Phase variation of type 1 fimbriae : a single cell investigation

Adiciptaningrum, A.M.

Direct Observation of Type 1 Fimbrial Switching

4

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

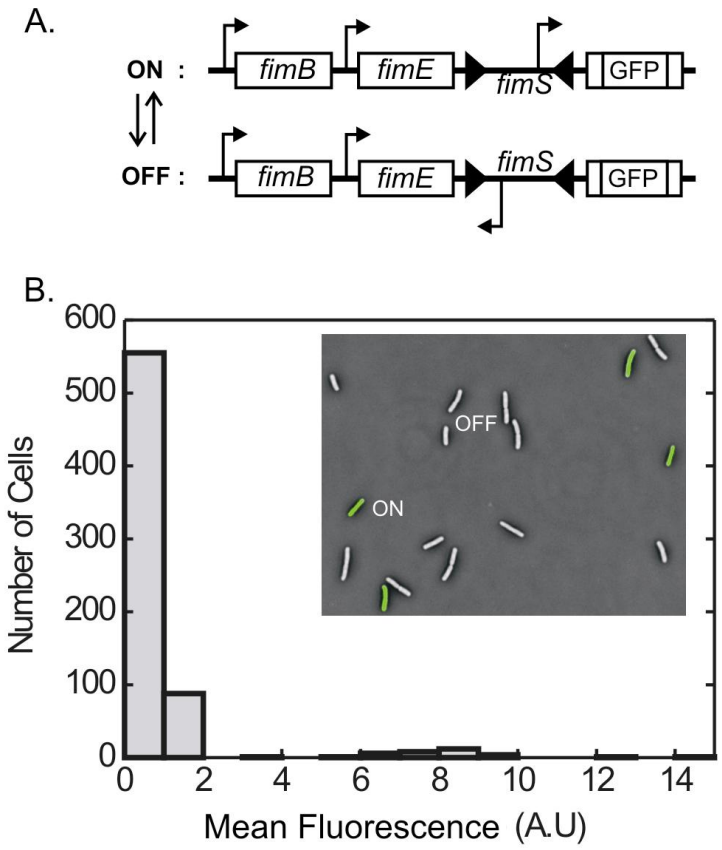

Figure 4.1 *fim* operon in the ON and OFF phase
(A) In the ON orientation, the promoter positioned within the invertible element *fimS* drives transcription of the *fim* operon. GFP has been inserted into *fimA*, the first gene of the *fim* operon, allowing single-cell monitoring of the *fim* operon 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

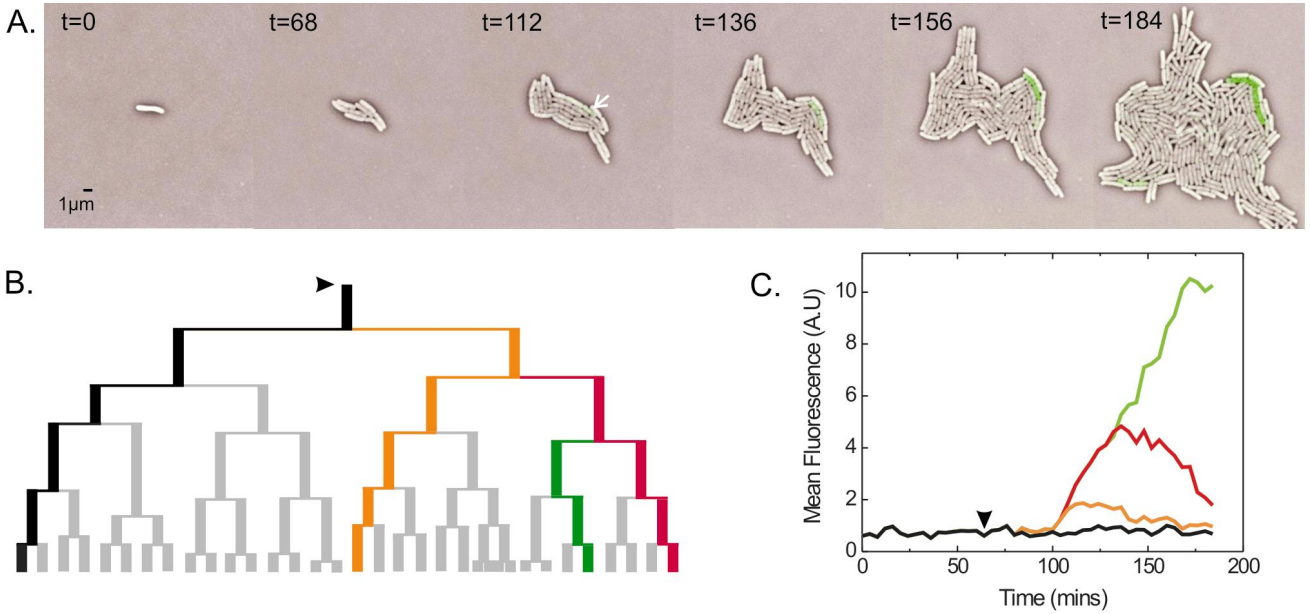

Figure 4.2 ON switching in a growing microcolony.
(A) Overlay of fluorescence and phase contrast images, showing a single OFF cell developing into a microcolony. A cell in this isogenic population starts to fluoresce spontaneously (white arrow), and develops into a brightly fluorescent sub-colony, indicating 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.

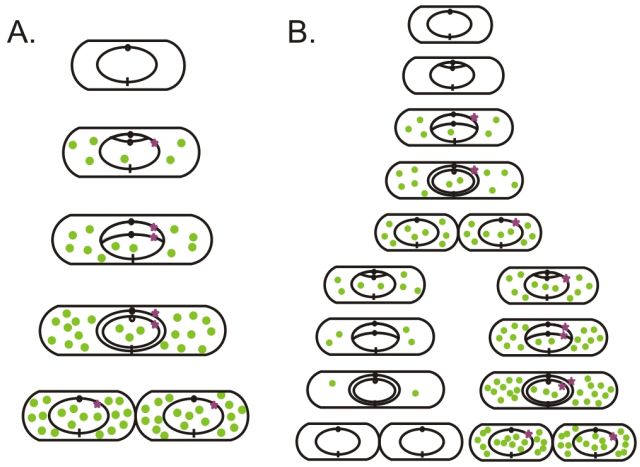

Figure 4.3 Fimbrial expression and DNA replication
Model for slow growing *E.coli* cells inheriting a single chromosome with an OFF *fim* switch. The chromosome is replicated in a bidirectional manner. At a certain moment, the switch spontaneously switches to the ON state (purple star).
(A) OFF to ON 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.

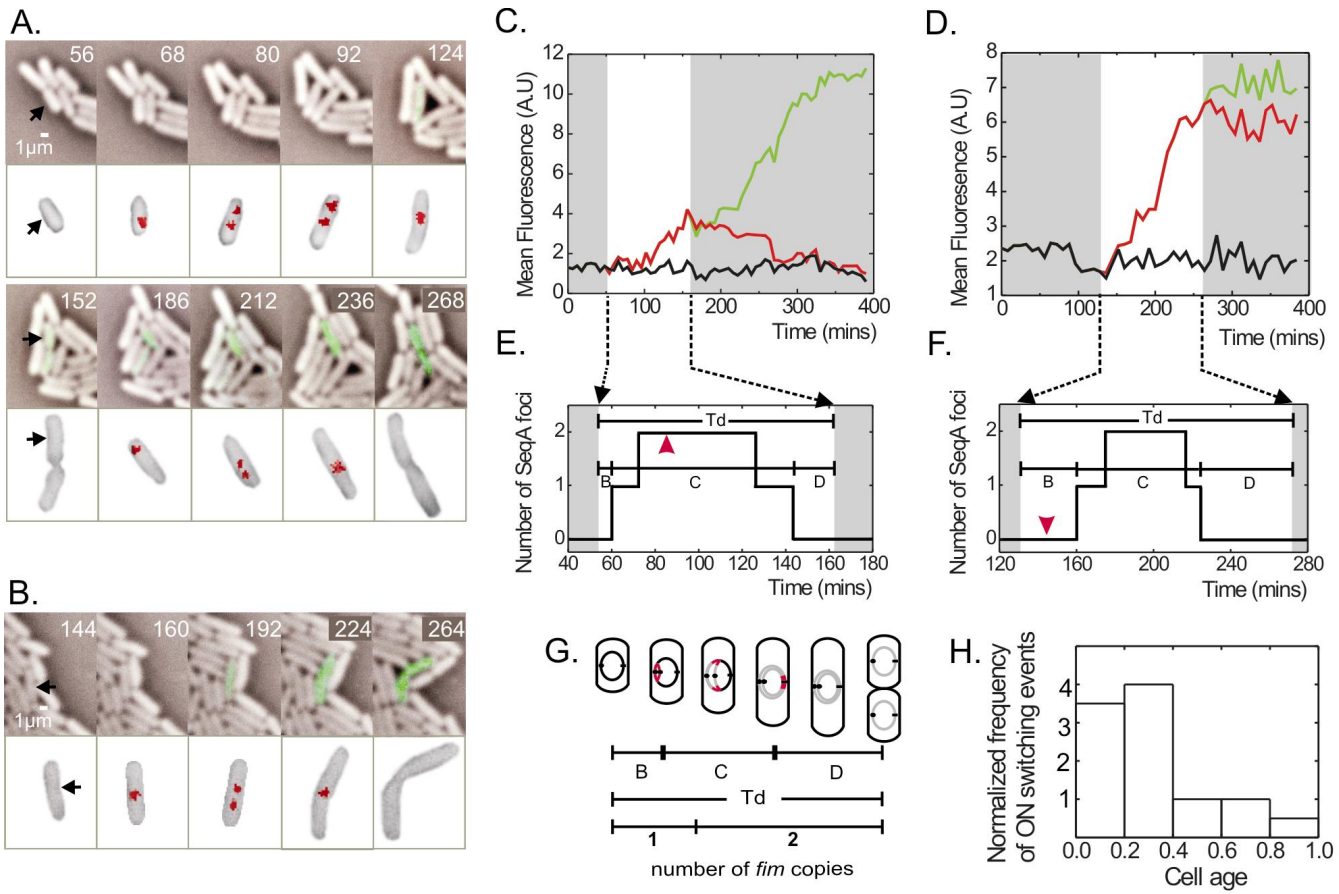
A.
56
68
80
92
124
1μm
152
186
212
236
268
B.
144
160
192
224
264
1μm
C.
Mean Fluorescence (A.U)
Time (mins)
D.
Mean Fluorescence (A.U)
Time (mins)
E.
Number of SeqA foci
Td
B
C
D
Time (mins)
F.
Number of SeqA foci
Td
B
C
D
Time (mins)
G.
B
C
D
Td
1
2
number of fim copies
H.
Normalized frequency of ON switching events
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.

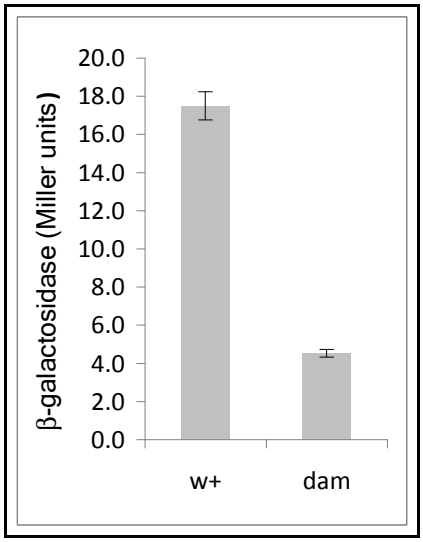

Figure 4.5 The effect of Dam on the β-galactosidase produced by a FimB-LacZ fusion. The fusion is situated in the chromosome at *fim* in a *ΔlacZYA* mutant of strain MG1655 [38]. β-galactosidase assays were conducted as described by Miller [84], following growth of the wild type or *dam* mutant strain 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.