the next round of replication initiation. In other words, the time after cell division highly influences the time of the next round of DNA replication initiation [7]. Our observation of a highly variable B period contradicts this proposal. Moreover, we find a mild anti-correlation between cell length and the duration of the B period, indicating the tendency of smaller cells to have longer B periods. A large spread in growth rates also results in a relatively small spread in cell length at the start of the C period. Taken together, the results provide support for models predicting that the start of DNA replication is triggered by the accumulation of a critical mass [33]. However, some variability in cell size already exists at the start of replication, indicating the limitations on the precision of this trigger.

Although great advances have been made in understanding the details of the DNA replication initiation process, the way it is linked to cell growth is still unclear. Currently, DnaA protein is believed to be the molecular entity that accumulates in parallel with cell mass. However, it is very likely that it is not the DnaA concentration itself that limits replication initiation, since it has been shown that DnaA is constitutively expressed during the cell cycle [102] and a high overexpression of DnaA only slightly affects the timing of replication initiation [2]. The ratio between the number of *oriC* and DnaA is also not the determinant of replication initiation since synchronous initiation is still observed at normal cell size for cells with extra *oriC* copies. The current hypothesis is that the ratio between the active form of DnaA (DnaA-ATP) and the inactive form (DnaA-ADP) is the critical entity, however the dynamics of the active and inactive forms of DnaA between successive replication rounds is not yet known [107]. It is of much interest to investigate how variation in the ratio of DnaA-ATP and DnaA-ADP *in vivo* would affect the conditions for DNA replication initiation in our system, as this could be the first step toward understanding how cell mass relates to DNA replication initiation at the molecular level.

# 4

# Direct Observation of Type 1 Fimbrial Switching

The defining feature of bacterial phase variation is a stochastic ‘all-or-nothing’ switching in gene expression. However, direct observations of these rare switching events have so far been lacking, thus obscuring possible correlations between switching events, as well as between switching events and other cellular events such as cell division and chromosome replication. We monitor the phase variation of type 1 fimbriae in real time in individual *Escherichia coli*, using GFP as a reporter. The chromosome replication process is tracked simultaneously by means of a SeqA-mCherry fusion. During switching we observe distinct expression patterns in multiple genealogically related lineages, with some displaying expression peaks, and others exhibiting monotonic increases. These patterns can be explained by a model that combines a single *fim* switching event with multiple genetic *fim* copies per cell in accordance with the Cooper-Helmstetter theory of chromosome replication. Analysis of the moment of switching at sub-cell cycle resolution shows a preference for switching at the beginning of the cell cycle. The results show that a mechanistic description of phase variation requires detailed information on the chromosome replication process. The observed correlation between *fim* switching and cell age challenges the traditional notion of phase variation as a random Poissonian phenomenon.

## 4.1 Introduction

Phase variation is characterized by a stochastic switching between discrete gene expression states [48,116]. This form of gene regulation contrasts with classical gene regulation, both in terms of the underlying molecular mechanisms and the effect on the population structure. Switching between phases typically involves a genetic or epigenetic mechanism, such as a DNA inversion or a loss or gain of DNA methylation, resulting in 'all-or-nothing' ON or OFF expression levels. As the switch state is heritable over multiple generations, it allows the population to differentiate and stably maintain sub-populations. Phase variant macromolecules are typically exposed on the cell surface, and play a central role in bacterial virulence and immune evasion.

The regulation of phase variation switching rates have been studied in detail for a number of systems [10]. However, the experimental methods used so far do not capture the switching event itself, but rather infer switching rates by monitoring the ratio between ON and OFF cells within a population [41]. For this reason, the most elementary questions on the correlations and dynamics of phase variation have remained unaddressed. For instance: is there a correlation between switching events in genealogically related cells? Does switching occur in particular phases of the cell cycle? What are the expression dynamics during switching? Here, we address these issues by following the expression of a phase-variable gene in single cells and in real time. In particular, we investigate the switching behavior of type 1 fimbrial expression in *Escherichia coli*.

The mechanism of *fim* switching is characterized by a DNA inversion event, in which a *cis* regulatory element that contains the promoter of the *fim* operon changes its orientation (fig. 4.1A). The DNA inversion process is thought to involve the following sequential sub-steps: a) recombinases (FimB and/or FimE) bind to the two inverted repeats that flank the invertible element [42,53], b) the two repeats and recombinases form a synaptic complex, thus forming a loop of the 314 bp invertable region, c) the actual inversion mediated by the tyrosine site-specific recombinases involves the breaking and rejoining of single DNA strands and a Holliday junction intermediate [21,43]. This process does not consume ATP and is reversible, and d) the synapse complex must likely be broken to stabilize the new DNA orientation and to allow transcription of the *fim* operon [87].

In order to detect the *fim* switching events in single cells, we clone the gene encoding for GFP in the chromosomal *fimA*: the first gene of the *fim* operon. Using time-lapse microscopy and image analysis, we follow cell growth and GFP expression level over time, and determine the genealogical relations within the population. These data allow us to reveal the fimbrial expression pattern during switching, pinpoint the moment of switching, and look for possible correlations between switching events in genealogically related individuals. We simultaneously monitor the chromosome replication process within the cell cycle using a fusion between the sequestering protein SeqA and the mCherry fluorescent reporter. SeqA has a binding affinity for hemimethylated DNA, as is present at nascent DNA near replication forks [52,125]. The number and position of the discrete SeqA foci indicate the replication progress [51]. These data allow us to correlate *fim* switching events with phases within the cell and chromosome replication cycle.

Our measurements reveal a generic fimbrial expression pattern within genealogically related lineages upon *fim* switching, which displays monotonicly increasing expression as well as expression increases followed by decreases. We show that these patterns depend on the growth rate and on the switching moment relative to replication of the *fim* operon, and explain them with a model that considers the number of *fim* operon copies during the cell cycle, as well as the epigenetic inheritance of expressed *fim* protein and RNA and their dilution by growth. The data indicates that the *fim* switching depends on the cell age, with a higher switching probability at the start, and a lower probability at the end of the cell cycle.

## 4.2 Results

### 4.2.1 *fim* phase variation in real-time at the single cell level

We first set out to characterize the distribution of the *fim* operon expression level at a single time point. Isogenic cells grown exponentially in defined rich medium, are spread onto agar and then imaged with phase contrast and fluorescence microscopy. The distribution of the mean GFP brightness per unit area indicate two distinct sub-populations, one with a low expression level (0.77±0.23 a.u.) and one with a high expression level (8.18±1.88 a.u.) (fig. 4.1B). These two subpopulations represent the OFF and ON states of *fimS*. The observed fraction of ON cells (4.6%) is in agreement with earlier studies [72]. The low level of expression in the OFF state, which is indistinguishable from cells lacking the GFP fusion, is consistent with the tight orientational specificity of the promotor.

With the aim to follow switching events in real time, we monitor the cells as they continue to grow and divide. Microcolonies of about 500 cells form over a period of about 5 hours on defined rich agar (generation time about 25 min). Some descendents of a non-fluorescent OFF progenitor spontaneously develop an increased fluorescence, resulting in a heterogeneous colony with patches of a

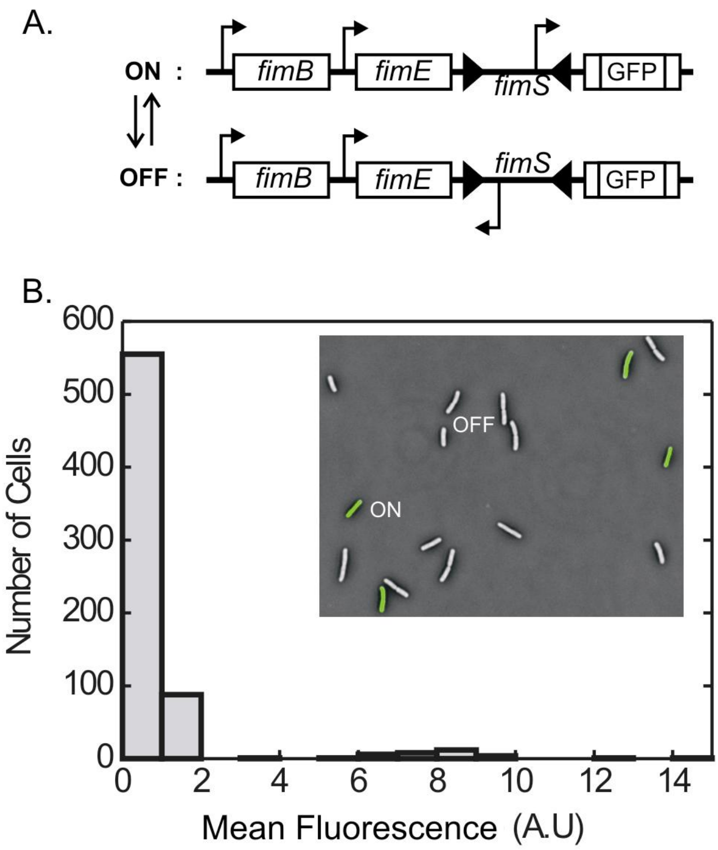

**Figure 4.1** *fim* operon in the ON and OFF phase
(A) In the ON orientation, the promoter positioned within the invertible element *fimS* drives transcription of the *fim* operon. GFP has been inserted into *fimA*, the first gene of the *fim* operon, allowing single-cell monitoring of the *fim* operon expression level. Transcription is turned OFF upon inversion of *fimS*, which positions the promoter in the opposite direction. This reversible inversion process is performed by two recombinases, FimB and FimE.
(B) Histogram of GFP fluorescence per cell. Image is a phase contrast micrograph in black and white, overlaid with the background corrected GFP fluorescence signal, shows cells deposited on agar. The distribution of mean fluorescence level per cell shows well differentiated ON and OFF states.

different phase (fig. 4.2A). Conversely, brightly fluorescent ON progenitors produce descendents with a decreasing fluorescence. In total, we monitored more than 200 progenitors producing over twenty thousand single-cell trajectories.

### 4.2.2 Distinct *fim* expression patterns are observed near switching

Using the phase-contrast and fluorescence images, we construct a family tree that details the genealogical relations within the microcolony, as well as the GFP expression history of each family member. Figure 4.2 shows such data for a typical OFF to ON switching event. Some lineages exhibit a monotonic increase in brightness, while other lineages display an increase followed by a decrease. These expression profiles develop over the course of several generations, which is significantly larger than the GFPmut2 maturation time of about 5.7 minutes (section 2.3 of this thesis). The maximum of the expression peaks appears to be variable. One may consider whether the observed decreases in brightness originate from ON to OFF switching events. However, because the bulk ON to OFF switching rate (of order 0.1 per cell per generation [41]) is too low to explain the data, one would have to invoke as-of-yet unknown correlations between switching events. Instead, we hypothesize that the observed expression pattern is caused by the chromosome replication process in combination with a single OFF to ON switching event.

In order to synthesize a sufficient number of chromosomes for all progeny, rapidly dividing bacteria are thought to have multiple nested replication forks operating simultaneously [27,28,126]. Before one chromosome is fully replicated, a new replication round already starts. Consequently, genes may be present in multiple copies, depending on their location on the chromosome. Genes positioned near the chromosome origin (*oriC*) are amplified in number, while those near the terminus (*terC*) are not. Within the *E.coli* chromosome, the *fim* system is located near *oriC* (at about 610.000 bps from *oriC* or about 13% of the whole chromosome), and should therefore be present in multiple copies.

How could multiple chromosomal *fim* copies explain the observed expression pattern? We consider a single OFF to ON event, in which initially all *fimS* copies within one cell are in the OFF state, and at a certain point a single DNA inversion event turns one *fimS* copy ON, resulting in GFP expression. In the data depicted in figure 4.2C, this moment may coincide with the first fluorescence increase seen in the orange lineage at about 100 minutes. Upon division, this single chromosomal ON copy is inherited by one of the daughters, which continues the GFP expression (red lineage). In contrast, the second daughter inherits chromosomal *fimS* copies in the OFF state, as well as GFP proteins and mRNA (orange lineage). Continued growth of the second daughter leads to dilution of the GFP protein and mRNA, resulting in a slow decrease in mean fluorescence. As the cell cycle of the red lineage progresses, its single ON copy is replicated, yielding two *fimS* copies in the ON state. After division, the lineage of one daughter (red) is seen to decrease in

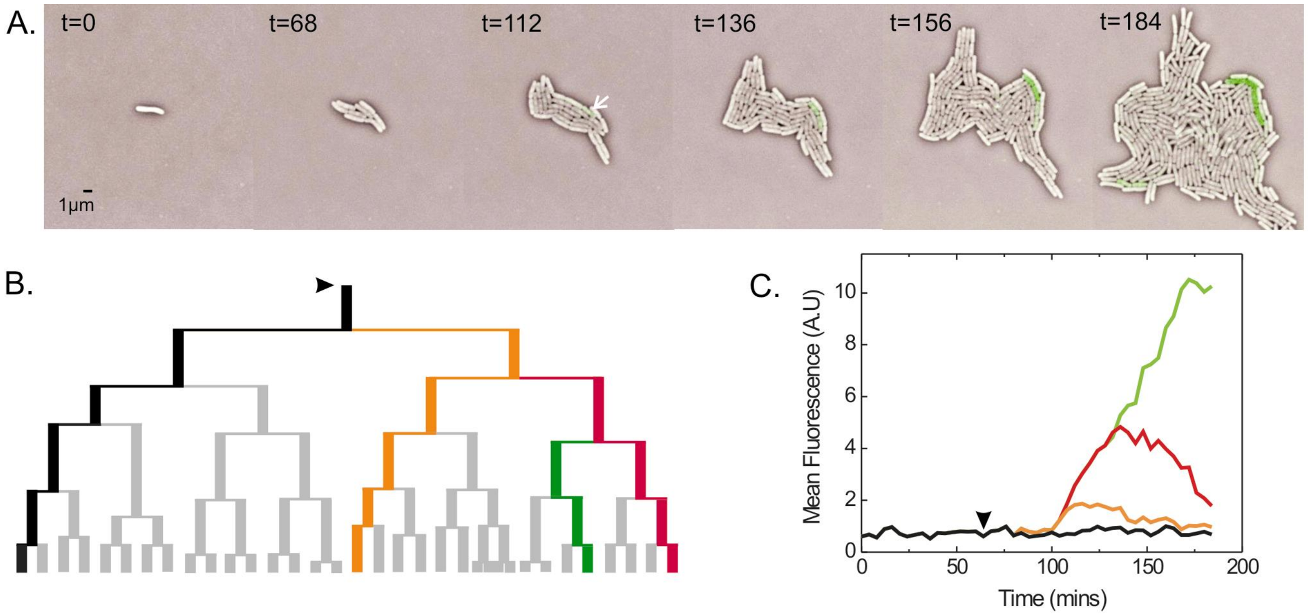

**Figure 4.2** ON switching in a growing microcolony.
(A) Overlay of fluorescence and phase contrast images, showing a single OFF cell developing into a microcolony. A cell in this isogenic population starts to fluoresce spontaneously (white arrow), and develops into a brightly fluorescent sub-colony, indicating an OFF to ON *fimS* inversion event. A later switching event is seen to occur in the bottom left corner of the microcolony.
(B) A section of the family tree representing the microcolony in panel A. The length of the bars represents the division time. The black arrow represents the start of the family tree. The fluorescence of the colored lineages is followed in panel C.
(C) GFP expression history of the four lineages labeled in panel B. The black lineage displays a low fluorescence level (mean fluorescence per unit area), which corresponds to the OFF phase. The orange and red lineages show first an increase in fluorescence, and then a decrease. The green lineage exhibits a monotonic expression increase.

brightness while the other daughter (green) continues to increase in expression, which suggests that the latter inherits both ON copies and the former only OFF copies. This grouped redistribution of *fimS* copies agrees with a nested arrangement of replication forks; where the most recently replicated DNA remains physically linked with the formerly replicated DNA. No further lineages with decreases are observed while the fluorescence level of the green lineage reaches steady state, suggesting that the green lineage inherited only two ON copies of *fim* and no OFF copies. An inheritance of two *fimS* copies at birth (and thus a total of four copies after replication) is in accordance with the Helmstetter-Cooper model of chromosome replication at this growth rate [27,28].

The above model describing *fim* switching in combination with chromosome replication provides a number of predictions. First, the GFP expression pattern should alter for different growth rates, as the chromosome replication scheme is expected to change with growth rate. For instance, according to the Helmstetter-Cooper model of DNA replication [27], slow growing cells with generation times of more than 60 minutes have one chromosome copy at birth, and two at division. This would imply that one lineage at maximum would show an expression peak, instead of two lineages for fast growing cells. Second, the expression pattern should depend strongly on the moment of switching within the cell cycle (fig.4.3). In slow growing cells, if switching occurs after birth but before replication of the switch *fimS*, then all descendents exhibit a monotonic fluorescence increase. If, on the other hand, switching occurs after *fimS* replication, only one of the two *fimS* copies is ON at division, resulting in one daughter lineage with an expression peak.

### 4.2.3 *fim* expression pattern depends on growth rate and cell cycle

To test the predictions outlined above we monitor *fimS* switching in slowly growing cells. In these experiments bacteria are grown on succinate minimal medium, which permits growth with a mean generation time of about120 min. In addition, we simultaneously imaged the replication fork, which allows us to determine the number of inherited chromosomes. We can then pinpoint the start and end of replication, and consequently the moment of *fimS* replication. The replication forks are visualized by means of a fusion between mCherry and SeqA, which is known to bind to newly replicated and hence hemimethylated DNA [19,40,109]. The delay in methylating the nascent DNA produces visible SecA foci in the wake of the replication fork [51]. In this fashion, one can correlate the moment of *fimS* replication, the moment of *fimS* switching, as well as the ensuing GFP expression pattern, within the cell cycle of individual cells.

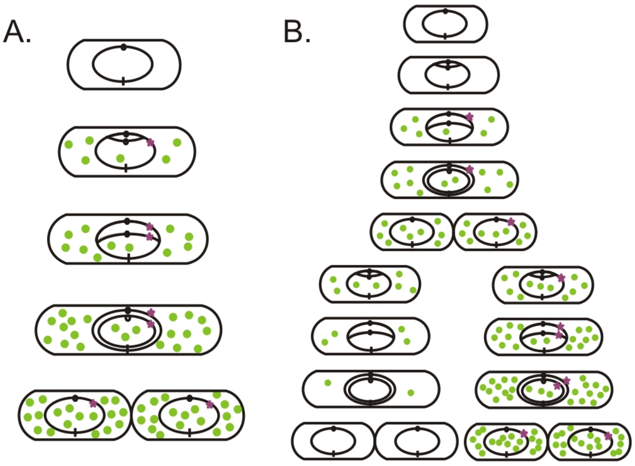

**Figure 4.3** Fimbrial expression and DNA replication
Model for slow growing *E.coli* cells inheriting a single chromosome with an OFF *fim* switch. The chromosome is replicated in a bidirectional manner. At a certain moment, the switch spontaneously switches to the ON state (purple star).
(A) OFF to ON switching occurring before *fim* replication. Both daughter cells inherit ON copies of *fim*, resulting in two lineages with monotonically increasing fluorescence.
(B) OFF to ON switching occurring after *fim* replication. One daughter cell inherits only GFP protein and mRNA, which is diluted by growth and thus decreases in concentration. The other daughter also inherits one chromosomal *fim* system in the ON state, resulting in a monotonic increase in fluorescence.

The slow growing cells display an altered expression pattern during switching compared to rapid growing cells. Two distinct classes of expression patterns are observed in the descendents of the cell displaying the initial expression increase. In the first pattern, one of the two daughters displays an expression peak, while the other shows a monotonically increasing expression (fig. 4.4C). In the second pattern, both daughters exhibit an expression increase (fig. 4.4D). The steady state GFP expression levels are similar in both patterns, though it is reached significantly faster in the second case, taking about 100 min. compared to about 200 min. in the first pattern. For ON to OFF switching, a similar but inverse GFP expression pattern is observed, which shows decreases in brightness followed by increases -back to the ON fluorescence level- (fig. 5.2B image b-f, sister cell of the pointed cell and fig. 5.2D), as well as monotonic decreases in the fluorescence level.

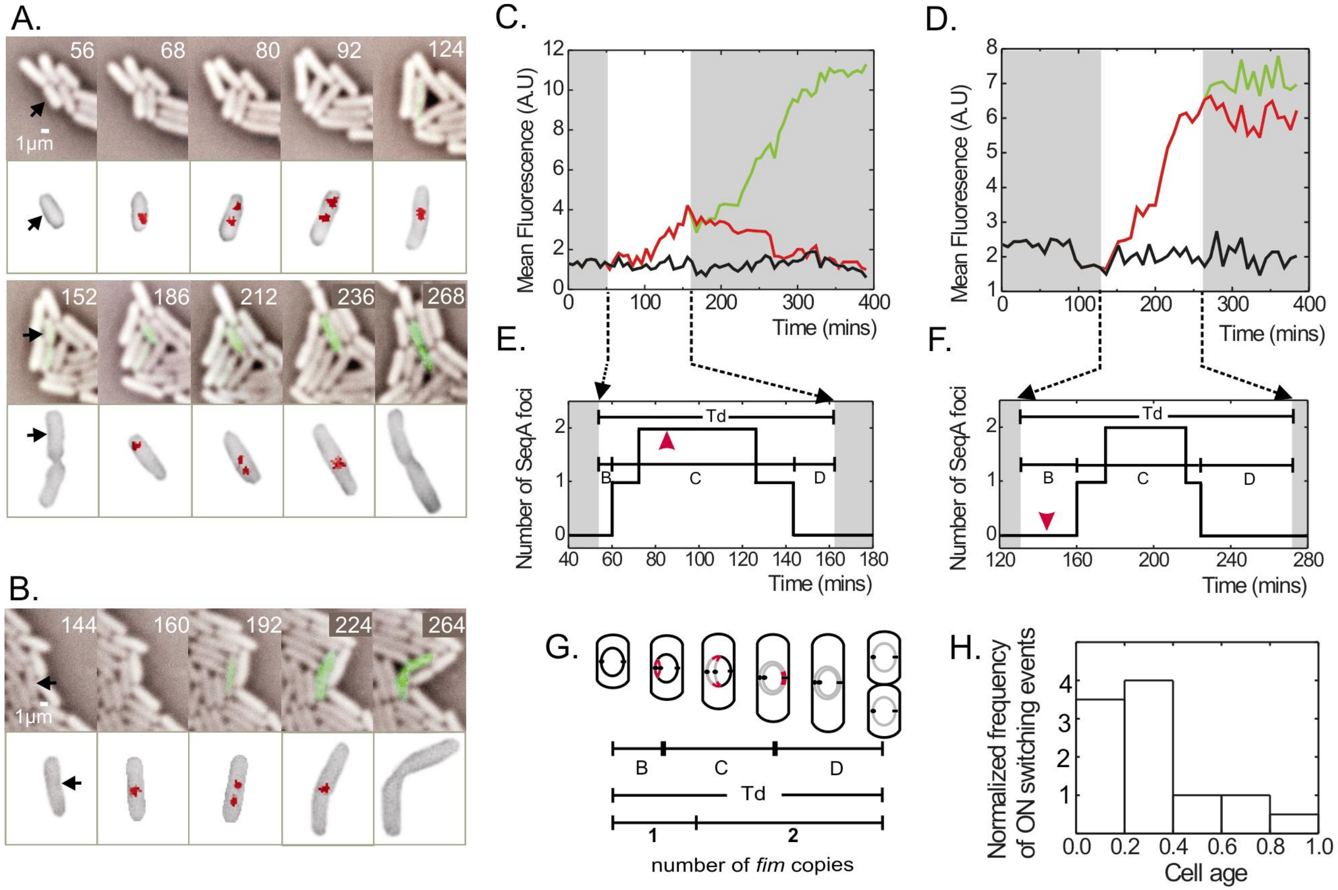
A.
56
68
80
92
124
1µm
152
186
212
236
268
B.
144
160
192
224
264
1µm
C.
Mean Fluorescence (A.U)
12
10
8
6
4
2
0
0
100
200
300
400
Time (mins)
D.
Mean Fluorescence (A.U)
8
7
6
5
4
3
2
1
0
100
200
300
400
Time (mins)
E.
Number of SeqA foci
Td
B
C
D
2
1
0
40
60
80
100
120
140
160
180
Time (mins)
F.
Number of SeqA foci
Td
B
C
D
2
1
0
120
160
200
240
280
Time (mins)
G.
B
C
D
Td
1
2
number of *fim* copies
H.
Normalized frequency of ON switching events
4
3
2
1
0.0
0.2
0.4
0.6
0.8
1.0
Cell age

**Figure 4.4** ON switching events in slow growing cells

(A) OFF to ON switch with transient expression increase. Top panel shows phase contrast images overlaid with GFP fluorescence signal reporting for *fim* expression. Bottom panel shows phase contrast images of *fim*-switching cells overlaid with mCherry fluorescence signal reporting for SeqA, which labels chromosome replication forks. Black arrow indicates the start of the lineage displayed at the lower panel.

(B) OFF to ON switch with only monotonic expression increases. Top panel shows phase contrast images overlaid with GFP fluorescence signal reporting for *fim* expression. Bottom panel shows phase contrast images of *fim*-switching cells overlaid with mCherry fluorescence signal reporting for SeqA, which labels chromosome replication forks. Black arrow indicates the start of the lineage displayed at the lower panel.

(C) and (D) GFP expression history of three lineages, from the measurements as displayed in panel A and B respectively. The black lineage displays a low fluorescence level (mean fluorescence per unit area), which corresponds to the OFF phase. The red lineage displays the first fluorescence increase, which either subsequently decreases (C), or continues to increase monotonically (D). The green lineage corresponds to a daughter branching away from the red lineage, and exhibits a monotonic expression increase in both C and D.

(E) and (F) Pattern of SeqA foci within the a cell cycle. The number of SeqA foci, which label the replication forks, is plotted as a function of time. Data is taken from the measurements displayed in panel A and B respectively. Red arrows indicate the start of fluorescence increase. The B period represents the time prior to the chromosome replication process (C period). The D period represents the time from the end of replication to the end of cell division. Td is the cell interdivision time.

(G) A schematic representation of the proposed chromosome replication process at slow growth, indicating the SeqA foci and *fimS* location along the chromosome. *fimS* is replicated at about 26 % of the C period.

(H) Frequency of observed OFF to ON switching events, normalized to the number of *fim* copies, as a function of cell cycle phase (N=13). The cell age is indicated along the x-axis, from birth (0) to division (1).

As previously shown in chapter 3, the number of SeqA foci within one cell cycle exhibits a regular 0-1-2-1-0 pattern (fig.4.4E and F) (chapter 3 this thesis). After an absence of SeqA foci at birth, the appearance of a single SeqA focus indicates the start of chromosome replication (fig. 4.4G). The single focus then splits into two separate foci, indicating a bidirectional movement of the two replication forks, which later merge again into one focus suggesting the forks neared termination. Near the end of the cell cycle the single remaining SeqA focus disappears, indicating a complete methylation of both chromosomes, and the end of replication. The delay in the methylation of nascent DNA has been reported to range from 1 to 13 minutes [23,76]. The moment of *fimS* replication is determined at 26% of the replication process, based on a bidirectional replication and a constant replication rate along the chromosome. The moment of *fimS* switching is determined by the intersection of the background fluorescence signal and the rising fluorescence signal, and by correcting for the GFP maturation time.

The data show that all *fimS* switching events that occur after *fimS* replication result in a single lineage with transient GFP brightness (fig. 4.4C and E). All *fimS* switching events occurring before *fimS* replication result in no lineages with transient brightness (fig. 4.4D and F). The results at slow growth verify the essential feature of the proposed model: the expression pattern is determined by the switching moment relative to the moment of *fimS* replication. Other details of the expression traces can also be understood within the framework of the replication model. For instance, because switching before replication quickly results in two *fimS* copies in the ON state, the expression rises faster, leading to a more rapidly established steady state expression level.

### 4.2.4 The OFF to ON *fim* switching rate is higher at beginning of cell cycle

Next, we investigate how the *fimS* switching events are distributed within the cell cycle. The null-hypothesis we aim to test is that for one given environmental condition, *fim* switching occurs randomly in time, with a constant switching probability per *fimS* throughout the cell cycle. We find that on average, *fimS* replication occurs at 0.36±0.11 of cell age (N=80 cell cycles). Given the doubling of *fimS* copy number upon replication, and the null hypothesis, the expected fraction of switching events that occur before *fimS* replication would be 0.22. Our data shows that the measured fraction is 0.43 (N=24 out of 55 OFF to ON switching events in slow growing cells), which is significantly higher than expected (chi-squared test, $P<0.001$). The distribution of switching probability in time (fig. 4.4H) shows that the switching probability is roughly twice as high in the first part of the cell cycle.

## 4.3 Discussion and outlook

Phase variation in bacteria is controlled by a variety of mechanisms, including those determined by changes in Dam methylation of specific regulatory sites and others by slipped-strand mispairing that results in the insertion or deletion of repeat DNA sequences. Phase variation can also be controlled by site-specific recombination, as is the case for type 1 fimbriation in *E.coli*. Here we develop techniques to follow phase variation in individual bacteria for the first time, using *fim* phase variation as our model system.

Unlike many epigenetic phase variation control mechanisms, the *fimS* inversion is not expected to require DNA replication as an essential step in the switching process [31,75]. However, a key finding of the present study is that OFF to ON *fim* switching occurs preferentially in the beginning of the cell cycle, before replication of the *fim* genes. What can be the cause of this observed switching bias? The known molecular mechanisms underlying *fimS* switching do not provide a clear answer as to how this occurs. One possible cause may be a variation in the level of recombinase *fimB*, which is required for OFF to ON phase variation. *fimB* expression is activated by regulators NanR and NagC, both of which bind to operators that include Dam methylation sites that are often fully unmethylated [111,112]. While the role of methylation in controlling NanR and NagC remains to be determined, *fimB* expression is diminished considerably in a *dam* mutant (fig. 4.5). Thus, the generation of unmethylated (or hemimethylated) DNA in the vicinity of *fimB* following passage of the replication fork, and a subsequent reduction in *fimB* expression, might be responsible for the effects observed. This scenario suggests a more active and regulated link between cell cycle and gene regulation.

Alternatively, the observed bias for switching at the beginning of the cell cycle may be caused by changes in DNA topology during replication. The site-specific recombination process that is at the basis of *fimS* switching has been shown to be sensitive to changes in DNA topology [35,61]. During DNA replication, the separation of complementary DNA strands by helicase results in a relaxed (or possibly positive supercoiled) region in front of the replication fork, and negative supercoiled region behind it [65]. The *fim* switching rates may thus be modulated by these changes in supercoiling density, or instead by the disruption of the recombinase synapse with the looped *fimS*, by the passing replication fork. A possible advantage of switching early in the cell cycle may be that it reduces the number of the transiently expressing, ambiguous lineages, thus limiting the waste of resources on unused *fim* protein and RNA.

By interrogating the behavior of single cells, this study reveals correlations between the cell cycle dynamics and *fim* switching behavior. It demonstrates that multiple genetic *fim* systems per cell ought to be taken into account in mechanistic descriptions of *fim* switching. The observed modulation of the *fim* switching rate with the cell cycle challenges the common notion that *fim* switching is a straightforward Poissonian process, in which the switching probability is constant in time for fixed external conditions. The methodologies presented here can be applied to further explore the boundary between randomness and correlations in phase variation, as well as to reveal the dynamics of other phase variation systems at the level of discrete events.

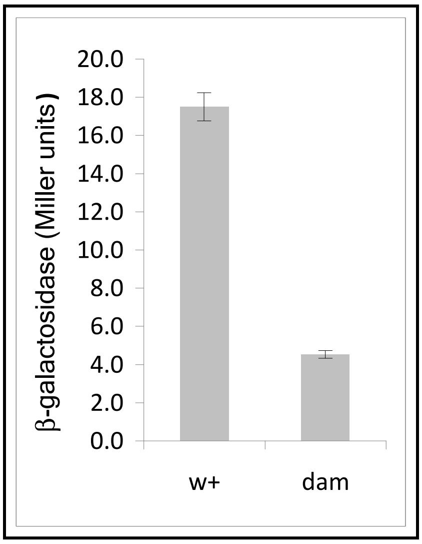

**Figure 4.5** The effect of Dam on the β-galactosidase produced by a FimB-LacZ fusion. The fusion is situated in the chromosome at *fim* in a *ΔlacZYA* mutant of strain MG1655 [38]. β-galactosidase assays were conducted as described by Miller [84], following growth of the wild type or *dam* mutant strain in RD glycerol medium at 37°C with rapid aeration to an $OD_{600}$ of 0.2. Experiments were repeated at least twice, and the values shown represent the mean of at least four samples with 95% confidence intervals included for each value. The work with *dam* mutants was performed by B.K. Sohanpal.

# The Effect of Cell History on the **OFF Switching Rate** of Type 1 Fimbriation in *E.coli*

5

Phase variation in bacteria is often considered to be a purely random Poissonian process, in which cells switch between distinct expression states. However, this assumption has been difficult to test directly in bulk experiments. In this study we investigate the statistics of *fim* phase variation, using GFP-labeling to follow switching in single cells in real time. Our data indicates that for our experimental conditions, OFF switching rate is constant over time and does not depend on the switching history, thus verifying the Poissonian nature of *fim* switching within several generations after an ON switching event.

## 5.1 Introduction

Following the work of Luria and Delbruck [77], phase variation studies have commonly assumed that phase variation is a purely random process that follows Poissonian statistics. Such studies include the examination of phase variation in flagellar expression in Salmonella [114] and type 1 fimbrial expression in *E.coli* [37]. However, this Poissonian process assumption itself has never been tested directly and would imply that the probability of switching is constant in time for a given environment.

In *E.coli*, the expression of type 1 fimbriae is phase variable, resulting in a heterogeneous population of fimbriate (ON state) and afimbriate cells (OFF state). The *fim* genetic switch, *fimS*, is a 314 bps invertible DNA segment containing a promoter sequence for the fimbrial structural genes. DNA inversion is performed by two different recombinases: FimB, which switches both ON and OFF and FimE, which is strongly biased to OFF switching.

Recent phase variation studies suggest various mechanisms that could lead to non-Poissonian switching behavior. One example is orientational control, in which the FimE recombinase that switches off the fimbrial expression is produced only when the switch is in the ON state [56,110]. As a consequence of the orientational control, it has been proposed that shortly after an ON event, the probability for a consequent OFF switching event may be lower than normal due to a lack of FimE at the beginning of the ON state [119,121]. In this scenario, the rate of OFF switching is not constant in time, resulting in a deviation from Poissonian switching statistics [119].

Traditional bulk methods are not suitable to address this question, as they measure the percentage of cells in a certain state, thus averaging the switching rate of many cells with different switching histories. In this study, we investigate *fim* switching statistics and the role of switching history. Specifically, we ask whether the time spent in the ON state has any effect on the OFF switching rate. Here we address this question by placing GFP under the control of the *fim* switch, and following the resulting fluorescence in single cells as they first switch OFF-to-ON, and later switch ON to OFF, in real time (fig. 5.1).

Our data provides the first evidence that the OFF switching rate is constant in time, and thus independent from its switching history. These results support the commonly used assumption that switching in type 1 fimbrial expression is a Poissonian process. Due to the limitation of growth on agar however, our observations were limited to intermediate time scales (between 2 to 7 generations).

Further investigations at both shorter and longer timescales are of much interest. Our approach can be used to study the stochastic properties and possible epigenetic effects in other phase variation systems.

## 5.2 Results

### 5.2.1 ON and OFF switching rates of the *fim* switch

Cells are grown exponentially in a MOPS medium supplemented with glucose and a low amount of leucine. They are spread on an agarose surface containing the same medium and observed with phase contrast and fluorescence microscopy. In this experiment we use glucose as the carbon source because the presence of glucose is known to reduce the level of PapB protein that represses the type 1 fimbrial expression [124]. Low concentrations of leucine are also added to the medium because it increases *fim* switching rates in bulk experiments [41].

Single cells are grown exponentially in microcolonies with an average division time of about 57 minutes. As mentioned earlier (chapter 4), fluorescence microscopy allows unambiguous discrimination of the ON and OFF states. During microcolony growth, multiple switching events are observed within one microcolony. A typical microcolony originates from a single OFF cell that previously had several ON switching is shown figure 5.1. Multiple ON switching events along the microcolony's growth results in observed patches of ON sub-microcolonies.

Single ON and OFF events are shown in figures 5.2A and C respectively. The corresponding changes in fluorescence level are indicated in figures 5.2B and D respectively. These data were extracted from the microscopy images using image-analysis software [100]. An ON switching event is usually readily observed within the first 10 minutes after the switching time. For example in figure 5.2A, the ON event begins to show in image 'b', where the switching cell starts to become brighter than its neighboring cells. The start of fluorescence increase is not the same as the switching time, but because the correction for GFP maturation time occurs in both time points, only the time difference needs to be considered. An OFF switch is indicated by a decrease in the fluorescence. The fluorescence decrease occurs mainly because of the slow dilution due to growth and division, so an OFF switching event is usually only observed about 1 generation after the switching time. Because we aim to isolate switching events, here we do not concentrate on the precise pattern of GFP expression as determined by the DNA replication process that was discussed in the previous chapter.

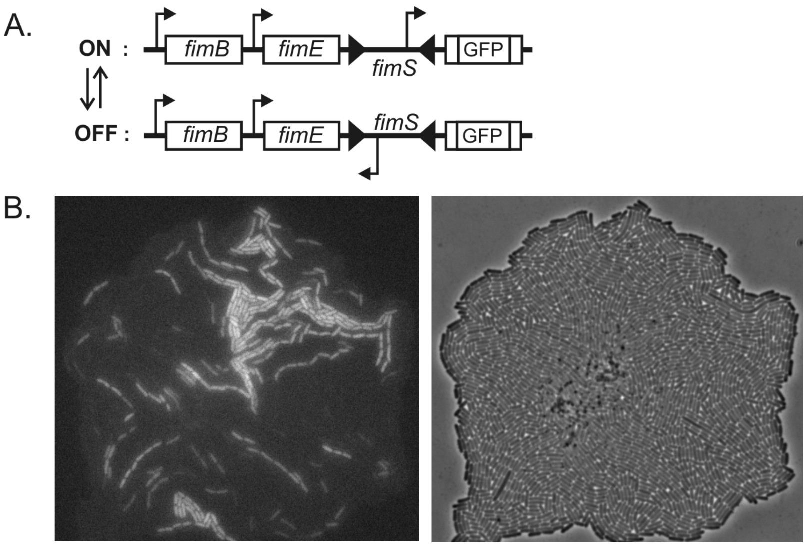

**Figure 5.1** Phase variation at the single cell level.
(A) In the ON orientation, the promoter that is positioned within the invertible element *fimS* drives transcription of the *fim* operon. GFP has been inserted into *fimA*, the first gene of the *fim* operon, allowing single-cell monitoring of the *fim* operon expression level. Transcription is turned OFF upon inversion of *fimS*, which positions the promoter in the opposite direction. This reversible inversion process is performed by two recombinases, FimB and FimE.
(B) A fluorescence (left) and phase contrast (right) image of a microcolony derived from a single OFF progenitor at the end of an experiment. Multiple ON switching events along the microcolony's growth results in observed patches of ON sub-microcolonies. There are more than 1000 cells shown in this image.

First, we aim to estimate the overall *fim* ON and OFF switching rates. Overall ON rates are computed by taking the total number of ON events that occur during colony growth, and dividing it by the total potential switchers for all time intervals. The number of total potential switchers is the number of OFF cells that could switch but do not. Overall OFF rates are determined in the same manner. We correct for the required time to observe a switching event, and for areas in the microcolony that cannot be properly observed, for example by excluding areas where two or more layers of cells grow on top of each other. The measured ON and OFF switching rates are $3.8\pm1.3*10^{-2}$ and $2.7\pm0.41*10^{-2}$ per cell per generation respectively. The measured OFF switching rate is comparable to values previously reported from bulk measurements, but our ON rate is higher [41]. Our result agrees

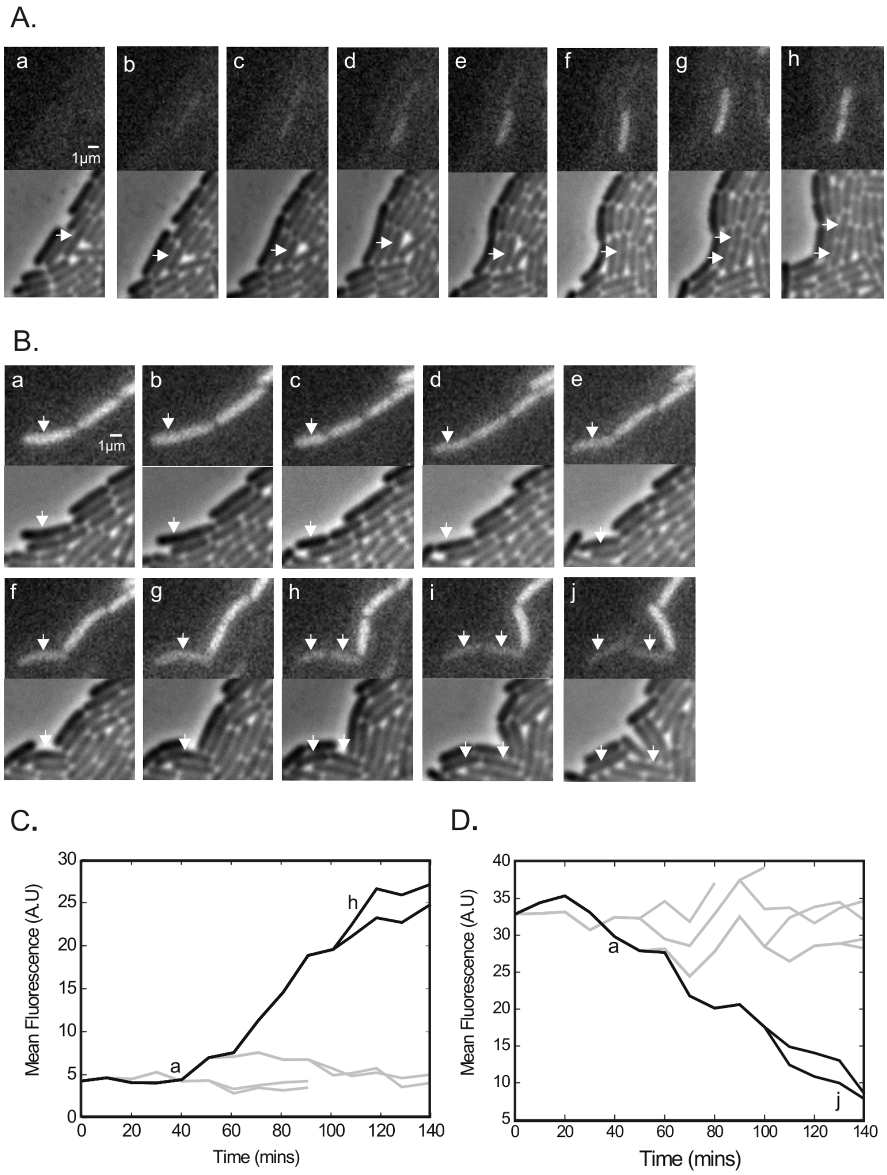
A.
a
b
c
d
e
f
g
h
1µm
B.
a
b
c
d
e
f
g
h
i
j
1µm
C.
D.
Mean Fluorescence (A.U)
Time (mins)
0
5
10
15
20
25
30
0
20
40
60
80
100
120
140
a
h
Mean Fluorescence (A.U)
Time (mins)
5
10
15
20
25
30
35
40
0
20
40
60
80
100
120
140
a
j

**Figure 5.2** ON and OFF switching event in growing microcolonies.
(A) Time-lapse fluorescence (upper panel) and phase contrast images (lower panel) of an ON switching event. The white arrow indicates the switching cell lineage. The time corresponds to the time in figure B as indicated by the letters a to h.
(B) Time-lapse fluorescence (upper panel) and phase contrast images (lower panel) of an OFF switching event. The white arrow indicates the switching cell lineage. The time corresponds to the time in figure D as indicated by the letters a to j.

(C) The quantified fluorescence level of cells in figure A with the switching lineage depicted as black lines. A single branching represents one cell division. Grey lines correspond to other cells within the family.
(D) The quantified fluorescence level of the cells in figure C with the switching lineage depicted as black lines. A single branching represents one cell division. Grey lines correspond to other cells within the family.

with previous work that states that the OFF reverse switching rate is important and should be accounted for in bulk rate measurements [41].

### 5.2.2 The effect of switching history on the OFF switching frequency

After estimating overall ON and OFF switching rates for the *fim* system, we set out to investigate the effect of cell history on the OFF switching rate. Our aim is to determine the ON time: the time between an ON and a subsequent OFF event (fig. 5.3). A large amount of OFF cells are followed in time and screened for ON switching events. After several cells switch ON are observed and their positions are marked, we then apply a smaller imaging timestep to follow the growth of the newly ON sub microcolonies. Typically, during the first few generations, cells were dividing in synchrony (up to the third generation). Over time, reverse switching events to the OFF state were observed within the newly formed ON sub-microcolonies (fig. 5.3B).

In total we recorded 152 ON-OFF traces, from about 1400 ON cells in 55 ON sub-microcolonies. Our data show a relatively constant switching rate with a mean of approximately $3.0 \pm 2.1 * 10^{-2}$ events per cell per generation between 100 and 400 minutes of ON time. The distribution of OFF switching rates shows a peaked shape (fig. 5.4B) and a Poissonian fit to the normalized data gives a correlation coefficient of about 0.8. The data in the last 50 minutes suggest a decrease in the switching rate (fig. 5.4A). However, although we cannot measure it directly, it does not seem likely that this downward trend continues beyond 400 minutes. First, the amount of data in this time regime is limited, making the measurements less precise. Second, the average of the data in figure 5.4A is similar to the switching rate for asynchronous ON cells, which have on average been in the ON state for a significantly longer time than 400 minutes is $2.7 \pm 0.41 * 10^{-2}$ per cell per generation (see dashed line). We therefore conclude that there is no significant deviation from the Poissonian distribution for OFF switching rates within the time window of our measurements. The precision of the distribution at the smaller time scales is lower due to the necessarily smaller number of cells for young sub-microcolonies. At the longest time-scales, the precision is lower because of the formation of multiple layers of cells that reduces the number of observable cells.

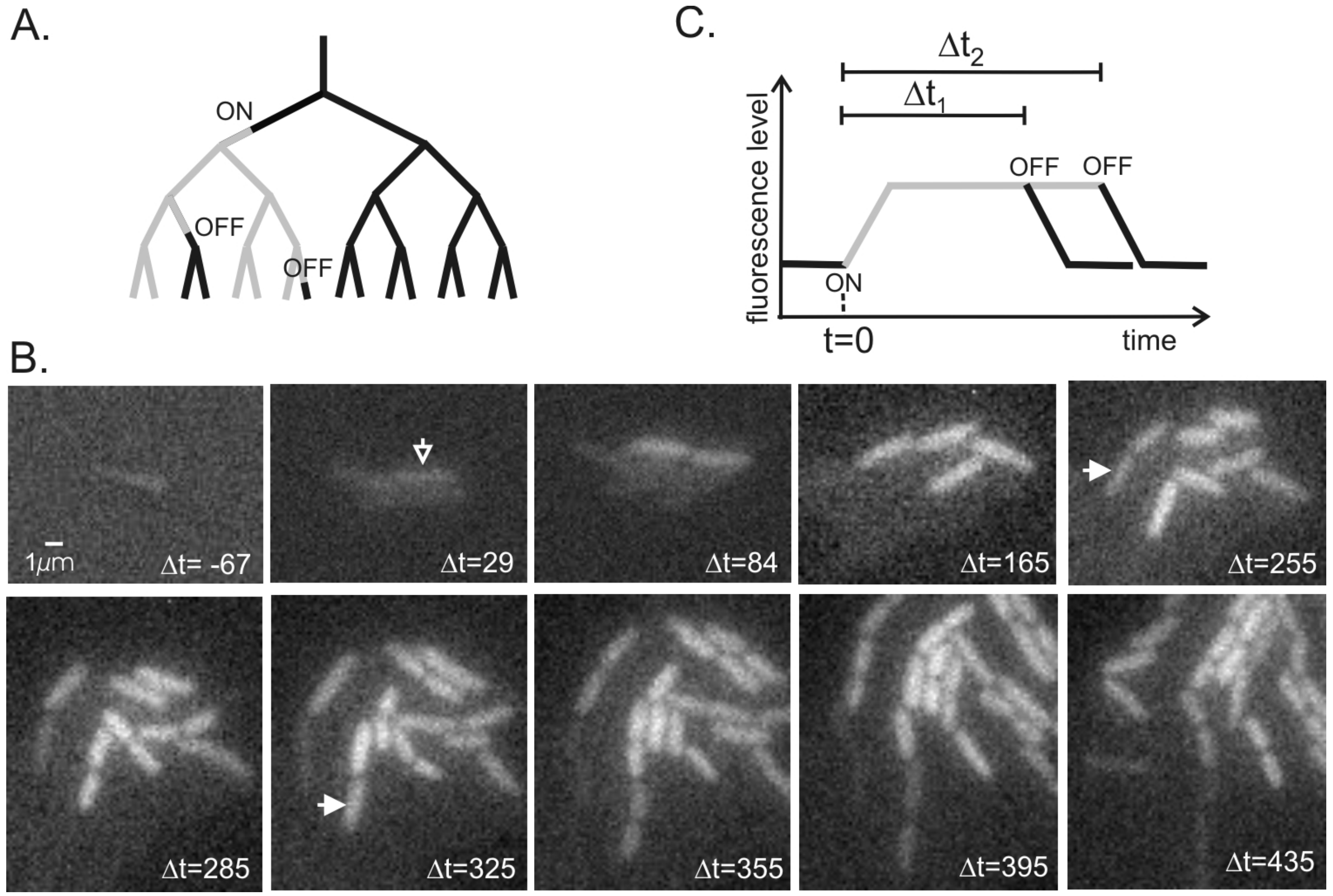A.
ON
OFF
OFF
C.
Δt2
Δt1
fluorescence level
OFF
OFF
ON
t=0
time
B.
1μm
Δt= -67
Δt=29
Δt=84
Δt=165
Δt=255
Δt=285
Δt=325
Δt=355
Δt=395
Δt=435

**Figure 5.3** OFF switching events in a newly ON sub-microcolony.

(A) A schematic family tree of the microcolonies followed in this study. An OFF cell acts as the progenitor of each colony. A single ON switching event indicates the start of an ON sub-microcolony (grey lines). Over time, multiple OFF events are observed within the ON sub-microcolony, resulting in OFF cell lineages (black lines).

(B) The open arrow points at the first ON cell of the sub-microcolony. After several divisions, multiple OFF switching events occurred. The solid arrows points at the newly formed OFF cells.

(C) The duration of a cell lineage being in the ON state before a certain OFF switching event occurs (Δt) is described in this graph. The switch ON is considered to be t=0. Images were generally taken every 10 minutes, except at the beginning a larger timestep was used.

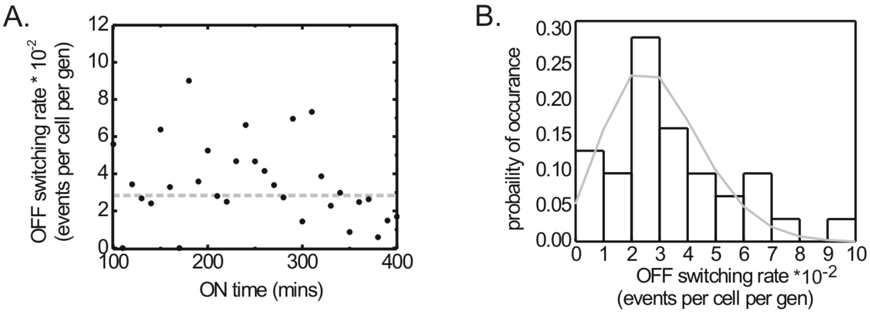

**Figure 5.4** Measured ON to OFF switching rates for ON sub-microlonies in time.

(A) OFF switching rates as a function of ON time per 10 minutes time bin. In total, 152 events were observed from 55 ON sub-microcolonies of different sizes aligned at t=0, which is the time of the ON switching event. Gray dotted lines represent the mean OFF switching rate for non-aligned ON cells (cells that have presumably been ON for a longer period of time).

(B) Normalized histogram of observed OFF switching rate for ON time between 100 and 400 minutes (the total sum is 1). The gray line represents a Poissonian probability mass function fitted to the experimental data (correlation coefficient = 0.8, mean 3.0, var=4.7).

The defining feature of a Poissonian stochastic behavior is the lack of 'memory'. This characteristic is expressed in our experiment as a constant switching probability in time. Models suggesting non-Poissonian behavior for the *fim* switch as a consequence of *fim* orientational control proposed a peak in the OFF switching probability as a function of time spent in the ON state [119,121]. This peak was suggested to occur within the first generation after a cell being switched ON. Due to poor statistics for the first 100 minutes, we cannot determine if there is a significant time period void of OFF switching events during the first moments after cells being turned ON. Such discrimination requires a very large number (several thousands) of ON sub-microcolonies to be followed in time.

In addition to the need for a large number of observed cells, there are more reasons why OFF switching events at early ON times is experimentally difficult to detect with our construct. Depending on when the ON switching occurs within the cell cycle, the epigenetic inheritance of GFP into cell lineages that has only OFF *fim* copies (described in chapter 4) masks some of the rare OFF switching events at early times. This particular case will arise as a result of an ON event that occurs before *fim* duplication and an OFF event after *fim* duplication. In another scenario, when both ON and OFF switching events occur after *fim* duplication, an early OFF event will appear as quick transient brightness, where a cell becomes slightly fluorescent and then loses its fluorescence again. A short imaging time-step is required to be able to detect such events reliably.

Another possibility to detect an early OFF event is the scenario when the ON event occurs after *fim* duplication in one cell cycle and the OFF event occurs after *fim* duplication in the next cell cycle. In this particular scenario, the number of daughter lineages that show transient brightness is increased by one. To illustrate such a scenario, consider a cell with a interdivision time of 60 minutes where DNA replication starts at the moment of division according to the Cooper-Helmstetter model of DNA replication. After duplication, there are a maximum of 2 copies of *fim*. An ON switching event after *fim* duplication will normally result in one daughter lineage with transient brightness. In the next cell cycle, after the ON copy of *fim* is duplicated, an OFF event after the duplication will cause one additional cell lineage with transient brightness.

## 5.3 Discussion and outlook

Phase variation has long been regarded as a purely random process in nature. However, direct evidence of this assumption has never been presented before. Bulk studies are not appropriate to address this question due to the averaging that takes place. In this study, we explore the possibility of phase variation as a non-Poissonian process, using *fim* phase variation as a model system. By using a single cell approach and recording cell lineage history over time as switching events occur, we provide the first results from such attempts. We follow newly formed ON sub-microcolonies and measure the OFF switching events over time. Our results indicate a constant OFF switching rate that does not depend on the switching history of the cells from about 100 minutes (approximately 2 interdivision times) to about 400 minutes. However, further investigations in other time regimes (very close to the starting ON event or much longer after) are required to fully address the posed questions.

In order to further study the switching behavior of the *fim* switch, investigations over different timescales would be of much interest. Due to the problems mentioned above, investigation of switching frequencies for shorter timescales (within the first 2 generations after an ON event) requires a slightly different approach. One possible approach is to construct a strain in which a second fluorescent protein is expressed when the *fimS* promoter is in the OFF orientation. This approach would allow the differentiation between transient brightness due to chromosome copy number effect and an OFF switching event. However, a trickier problem to solve is the statistical requirements of a large amount of newly switched ON cells. This requirement implies another requirement for an even larger number of OFF cells from which the ON cells originated. To reduce this number, an experimental condition that increases *fim* switching rates is then preferred, or the use of a mutant that has an increased ON switching rate is of much interest. A high density of cells packed in micro-structures might act as an alternative approach.

Moreover, some of the open questions related to *fim* switching behavior might be addressed by following the dynamics of the recombinase level itself in time. Since recombinases are expressed in relatively low levels *in vivo*, a different approach of labelling from the one applied here for *fimS* is required. A membrane bound fluorescent protein has recently been used to probe low gene expression in live cells [127], and thus might provide a way to follow FimE expression dynamics in *E.coli*.

## References

1. Adachi S, Fukushima T, Hiraga S (2008) Dynamic events of sister chromosomes in the cell cycle of Escherichia coli. Genes Cells 13: 181-197.

2. Atlung T, Lobner-Olesen A, Hansen FG (1987) Overproduction of DnaA protein stimulates initiation of chromosome and minichromosome replication in Escherichia coli. Mol Gen Genet 206: 51-59.

3. Avery SV (2006) Microbial cell individuality and the underlying sources of heterogeneity. Nat Rev Microbiol 4: 577-587.

4. Bach T, Krekling MA, Skarstad K (2003) Excess SeqA prolongs sequestration of oriC and delays nucleoid segregation and cell division. Embo J 22: 315-323.

5. Bach T, Krekling MA, Skarstad K (2003) Excess SeqA prolongs sequestration of oriC and delays nucleoid segregation and cell division. Embo J 22: 315-323.

6. Barcena M, Ruiz T, Donate LE, Brown SE, Dixon NE, et al. (2001) The DnaB.DnaC complex: a structure based on dimers assembled around an occluded channel. Embo J 20: 1462-1468.

7. Bates D, Kleckner N (2005) Chromosome and replisome dynamics in E. coli: loss of sister cohesion triggers global chromosome movement and mediates chromosome segregation. Cell 121: 899-911.

8. Bayne-Jones S, Adoplh EF (1932) Growth in size of micro-organisms measured from motion pictures. Journal of Cellular and Comparative Physiology 1: 387-407.

9. Benzer S (1953) Induced synthesis of enzymes in bacteria analyzed at the cellular level. Biochim Biophys Acta 11: 383-395.

10. Blomfield I (2001) The Regulation of Pap and Type 1 Fimbriation in Escherichia coli: Academic Press. 1-49 p.

11. Blomfield IC, McClain MS, Princ JA, Calie PJ, Eisenstein BI (1991) Type 1 fimbriation and fimE mutants of Escherichia coli K-12. J Bacteriol 173: 5298-5307.

12. Blomfield IC, Vaughn V, Rest RF, Eisenstein BI (1991) Allelic exchange in Escherichia coli using the Bacillus subtilis sacB gene and a temperature-sensitive pSC101 replicon. Mol Microbiol 5: 1447-1457.

13. Boye E, Stokke T, Kleckner N, Skarstad K (1996) Coordinating DNA replication initiation with cell growth: differential roles for DnaA and SeqA proteins. Proc Natl Acad Sci U S A 93: 12206-12211.

14. Boye E, Nordstrom K (2003) Coupling the cell cycle to cell growth. EMBO Rep 4: 757-760.

15. Brehm-Stecher BF, Johnson EA (2004) Single-cell microbiology: tools, technologies, and applications. Microbiol Mol Biol Rev 68: 538-559.

16. Bremer H, Churchward G (1977) Deoxyribonucleic acid synthesis after inhibition of initiation of rounds of replication in Escherichia coli B/r. J Bacteriol 130: 692-697.

17. Bremer H, Churchward G (1978) Age fractionation in bacteria by membrane elution: relation between age distribution and elution profile. J Theor Biol 74: 69-81.

18. Bremer H, Chuang L (1981) The cell cycle in Escherichia coli B/r. J Theor Biol 88: 47-81.

19. Brendler T, Sawitzke J, Sergueev K, Austin S (2000) A case for sliding SeqA tracts at anchored replication forks during Escherichia coli chromosome replication and segregation. Embo J 19: 6249-6258.

20. Bullitt E, Makowski L (1995) Structural polymorphism of bacterial adhesion pili. Nature 373: 164-167.

21. Burns LS, Smith SG, Dorman CJ (2000) Interaction of the FimB integrase with the fimS invertible DNA element in Escherichia coli in vivo and in vitro. J Bacteriol 182: 2953-2959.

22. Cairns J (1963) The bacterial chromosome and its manner of replication as seen by autoradiography. J Mol Biol 6: 208-213.

23. Campbell JL, Kleckner N (1990) E. coli oriC and the dnaA gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork. Cell 62: 967-979.

24. Chandler M, Bird RE, Caro L (1975) The replication time of the Escherichia coli K12 chromosome as a function of cell doubling time. J Mol Biol 94: 127-132.

25. Chu D, Blomfield IC (2007) Orientational control is an efficient control mechanism for phase switching in the E. coli fim system. J Theor Biol 244: 541-551.

26. Churchward G, Estiva E, Bremer H (1981) Growth rate-dependent control of chromosome replication initiation in Escherichia coli. J Bacteriol 145: 1232-1238.

27. Cooper S, Helmstetter CE (1968) Chromosome replication and the division cycle of Escherichia coli B/r. J Mol Biol 31: 519-540.

28. Cooper S (1991) Bacterial Growth and Division. San Diego: Academic Press, Inc.

29. Cooper S (2006) Regulation of DNA synthesis in bacteria: Analysis of the Bates/Kleckner licensing/initiation-mass model for cell cycle control. Mol Microbiol 62: 303-307.

30. Cormack BP, Valdivia RH, Falkow S (1996) FACS-optimized mutants of the green fluorescent protein (GFP). Gene 173: 33-38.

31. Davidsen T, Tonjum T (2006) Meningococcal genome dynamics. Nat Rev Microbiol 4: 11-22.

32. den Blaauwen T, Aarsman ME, Wheeler LJ, Nanninga N (2006) Pre-replication assembly of E. coli replisome components. Mol Microbiol 62: 695-708.

33. Donachie WD (1968) Relationship between cell size and time of initiation of DNA replication. Nature 219: 1077-1079.

34. Donachie WD (1969) Control of cell division in Escherichia coli: experiments with thymine starvation. J Bacteriol 100: 260-268.

35. Dove SL, Dorman CJ (1994) The site-specific recombination system regulating expression of the type 1 fimbrial subunit gene of Escherichia coli is sensitive to changes in DNA supercoiling. Mol Microbiol 14: 975-988.

36. Dybvig K (1993) DNA rearrangements and phenotypic switching in prokaryotes. Mol Microbiol 10: 465-471.

37. Eisenstein BI (1981) Phase variation of type 1 fimbriae in Escherichia coli is under transcriptional control. Science 214: 337-339.

38. El-Labany S, Sohanpal BK, Lahooti M, Akerman R, Blomfield IC (2003) Distant cis-active sequences and sialic acid control the expression of fimB in Escherichia coli K-12. Mol Microbiol 49: 1109-1118.

39. Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. Science 297: 1183-1186.

40. Fujikawa N, Kurumizaka H, Nureki O, Tanaka Y, Yamazoe M, et al. (2004) Structural and biochemical analyses of hemimethylated DNA binding by the SeqA protein. Nucleic Acids Res 32: 82-92.

41. Gally DL, Bogan JA, Eisenstein BI, Blomfield IC (1993) Environmental regulation of the fim switch controlling type 1 fimbrial phase variation in Escherichia coli K-12: effects of temperature and media. J Bacteriol 175: 6186-6193.

42. Gally DL, Leathart J, Blomfield IC (1996) Interaction of FimB and FimE with the fim switch that controls the phase variation of type 1 fimbriae in Escherichia coli K-12. Mol Microbiol 21: 725-738.

43. Grindley ND, Whiteson KL, Rice PA (2006) Mechanisms of site-specific recombination. Annu Rev Biochem 75: 567-605.

44. Guarne A, Brendler T, Zhao Q, Ghirlando R, Austin S, et al. (2005) Crystal structure of a SeqA-N filament: implications for DNA replication and chromosome organization. Embo J 24: 1502-1511.

45. Hahn E, Wild P, Hermanns U, Sebbel P, Glockshuber R, et al. (2002) Exploring the 3D molecular architecture of Escherichia coli type 1 pili. J Mol Biol 323: 845-857.

46. Hanks MC, Masters M (1987) Transductional analysis of chromosome replication time. Mol Gen Genet 210: 288-293.

47. Helmstetter CE, Pierucci O (1976) DNA synthesis during the division cycle of three substrains of Escherichia coli B/r. J Mol Biol 102: 477-486.

48. Henderson IR, Owen P, Nataro JP (1999) Molecular switches--the ON and OFF of bacterial phase variation. Mol Microbiol 33: 919-932.

49. Herrick J, Kohiyama M, Atlung T, Hansen FG (1996) The initiation mess? Mol Microbiol 19: 659-666.

50. Hinde P, Deighan P, Dorman CJ (2005) Characterization of the detachable Rho-dependent transcription terminator of the fimE gene in Escherichia coli K-12. J Bacteriol 187: 8256-8266.

51. Hiraga S, Ichinose C, Niki H, Yamazoe M (1998) Cell cycle-dependent duplication and bidirectional migration of SeqA-associated DNA-protein complexes in E. coli. Mol Cell 1: 381-387.

52. Hiraga S, Ichinose C, Onogi T, Niki H, Yamazoe M (2000) Bidirectional migration of SeqA-bound hemimethylated DNA clusters and pairing of oriC copies in Escherichia coli. Genes Cells 5: 327-341.

53. Holden N, Blomfield IC, Uhlin BE, Totsika M, Kulasekara DH, et al. (2007) Comparative analysis of FimB and FimE recombinase activity. Microbiology 153: 4138-4149.

54. Jacob-Dubuisson F, Heuser J, Dodson K, Normark S, Hultgren S (1993) Initiation of assembly and association of the structural elements of a bacterial pilus depend on two specialized tip proteins. Embo J 12: 837-847.

55. Jones CH, Pinkner JS, Roth R, Heuser J, Nicholes AV, et al. (1995) FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. Proc Natl Acad Sci U S A 92: 2081-2085.

56. Joyce SA, Dorman CJ (2002) A Rho-dependent phase-variable transcription terminator controls expression of the FimE recombinase in Escherichia coli. Mol Microbiol 45: 1107-1117.

57. Kaguni JM (2006) DnaA: controlling the initiation of bacterial DNA replication and more. Annu Rev Microbiol 60: 351-375.

58. Kang S, Lee H, Han JS, Hwang DS (1999) Interaction of SeqA and Dam methylase on the hemimethylated origin of Escherichia coli chromosomal DNA replication. J Biol Chem 274: 11463-11468.

59. Kang S, Han JS, Kim KP, Yang HY, Lee KY, et al. (2005) Dimeric configuration of SeqA protein bound to a pair of hemi-methylated GATC sequences. Nucleic Acids Res 33: 1524-1531.

60. Kawakami H, Su'etsugu M, Katayama T (2006) An isolated Hda-clamp complex is functional in the regulatory inactivation of DnaA and DNA replication. J Struct Biol 156: 220-229.

61. Kelly A, Conway C, T OC, Smith SG, Dorman CJ (2006) DNA supercoiling and the Lrp protein determine the directionality of fim switch DNA inversion in Escherichia coli K-12. J Bacteriol 188: 5356-5363.

62. Keyamura K, Fujikawa N, Ishida T, Ozaki S, Su'etsugu M, et al. (2007) The interaction of DiaA and DnaA regulates the replication cycle in E. coli by directly promoting ATP DnaA-specific initiation complexes. Genes Dev 21: 2083-2099.

63. Koppes LH, Woldringh CL, Nanninga N (1978) Size variations and correlation of different cell cycle events in slow-growing Escherichia coli. J Bacteriol 134: 423-433.

64. Koppes LJ, Meyer M, Oonk HB, de Jong MA, Nanninga N (1980) Correlation between size and age at different events in the cell division cycle of Escherichia coli. J Bacteriol 143: 1241-1252.

65. Kornberg A, Baker J (1992) DNA Replication. New York: W.H Freeman & Co.

66. Kubitschek HE, Freedman ML (1971) Chromosome replication and the division cycle of Escherichia coli B-r. J Bacteriol 107: 95-99.

67. Kubitschek HE (1974) Estimation of the D period from residual division after exposure of exponential phase bacteria to chloramphenicol. Mol Gen Genet 135: 123-130.

68. Kubitschek HE, Newman CN (1978) Chromosome replication during the division cycle in slowly growing, steady-state cultures of three Escherichia coli B/r strains. J Bacteriol 136: 179-190.

69. Kubitschek HE, Woldringh CL (1983) Cell elongation and division probability during the Escherichia coli growth cycle. J Bacteriol 153: 1379-1387.

70. Kussell E, Leibler S (2005) Phenotypic diversity, population growth, and information in fluctuating environments. Science 309: 2075-2078.

71. Lane HE, Denhardt DT (1975) The rep mutation. IV. Slower movement of replication forks in Escherichia coli rep strains. J Mol Biol 97: 99-112.

72. Leathart JB, Gally DL (1998) Regulation of type 1 fimbrial expression in uropathogenic Escherichia coli: heterogeneity of expression through sequence changes in the fim switch region. Mol Microbiol 28: 371-381.

73. Longo D, Hasty J (2006) Dynamics of single-cell gene expression. Mol Syst Biol 2: 64.

74. Louarn J, Funderburgh M, Bird RE (1974) More precise mapping of the replication origin in Escherichia coli K-12. J Bacteriol 120: 1-5.

75. Low DA, Weyand NJ, Mahan MJ (2001) Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. Infect Immun 69: 7197-7204.

76. Lu M, Campbell JL, Boye E, Kleckner N (1994) SeqA: a negative modulator of replication initiation in E. coli. Cell 77: 413-426.

77. Luria SE, Delbruck M (1943) Mutations of Bacteria from Virus Sensitivity to Virus Resistance. Genetics 28: 491-511.

78. Maaloe O, Hanawalt PC (1961) Thymine deficiency and the normal DNA replication cycle. I. J Mol Biol 3: 144-155.

79. Maamar H, Raj A, Dubnau D (2007) Noise in gene expression determines cell fate in Bacillus subtilis. Science 317: 526-529.

80. Manor H, Deutscher MP, Littauer UZ (1971) Rates of DNA chain growth in Escherichia coli. J Mol Biol 61: 503-524.

81. Marunouchi T, Messer W (1973) Replication of a specific terminal chromosome segment in Escherichia coli which is required for cell division. J Mol Biol 78: 211-228.

82. Meury J, Bahloul A, Kohiyama M (1995) Importance of the replication origin sequestration in cell division of Escherichia coli. Biochimie 77: 875-879.

83. Michelsen O, Teixeira de Mattos MJ, Jensen PR, Hansen FG (2003) Precise determinations of C and D periods by flow cytometry in Escherichia coli K-12 and B/r. Microbiology 149: 1001-1010.

84. Miller J (1972) Experiments in Molecular Genetics. New York: Cold Spring Harbor Laboratory Press.

85. Molina F, Skarstad K (2004) Replication fork and SeqA focus distributions in Escherichia coli suggest a replication hyperstructure dependent on nucleotide metabolism. Mol Microbiol 52: 1597-1612.

86. Mott ML, Berger JM (2007) DNA replication initiation: mechanisms and regulation in bacteria. Nat Rev Microbiol 5: 343-354.

87. Mumm JP, Landy A, Gelles J (2006) Viewing single lambda site-specific recombination events from start to finish. Embo J 25: 4586-4595.

88. Newman CN, Kubitschek HE (1978) Variation in periodic replication of the chromosome in Escherichia coli B/rTT. J Mol Biol 121: 461-471.

89. Nielsen O, Lobner-Olesen A (2008) Once in a lifetime: strategies for preventing re-replication in prokaryotic and eukaryotic cells. EMBO Rep 9: 151-156.

90. Norris TL, Baumler AJ (1999) Phase variation of the lpf operon is a mechanism to evade cross-immunity between Salmonella serotypes. Proc Natl Acad Sci U S A 96: 13393-13398.

91. Onogi T, Niki H, Yamazoe M, Hiraga S (1999) The assembly and migration of SeqA-Gfp fusion in living cells of Escherichia coli. Mol Microbiol 31: 1775-1782.

92. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A (2002) Regulation of noise in the expression of a single gene. Nat Genet 31: 69-73.

93. Painter PR (1974) The relative numbers of different genes in exponential microbial cultures. Genetics 76: 401-410.

94. Pardee AB (1974) A restriction point for control of normal animal cell proliferation. Proc Natl Acad Sci U S A 71: 1286-1290.

95. Powell EO, Errington FP (1963) Generation times of individual bacteria: some corroborative measurements. J Gen Microbiol 31: 315-327.

96. Pritchard RH, Barth PT, Collins J (1969) Control of DNA synthesis in bacteria. SympSocGenMicrobiol 19 263-297.

97. Pritchard RH, Zaritsky A (1970) Effect of thymine concentration on the replication velocity of DNA in a thymineless mutant of Escherichia coli. Nature 226: 126-131.

98. Ramsey S, Ozinsky A, Clark A, Smith KD, de Atauri P, et al. (2006) Transcriptional noise and cellular heterogeneity in mammalian macrophages. Philos Trans R Soc Lond B Biol Sci 361: 495-506.

99. Remaut H, Tang C, Henderson NS, Pinkner JS, Wang T, et al. (2008) Fiber formation across the bacterial outer membrane by the chaperone/usher pathway. Cell 133: 640-652.

100. Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB (2005) Gene regulation at the single-cell level. Science 307: 1962-1965.

101. Ryan VT, Grimwade JE, Nievera CJ, Leonard AC (2002) IHF and HU stimulate assembly of pre-replication complexes at Escherichia coli oriC by two different mechanisms. Mol Microbiol 46: 113-124.

102. Sakakibara Y, Yuasa S (1982) Continuous synthesis of the dnaA gene product of Escherichia coli in the cell cycle. Mol Gen Genet 186: 87-94.

103. Sambrook J, Russell DW (2001) Molecular Cloning A Laboratory Manual USA: Cold Spring Harbor Laboratory Press.

104. Schaechter M, Maaloe O, Kjeldgaard NO (1958) Dependency on medium and temperature of cell size and chemical composition during balanced grown of Salmonella typhimurium. J Gen Microbiol 19: 592-606.

105. Schaechter M, Williamson JP, Hood JR, Jr., Kochal (1962) Growth, cell and nuclear divisions in some bacteria. J Gen Microbiol 29: 421-434.

106. Schilling JD, Mulvey MA, Hultgren SJ (2001) Structure and function of Escherichia coli type 1 pili: new insight into the pathogenesis of urinary tract infections. J Infect Dis 183 Suppl 1: S36-40.

107. Sekimizu K, Bramhill D, Kornberg A (1988) Sequential early stages in the in vitro initiation of replication at the origin of the Escherichia coli chromosome. J Biol Chem 263: 7124-7130.

108. Skarstad K, Steen HB, Boye E (1985) Escherichia coli DNA distributions measured by flow cytometry and compared with theoretical computer simulations. J Bacteriol 163: 661-668.

109. Slater S, Wold S, Lu M, Boye E, Skarstad K, et al. (1995) E. coli SeqA protein binds oriC in two different methyl-modulated reactions appropriate to its roles in DNA replication initiation and origin sequestration. Cell 82: 927-936.

110. Sohanpal BK, Kulasekara HD, Bonnen A, Blomfield IC (2001) Orientational control of fimE expression in Escherichia coli. Mol Microbiol 42: 483-494.

111. Sohanpal BK, El-Labany S, Lahooti M, Plumbridge JA, Blomfield IC (2004) Integrated regulatory responses of fimB to N-acetylneuraminic (sialic) acid and GlcNAc in Escherichia coli K-12. Proc Natl Acad Sci U S A 101: 16322-16327.

112. Sohanpal BK, Friar S, Roobol J, Plumbridge JA, Blomfield IC (2007) Multiple co-regulatory elements and IHF are necessary for the control of fimB expression in response to sialic acid and N-acetylglucosamine in Escherichia coli K-12. Mol Microbiol 63: 1223-1236.

113. Stewart EJ, Madden R, Paul G, Taddei F (2005) Aging and death in an organism that reproduces by morphologically symmetric division. PLoS Biol 3: e45.

114. Stocker B (1949) Measurements of rate of mutation of flagellar antigenic phase in Salmonella typhi-murium. J Hyg 47: 398-413.

115. Thattai M, van Oudenaarden A (2004) Stochastic gene expression in fluctuating environments. Genetics 167: 523-530.

116. van der Woude MW, Baumler AJ (2004) Phase and antigenic variation in bacteria. Clin Microbiol Rev 17: 581-611.

117. Veening JW, Smits WK, Kuipers OP (2008) Bistability, epigenetics, and bet-hedging in bacteria. Annu Rev Microbiol 62: 193-210.

118. Veening JW, Stewart EJ, Berngruber TW, Taddei F, Kuipers OP, et al. (2008) Bet-hedging and epigenetic inheritance in bacterial cell development. Proc Natl Acad Sci U S A 105: 4393-4398.

119. Visco P, Allen RJ, Evans MR (2008) Exact Solution of a Model DNA-Inversion Genetic Switch with Orientational Control. Physical Review Letters 101: 11810041-11810044.

120. Wold S, Skarstad K, Steen HB, Stokke T, Boye E (1994) The initiation mass for DNA replication in Escherichia coli K-12 is dependent on growth rate. Embo J 13: 2097-2102.

121. Wolf DM, Arkin AP (2002) Fifteen minutes of fim: control of type 1 pili expression in E. coli. Omics 6: 91-114.

122. Wolf DM, Vazirani VV, Arkin AP (2005) Diversity in times of adversity: probabilistic strategies in microbial survival games. J Theor Biol 234: 227-253.

123. Wolf DM, Vazirani VV, Arkin AP (2005) A microbial modified prisoner's dilemma game: how frequency-dependent selection can lead to random phase variation. J Theor Biol 234: 255-262.

124. Xia Y, Gally D, Forsman-Semb K, Uhlin BE (2000) Regulatory cross-talk between adhesin operons in Escherichia coli: inhibition of type 1 fimbriae expression by the PapB protein. Embo J 19: 1450-1457.

125. Yamazoe M, Adachi S, Kanaya S, Ohsumi K, Hiraga S (2005) Sequential binding of SeqA protein to nascent DNA segments at replication forks in synchronized cultures of Escherichia coli. Mol Microbiol 55: 289-298.

126. Yoshikawa H, O'Sullivan A, Sueoka N (1964) Sequential Replication of the Bacillus Subtilis Chromosome. 3. Regulation of Initiation. Proc Natl Acad Sci U S A 52: 973-980.

127. Yu J, Xiao J, Ren X, Lao K, Xie XS (2006) Probing gene expression in live cells, one protein molecule at a time. Science 311: 1600-1603.

128. Zaritsky A (1975) Rate stimulation of deoxyribonucleic acid synthesis after inhibition. J Bacteriol 122: 841-846.

## Summary

The first investigations of cell-to-cell heterogeneity in isogenic populations date back decades, with the early investigation of bi-modal β-galactosidase expression in batch cultures being one example. As a result of recent technological developments such as time-lapse microscopy, fluorescent labeling, and image-analysis tools, cell-to-cell heterogeneity has recently received an increased interest.

In this thesis we report the results from our investigation of two different processes. First, we examined the cell-to-cell heterogeneity of *Escherichia coli* cell cycle and DNA replication process. Second, we investigated the phase variation of *Escherichia coli* type 1 fimbrial expression.

In chapter 2 we report the experimental procedures that were developed to obtain the results presented in later chapters. To follow DNA replication within the cell cycle, SeqA protein was fused to fluorescent mCherry protein. SeqA is a DNA sequestering protein involved in negative regulation of DNA replication initiation by binding to nascent and hemimethylated DNA. *In vivo*, SeqA forms visible discrete foci when fluorescently labeled, and allows one to track the location of DNA replication forks. To follow the phase varying type 1 fimbrial expression dynamics, we place *gfpmut2* gene under the control of the chromosomal *fim* switch. Time-lapse microscopy combined with fluorescent labeling allowed us to follow the growth of single cells into microcolonies as well as their DNA replication and *fim* switching dynamics.

In chapter 3 we focus on cell-to-cell heterogeneity of the division cycle and on DNA replication process in individual slow growing *E.coli* cells. In *E.coli*, DNA replication starts at the origin of the replication site *oriC*, and finishes at the termination site *ter*. Traditionally, individual events within the division and DNA replication cycle are studied separately and correlations between the two processes are inferred from modelling. This approach lead to the concept of initiation mass, in which DNA replication initiation is proposed to occur at a constant critical mass per origin of replication. However, separate investigations do not allow one to correlate two or more events or quantities pertaining to a single cell cycle. The ability to follow DNA replication and cell growth simultaneously offers a unique opportunity to investigate such correlations directly.

The investigation of SeqA dynamics in slow growing cells has shown both the appearance and disappearance of SeqA foci within one cell cycle according to the following regular pattern in the number of foci: 0-1-2-1-0. This pattern is consistent with the absence of replication at the start and end of the cell cycle as

well as bidirectional movements of two replication forks in between. We concluded that the ratio of SeqA foci to the number of replication fork is 1:1.

We observed a strong anticorrelation between cell interdivision time and its elongation rate; a relationship that has not, to our knowledge, been previously reported. Interestingly, we find that the larger cells in the population predominantly exhibit smaller interdivision times, while the smaller cells predominantly exhibit larger interdivision times. Furthermore, there is only a small variance in cell size at the start of DNA replication. Together with the observation of large variation in the length of the B period (the period between cell birth and the start of replication) and a correlation between the B period and interdivision time, our results support the critical mass model. Finally, we observed that the C period is relatively invariant, as opposed to the very variable B period and D period (the period between the end of replication and cell division).

In chapter 4 we report the results of the first real-time study of type 1 fimbrial phase variation in *E. coli* at the single cell level. Phase variation of antigens and of other proteins expressed on the cell surface largely occurs in pathogens. Classically viewed as a mechanism to avoid host immune system, phase variation involves a heritable and reversible ‘all-or-nothing’ expression of one or multiple sets of genes. Switching between states is considered to be stochastic in nature, resulting in a mixed population of different states.

In *E.coli*, the expression of type-1 pili (fimbriae) is phase variable. The promotor for *fim* expression is positioned in the invertible 314 bps DNA fragment (*fimS)*. Upon inversion, the promoter sequence in *fimS* changes orientation, and hence acts as a reversible genetic switch. In the ON orientation, *fimS* drives the expression of multiple fimbrial structural genes. In the inverted OFF orientation *fim* genes are not expressed. The inversion of *fimS* is performed by two site-specific recombinases, FimB and FimE. While FimB is able to switch in both directions, FimE has an extreme bias to switch towards the OFF orientation. Bulk studies have produced a detailed understanding of the regulation of the *fim* system. However, this approach cannot capture the switching dynamics nor the correlation between switching events and other cell processes (such as DNA replication).

Bulk experimental methods on the *fim* system monitor the ratio of the *E.coli* population in different states to infer the occurrence of switching events. Here, we place GFP (Green Fluorescent Protein) under the control of the *fim* switch, and perfome timelapse phase contrast and fluorescence microscopy of growing microcolonies. We directly followed individual switching events and observed distinct GFP expression patterns following such an event among genealogically related cell lineages. In addition to the expected monotonic increase in brightness, we also see lineages with transient brightness. By simultaneously monitoring the progress of replication, we show that these patterns arise from the inheritance of different *fim* states by daughter cells. Our data is consistent with multiple replication forks as described by the Cooper-Helmstetter model. Furthermore, we find a dependency of OFF to ON switching on cell age, with a higher probability of switching at the beginning of the cell cycle, suggesting a possible correlation between *fim* switching and chromosome replication.

In chapter 5 we explore how *fim* switching is influenced by the cell history. Early investigations on phase variation have assumed that phase variation is a purely random process that follows Poissonian statistics. However, this assumption has never been tested directly. Recent phase variation studies suggest various possible mechanisms that would lead to non-Poissonian switching behaviour. One example is the dependence of FimE expression on the state of *fimS*. Such regulation has been postulated to affect the OFF switching probability in time due to the time required for FimE to achieve steady state level.

We followed a single cell in a growing microcolony that just switched to the ON state and measured the time it remained ON. The data indicates that the OFF switching rate is constant over time and is independent of the switching history, thus verifying the Poissonian nature of *fim* switching.

## Samenvatting

De eerste ideeën over heterogeniteit tussen cellen in isogene populaties zijn al tientallen jaren oud, bijvoorbeeld uit het eerste onderzoek naar bimodale β-galactosidase-expressie in een *batch culture*. Deze heterogeniteit is opnieuw onder de aandacht gekomen dankzij recente technologische ontwikkelingen als *time-lapse* microscopie, fluorescente labeling en beeldanalysesoftware.

In dit proefschrift rapporteren we de resultaten van ons onderzoek naar twee verschillende processen in *Escherichia coli*. Het eerste is de cel-tot-cel heterogeniteit van de celcyclus en het DNA-replicatieproces en het tweede is de fasevariatie van type 1 fimbrische expressie.

In hoofdstuk 2 beschrijven we de experimentele methoden die zijn ontwikkeld om de resultaten in de latere hoofdstukken te vergaren. Om de replicatie van DNA tijdens de celcyclus te volgen is het fluorescente *mCherry* eiwit gefuseerd met *SeqA* eiwit. *SeqA* is een DNA-sekwestratie eiwit dat betrokken is bij de negatieve regulatie van de aanvang van DNA-replicatie door binding aan vers en hemi-gemethyleerd DNA. *In vivo* verzamelt fluorescent gelabeld *SeqA* zich zichtbaar op enkele punten met een hoge concentratie, waardoor de locatie van de replicatievertakking waargenomen kan worden. Om de fase-afhankelijke type 1 fimbrische expressie te volgen hebben we het *gfpmut2* gen onder de invloed van de chromosomale *fim*-schakelaar geplaatst. De combinatie van fluorescente labeling en *time-lapse* microscopie heeft ons in staat gesteld om tegelijkertijd de groei van een enkele cel tot een microkolonie en de DNA-replicatie en dynamica van de *fim*-schakeling te volgen.

In hoofdstuk 3 richten we ons op de de cel-tot-cel heterogeniteit van de celdelingscyclus en het DNA-replicatieproces in individuele langzaam groeiende *E.coli*-cellen. In *E.coli* vangt de DNA-replicatie aan bij de replicatieoorsprong *oriC* en eindigt deze bij het eindpunt *ter*. Gewoonlijk worden afzonderlijke gebeurtenissen in de celdelings- en DNA-replicatiecycli afzonderlijk bestudeerd en worden correlaties tussen beide afgeleid uit modellen. Deze aanpak heeft geleid tot het concept van de initiatiemassa, waarbij de aanvang van de DNA-replicatie zou plaatsvinden bij een constante kritische massa per replicatieoorsprong. De onafhankelijke metingen geven echter geen inzicht in de correlatie tussen gebeurtenissen of hoeveelheden binnen eenzelfde celcyclus. De mogelijkheid om DNA-replicatie en celgroei gelijktijdig te observeren biedt een unieke kans om zulke correlaties direct te onderzoeken.

Het onderzoek naar *SeqA*-dynamica in langzaam groeiende cellen heeft aangetoond dat het aantal locaties met een hoge concentratie in de loop van de celcyclus altijd het volgende patroon volgt: 0-1-2-1-0. Dit patroon is consistent met het ontbreken van replicatie aan het begin en einde van de celcyclus en bidirectionele beweging van de replicatievertakkingen daartussenin. We concluderen dat elke lokaal verhoogde concentratie van *SeqA* overeenkomt met een replicatievertakking.

We hebben een sterke anticorrelatie waargenomen tussen de groeisnelheid van de cel en de tijd tussen celdelingen – een verband dat naar ons weten nog niet eerder is gerapporteerd. We hebben gezien dat de grotere cellen uit een populatie overwegend minder tijd tussen celdelingen hebben dan de kleinere cellen. Bovendien hebben we slechts een kleine variatie in celgrootte waargenomen bij aanvang van de DNA-replicatie. Samen met de grote variatie in de duur van de B-periode (de tijd tussen celdeling en de aanvang van replicatie) en een correlatie tussen de B-periode en de tijd tussen celdelingen ondersteunen onze waarnemingen het kritische-massamodel. Tenslotte hebben we gezien dat de duur van de C-periode (de duur van de replicatie zelf) relatief constant is, in tegenstelling tot de zeer variabele B-periode en D-periode (de tijd tussen het voltooien van de replicatie en de celdeling).

In hoofdstuk 4 beschrijven wij de resultaten van de eerste *real-time* studie naar de fasevariatie van de type-1 fimbriae in *E. coli* op ééncellig niveau. Fasevariatie van antigenen en andere eiwitten die op het celoppervlak tot expressie komen, is wijd verbreid onder ziekteverwekkers. Fasevariatie wordt traditioneel gezien als een mechanisme om het gastheerimmuunsysteem te vermijden en betreft een erfelijke en omkeerbare ‘alles-of-niets’ expressie van één gen of een reeks genen. Het schakelen tussen de staten wordt beschouwd als een stochastisch proces, dat resulteert in een heterogene populatie van verschillende staten.

In *E.coli* is de expressie van het type 1 pili (fimbriae) fasevariabel. De promotor voor *fim* expressie bevindt zich in een omkeerbaar DNA fragment van 314 bps (*fimS*). Wanneer omkering plaatsvindt, verandert de promotor in *fimS* van orientatie en fungeert daarmee als genetische schakelaar. In de ‘ON’-oriëntatie, stuurt *fimS* de expressie aan van een aantal structurele genen van het *fim* systeem. In de ‘OFF’-oriëntatie worden de *fim*-genen niet tot expressie gebracht. De inversie van *fimS* staat onder controle van twee locatie-specifieke recombinases, FimB en FimE. Terwijl FimB in beide richtingen kan schakelen, heeft FimE een grote voorkeur om te schakelen in de ‘OFF’-oriëntatie. Bulk-studies hebben een gedetailleerd inzicht verschaft in de regulatie van het *fim*-systeem. Echter, deze benadering is niet geschikt voor de exploratie van de schakel-dynamica en van de correlatie tussen schakel-gebeurtenissen en andere celprocessen (zoals de replicatie van DNA).

Het optreden van schakel-gebeurtenissen wordt in bulk-studies van het *fim*-systeem bepaald door het volgen van de verhouding van de *E.coli*-populaties in verschillende staten. Hier hebben wij GFP (Green Fluorescent Protein) onder controle van de *fim*-schakelaar geplaatst en met *timelapse*-fasecontrast- en fluorescentie-microscopie gekeken naar groeiende microkolonies. Individuele schakel-gebeurtenissen konden we direct volgen en we observeerden verschillende GFP-expressiepatronen in genealogisch verwante cellijnen. Naast de verwachte monotone stijging van het fluorescentieniveau, zagen we ook cellijnen met een tijdelijke fluorescentie-toename. Door gelijktijdig de voortgang van de replicatie te volgen, tonen wij aan dat deze patronen het gevolg zijn van de overerving van verschillende *fim*-staten door de dochtercellen. Onze data zijn consistent met het optreden van meerdere replicatievorken zoals door het Cooper-Helmstetter-model wordt beschreven. Voorts vonden wij dat 'OFF'-naar-'ON'-schakeling afhankelijk is van de celleeftijd, met een hogere schakel-kans aan het begin van de celcyclus, wat een mogelijke correlatie tussen *fim*-schakeling en chromosoomreplicatie suggereert.

In hoofdstuk 5 onderzoeken wij hoe de *fim*-schakeling door het celverleden wordt beïnvloed. In eerder onderzoek naar fasevariatie werd verondersteld dat dit een zuiver willekeurig proces is dat een Poisson-statistiek volgt. Echter, deze aanname is nooit direct getest. In recente fasevariatie-studies worden verschillende mogelijke mechanismen gesuggereerd die zouden kunnen leiden tot niet-Poisson schakel-gedrag. Een voorbeeld is de afhankelijkheid van *fimE*-expressie van de staat van *fimS*. Er is geopperd dat een dergelijke regulatie de 'OFF'-schakel-kans tijdsafhankelijk maakt vanwege de tijd die FimE nodig heeft om een *steady-state* niveau te bereiken.

Wij volgen cellen in een groeiende microkolonie en maten de tijd die een cel in de 'ON'-staat blijft, nadat hij naar de 'ON'-staat geschakeld is. De gegevens laten zien dat de 'OFF'-schakel-snelheid constant is in de tijd en onafhankelijk van het schakel-verleden, waarmee het Poisson-karakter van de *fim*-schakeling wordt bevestigd.

## Publications

This thesis is based on the following publications:

Direct Observation of Type 1 Fimbrial Switching
    Aileen M. Adiciptaningrum, Ian C. Blomfield, Sander J. Tans
    (submitted to EMBO Reports, in revision)

SeqA Dynamics in Slow Growing Cells
    Aileen M. Adiciptaningrum, Sander J. Tans
    (in preparation)

The Effect of Cell History on the OFF Switching Rate of Type 1 Fimbriation in *E.coli*
    Aileen M. Adiciptaningrum, Robert Hauer, Kostya Shundyak, Sander J. Tans
    (in preparation)

## Acknowledgements

AMOLF is a really great and special place to work. Not just because its extremely conducive scientific atmosphere, but also because the extraordinary people it is made of. There, one can always discuss about anything, and I mean anything. From science to politics to why men should not wear skirts while there are no real physical constraints to it. It is a place where the discussion was taken seriously enough to result in at least 3 different theories as probable explanations. So you see, this place is great!

During my exciting time at AMOLF, I received so many support and help from everybody and I would like to take some space here to thank you all. First, I would like to thank you, Sander. I really enjoyed myself to have worked in your group. You are always very supportive, extremely creative, and encouraging. Your extreme clear-headedness amazes me every time we have a discussion and seems to be able to find a light at the end of every tunnel. You have shown me many scientific traits at a level I hope to one day reach them myself. I really thank you for this.

Next, I would like to thank my promotor Daan Frenkel, and other group leaders in the 'overloop' wing, Pieter Rein, Bela, Marileen and Gijsje who always show constant scientific support for each PhD student in AMOLF and always be there as role models we all can look up to. I also thank Conrad, Tanneke, Svetlana and our collaborator, Ian Blomfield, who are generous discussion partners in things I have little knowledge about.

During the process of thesis writing, many people have helped me. Paige, Jocelyn and Justine for English corrections, Simon and Frank for Dutch translations. Thank you so much for spending so much time reading all those words. Where would I be without your help. The great cover is co-made by Julien, thanks a lot and I'm curious about your new pieces. Kostya, who I hijacked a bit from Bela (thanks again!), thanks for your huge help and interesting discussions that we had. Of course this thesis will not be here without the support many other people in AMOLF: reception, E&I, workshop, library, secretary, purchase, PV, management, administration, human resources, travel & housing and canteen who has helped me along the way during my PhD lifetime.

What a blessing to have really exceptional people as fellow group members. Thank you all for making my time at AMOLF really enjoyable. Frank (my longest running office mate) we shared many many experiences 'during' and 'after' AMOLF period- I'll never forget that, Daan (with his free candies) good luck with your Mendelian plant boxes version2, Roland and Kim (what a friendship you offer me) I owe all the morning dilutions to you, Marjon (my paranimf & much more) your daily support & cheering up is truly irreplaceable, Manju (the cool father) you're also a cool office mate, Alex (secretly already group member at heart before he even knew it), Laurens (the winner of group lasergame) and Michael (please hate us more during the game). The ones who had left: Eva – for all our girltalks, Ruud – for your company right there next to the KillerMike, Thomas – the optics master, Matt – awesome BBQ, Philip, Ienas, Jerien, Ndika for bringing special atmosphere to the group and my students Robert and Merlijn, for your hard work.

Next, I would like to thank my other great paranimf- Laura (I wish you all the best for your future plan with Frank; and climbing, obviously), and my fellow poker maniacs. Very fortunately, majority of people in Tans group are natural participants in the wonderful Monday poker in the canteen, plus others that contribute equally: Liedewij (queen of PV poker tournament) thanks for organizing a memorable Queen's day shopping, Gertjan (your face always shows your cards) good luck with the soon coming fatherhood, Christian (the enthusiastic newcomer) thanks for lending me your party capital, and Niels (a true gamer) for your forever interesting card analysis- I am so going to miss that!

Because each and everyone at AMOLF has their own wonderful, unique, special way contributing to the very supportive and enjoyable atmosphere, I would like to also thank co-lab-occupiers and other 'overloop-ers': Nienke, Simon, Koos, Ana, Wiet (who all looked as ridiculous –read: awesome- as me in costumes), Paige (you go girl!), Behnaz (thanks for all the tips), Eva D, Sanne, Rutger, Marina, Svenja, Iza, Thorsten, Bas, Lieve, Saskia, Mel, Philip, Thomas. I would also like to thank many people that have left AMOLF, Patrick, Siebe, Nefeli, Suckjoon, Rosalind, Tatiana, Chantal, Rhoda, Marco, Sorin, Enrico, Jacob and all others.

Last but not least, I would like to thank Andre, for our time together making me the person I am today, and my indescribable thanks to my parents and sisters, for the unconditional love and support available to me at all times.

## Curriculum vitae

Aileen Martinia Adiciptaningrum was born in Semarang, Indonesia, on March 10th, 1980. From 1997 to 2001 she studied Chemistry at the Bandung Insitute of Technology in Bandung. For her bachelor degree, she did research at the Physical Chemistry Department resulting in a thesis entitled ‘Thermodynamical Studies and Thermal Analysis of Chlorination Reaction of Rice Husks into Silicontetrachloride’. In 2001 she started a Master’s Programme in Biotechnology at the Wageningen University, and did an internship at Numico Research BV. Her work with lactic acid bacteria resulted in a thesis entitled ‘Eligibility of Bacterial Fermentations for the Generation of New Compounds from Various Raw Material’ in 2003. In February 2003 she started her PhD research in the group of dr. Sander Tans at the FOM Institute for Atomic and Molecular Physics (AMOLF) in Amsterdam. At the beginning of her PhD she investigated *in vivo* mutation and evolution of *E.coli*, and in 2005 she started to work with type 1 fimbrial expression system in *E.coli*. The result of her work with the *fim* system is described in this thesis.