Investigations into the Cellular Localization of the Mismatch Repair Protein MutH

Katrine Bjune

Norwegian University of Life Sciences
Department of Chemistry, Biotechnology and Food Science
Master Thesis 60 credits 2013

Abstract

The semiconservative replication of DNA in *Escherichia coli (E. coli)* gives rise to hemimetylated GATC sites behind the replication forks. Three proteins have been shown to bind to such sites; the negative regulator of chromosomal DNA replication SeqA, the mismatch repair protein MutH and Dam methyltransferase. The latter methylates the GATC sites, turning them into a non-target for the two other proteins. Earlier studies have indicated that SeqA will bind right behind the replication forks, however new studies estimate that at least 1000 bp on average separates the SeqA complex from the replisome. Studies performed in this thesis were conducted to find out if MutH binds to the hemimethylated GATC sites between the replisome and the SeqA complex.

Attempts were made to tag *mutH* at the C-terminal end on the chromosome, but in this process the promoter of the *ygdQ* was deleted resulting in a cell that did not express the YgdQ protein. Little is known about the YgdQ protein except that it is an inner membrane protein. The delta *ygdQ* strain had a decreased growth rate, a DNA content equivalent to more than 10 *E. coli* cells, a round shape and an increased mass. The cells seem to struggle to divide. The phenotype of cells in this strain corresponds to cells with a deficient peptidoglycan layer, making us believe that the YdgQ protein is important for the peptidoglycan layer and the cell shape.

New strains were constructed from the strain SF149 with SeqA-YFP and SSB-CFP, one with delta *mutH* and one with a *mutD5* mutation in the epsilon subunit. We found that the mutation ratio was higher in the cells lacking proofreading compared to the strain lacking mismatch repair (MMR), which we believe came from mismatch repair saturation. The average length between SeqA and SSB foci decreased almost 50 nm when the *mutH* genewas removed compared to the corresponding wild type The strain with the defect proofreading subunit

(mutD5 mutation) had an increased the distance between SeqA and SSB of about 50 nm. These results suggest that the MMR proteins work in a short window between the replisome and SeqA. However, we cannot say from these results if it is the lack of MutH or the lack of a functional MMR that causes the decrease of distance between the SeqA and the SSB foci.

Sammendrag

I *Escherichia coli* (*E. coli*) gir den semikonservative replikasjonen av DNA opphav til hemimetylerte GATC-seter bak replikasjonsgaffelen. Tre proteiner har vist seg å binde seg til slike seter: SeqA, som er en negativ regulator av kromosomal DNA-replikasjon, reparasjonsproteinet MutH og Dam methyltransferase. Dam methyltransferase metylerer GATC-setene og gjør at de to førstnevnte proteinene ikke lenger kan binde seg til disse setene. Tidligere studier har indikert at SeqA binder seg til DNA like bak replikasjonsgaffelen, men nyere studier estimerer at det gjennomsnittlig er 1000 bp mellom SeqA-komplekset og replisomet. Studiene i denne oppgaven ble utført for å finne ut om MutH binder seg til de hemimetylerte GATC-setene mellom replisomet og SeqA komplekset.

mutH ble forsøkt tagget på den C-terminale enden på kromosomet, i denne prosessen ble promotoren til *ygdQ*-genet slettet. Dette resulterte i celler som ikke lenger uttrykte YgdQ-proteinet. Det er foreløpig lite kunnskap om YgdQ-proteinet, bortsett fra at det er et indremembranprotein. Cellene uten YgdQ hadde en nedsatt veksthastighet, et DNA-innhold tilsvarende mer enn 10 *E. coli* celler, en rund form og en økt masse. Det virket som cellene hadde problemer med å dele seg. Fenotypen til cellene uten YgdQ ligner celler med et ødelagt eller et dårlig fungerende peptidoglycanlag. Dette får oss til å mistenke at YgdQ-proteinet kan være viktig for peptidoglycanlaget og celleformen.

Nye bakteriestammer ble laget fra SF149-stammen med SeqA-YFP og SSB-CFP, en stamme uten *mutH* og en stamme med en *mutD5*-mutasjon i epsilon-subenheten. Vi fant ut at mutasjonsraten var høyere i stammen uten reparasjonssenteret i polymerasen (*mutD5*-mutasjonen) enn i cellene uten reparasjonsproteiner. Vi tror dette skylles at reparasjonsproteinsystemet er mettet. Den gjennomsnittlige lengden mellom SeqA- og SSB-proteinene i stammen som manglet MutH sank med ca 50 nm i forhold til den korresponderende villtype-cellen.

Stammen med *mutD5*-mutasjonen hadde en økt gjennomsnittsavstand på 50 nm mellom SeqA og SSB. Disse resultatene tyder på at reparasjonsproteinene har sin oppgave på DNAet et sted mellom SeqA og replisomet. Vi kan derimot ikke si om avstanden forandrer seg på grunn av manglende MutH-binding, eller om det skyldes mangelen på et fungerende reparasjonsmaskineri.

Acknowledgements

This work was carried out at the Department of Cell Biology, Institute for Cancer Research, The Norwegian Radium Hospital from August 2012 to December 2013. I would like to thank my two supervisors Solveig Fossum-Raunehaug and Kirsten Skarstard, the latter who also is the head of the group, for giving me the opportunity to work with this interesting field of research. I am sincerely grateful for the excellent guidance throughout the work with this thesis.

I would also like to thank Anne Wahl and Emily Helgersen for priceless help in the laboratory, helping me by introducing me to new techniques.

Finally I would thank Jussi Mikael Ånestad for being my rock, helping me see the thesis through new eyes and my family including Fredrik André whose name cannot go unmentioned.

Oslo, November 2013

Katrine Bjune

Nomenclature

Abbreviation	Description
Aa	Amino acid
Amp	Ampicillin
ATP	Adenosine Triphosphate
BLAST	Basic Local Alignment Search Tool
BSA	Bovine Serum Albumin
Bp	Base pairs
CAA	Ceso Amino Acids
Cam	Chloramphenicol
Chr	Chromosome
CFP	Cyan Fluorescent Protein
Cl	Chloride
dH_2O	Distilled H_2O
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate (ATP, GTP, CTP, TTP)
et al.	From latin, with others
E. coli	Escherichia coli
F	Farad (capacitance)
GFP	Green Fluorescent Protein
Glu	Glucose
In vivo	Process acting in a living cell
In vitro	Process acting in a reacting tube
IPTG	Isopropyl /textbeta -D-1-thiogalactopyranoside
Kb	Kilobases
Km	Kanamycin

LB	Luria Bertani
Mg	Magnesium
MMR	Mismatch repair
MilliQ-water	Distilled H_2O
Na	Sodium
OD	Optical density
PCR	Polymerase Chain Reaction
PSF	Point Spread Function
RNA	Ribonucleic acid
Rpm	Revolutions per minute
Tet	Tetracycline
Tm	Melting temperature
U	Enzyme unit
UV	Ultraviolet
V	Volt
Wt	Wild type
X-gal	5-bromo-4-chloro-3-indolyl-/textbeta -D-galactopyranoside
YFP	Yellow Fluorescent Protein
λ	Wavelength
℃	Degrees Celsius
Δ	Deletion
µ	micro
e.g	From latin exempli gratia, meaning: for example
g	gram
m	milli
n	nano
τ	Generation time

Contents

1. Introduction

One of the key questions in biology is to understand how cells duplicate their cellular contents and divide. Despite significant differences between eukaryotic and prokaryotic cells, most cellular processes (e.g., metabolism, DNA replication, cell growth, and division) are highly conserved. This allows for use of model organisms such as *Escherichia coli (E. coli)* with desirable properties to ease the facilitation of research, while the findings remain transferable to other organisms and cells.

1.1. *Escherichia coli*

E. coli is part of the *Enterobacteriaceae* family and is classified as Gammaproteobacterias (Miklos Degre [2008]). The family of *Enterobacteriaceae* is large and diverse, and includes 157 species grouped into 29 genera. All family members share a similar structure; they are gram-negative, non-spore forming, rod-shaped bacteria with a new born size of about 1-2 × 0.5 micrometers (Miklos Degre [2008]). *E. coli* is facultative anaerobic, meaning its need for oxygen to produce ATP changes in response to the environment. The bacterium is very diverse and has been found to adhere to among other the mucus layer of the lower intestine. More specifically, wild type *E. coli* has a negative supercoiled genome, $4.6 \cdot 10^6$ base pairs long (Blattner et al. [1997]). These properties, as well as its high growth rate and the fact that *E. coli* has been intensively studied has made it into a preferred model organism, and was one of the first prokaryotic cells with its whole genome entirely sequenced (Blattner et al. [1997])

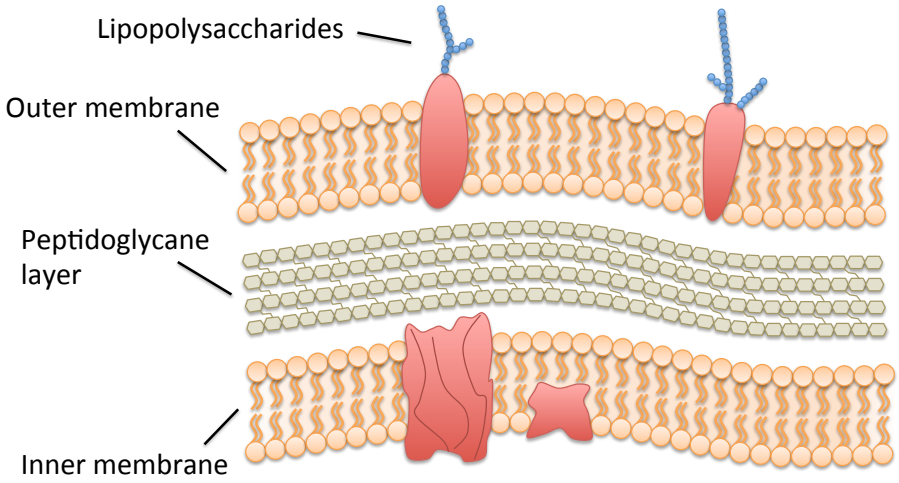


Figure 1: *An illustration of the gram negative cell wall in* E.coli. *(All figures in this thesis are made by the author exclusively for this thesis, unless specified otherwise)*

1.2. The Cell Cycle of *E. coli*

Exponentially growing bacteria cells have a doubling time depending on the growth medium and divide into two equally sized daughter cells. The cells reproduce asexually, and divide through a process called binary fission. In order to produce two identical daughter cells, DNA replication has to occur. The replication must be accurate and the two chromosomes must be segregated equally so that each daughter cell gets its own copy of the entire genome.

The *E. coli* cell cycle is divided into three periods during slow growth; the B-, C- and D-period, and correspond to the eukaryotic G1, S and G2 phase. (Cooper and Helmstetter [1968], Boye et al. [1996]). The B-period starts as the two sister cells are born and continues until initiation of DNA replication. The C-period includes initiation at the origin, elongation along the DNA and finally termination at the terminus. The D-period lasts from the time of termination of replication until cell division. The length of the C- and D-periods is dependent on the growth rate, and increases with increasing

doubling time (Skarstad et al. [1985])

Slow growing *E. coli* cells have doubling times (τ) of about 60 minutes. In such cells the C- and D- periods do not exceed the doubling time (τ > C + D). When *E. coli* is grown in a rich medium, the cell cycle is more complex. The C and D period will then exceed the doubling time (τ <C + D). This means that the cells start a new round of DNA replication before the old one has finished. When the doubling time is less than (C + D), the replication starts in the mother cell. Under extremely fast growth (τ < 1/2(C + D)) the replication starts already in the grandmother cell.

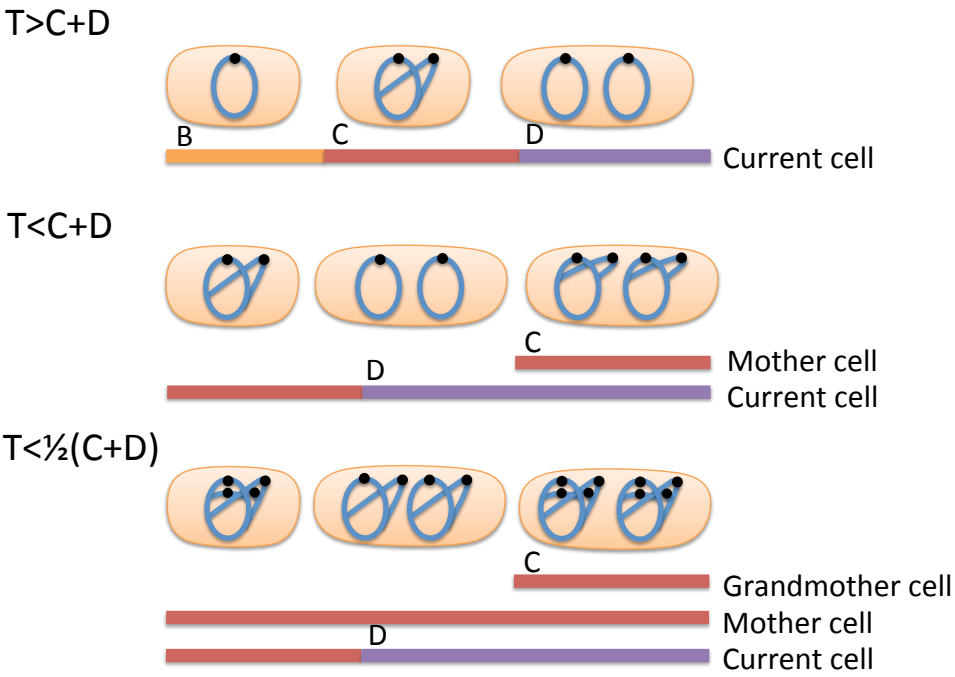


Figure 2: *a) Examples of* E. coli *cell cycles. a) Slow growing cell without overlapping replication cycles. b) A rapidly growing cell with overlapping replication cycles, where initiation of replication starts in the mother cell. c) A rapidly growing cell with overlapping replication cycles. Initiation of replication starts in the grandmother cell.*

1.3. Initiation of Replication at the Origin

Replication is a duplication of the entire genome starting at an origin, from where two forks move bidirectionally toward the terminus. In *E. coli* the origin is called *oriC*, and consists of at least 245 bp (Oka et al. [1980]). Strand separation occurs far left on *oriC*, in an area rich on AT base pairs (13-mer sites) (Bramhill and Kornberg [1988], Kowalski and Eddy [1989],Asai et al. [1990]). In addition to the high amount of AT base pairs, *oriC* contains 9-mer sites, and a large number of GATC sites in comparison to the rest of the chromosome (Zyskind and Smith [1986]). The 9-mer sites are binding boxes for the initiation protein DnaA, and are divided into R1-5 consisting of 5'TTATC/ACAC/AA consensus sequences.

The DNA replication is initiated by the binding of DnaA proteins to four of the five R sites in *oriC* (Cassler et al. [1995], Samitt et al. [1989]). This binding promotes a conformational change bending the DNA 40° (Messer et al. [2001], Roth and Messer [1995]). This bending of the DNA allows the binding of DnaA to the last low-affinity R box, which in turn triggers a number of other events that eventually end in an open complex at the AT rich region (Speck and Messer [2001]). Several factors, like the proteins IHF and HU, aid DnaA in the strand opening process (Dixon and Kornberg [1984], Skarstad et al. [1990]). The open complex promotes the binding of helicase (DnaB) by the help of the helicase loader (DnaC) (Carr and Kaguni [2001]; Mott et al. [2008]). As the DnaC leaves the complex, activation of DnaB and primase (DnaG) occurs (Fang et al. [1999]). DnaG starts synthesizing primer as soon as it is connected to DnaB, this is followed by the binding of DNA polymerase III holoenzyme. This complex is called the replisome, and the chromosomal replication is in progress.

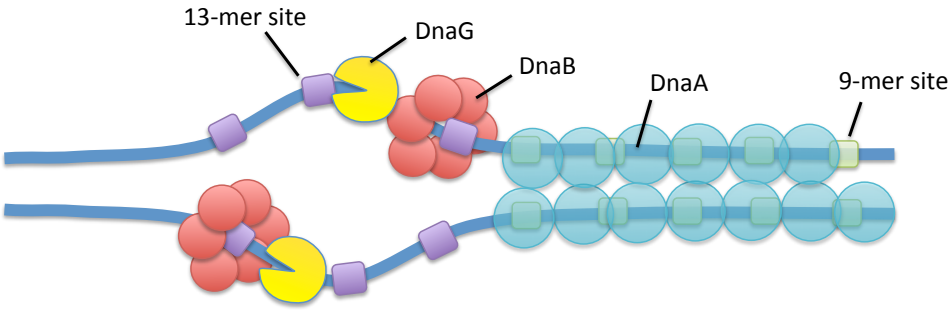


Figure 3: *Initiation of replication at* oriC *in* E. coli*. DnaA binds to the R boxes and triggers an opening in the AT rich region, allowing DnaB and DnaG to bind.*

1.4. DNA Polymerase

After the open complex formation, accumulation of proteins including helicase, single strand binding proteins (SSB), topoisomerase, ligase and DNA polymerase III holoenzyme occurs at *oriC*. DNA polymerase III core enzymes are the parts of the holoenzyme that performs the replication process where the old DNA strand functions as a template for the synthesis. The replication process has an extremely high degree of fidelity and is capable of adding as many as 1000 bp per second. Despite a thorough matching of dNTPs through selective hydrogen-bonding between the A-T, and the C-G bases, wrong bases, insertions or deletions are on average incorporated into the new strand in every 10^5 base (Kunkel and Bebenek [1988]) The polymerase can correct some of these mistakes through a proofreading mechanism where the enzyme can backup and undo the mismatch (Friedberg [1995]). (James D Watson [2008])

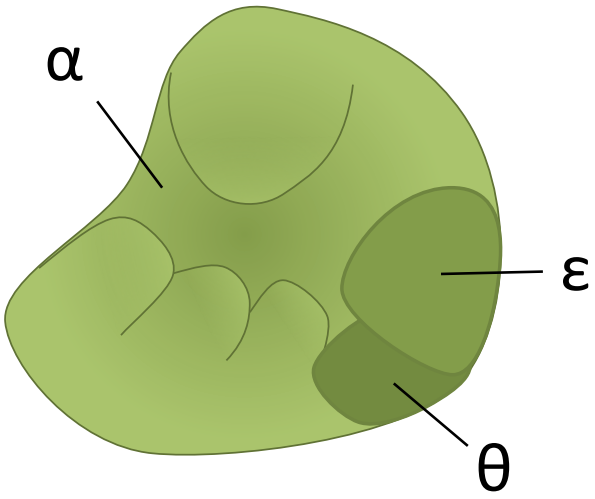


Figure 4: *A part of the DNA polymerase III holoenzyme is the polymerase core enzyme. It consists of three subunits: the α protein which correctly pairs the bases with the template strand, the ε subunit (dnaQ) which operates as the polymerase proofreading center. This subunit has a 3' to 5' exonuclease activity, where mismatched bases are removed. Finally, the theta subunit (holE) which is believed to affect the exonuclease activity of the ε subunit (Taft-Benz and Schaaper [2004]).*

The core enzyme in the DNA polymerase III holoenzyme resembles a partially closed right hand, where the DNA to be copied is held in a cleft in the large protein. The palm domain located in the center of the α subunit is composed of beta-sheets (Ollis et al. [1985]). This part of the protein contains the primary elements of the catalytic site, but also features for monitoring the base pairing between the most recently added nucleotides. Non-base pair specific hydrogen bonds are continuously made between the minor groove of the newly synthesized DNA helix and the polymerase. A mismatched base pair, insertion or deletion will interfere with this contact and the polymerase loses its affinity for the DNA strand. The OH-group at the far end of the newly synthesized strand gets geometrically altered, leading to a reduction in nucleotide addition rate. The exonuclease activity center (the ε subunit) on the polymerase has a tenfold higher

affinity for these kinds of distortion. Therefore, when a mismatch occurs, the newly synthesized strand is moved down to the exonuclease active site. The removal of the mispaired bases enables the DNA synthesis to continue. The proofreading center decreases the error rate from one in 10^5 to one in 10^7, thus some mismatches still escape detection (Schaaper [1993], Friedberg [1995]). In *E. coli,* with its $4.6 \cdot 10^6$ base pairs, the polymerase will on average incorporate an incorrect nucleotide every second replication cycle. (James D Watson [2008])

1.5. Controlling Initiation of Replication

It is very important that initiation of replication only occurs once per cell cycle, DNA replication is therefore a highly regulated process with several mechanisms to prevent newly replicated origins from reinitiating too early (Boye et al. [2000]). There are four important mechanisms. (I) The origin and the DnaA promoter is inactivated by a process called sequestration. The SeqA protein prevents transcription of DnaA and binding of DnaA to the origin region (Campbell and Kleckner [1990], Waldminghaus and Skarstad [2009]) (see section 1.5.1). (II) The *datA* locus is located close to the *oriC* and contains a high number of DnaA binding sites (Roth and Messer [1995]). When the chromosome is copied, the number of these binding sites doubles, and titrates large amounts of DnaA-ATP protein (Kitagawa et al. [1998]). (III) Only the active form of DnaA, DnaA-ATP acts as an initiator protein. After initiation, hydrolysis of ATP to ADP occurs through two processes called RIDA (regulatory inactivation of DnaA) and DDAH (*datA*-dependent DnaA-ATP hydrolysis). The RIDA process is the hydrolysis of DnaA-ATP by the protein Hda in complex with the β -clamp behind the replisome (Katayama et al. [2010]). DDAH is a process where the *datA* locus and the protein IHF have an active role in the hydrolysis of DnaA-ATP (Kasho and Katayama [2013], Katayama et al. [1998], Kato and Katayama [2001]). (IV) Autoregulation of the *dnaA* gene promoter. The *dnaA* gene is negatively autoregulated, the active form of

the DnaA protein binds to its own promoter and inhibits transcription (Polaczek and Wright [1990]).

1.5.1. Sequestration

DNA is replicated in a semiconservative fashion, meaning that every round of replication produces DNA helices with one new and one old strand. The old strands are fully methylated, such that a CH_3 group is attached to the N_6 position on the adenine nucleotide on every GATC sequence. The new DNA strand remains unmethylated for a short time. Thus, the two newly formed DNA helices are hemimethylated. The new DNA strands are subjected to methylation by Dam methyltransferase about 1-2 min after replication (Campbell and Kleckner [1990], Waldminghaus and Skarstad [2009].

The *oriC* contains an unusually high amount of GATC sites compared to the rest of the chromosome. The GATC sites in *oriC* and also the *dna*A promoter region remain hemimethylated for approximately one third of the cell cycle (Campbell and Kleckner [1990], Waldminghaus and Skarstad [2009]). It was found that *in vivo* replication only occurs on fully methylated origins, and that a factor named SeqA binds and protects the hemimethylated DNA from Dam methyltransferase (Messer et al. [1985], Lu et al. [1994], von Freiesleben et al. [1994]). SeqA binds and sequesters the hemimethylated origin and thereby prevents a rapid return to a fully methylated state and reinitiation.

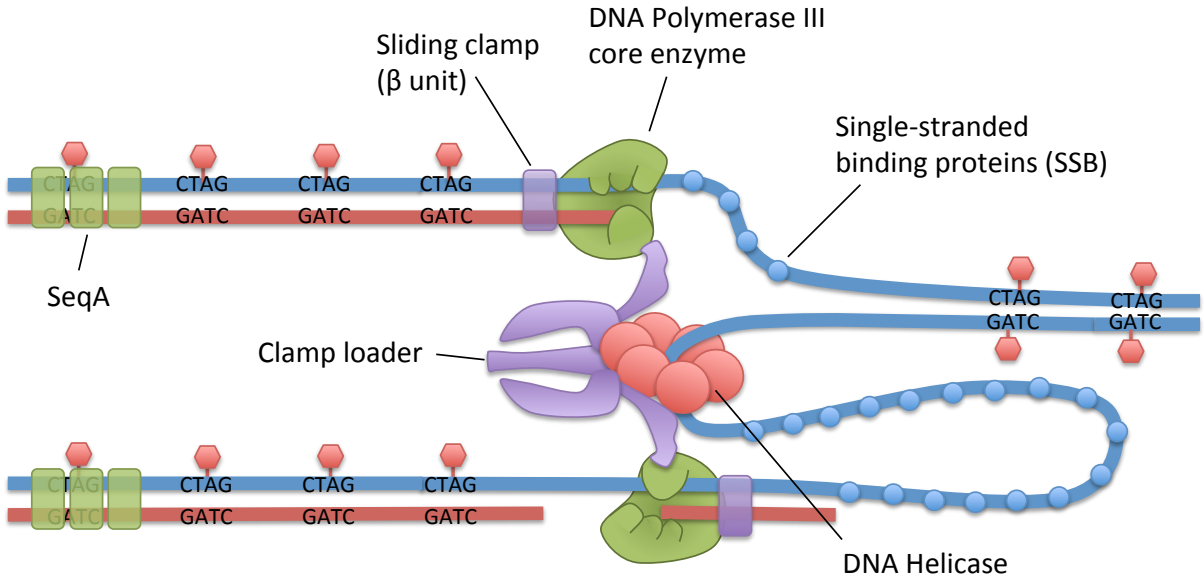


Figure 5: *SeqA binds hemimethylated GATC, which occurs behind the replisome.*

1.6. The SeqA Protein

SeqA is a protein identified as a negative regulator of initiation (Lu et al. [1994], von Freiesleben et al. [1994]). The protein has a size of 21kDa, is non-essential and appears in approximately 1000 copies per cell (Slater et al. [1995]). SeqA has two functional domains, the N- and C-terminal domain. The N-terminal domain folds into two α-helices and one β-strand and is required in multimerization with other SeqA proteins (Odsbu et al. [2005], Guarne et al. [2005]). The C-terminal domain folds into seven alpha-helices and three antiparallel beta-sheets. This domain is necessary for interaction with DNA, mainly the major groove of the hemimethylated GATC sequence (Guarne et al. [2002], Fossum et al. [2003]). The multimerization properties of SeqA are known to have a cooperative effect on nearby sites (Slater et al. [1995], Fossum et al. [2003]). This effect might explain findings where more than one SeqA protein is bound per GATC site (Slater et al. [1995], Lee et al. [2001]). Deletion of SeqA results

in cells that reinitiate DNA replication, and initiation is therefore often unsynchronized in these cells (Lu et al. [1994], von Freiesleben et al. [1994]). Mutations in SeqA that interrupt binding to DNA or multimerization of the SeqA proteins also result in cells with origins that are reinitiated (Fossum et al. [2003], Odsbu et al. [2005]).

The SeqA protein has also been shown to be involved in organization and/or segregation of the newly replicated genome (Lu et al. [1994], Klungsoyr and Skarstad [2004], Hiraga et al. [1998], von Freiesleben et al. [2000], Fossum et al. [2007], Molina and Skarstad [2004], Waldminghaus et al. [2012], Bach et al. [2003]). SeqA binding to DNA has been demonstrated to change DNA topology, and properties in restraining negative supercoiled DNA (Torheim and Skarstad [1999], Klungsoyr and Skarstad [2004]).

1.7. The Mismatch Repair System

The mismatch repair system (MMR) detects and repairs errors that are not corrected by the DNA polymerase. Initially the MMR system distinguishes a mismatch from the correctly paired bases by scanning the genome. However, since both bases in a mismatch are chemically normal, enzymes cannot simply scan the DNA for abnormal chemical structures. The mismatch repair system must specifically identify incorrectly incorporated bases in the newly synthesized daughter strand. Mismatches that do not get repaired before the next round of replication result in a mutation. The MMR system decreases the error rate 100- to 1000-fold (Zahrt et al. [1994]). Thus, mutations in *E. coli* only will occur for every 10^{10} nucleotide added. (James D Watson [2008])

1.7.1. Mismatch Detection

MMR is initiated once a MutS homodimer detects a mismatch or an insertion-deletion loop (Su and Modrich [1986], Parker and Marinus [1992]). MutS scans the DNA

and recognizes mismatched bases. The part of the DNA that contains an error gets embraced by the MutS dimer, which in turn undergoes a conformational change upon binding of ATP. This induces a kink in the helix (Lamers et al. [2000],Obmolova et al. [2000], Junop et al. [2001], Natrajan et al. [2003]). The DNA-MutS complex recruits MutL, which interacts physically with MutS and enhances mismatch recognition. MutL functions as a mediator between MutS and the endonuclease MutH, hence MutS activates MutL, which in turn activates MutH. Like MutS, MutL possesses ATP binding, and both proteins use hydrolysis of phosphate groups in their proofreading role (Shimada et al. [2013]). Studies have shown that using a nonhydrolyzing analog of ATP enhances MutL's capacity to activate MutH (Ban and Yang [1998], Yang [2000]). This indicates that it is not the hydrolysis that is essential for MutL-MutH activation, rather the binding itself. James D Watson [2008]

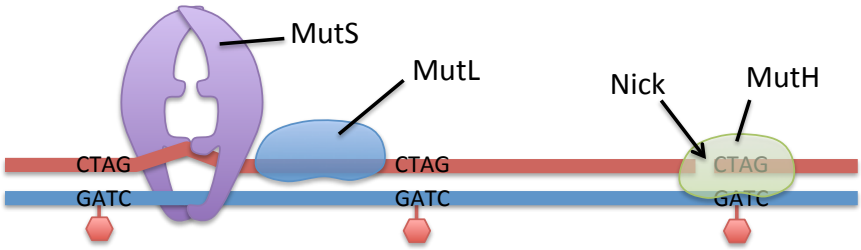


Figure 6: *Illustration of mismatch repair in* E. coli*. MutS detects a distortion in the DNA backbone which arises from mispaired bases. Mismatch bound MutS activates MutL, which in turn activates MutH. The activated MutH induces a nick in the unmethylated strand at the GATC site.*

As mentioned in chapter 1.5.1, newly replicated DNA strands lack methylated adenine nucleotides in GATC sequences. MutH binds to such hemimethylated GATC and discriminates between the newly synthesized strand and the parental strand. MutH normally has a low endonuclease activity, but this feature is increased 50 fold in the presence of a MutS-MutL complex at a nearby mismatch site (Au et al. [1992]). The site of the error has in some cases proven to be several hundred base pairs away from

the discrimination site (Modrich and Lahue [1996], Mendillo et al. [2007]). How the interaction between the MutS-MutL complex and MutH takes place is still uncertain. Three theories are suggested which likely could explain the interactions.

1.7.2. Activation Models of MutH

The stationary model proposes that interaction between mismatch repair proteins induces bending or looping of DNA that brings the two distant sites to proximity. MutS remains bound to the mismatch site at all time (Guarne et al. [2004], Junop et al. [2001]) The MutS ATPase activity acts in a proofreading role, where it is used to verify the binding at the mismatch and authorize downstream commissioning. Support of this theory came from the experiment of Junop et al. [2001], where it is shown that MutH cleaves GATC sites located on a separate DNA molecule without a MutS bound mismatch.

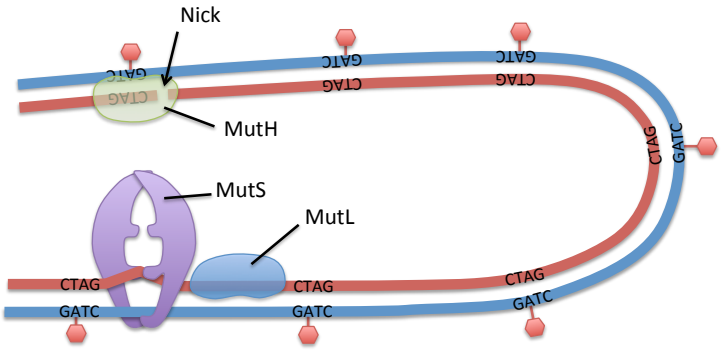


Figure 7: *The stationary model suggests that MutS remain bound at mismatch site, and that DNA looping brings the two sits in proximity.*

The *cis* or moving model suggests that the MutS-MutL complex binds at the mismatch site and later moves away from the mismatch in search of a discrimination site where MutH can be recruited for its exonuclease activity. There are two moving models: the translocation model and the molecular switch model. The translocation model proposes that MutS displays a reduced binding affinity towards the mismatch site upon

ATP binding, and that ATP hydrolysis drives unidirectional translocation of the protein complex along the DNA helix (Allen et al. [1997]). The DNA is threaded through the MutS dimer until it reaches the strand discrimination site.

The molecular switch model suggests that ADP bound MutS binds mismatched DNA. The binding of MutS triggers a conformational change that allows an exchange to ATP and formation of a sliding clamp. The theory states that it is the actual binding of ATP to MutS, not the hydrolysis that promotes sliding of the complex along the helix to the discrimination site (Fishel [1998], Gradia et al. [1997], Acharya et al. [2003], Jiang et al. [2005], Qiu et al. [2012]).

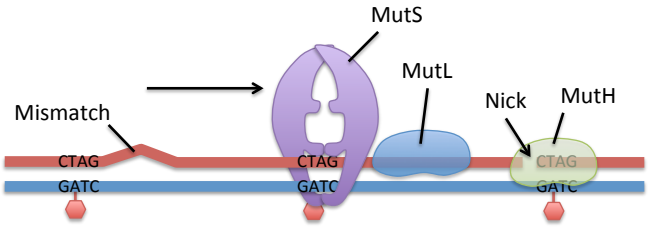


Figure 8: *The moving model proposes that MutS is activated and slides along the DNA from the mismatched base pair to the discrimination site.*

Despite numerous papers on the *in vitro* findings on the latter topic, a third and final model of MutL polymerization is proposed. This model suggests that MutS binds to the mismatch site and recruits MutL, which polymerizes along the helix from the mismatch to the discrimination site where it activates MutH (Modrich [1987]). This last model is supported by articles, which point out that bound MutS protects the DNA from DNase I digestion, while the addition of MutL dramatically extends the footprint (Selmane et al. [2003],Grilley et al. [1989]). Other supporting articles demonstrate that the cellular amount of MutL is the limiting factor in the repair of numerous mismatches in a cell, despite the fact that MutS and MutL are present in equimolar concentrations in wild type cells (Damagnez V and M [1989], Ivan Matic and Radman [2003], Negishi et al. [2002], Feng et al. [1996]). One of the newest articles on this subject emphasizes that

MutS and MutL colocalize on unrepaired mismatch sites and form fluorescent foci. On average the MutL foci are 2.7 times more intense than the MutS foci, supporting the MutL polymerization theory (Elez et al. [2012]).

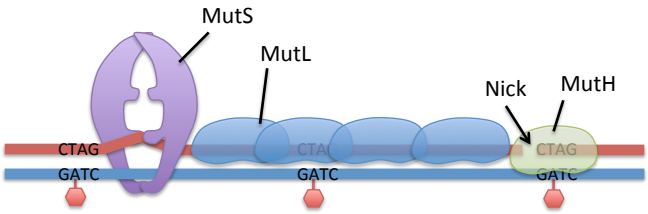


Figure 9: *The third model suggests that MutL polymerizes along DNA, between the activated MutS site and the discrimination site to activate MutH.*

1.7.3. Finishing Mismatch Repair

After MutH is activated by MutL, MutH induces a nick in the unmethylated strand (Welsh et al. [1987]). The DNA helix is unwound by the helicase UvrD. Its action starts at the discrimination site and moves in the direction of the mismatch. One of the exonucleases Exo VII, RecJ or Exo I digests the displaced newly replicated strand to a point beyond the mismatch (Matson [1986], Viswanathan and Lovett [1998]). The assigned exonuclease depends on which side of the mismatch the nick occurs, Exo VII and RecJ remove nucleotides from 5' to 3', while Exo I removes nucleotides in the opposite direction (3' to 5'). This action produces a single stranded gap, which is protected by single stranded binding proteins until DNA polymerase III binds and fills the gap with correctly paired bases before sealing is done by DNA ligase (Lahue et al. [1989]).James D Watson [2008]

1.8. Aims of the Study

There are three different proteins that have been shown to bind to hemimethylated GATC sites: SeqA, MutH and Dam methyltransferase. They all have cellular tasks related to newly replicated DNA.

Studies have shown how SeqA binds in large compact structures. SeqA binds dynamically, meaning that the protein continuously binds to newly replicated hemimethylated GATC sites constructed right downstream of the replisome. It is thought that this dynamic binding allows Dam methyltransferase occasionally access to the GATC sits. All the GATC sites are methylated within 1-2 min after the replication forks have passed (Waldminghaus and Skarstad [2009]). A study using an immunofluorescent method with the color label BrdU showed that SeqA binds newly synthesized DNA, and allowed the assumption that SeqA would bind hemimethylated GATCs right downstream of the replisome. (Molina and Skarstad [2004], Adachi et al. [2005]). However, new studies carried out by Solveig Fossum-Raunehaug show with the strain SF149, containing the SeqA and SSB proteins connected to the GFP (green fluorescent protein) like-proteins YFP (yellow fluorescent protein) and CFP (cyan fluorescent protein) respectively, that SeqA is not colocalized with the replisome, but rather partly colocalized. The distance between the partly overlapping foci vary around 200-250 nm. When DNA is stretched out, i.e. not bound by nucleoid associated proteins, the distance corresponds to about 1000 bp. Since a GATC site occurs on average once every 256 bp, four GATC sites are likely to appear unbound between the SeqA and SSB foci. It is possible that a small portion of the cellular SeqA proteins bind to these sites too, but it is also possible that these sites are occupied by another protein, for example the MutH protein. The mismatch repair proteins have an important role in the replication fidelity process, and probably have its function on the newly synthesized DNA right behind the replisome. The question in this thesis is where MutH binds on hemimethylated DNA compared to SeqA.

2. Methods

In this chapter, information is given about the different methods. For detailed protocols, refer to section B.

2.1. Growth Conditions

In order to grow *E. coli* cells it is necessary to use media containing essential nutrients. Different types of media have different contents of nutrients which in turn regulates the bacterial growth rate. When making overnight cultures (ONC) bacterial cells are inoculated from a glycerol stock (1:1 ratio of bacteria culture and 87% glycerol) or a colony grown on a plate. The ONC are made by inoculating cells in 1.5 ml medium and grown for 16 to18 hours with vigorous shaking (~1000 rpm). In this thesis LB and ABB_1 Glu Caa media are used for bacterial growth (see table 14 , 13).

Bacterial growth is measured using a spectrophotometer, which estimates the bacterial concentration in the cell suspension by measuring the sample's optical density (OD). The OD is measured at 450 nm for ABB_1 Glu Caa medium and 600 nm for LB medium. OD measurements are conducted by illuminating the sample with monochromatic light beams, which pass through the cell suspension. The light which passes through the cuvette gets detected, while the rest of the light is scattered in all directions, the more bacteria in the sample, the more scattered light. The absorbance from the cell suspension is measured by a spectrophotometer periodically, and the doubling time can be calculated by the increase in cell density. (See section B.1)

2.2. Purification of DNA

2.2.1. Purification of DNA Fragments from Gel or PCR

Purification of DNA from agarose gels and PCR reactions is done using the Wizard gel and PCR clean-up system. The agarose gel is dissolved in the presence of guanidine isothiocyanate, whereas the PCR product is directly purified by the same substance. The principle of purification is based on DNA's ability to bind to silica after the DNA molecule has been exposed to chaotropic salt. The chaotropic salt interferes with the stabilizing intramolecular interaction in the DNA molecule, thereby linearizing it. The sample is centrifuged through the silica membrane column where the DNA binds to the membrane and is eluted after addition of nuclease free water. (Promega [2013a])

2.2.2. Purification of Plasmid DNA from Bacterial Cells

The principle behind purification of DNA is based on alkaline lysis of bacterial cells followed by absorption of DNA. Both the Quiagen miniprep purification kit and the JetStar midiprep purification kit are based on the same principle of plasmid DNA purification.

Cells are harvested by centrifugation and resuspended in TrisHCL containing RNaseA to maintain a stable pH and to cut the RNA molecules in the sample. The solution contains EDTA which binds cations like Mg^{2+}, making the cells unstable. The cells are lysed after addition of NaOH/SDS. NaOH loosens the cell membrane, while SDS dissolves components of the cell membrane, leading to a cell lysis that releases plasmid DNA into the solution. The lysis is neutralized by increasing the salt concentration. The high salt concentration causes the SDS to precipitate with denatured proteins and chromosomal DNA. The sample is then purified with the use of a column and the DNA plasmids remains in the column until elution with nuclease free water. (Joseph Sambrook [2001], QIAGEN [2012]) (See section B.4.1 and B.4.2).

2.2.3. Purification of Chromosomal DNA from Bacterial Cells

Purification of chromosomal DNA can be done by using the Wizard Genomic DNA purification Kit. The first step is to collect cells and destroy the cell walls by cell lysis. RNaseA is added to remove all RNA molecules in the sample. Cellular proteins are removed by salt precipitation, which precipitate proteins but leaves the high molecular weight genomic DNA in the solution. Lastly, isopropanol precipitation of the genomic DNA is performed. The DNA is resolved in nuclease free water. (See section B.5)

2.3. DNA Concentration Measurements

Nanodrop from Thermo Scientific is a spectrophotometer which calculates the DNA concentration by emitting light within a spectrum of different wavelengths and measures the absorbance of light in the sample. The information from the different absorbance curves is used to calculate the DNA concentration in the sample. In the measuring process Nanodrop measures DNA, RNA, proteins and degraded nucleic acids. The DNA concentration is estimated in a wide range from 0.4 to 15,000 ng/µ l. The Nanodrop spectrophotometer is therefore only used on high concentration samples with a tolerance for small inaccuracies.

2.4. DNA Precipitation

DNA is polar due to its highly negatively charged phosphate backbone. As a result, H_2O molecules make hydration shells around the DNA, allowing the DNA molecules to spread through the solution. It is sometimes necessary to increase the DNA concentration. This is done by precipitating the DNA. When a calculated amount of positively charged ions is added together with a less polar liquid like ethanol or isopropanol, the water-DNA protection shield is disturbed. If enough polar liquid is added, ion bonds between

phosphate groups and positively charged ions become strong enough for DNA precipitation. (Joseph Sambrook [2001])

2.5. Agarose Gel Electrophoresis

Agarose gel electrophoresis is a method for separating DNA fragments over an electric field on the basis of size and shape. DNA fragments from 0.5 kb to 25kb may be distinguished using this method. The rate of migration depends on the applied voltage, the buffer's ionic strength and the size of channels. The channel size is regulated by the agarose concentration. The agarose gel is placed in a TAE buffer with a pH of 8, maintaining a negatively charged DNA. The DNA will therefore migrate towards the positive pole in the electrical field. Gel electrophoresis may be used for identification, but also purification of DNA fragments.

2.5.1. Gel Red Staining

After agarose gel electrophoresis, the gel is stained in GelRed for visualization of the DNA fragments. GelRed stain binds to both single and double stranded DNA, but also RNAs. The direct binding to DNA makes the dye potentially mutagen inducing. GelRed was designed to replace the more toxic ethidium bromide (EtBr) and is constructed to be unable to pass the cell membrane and therefore is less harmful. In this thesis the staining was performed after electrophoresis, but the GelRed staining can be performed both before and after electrophoresis. The stain consists of a fluorophore that binds to DNA and emits lights when exposed to UV-light.

2.6. Polymerase Chain Reaction

Polymerase chain reaction (PCR) is an automated and rapid method used to amplify specific regions of DNA, and allowing copy numbers up to a billion fold. PCR takes

advantage of the hydrogen bond's melting point. When the temperature increases, the hydrogen bonds break and the two strands separate, allowing primers and polymerase to bind and copy a chosen part of the sequence. Regulation of the temperature in the solution controls the availability, the binding and the initiation of the amplification. The denaturing temperature is 95℃, the temperature is then decreased to a temperature 5℃ below the primer's melting point (typically 50-60℃), permitting the primers to bind complementary single stranded DNA. The temperature is again raised, this time to 72℃, which is the ideal temperature for the DNA polymerases. The DNA polymerase binds to the primers and adds corresponding nucleotides to the template. These steps are repeated 20-30 times in cycles. Joseph Sambrook [2001]

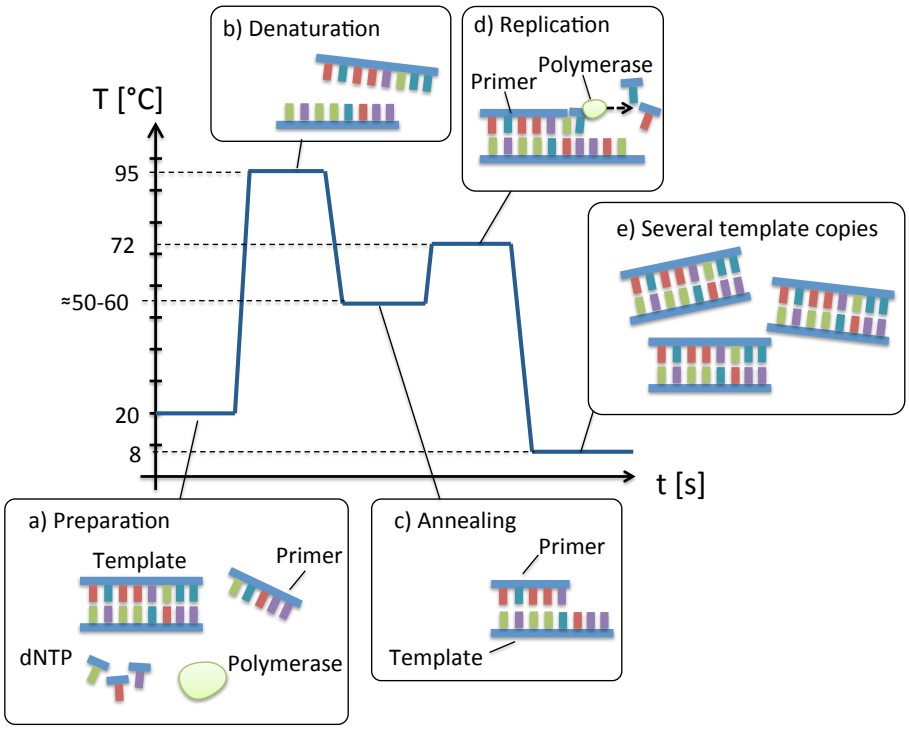


Figure 10: *A schematic drawing of the reactions at the different temperatures in a PCR.*

2.6.1. Primer Construction

When PCR is performed, custom made primers are needed. The primers typically consists of 20-25 bp upstream and downstream of the sequence of interest. The left primer is complementary to the 5' to 3' DNA strand. The right primer is reverse complementary. It is important that the primers have approximately the same melting temperature. The melting temperature depends on the length of the DNA molecule and the nucleotide composition.

Primers can also add nucleotides to the original sequence. Removal of stop or start codons or addition of enzyme restriction sites can be carried out by adding or removing sequences in the primers. Primers used to make the PCR fragment for the Red/ET cloning, are made with homologous tails to a specific part of the *E. coli* chromosome. Its gives the fragment an opportunity to engage in homologous recombination and get inserted into a specific area on the chromosome 2.9.

2.7. Cloning in *E. coli*

2.7.1. Poly A-tailing of Blunt Ended DNA

Blunt ended PCR fragments can often be hard to ligate into a vector. As an alternative, T-tailed vectors like pGEM-T Easy may be used. Non-proofreading Taq DNA polymerase is used to incorporate adenine nucleotides onto the 3' end of DNA fragments. The result is adenine tails of about 150 nucleotides is incorporated on the PCR fragment. The sequence can thereby be ligated into open vectors with thymine-tails. (Joseph Sambrook [2001])

2.7.2. Cloning into the pGEM-T Easy Vector

pGEM-T Easy is a double stranded circular plasmid consisting of 3015bp. The pGEM-T Easy vector has an ampicillin resistance gene and two thymine tails in the middle of a *lac*Z gene, which can be used to insert genes with adenine tails (see figure 11). The *lac*Z gene encodes for Beta-galactosidase, an enzyme that normally cleaves lactose into galactose and glucose. However when bacteria with this gene are grown in the presence x-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside), a synthetic analog of lactose, the Beta-galactosidase enzyme cleaves x-gal into galactose and 5-bromo-4-chloro-3-hydroxyindole. The latter spontaneously dimerize, and is oxidized into 5.5-dibromo-4.4-dichloro-indigo, an insoluble blue product.

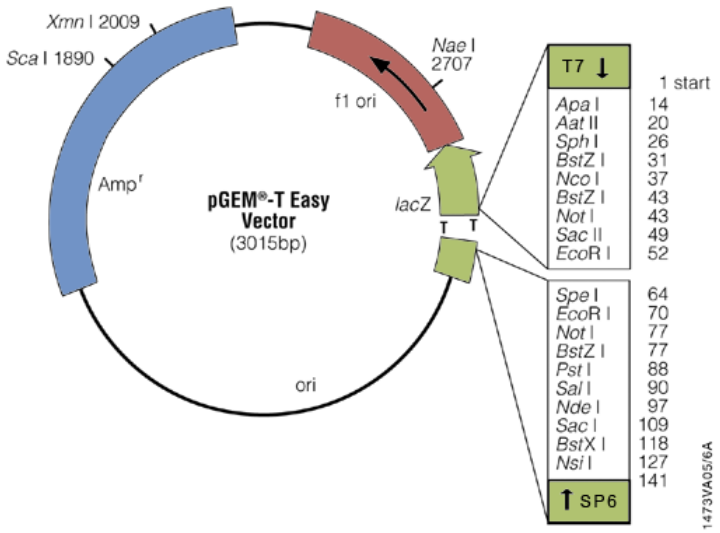


Figure 11: *The pGEMT-T Easy vector Promega [2013b]*

DNA fragments with adenine tails are ligated into the pGEM-T easy vector and transformed into the bacterial cell JM109 (from Promega). JM109 is competent *E. coli* cells lacking functional *lac*Z gene on the chromosome. JM109 also carry the recA1 mutation and lack the *E. coli* K restriction system. The strain is therefore especially recommended because they avoid undesirable recombination between chromosomal

DNA and plasmid DNA (Bac [2013]). The bacterial cells are grown on an LB agar plate with ampicillin, IPTG (isopropylthio-β -galactoside) and x-gal. IPTG is a non-hydrolyzable analog of allolactose which binds to the *lac* repressor. The *lac* repressor then change conformation and allow transcription of the *lac* operon. Blue colonies indicate a functional *lac*Z gene, hence no insertion. White colonies lack a functional lacZ gene, meaning that they have a pGEM-T Easy vector with the gene of interest inserted into the polylinker region.

2.7.3. Restriction Enzyme Digestion

Restriction enzymes recognize specific sequences leading to a cleavage of the DNA strands. The digested DNA ends up with sticky or blunt ends, depending on the restriction enzyme. DNA with sticky ends has short stretches (2-4bp) of single stranded DNA that possess the property of self ligation or ligate to regions with complementary nucleotides. Blunt ends on the other hand are universally compatible all other blunt ended DNA stand. (See section B.7)

2.7.4. Dephosphorylation of Vectors

During ligation of an insert into a vector cut with only one enzyme, problems with vector religation may be encountered. To solve this problem dephosphorylation of vector may be performed. Phosphatase catalyzes the removal of the 5' phosphor group from both DNA and RNA. Dephosphorylated plasmids lack the required 5' phosphoryl termini and can therefore not self-religate. (Joseph Sambrook [2001]) (See section B.8)

2.7.5. Ligation

Ligation is a process where catalyzation of the phosphodiester bond between the 5' phosphate group and the 3' hydroxyl terminal is executed by DNA ligase. These

covalent bonds are made by ATP hydrolysis between sticky or blunt ends, as well as at single stranded nicks in DNA and RNA.

$$\frac{\text{X ng vector} \cdot \text{bp insert} \cdot 3}{\text{bp vector}} = \text{X ng insert}$$

The molar ratio between the vector and the DNA insert should be 1:3 in the ligation reaction mix. (See section B.9)

2.8. Horizontal Gene Transfer

Horizontal gene transfer refers to the transfer of genetic material from one cell to another without reproducing itself. Several research techniques in molecular biology requires foreign genes to be inserted into host cells. There are three types of horizontal gene transfer. (I) Conjugation, transfer of genetic material between cells in direct contact. (II) Transduction, injection of foreign DNA with the aid of a bacteriophage. (III) Transformation, uptake of exogenous DNA from the surroundings.

2.8.1. Transformation

Cells with the ability to take up foreign DNA from the surroundings are called competent. Competence occurs only as a time limiting response to environmental conditions such as starvation or cell density. Some cells possess the property of natural competence, while other can be made competent. *E. coli* is an example of cells that must be made competent. There are two types of competent cells; electrocompetent and chemically competent cells.

Transformation into electrocompetent cells is performed through electroporation. Electroporation is a technique where voltage is applied over a field containing a high concentration of bacterial cells, making a short pulse which creates transitional pores

in the cell membrane.The electrical pulse causes rapid rearrangement of the lipids in the lipid bilayer, forming water filled holes where exogenous DNA molecules can enter. Joseph Sambrook [2001]

Chemically competent cells are made competent by the use of multiple cations which alter the permeability of the cell membrane allowing DNA to cross over into the cell. The exact mechanism is unknown, but it is believed that $CaCl_2$ neutralizes the unfavorable reaction between the negatively charged DNA and the cell membrane. Treatment through heat opens the cell membrane pores allowing exogenous DNA to enter the cell. Joseph Sambrook [2001]

2.8.2. P1 Transduction

P1 transduction is a method of transferring genetic material from a donor cell to a recipient cell via bacterial virus called a phage. In this process the phage attacks the cell wall of the donor cell and injects its viral DNA. The phage DNA acts as template for replication and transcription, and phage proteins are synthesized. The phage enzymes will than digests the chromosome DNA in the donor cell. Some of the pieces of the bacterial DNA will mistakenly be packed into the phages, resulting in phage particles carrying donor cell DNA instead of phage DNA. The donor cell is lysed and the phages are released. These phage particles can later infect a new population of bacteria, where the DNA from the donor bacteria will be transferred into this recipient cell. Transduction of cellular DNA may lead to recombination between the donor DNA and the recipient DNA. Recombination is necessary otherwise the transduction is not successful (see figure 12). (Gerard J. Tortora. Berdell R. Funke [2012])

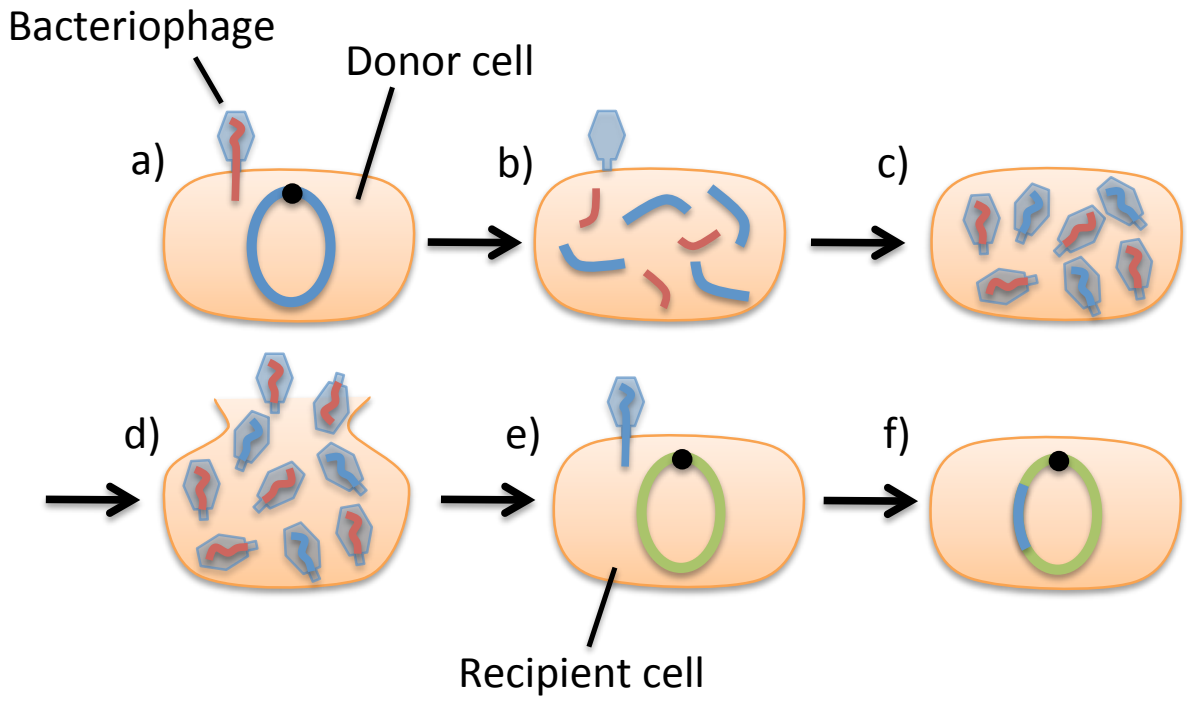


Figure 12: *The life cycle of a bacteriophage. a) The phage inserts its viral DNA into the bacteria. b) The phage enzymes digest the* E. coli *chromosome c) DNA are packed into protein coats d) The* E. coli *donor cell is lysed e) A phage containing donor cell DNA infect the recipient cell f) Homologous recombination between the DNA from the donor cell and the recipient cell*

2.9. Red/ET Cloning

Red/ET recombination also referred to as λ-mediated recombination. The Red/ET recombination technique is used to incorporate genes or DNA fragments into the chromosome by utilizing the principle of homologous recombination. This method was originally developed to inactivate chromosomal genes and is described in the article by Datsenko and Wanner (Datsenko and Wanner [2000]). The method can also be used for insertion, deletion, replacement of DNA sequences and point mutation. In this thesis Red/Et cloning was used to insert a fluorescence gene (the gene of

a GFP like-protein) behind the gene sequence of a protein of interest. This is also referred to as tagging of proteins, because the sequence of the fluorescence gene will be translated with the original gene giving rise to a protein with a fluorescent protein marker.

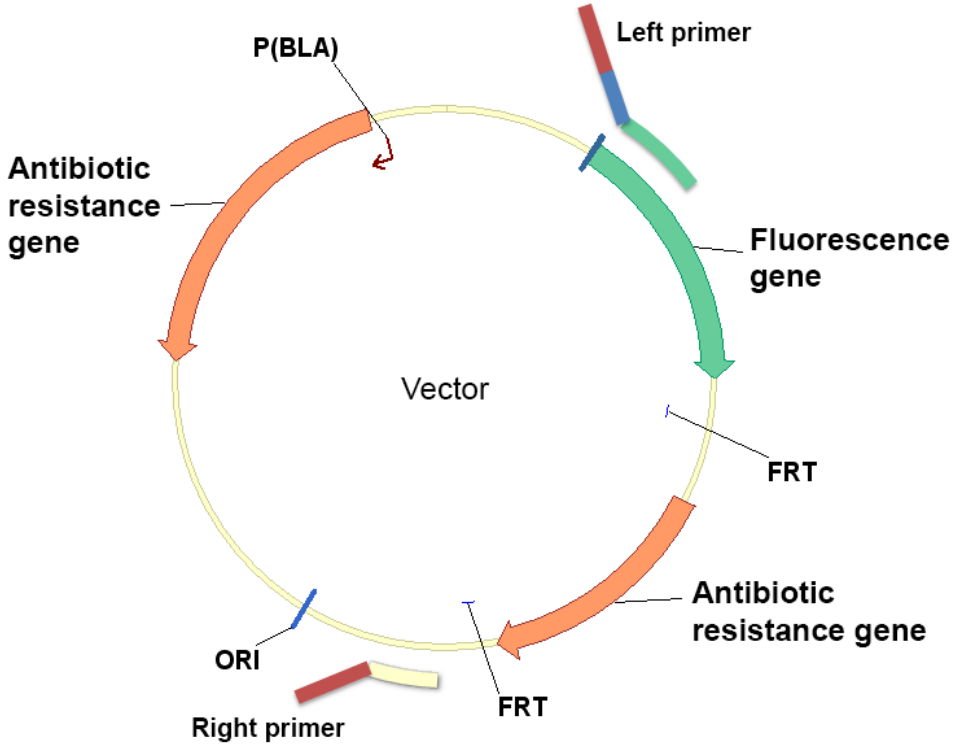


Figure 13: *The base needed to insert a fluorescent gene on to a specific part of the chromosome by the Red/ET cloning method is a plasmid with the gene of interest and an antibiotic resistance gene.*

The starting point is a plasmid with the fluorescence gene and an antibiotic resistance gene with FRT sites on both sides. The FRT sites are included so that the antibiotic resistance marker can be cut out from the chromosome at a later stage if desired. The sequence is amplified using PCR. The primers used in the PCR reaction have two parts, one which is complementary to the plasmid containing the gene sequence, and the other part consist of a homologue tail complementary to the sequence on

the genome at the desired point of insertion. A linker sequence can be added to the left primer, resulting in an amino acid linkage between the protein of interest and the protein tag when the gene is expressed inside the cell.

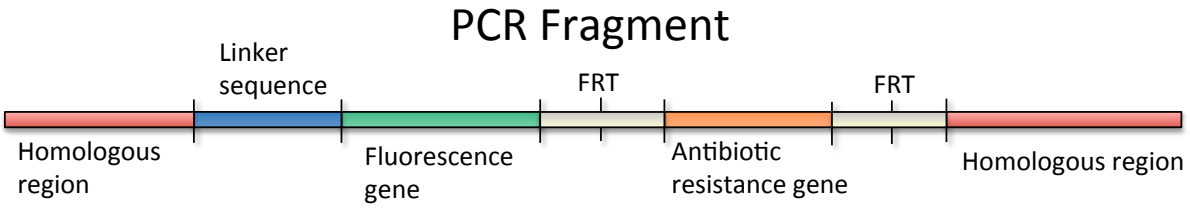


Figure 14: *The PCR fragment achieved from a PCR reaction of the vector and the primers illustrated in figure 13.*

The PCR fragment contain from left a homologous region complementary to the chromosome where the gene is desired inserted, a linker sequence, the gene of the fluorescent protein, the antibiotic resistance gene and a homologous region complementary to the chromosome. This DNA fragment is electroporated into pre-prepared electrocompetent AB1157 cells with pRed/ET. (GeneBridges [2013])

AB1157 cells containing pRed/ET is used in Red/ET cloning. The plasmid contains a tetracycline resistance gene and genes of two proteins (Redα /Redβ) important in homologous recombination behind an arabinose promoter. Redα is exonuclease which digests DNA in 5' to 3' direction, Redβ is a DNA annealing protein. The functional interaction between Redα and Redβ is necessary in order to catalyze the homologous recombination reaction. The recombination occurs between the homology regions of the PCR fragment and chromosome. (GeneBridges [2013])

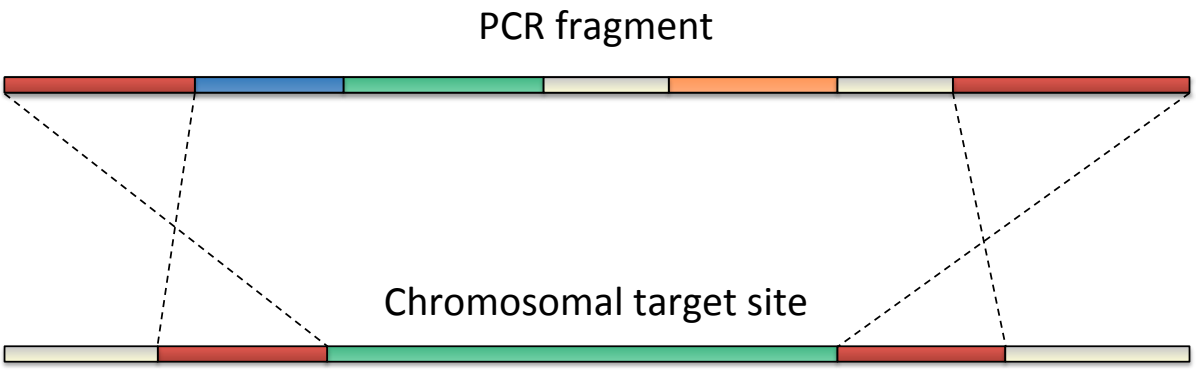


Figure 15: *A schematic drawing of how the PCR fragment is inserted at the target site on the chromosome. A part of the chromosome is deleted in the process of inserting the PCR fragment.*

2.10. Microscopy

The theory in this section is from Rudi Rottenfusser (Rottenfusser et al.) unless specified otherwise.

A light microscope consists of an ocular lens, an objective lens, a condenser and a light source. The sample is placed under the ocular lens behind the objective. The light from the light source under the sample is focused by the condenser so that only the sample is illuminated. The objective forms a picture of the samples, which the lens enhances in the ocular.

Light microscopy faces a fundamental limit when it comes to resolution. The highest achievable resolution is governed by a physical limit known as the diffraction limit, which is an absolute limit resulting from the wavelike nature of light. When light rays pass through a small aperture, they will begin to diverge and interfere.

The diffraction limit is noticeable when observing light from a point source in a microscope. The point will appear as a blurry disk, a so-called *Airy disk* (after its discoverer George Airy). The size of the disk is governed by the *point spread function*

(PSF), which describes how light from a point source is spread spatially when observed in a microscope.

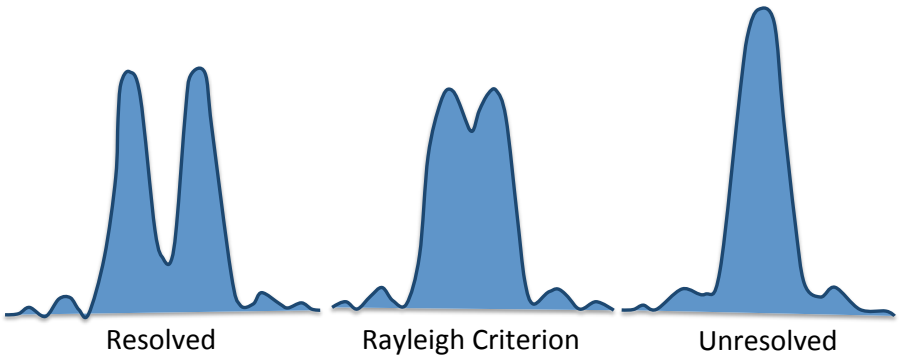


Figure 16: *The point spread functions of two emitting point sources. Left: Two resolvable point sources. Middle: Two point sources at the diffraction limit (Rayleigh criterion), this is the shortest distance two point sources can have while still being resolvable. Right: Two indistinguishable point sources.*

When two point sources are near each other, their point spread functions start overlapping, and when they are sufficiently close, they become indistinguishable. The Rayleigh criterion is a widespread criterion for the minimum resolvable detail. It states that two point sources are just resolved when the first minimum of one point spread function coincides with the principal maximum of the other. If they are any closer, they are not resolvable. This critical distance is known as the diffraction limit, and for a microscope it may be approximated by:

$$d = \frac{1.22\lambda}{\mathrm{n}_{\mathrm{condenser}} \sin(\theta_{\mathrm{condenser}}) + \mathrm{n}_{\mathrm{objective}} \sin(\theta_{\mathrm{objective}})} = \frac{1.22\lambda}{\mathrm{NA}_{\mathrm{condenser}} + \mathrm{NA}_{\mathrm{objective}}}$$

where λ is the wavelength of the light, θ is a number dependent upon the diameter and focal length of the lens, n is the refractive index of the medium between the objective and the cover glass on the sample and NA is the numerical aperture of the microscope, and is a measurement of the ability of the microscope to gather light and resolve detail

at a fixed object distance. (Abramowitz [2004]). d is the minimum distance two point sources can have, while still being distinguishable. This is also referred to as the *resolution* of the microscope.

The practical limit for θ is about 70°-80°. Using oil with a refractive index 1.51, the maximum NA in the objective is typically 1.46. In an air condenser (air has a refractive index 1), the maximum NA is around 0.97. The shortest wavelength of visible light is about 400 nm (violet), so the maximum resolution of a light microscope is thus:

$$\frac{1.22 \cdot 400\,\mathrm{nm}}{1.46 + 0.97} = 201\,\mathrm{nm}$$

However the wavelengths used are not always 400 nm, especially in fluorescent microscopy when the wavelength depends on the fluorochromes in the sample.

2.10.1. Fluorescence Microscopy

A fluorescent microscope works by the same principle as the light microscope, but exploits the fluorescence of the object. Fluorescent proteins are used to tag proteins of interests by inserting the gene of a fluorescent protein into the target proteins reading frame (see section 2.9). The fluorescent proteins in the sample re-emits light with a shorter wavelength upon illumination, and can be visualized by microscopy snap-shot imaging. The biological macromolecule green fluorescent protein (GFP) was the first discovered in the jellyfish *Aequorea victoria* (Shimomura [1962]). The GFP consists of 238 amino acids forming an extremely rigid β-barrel fold surrounding a central α-helix (Prasher et al. [1992], Ormo et al. [1996], Yang et al. [1996]). The α-helix is the chromophore and is formed during post translational modification of the protein in the presence of oxygen, resulting in maturation of the GFP to its fluorescent form (Heim et al. [1994], Cubitt et al. [1995]). The wild type GFP protein emits light in the spectrum around 500 nm, which is in the low green portion of the visible spectrum. After GFP

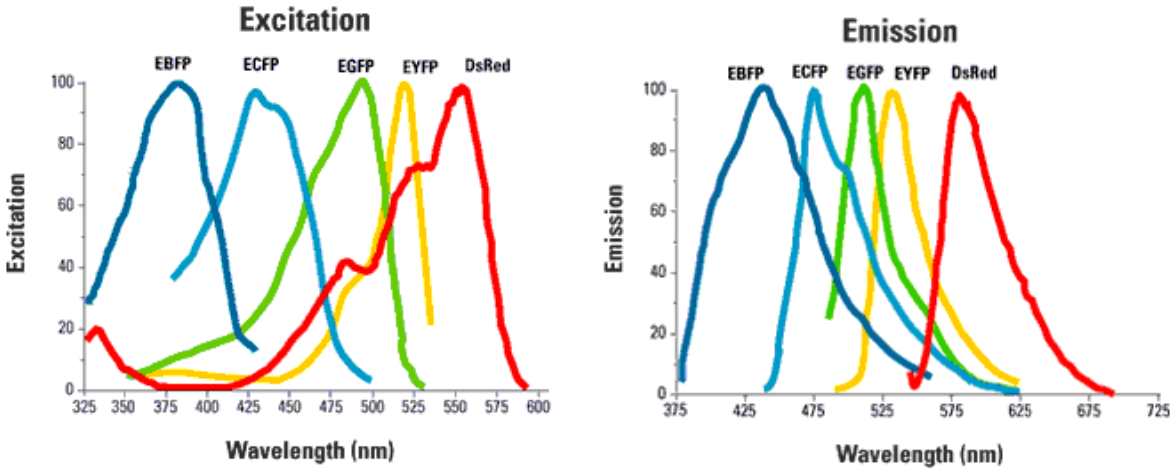


Figure 17: *Exitation and emission spectrum of the mentioned fluorescent proteins Duke-University [2012].*

was first discovered, several different mutants of the protein have been engineered to give different characteristic color signals (Tsien [1998], Matz et al. [2002]). Among the color mutants made from the GFP are BFP (blue fluorescent protein), CFP (cyan fluorescent protein) and YFP (yellow fluorescent protein), all emitting light in the lower part of the visible spectrum. GFP-like molecules emitting light in the red part of the visible spectrum (DsRed) were eventually found in corals (Matz et al. [1999], Matz et al. [2002]). Mutations changing the proteins' folding, improving brightness and photo-stability of the proteins have been performed to optimize the fluorescent proteins.

Table 2: *The wavelenght of the different filters in the Leica DM6000*

Filtercube	Excitation	Emission	Fluorescens
CFP	BP436/20	BP480/40	CFP
GFP	BP470/40	BP525/50	GFP
JP2	510/20	560/40	YFP
Y3	545/30	BP610/75	Cy3

Calibration of the microscope filters for multicolor imaging of colocalization studies was performed using TetraSpeck™ Microspheres (0.2 and 0.5 µm molecular probes). A microscope slide containing microspheres stained with four fluorescent dyes was used to make sure that a shift did not occur when switching between filters in the microscope. LifeTechnologiesCorporation [2013]

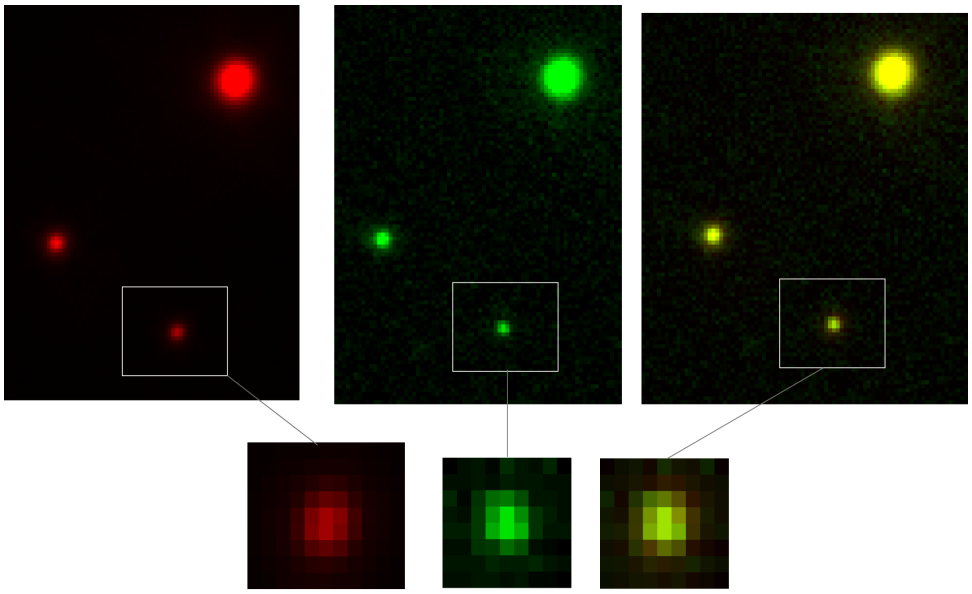
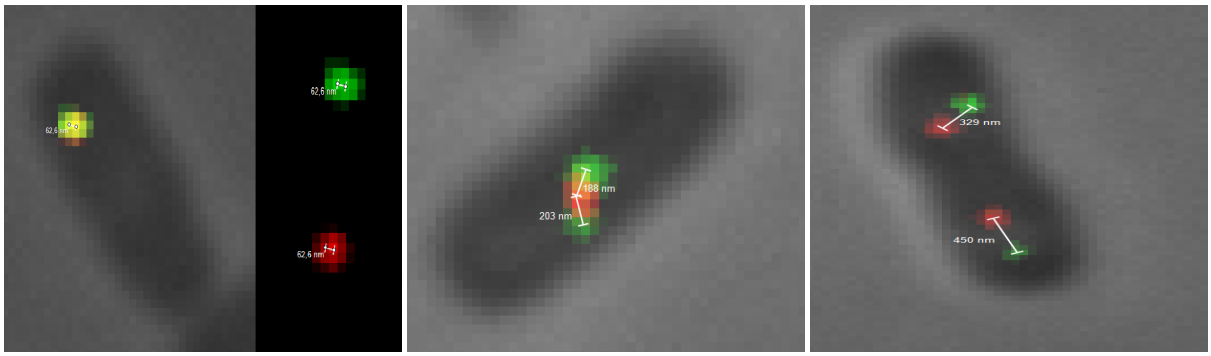


Figure 18: *Microscopy images from filter calibration using TetraSpeck™ micromolecular probes. The images of the green and red foci are taken with two different filters, and when the two images are laid on top of each other they are completely overlapping (yellow foci).*

2.10.2. Measuring Foci Distance in Microscopic Samples

The microscopy images in this thesis are prepared in the microscope Leica DM6000. The camera DFC350 FX installed on the microscope takes the fluorescent pictures separately and the different pictures are laid on top of each other, making an image of the foci and their location in the cell. Image analysis is conducted in the program

LasAF. The contrast is adjusted for each cell so that we clearly see the brightest pixel, and the measurements of distances between the foci are performed in a straight line between the brightest pixels in the center of each focus. If there is more than one focus of one of the colors, two measurements are performed, and if the foci are colocalized, the two pictures are looked at individually and the center point of each focus is found. The distance is measured from the two chosen points.



(a) **(b)** **(c)**

Figure 19: *Examples of foci measurements; a) Two colocalized foci, the center of the foci are found by analyzing the two pictures individually. b) Two green foci on each side of a red focus, both distances are measured. c) Two paired foci on each end of the cell.*

2.11. Flow Cytometry

In vivo analysis of DNA content and cell size of individual cells can be determined rapidly and precise by flow cytometry (Boye and Løbner-Olesen [1991], Skarstad et al. [1995]). The cells are stained with a DNA-specific (Hoechst 33258) and a protein-specific dye (fluorescein thiosocyanate, FITC), and the cells illuminated one by one with a laser beam. The flow cytometer exploits dye molecules' ability to absorb light of one wavelength and emit light of another wavelength. The emitted light is proportional to the DNA content and cell size. The method has a very high sensitivity

and provides information about the cell cycle, the timing of initiation of replication, the speed of replication fork movement, cell size and DNA content at the time of initiation (Skarstad et al. [1995]).

2.11.1. Preparing Exponential and Run Out Samples

The cell culture is grown exponentially until OD 0.15. The sample is split into two parts, the exponentially grown cells and another part which is treated with two different antibiotics (rifampicin and cepalexin). The samples treated with antibiotics are called replication run-out sample. Rifampicin binds to the β-subunit of RNA polymerase and thereby inactivating it. The initiation of DNA replication is dependent on RNA polymerase, but not the elongation, hence ongoing replication will finish. Cephalexin hinders the formation of cell septum and thereby prevents cell division. The run out sample will therefore only contain fully replicated chromosomes. By analyzing the run out sample we can decide when initiation of replication occurs and if the cells initiate replication in synchrony. Cells with synchronous initiation of replication will have 2^n chromosomes, while cells with asynchronous initiation will contain a chromosomal number$\neq 2^n$. (Skarstad et al. [1986]; Boye and Løbner-Olesen [1991])

2.11.2. Staining Flow Samples

The fluorochrome dyes, FITC (fluorescein-isothiocyanate) and Hoechst 33258, are added to both the samples. FITC makes covalent binding to proteins, while Hoechst is an equilibrium color which binds to DNA. A culture of CM735 with a doubling time of 3.5 hours and just one or two chromosomes per cell is used as an internal standard to identify the DNA content in the analyzed samples. The standard cells are separated into two parts before FITC staining. One of them is stained with FITC and thereby called FITC-positive, while the other one is called FITC-negative. Both of the two standard samples are dyed with Hoechst. Subsequently, after staining of samples,

a small amount of FITC-negative cells is added to each sample. During analysis of cell samples in the flow cytometer the standard cells are recognized as a population without FITC.

The distribution of DNA or cell size is found by the help of the flow cytometer LSR II. The flow cytometer will generate histograms over the DNA and protein distribution, which again can be used to calculate the cell cycle pattern.

2.12. Cell Cycle Analysis

The cell cycle pattern can be calculated using the DNA histogram obtained in the flow cytometer and the cell's doubling time. The number of newborn cells is always twice the number of dividing cells since one dividing cell gives rise to two new born cells.

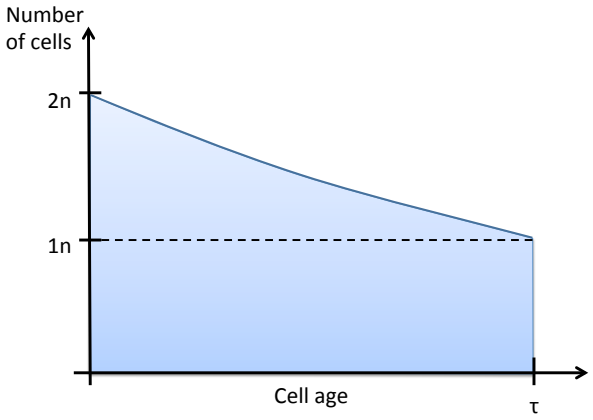


Figure 20: *The distribution of cells in the population by age*

The equation for the growing population, where τ is the doubling time, n the number of cells and t the time:

$$n(t) = n(0) \cdot 2^{-\frac{t}{\tau}} \tag{1}$$

The equation is plotted in figure 20 from age 0 to τ. This number of cells decreases

by time. By normalizing, i.e. dividing this expression by the area under the curve from t=0 to t=τ, the following expression was obtained:

$$\mathrm{p(t)} = \frac{2\ln 2}{\tau} \cdot 2^{-\frac{\mathrm{t}}{\tau}} \tag{2}$$

The area under this curve from t=0 to t=τ is 1, thus the function p represents the probability distribution.

The age distribution can be used statistically to calculate the cell cycle parameters and the distribution of replication forks in a population. To be able to say more about the cell cycle parameters, we need to compare DNA histograms from replication run out samples with the FITC negative standard. The DNA histograms of the replication run out samples normally have two peaks. The first peak represents the cells that did not initiate a new round of replication when the antibiotics rifampicin and cephalexin were added to the sample. The second peak represents the cells that had already initiated replication. By looking at the run out histogram one can also decide where initiation of replication starts (see figure 21). If the DNA histograms show a two and four origin peak, the initiation of replication starts in the mother cell, four and eight origin peak indicates that it starts in the grandmother cell. And if the replication run out histogram shows eight and sixteen origin peaks, initiation of replication occurs in the great grandmother cell.

The initiation age a_i is the time when the cell initiates DNA replication, i.e. it is the start of the C-period. Integration of the age distribution (2) between the age 0 to a_i we get the equation:

$$\mathrm{F} = 2 - 2^{\frac{\tau - \mathrm{ai}}{\tau}} \tag{3}$$

F is the number of cells between the ages 0 and a_i, and can be found in the DNA histogram. This is equivalent to the number of cells in the first peak in the run out

samples. When rewriting the equation with respect to a_i, we can find the timing of initiation:

$$\mathrm{ai} = \tau - \frac{\ln(2 - \mathrm{F})}{\ln 2\tau} \tag{4}$$

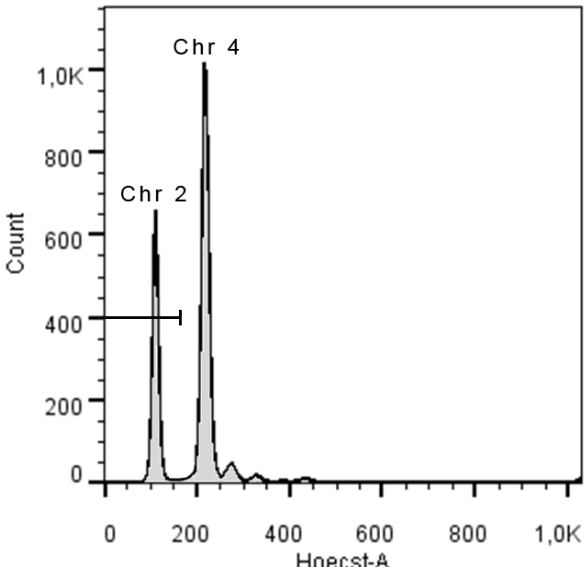


Figure 21: *The F value equals the % of cells in the first chromosome peak (chr 2)*

We can decide the length of the C-period and the termination age (a^t) if both a_i and τ is known, with the help of the DNA content in the cells. a^t is the cell age at termination of replication, when the C-period stops and the D-period starts. The length of the C-period can be determined using data from the DNA histogram of the exponential sample. The width of the DNA histogram range from one value to twice the value, the DNA content in the histogram is equivalent to the range of DNA from a newborn cell to a dividing cell. The first value, K1, represent the DNA content of newborn cells, K2 the DNA content of dividing cells.

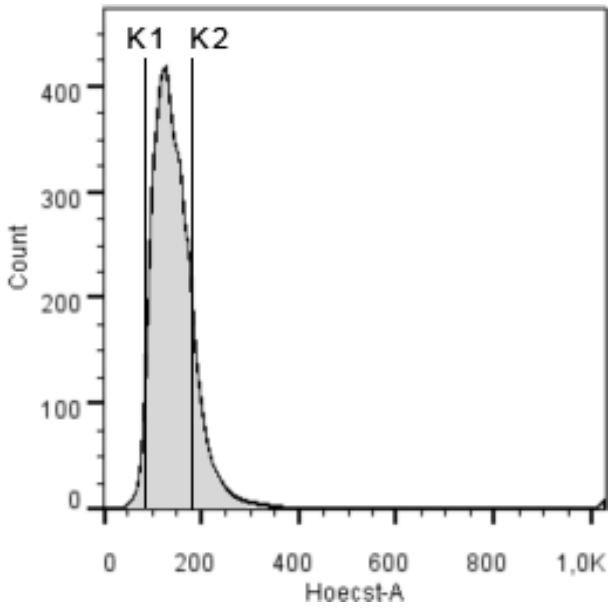


Figure 22: *A DNA histogram of exponentially grown cell showing K1 and K2*

The first value, K1, represent the DNA content of newborn cells, K2 the DNA content of dividing cells. The equation with respect on K1, for cells that initiate in the mother cell:

$$\mathrm{K1} = 1 + \frac{\tau - \mathrm{ai}}{\mathrm{C}} \tag{5}$$

The equation (5) applies only if the cells initiate replication in the mother cell, other equations are used to find the cell cycle parameters for cells with other types of initiation patterns. For more information refer to the paper by Caroline Stokke (Stokke et al. [2012]).

The termination age is found by subtracting the doubling time from the initiation age a_i and the C-period. When DNA replication is initiated in the mother cell, the cell cycle extend over two generations. C+D equals therefore the doubling time times two minus the initiation age a_i. The D-period can be calculated by:

$$\mathrm{D} = 2\tau - \mathrm{ai} - \mathrm{C} \tag{6}$$

The cell cycle simulation program developed by Caroline Stokke (Stokke et al. [2012]) is used to simulate the cell cycle parameters by finding the best fit between the theoretical and experimental DNA histogram. The number of cells per channel of the experimental histogram is obtained by analyzing the flow sample in FlowJo 7.6.5. The theoretical histogram is estimated in the simulating program by calculating the cell cycle parameters and the age distribution.

2.13. Measurements of Spontaneous Mutations

The resistance against rifampicin occurs after a single point mutation in the RNA polymerase gene on the *E. coli* chromosome. This knowledge can be used to measure the ratio of spontaneous mutations which occurs in a strain. Spontaneous mutations occur in the absence of exogenous DNA damaging agents, but arise by errors linked to DNA replication. The mutation ratio is measured by counting the number of mutants on a 100 µg/ml rifampicin agar and dividing it on number of cells plated.

$$\frac{\text{Number of cells on the rifampicin agar}}{\text{Number of cells plated on the agar}} = \text{Mutation ratio}$$

This method is used to measure how many of the cells in a population have a specific mutation, which again reflect the mutation frequency.

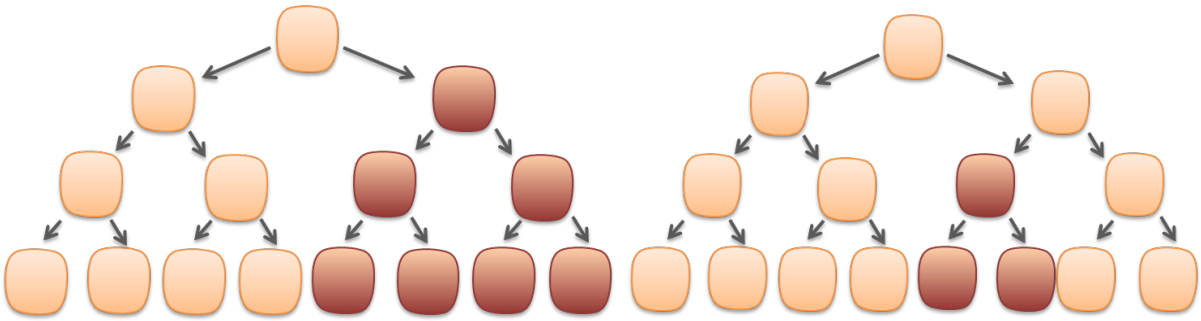


Figure 23: *The figure illustrates rifampicin resistance. Strains with high mutation frequency will more likely mutate a cell in the early stages and have more mutants than a strain with low mutation frequency.*

3. Results

To study the localization of MutH relative to SeqA on hemimethylated DNA *in vivo*, new *E. coli* strains needed to be constructed. Two different strategies were used.

The first was trying to visualize MutH with a fluorescent tag. We decided to construct a strain with a fluorescent protein fused to the C-terminal end of the MutH protein by the Red/ET cloning method. This was justified by the assumption that an alteration of the C-terminal end does not affect MutH endonuclease activity or its ability to interact with MutL. This is motivated by the result in Loh et al. [2001], where deletion of five amino acids on the far end of the C-terminal of the protein did not affect MutH's function *in vitro* or *in vivo*. The second strategy was based on the strain SF149 with the fluorescent proteins YFP and CFP fused to SeqA and SSB, respectively. The idea was to study if a *mutH* deletion or a mutant *dnaQ* (a non-functional μ subunit) could affect the distance between these two protein complexes.

3.1. Cloning the Gene of the Fluoroescent Protein mCerulean3 into the Vector pSF36

We wanted to tag MutH with the fluorescent protein mCerulean3 using Red/ET cloning. mCerulean3 is an optimized mutant of the CFP with the same emission spectrum. First a vector was constructed that contained the mCerulean3 gene upstream of the chloramphenicol resistance gene. The resistance gene was necessary as a marker in the tagging process. mCerulean3 was therefore subcloned via pGEM-T Easy and then inserted upstream of the chloramphenicol resistance gene in the pSF36 vector.

3.1.1. Amplification of mCerulean3 with PCR

The gene of the fluorescent protein mCerulean3 was amplified from the plasmid pmCerulean3 using PCR (Markwardt ML [2011]). The primers that were constructed to amplify the mCerulean3 gene from the pmCerulean3 were named Right/Left_mCerulean3 (see table A.3). The left primer was constructed so that a BamHI cutting site was inserted upstream of the mCerulean3 gene for later gene direction identification, and the start codon of the gene was removed, because the translation will continue from *mut*H to mCerulean3.

The PCR was performed as described in section B.3, with the polymerase Pfu Ultra II High Fidelity. A part of the PCR reaction was checked with 0.8% agarose gel electrophoresis to make sure that the sequence had the correct size B.2.

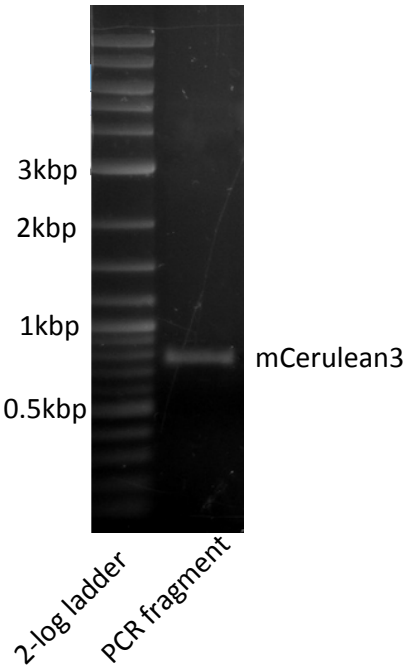


Figure 24: *Agarose electrophoresis gel of a PCR reaction of pmCerulean3 with Left/Right_mCerulean3 primers. A 2-log ladder was used as reference*

The PCR fragment was approximately 750 bp, which corresponds with the size of mCerulean3. The rest of the PCR product was purified using PCR clean up (B.6).

3.1.2. Cloning the mCerulean Gene into pGem-T Easy

The mCerulean3 gene was subcloned via the pGEM-T Easy vector (see 2.7.2). Adenine nucleotides were added to the 3' ends by the process of poly-A tailing (see B.10). The mCerulean3 gene was then ligated into the pGEM-T Easy vector in the ratio 3:1, respectively (see B.9). The ligation mix was transformed into the chemically competent JM109 cells, and plated on an ampicillin/IPTG/x-gal agar plate (see B.12.1). Only the cells that had the pGEM-T Easy plasmid survived on the ampicillin agar, and only the colonies of cells with an nonfunctional *lacZ* gene (i.e. have mCerulean3 inserted) appeared white on the agar (see 2.7.2).

A positive and a negative control were included in the transformation. The positive control measured the transformation efficiency of the chemical competent cells JM109 by transforming them with the supercoiled vector pUC19 containing an ampicillin resistances gene. The negative control contained only the competent cells on an ampicillin agar to monitor the ampicillin resistance.

Table 3: *The result from the transformation of pGEM-T Easy with mCerulean3*

Reaction	Transformation	Transformants
Negative control	JM109 cells	0
Positive control	JM109 + pUC19	418
Ligation mix	JM109 + pGEM-T Easy with mCerulean3	97

It was hard to identify if the colonies from the ligation mix were blue or white, so nine of the colonies which appeared white were chosen and reinoculated on a new ampicillin/IPTG/x-gal plate. One of the nine colonies proved to be white, i.e. having a mCeruelan3 gene inserted into the *lacZ* gene in the pGEM-T Easy vector. This

plasmid was named pKB01, and the strain KB01.

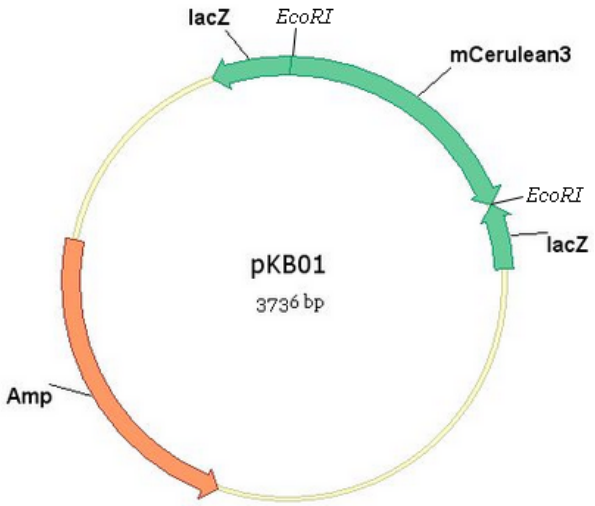


Figure 25: *The KB01 vector was made by ligating the mCerulean3 fragment into the* lac*Z gene of the pGEM-T Easy vector. The* lacZ *gene is thus nonfunctional. The direction of mCerulean3 in the vector is not known. This figure is an illustration of one of the two potential directions of the gene.*

3.1.3. Cloning into the pSF36 Vector

The pSF36 plasmid consists of 3706 bp and contains resistance genes for ampicillin and chloramphenicol. The mCerulean3 gene was cut out of the pKB01 with EcoRI and inserted into pSF36 cut with the same restriction enzyme (B.7).

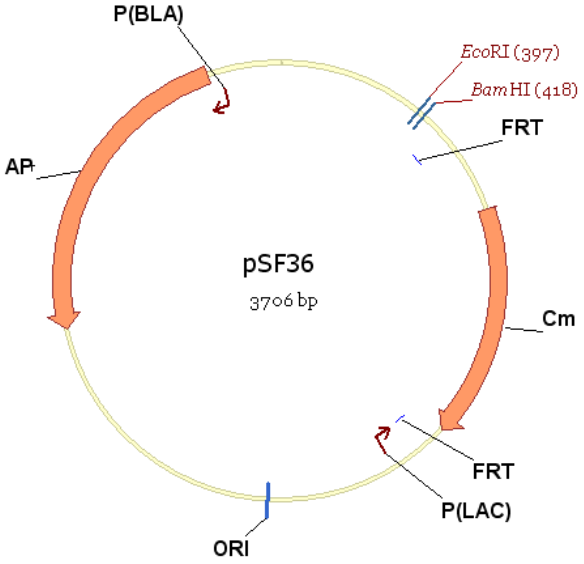


Figure 26: *The vector pSF36*

pKB01 (pGEM-T Easy + mCerulean3) and pSF36 were purified with midiprep plasmid purification (see section B.4.2). EcoRI cutting sites are located on both sides of the inserted gene in the pKB01 vector. The EcoRI cut pKB01 was separated using 0.8 % agarose gel electrophoresis, giving bands at approximately 3000 and 750 bp (see figure 27).

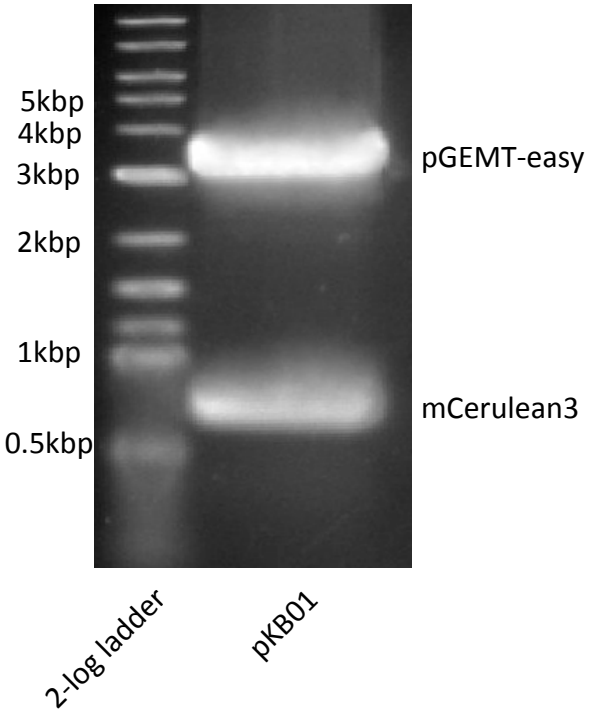


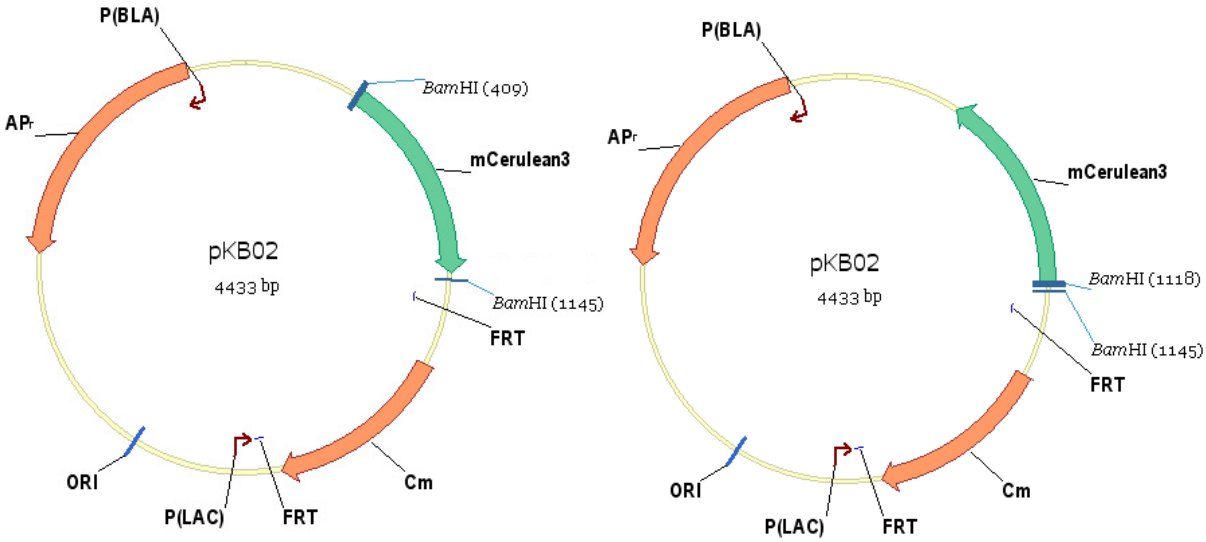
Figure 27: *Agarose electrophoresis gel of KB01 cut with EcoRI.*

The EcoRI cut mCerulean3 was purified from the gel (see section B.6). The phosphor group on each end of the EcoR1 cut pSF36 was removed by dephosphorylation (see B.8) to prevent the vector from self religation. A ligation mix was then prepared containing the dephosphorylated pSF36 and the purified mCerulean3 in ratio 1:3, respectively (see B.9). The ligation mix was transformed by electroporation into the electrocompetent DH10B cells (B.12.2). The transformants were selected on LB plates with ampicillin and chloramphenicol. The same control experiments were performed as in section 3.1.2, with both a positive and negative control of the transformation.

Table 4: *Results from transformation of pSF36 with mCerulean3*

Reaction	Transformation	Transformants
Negative control	Competent DH10B cells	0
Positive control	Competent DH10B cells +pUC19	629
Ligation mix	Competent DH10B cells + pSF36 with mCerulean3	48

Ligation of mCerulean3 into the EcoRI site of the pSF36 resulted in two potential plasmids. One in which the mCerulean3 gene was inserted in the same direction as the chloramphenicol resistance gene and one in which the gene was inserted the opposite way (see figure 28). Later use of the plasmid requires the vector to have the mCerulean3 gene in the same direction as the chloramphenicol resistance gene.



(a) *mCerulean3 inserted into pSF36 in the desired direction*

(b) *mCerulean3 inserted into pSF36 in the opposite of the desired direction*

Figure 28: *The two potential vectors which can occur from ligation of mCerulean3 into EcoR1 cut pSF36*

Five colonies were chosen from the agar plate and the plasmids from the cells were harvested and purified using miniprep (B.4.1). The five chosen colonies were temporarily named from 1 to 5. To control if the mCerulean3 gene was inserted into pSF36 and to check the direction of the gene in the plasmid, the plasmids were cut with the restriction enzyme BamHI. mCerulean3 has a BamHI cutting site on the 5'end of the gene, and the plasmid has a BamHI cutting site upstream of the chlorampenicol gene. Plasmids with mCerulean3 inserted in the same direction as the chloramphenicol resistance gene will have two DNA fragments in an agarose gel, one

at approximately 750 bp (mCerulean3) and one at 3600 bp (vector). Plasmids with mCerulean3 inserted in the other direction will only have one DNA fragment at about 4400 bp. Figure 29 shows that three plasmids proved to have mCeruelan3 inserted in the same direction as the chloramphenicol resistance gene (clone 3-5).

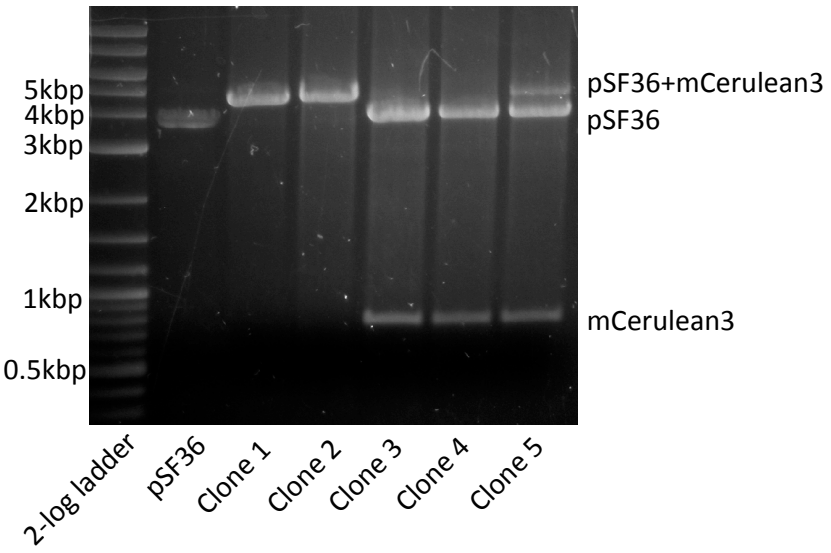


Figure 29: *Agarose gel electrophoresis of pSF36 with mCerulean3 inserted into EcoRI site. Lane 1: 2-log ladder, lane 2: uncut pSF36, lane 3-7: plasmids from clone 1-5 cut with BamHI.*

One plasmid in which mCerulean3 transcribed in the opposite direction as the chloramphenicol gene was named pKB02, and one plasmid in which mCerulean3 was transcribed in the same direction as the chloramphenicol gene was called pKB03. The KB02 and the KB03 which contained pKB02 and pKB03 respectively, were kept and prepared for deep freeze storage (see B.1.2).

3.2. Tagging the C-terminal end of Chromosomal MutH with mKate2

It turned out to be more time consuming than expected to insert the mCerulean3 gene into the pSF36 vector, and a different strategy had to be taken up in parallel. pEH04, made by Emily Helgersen, was therefore chosen. This plasmid is a modified pSF36 with the red fluorescent protein mKate2 (a color mutant of the fluorescent protein DsRed) (without the start codon) inserted into the EcoRI restriction site. pEH04 has 4411kb and the mKate2 gene goes in the same direction as the chloramphenicol resistance gene.

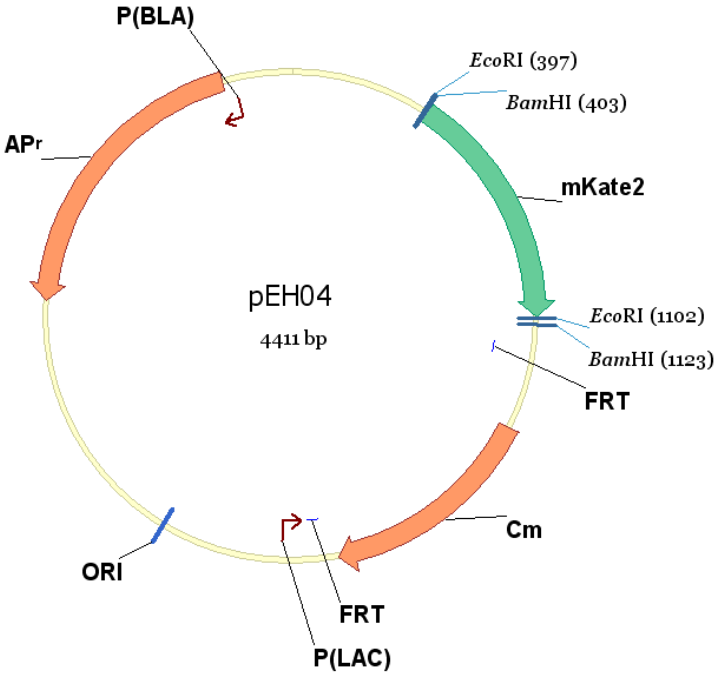


Figure 30: *pEH04 made by Emily Helgersen*

The plasmid pEH04 is preserved in the strain with the name EH04, which was grown in LB overnight and harvested the next day by miniprep B.4.1.

3.2.1. Primers for the PCR Fragment used in Red/ET Cloning

The left primer (Left_mKate2_MutH) was designed so that the 50 first base pairs were homologous to the sequence at the end of the *mutH* gene, not including the stop codon. The next part of the left primer consists of 36 bp of linker sequence which after Red/ET cloning results in a 12 aa linker of arginine and proline, connecting MutH and mKate2 at the protein level. The last part of the left primer contains 23 bp homologous to the beginning of mKate2 (on the plasmid). The right primer (Right_mKate2_30bpMutH) was made in a 3' to 5' direction, reverse complementary, consisting of 50 bp with homology to the area 30 bp downstream of *mutH* and 24 bp with homology to pEH04 downstream of the chloramphenicol resistance gene (the sequence can be found in table A.3) (see also 2.6.1).

In contrast to other primers, the whole primer sequence was not taken into account when deciding the melting temperature, only the parts that were complementary to pEH04.

3.2.2. Tagging the C-terminal end of MutH with mKate2 by Removal of 30 bp Downstream of *mutH*

A large PCR reaction containing the primers Left_mKate2_MutH and Right_mKate2_30bpMutH, the plasmid pEH04 and Pfu Ulatra II fusion HS DNA polymerase was prepared equivalent to 12 reactions (see B.3). The PCR reaction was purified (see B.6) and the DNA concentration was measured to 120 ng/µl. To conduct a successful Red/ET cloning we need to add between 400-800 ng DNA into the transformation sample, but only 2 µl can be added. If we add more we encounter problems with either too much salt in the solution, causing an electrical discharge during electroporation or affecting the cell density. Both result in no transformation. The sample was therefore precipitated to increase the DNA concentration (see B.11), the concentration was raised to 260

ng/ml. 20 µl of the PCR reaction was examined by agarose gel electrophoresis. The fragment was expected to contain the homologous tails, the mKate2 gene and the chloramphenicol resistance gene with a lenght of about 1500 bp.

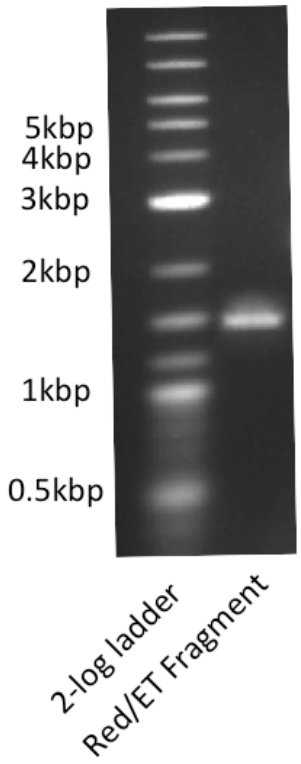


Figure 31: *Agarose gel electrophoresis of the mKate2-cam PCR fragment used in the Red/ET cloning*

One band of approximately 1500 bp was observed in the gel, which corresponds to the PCR fragment of interest. The goal of Red/ET cloning was to insert the PCR product downstream of *mutH* on the chromosome. To obtain this, the strain AB1157 with pRed/ET (see 2.9) was used (the Red/ET-cloning protocol B.14) was followed. 2 µl (520 ng) of PCR product was transformed into the AB1157 pRed/ET cells. The sample was incubated in LB for 3 hours at 37℃ to allow the recombination process to occur, before the sample was plated on a LB agar containing chloramphenicol.

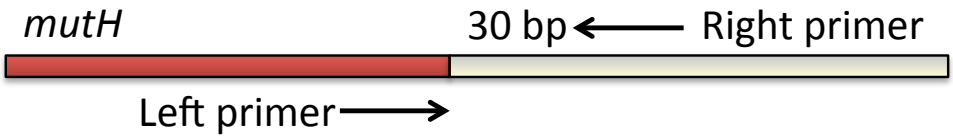
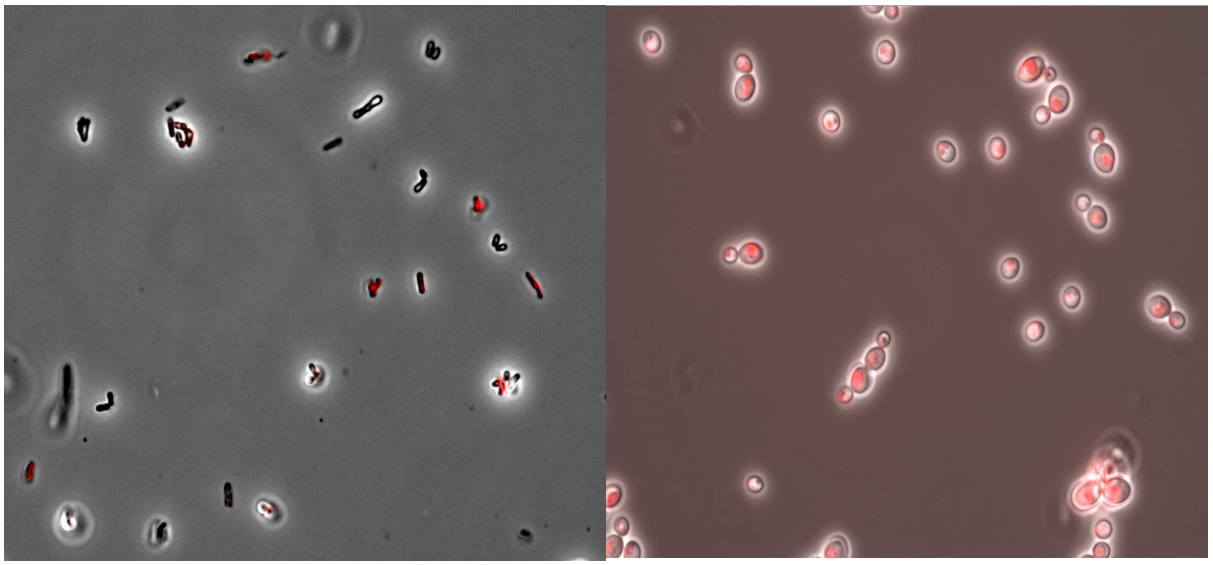


Figure 32: *The figure illustrates where the homologous part of the primers are complementary to the chromosome. 30 bp downstream of the* mutH *gene are removed in the homologous recombination process.*

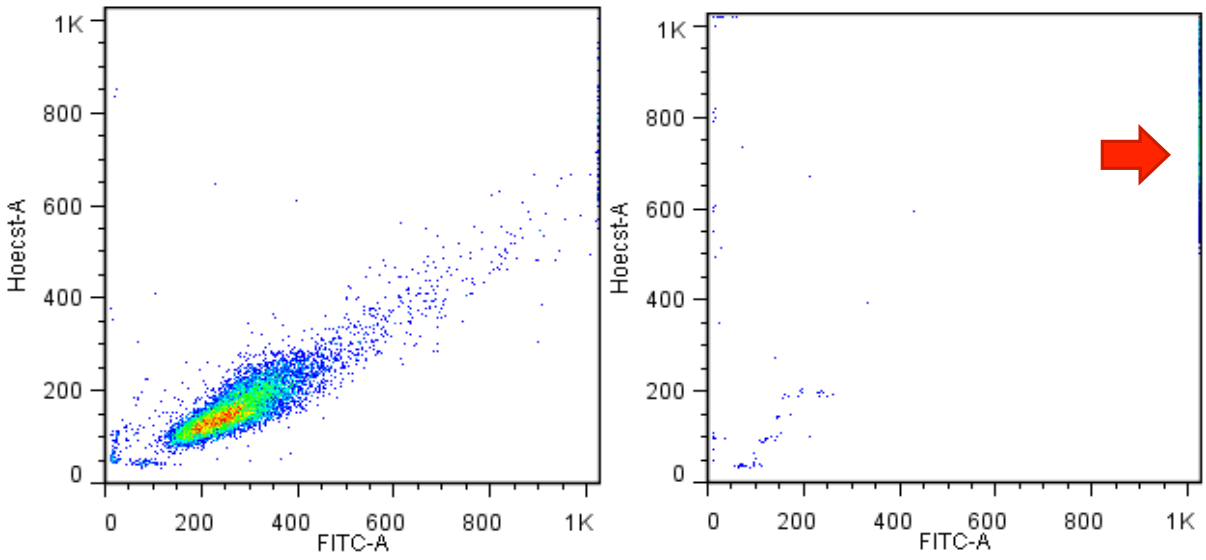
Only very small colonies of bacterial cells were visible to the naked eye the next day. The plate was therefore incubated another day. The third day one of the colonies was chosen and named KB04. One part of the chosen colony from the LB agar was used to prepare an ONC. The other part was prepared for microscopy by resolving it in 10 µl nuclease free water and placing it on an glass slide with a cover glass added on top (see figure 33a). The next day the same thing was performed with 10 µl of the overnight culture. Microscopy and investigation of fluorescence were conducted i Leica DM6000 with the filter Y3 which is complementary with the excitation and emission spectra of mKate2 (see table 2) (Corporation [2012]).



(a) *A colony from the LB agar* **(b)** *Overnight culture in LB*

Figure 33: *Microscopy images of the strain KB04 from the Red/ET-cloning under two different conditions. a) The first sample was prepared directly from a colony on the LB agar. b) The microscopy sample was prepared from an LB overnight culture. These cells were large and had a spherical shape. The cells seem to have lost their ability to form a normal rod shaped structure, and the cells appear to have expanded greatly.*

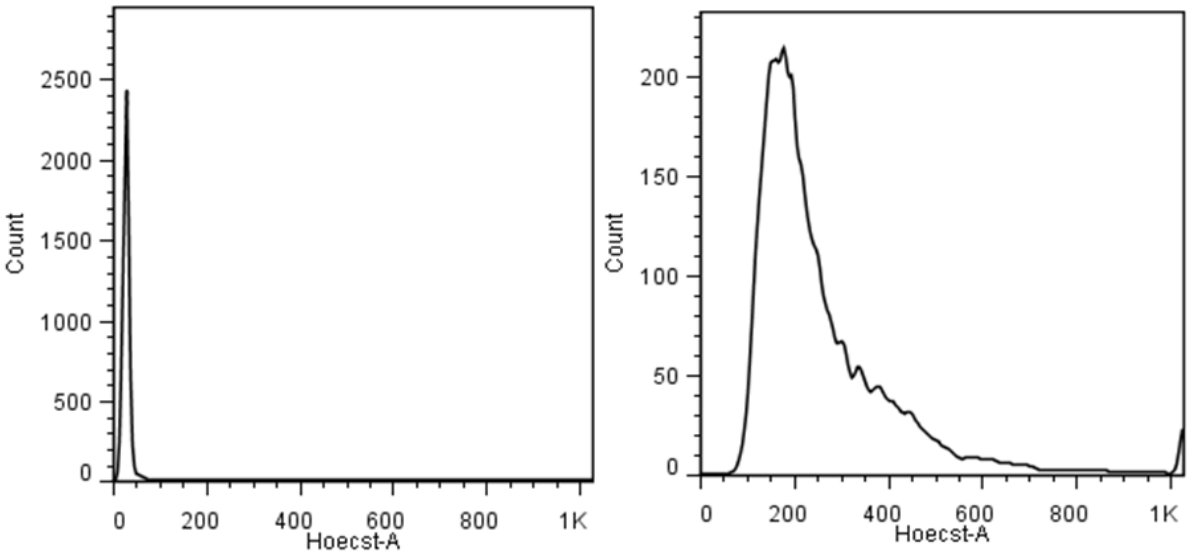
Flow cytometry analysis and investigation of growth rate was performed of KB04 cells in LB medium at 37℃. The doubling time was found to be 83 min. The flow samples were analyzed by LSR II, and show the same as what we saw in the microscope, that the KB04 cells are very large. We saw in the Hoechst/FITC plot (DNA content against cell size) that the cell size and DNA content were much larger than the cells in the wt strain AB1157.



(a) *AB1157* **(b)** *KB04*

Figure 34: *The plot show the DNA content (y-axis) against the cell size (x-axis) for exponentially grown AB1157 and KB04. The KB04 cells have both a DNA content and a cell size larger than what the plot show.*

The next time KB04 was analysed with flow cytomerty the Hoechst threshold was down regulated so that the one chromosome peak in the standard cells was in channel 5 (usually the one chomosome peak lies in channel 15) (see figure 35a). The standard strain has both a one- and a small two-chromosome peak, but when the Hoechst threshold was down regulated as much as we had to do her, the peaks merged into each other. This was done to get the whole DNA content of the KB04 cell into the histogram. The Hoechst histogram of the standard cells shows that the one chromosome peak lies in gate 20, giving approximately 0.23 mbp per gate (4.6 mbp/20)..



(a) *The DNA content from the CM735 standard cell*

(b) *The DNA content from the exponentally grown KB04*

Figure 35: *The Hoechst histograms for the standard strain CM735 (analyzed in channel 5) and the KB04 strain. The KB04 strain has a very large DNA content*

The exponentially grown KB04 cells have a DNA content around gate 200, meaning that the cells have about 46 mbp per cell (0.23 mbp x 200), a DNA content equivalent to 10 *E. coli* cells.

Data from the Profiling of *Escherichia coli* chromosome database (the E.coli Chromosome [2013]) suggests that the promoter of the gene downstream of *mutH*, the *ygdQ* gene, starts 6 bp after the *mutH* stop codon. In the Red/ET cloning, 30 base pairs behind *mutH* were removed in the homologous recombination process, and we therefore believe that it is safe to assume that the KB04 does not express the *ygdQ* gene. Thus, it cannot be used in further investigation of MutH foci.

3.2.3. Tagging the C-terminal end of MutH with mKate2 by Removal of 6 bp Downstream of *mutH*

In order to avoid removal of the *ygdQ* promoter region, tagging of *mutH* with mKate2 was performed by constructing a primer that was homologous to the chromosome 6 bp downstream of the *mutH* gene (Right_mKate2_6bpMutH, the sequence is shown in table A.3). The left primer was not changed. PCR and Red/ET cloning was performed as previous described (see 3.2.2) (where the complementary part of the primers are located on the chromosome is illustrated in figure 36). The transformation mixture was plated on a LB agar plate containing chloramphenicol and incubated in 37℃. Five different attempts on the Red/ET cloning had to be performed before finally two colonies appeared on one of the LB agar plates. They were named KB05-1 and KB05-2 and overnight cultures of the two colonies were prepared.

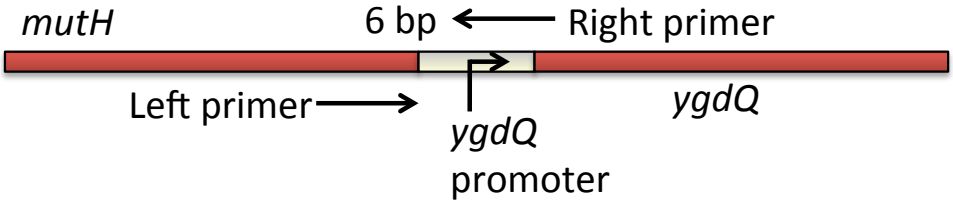


Figure 36: *The figure illustrates where the homologous part of the primers are complementary to the chromosome. Only 6 bp downstream of the* mutH *gene are now removed in the homologous recombination process.*

Microscopy samples of the strains were prepared directly from the overnight culture of the two strains. In the images obtained from the microscope (data not shown) there was some fluorescence, but the fluorescence was not concentrated in the cells. Two new primers were therefore synthesized (Left/Right_insert_behind_MutH sequence is shown in table A.3) to analyze with PCR if these two colonies had the fluorescent protein mKate2 inserted downstream of the *mutH* gene. The left primer was placed in the end of the *mutH* gene, and the right primer was placed in front of the *ygdQ* gene.

Chromosomal DNA from the two colonies, and a control genome from AB1157 (wild type) were purified (see B.5). PCR was performed as described (B.3), and the PCR products were analyzed with agarose gel electrophoresis (B.2).

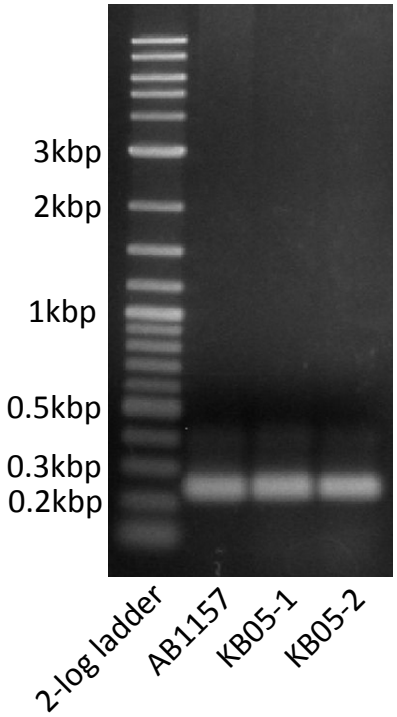


Figure 37: *The PCR fragment from the KB05-1 and KB05-2 were the same size as the AB1157* E. coli *wild type cell, and therefore had no insertion behind* mutH

If MutH was tagged with mKate2, the fragment should be about 900 bp long. We found that the two colonies contained DNA fragment of the same size as the wild type (~250 bp), meaning that mKate2 was not inserted downstream of *mutH*. It was at this point desided to investigate the localization of MutH more indirectly.

3.3. Construction and Studies of new *E. coli* Strains from the SF149 Strain

The strain SF149 contains SeqA protein tagged with YFP and SSB protein tagged with CFP. The YFP gene lies right downstream of the chromosomal SeqA gene, and is translated with SeqA. The SSB-CFP is (unlike SeqA-YFP) non-functional, so the cell still requires the wt SSB. The connected SSB-CFP genes are located in the *lamB* site, upstream of the wild type SSB gene. The SSB-CFP construct is located in the *lamB* site and is transcribed from is own SSB promoter**.** It is therefore assumed that there is an equal distribution of SSB wt and SSB-CFP. Fluorescence widefield microscopy of the SF149 shows the presence of distinct SeqA and SSB foci in the cells. These foci were found to be close and sometimes partly colocalized (Solveig Fossum-Raunehaug, unpublished data). Here SF149 was used to verify the average distance between SeqA and SSB. To determine if the MutH binding affects the SeqA binding, two different strains were constructed from the SF149 cells; *mutH* knock out and mutant *dnaQ.* We wanted to study whether the deletion of the *mutH* gene or if the *mutD5* mutation in the *dnaQ* gene (ε subunit in the polymerase) would affect the distance between SeqA and SSB.

3.3.1. Deletion of the *mutH* Gene in SF149

Δ*mutH* was transferred from ME125 (Elez et al. [2012]) (which has a kanamycin resistance gene instead of the *mutH* gene), into SF149 cells using P1 transduction (see section B.13). The strain was named KB09. Deletion of the *mutH* gene was checked with PCR using the primers Left/Right_deltaMutH (sequence is found table A.3) (see B.3 for the PCR reaction). These primers are complementary to the area 30 bp upstream and 30 bp downstream of *the mutH* gene. The strains ME125 and wild type AB1157 was included as controls.

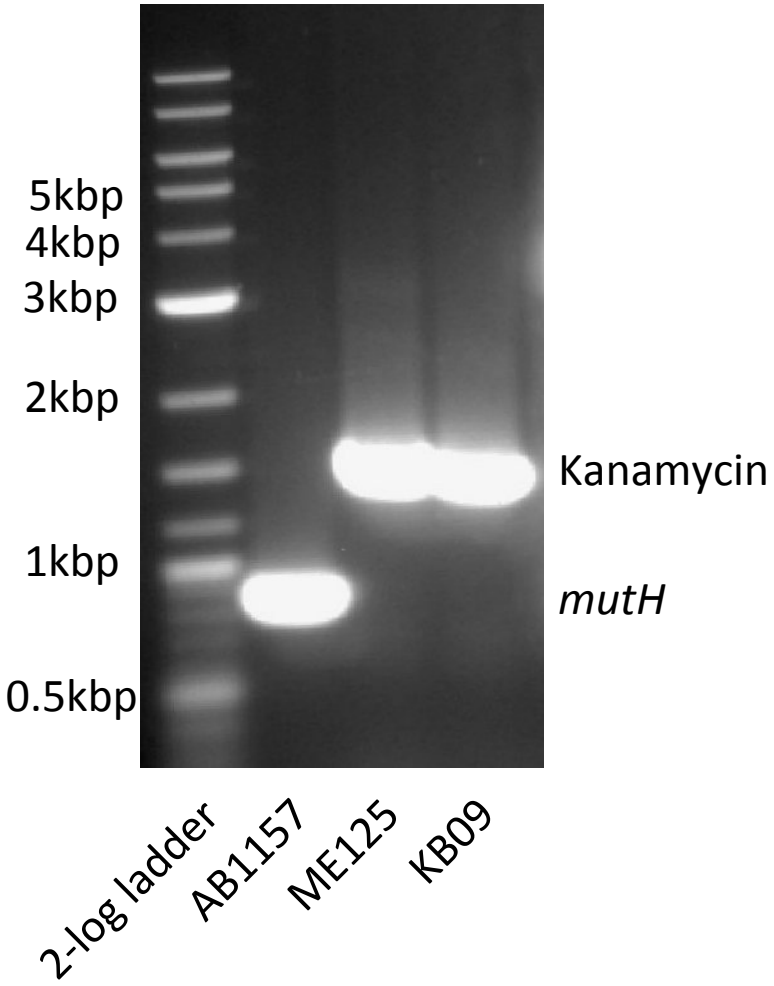


Figure 38: *The electrophoresis gel showing the DNA fragments from PCR with primers 30 bp upstream of* mutH *to 30 downstream of the gene.*

The kanamycin resistance gene has about 1500 bp, while the *mutH* gene has approximately 750 bp. The AB1157 has a DNA fragment equivalent to the *mutH* gene. Both KB09 and ME125 strains have a fragment of about 1500 bp, corresponding to the kanamycin resistance gene. KB09 and ME125 therefore lack the *mutH* gene.

3.3.2. Mutation of the Epsilon Subunit in the SF149 Strain

The strain FR680 (Elez et al. [2010]) has the *mutD*5 mutation in the *dnaQ* gene downstream of a tetracycline resistance gene. The strain was used as a donor strain, to P1 transduce the mutation over to SF149. Selection was performed on a tetracycline agar, and the strain was named KB12.

To be certain that KB12 had the mutation in the *dnaQ* gene, the gene was sequenced. The gene was copied by PCR with the primers Left/Right_MutD5 (see A.3) which are complementary to the area 30 bp upstream and 30bp downstream of the *dnaQ* gene.

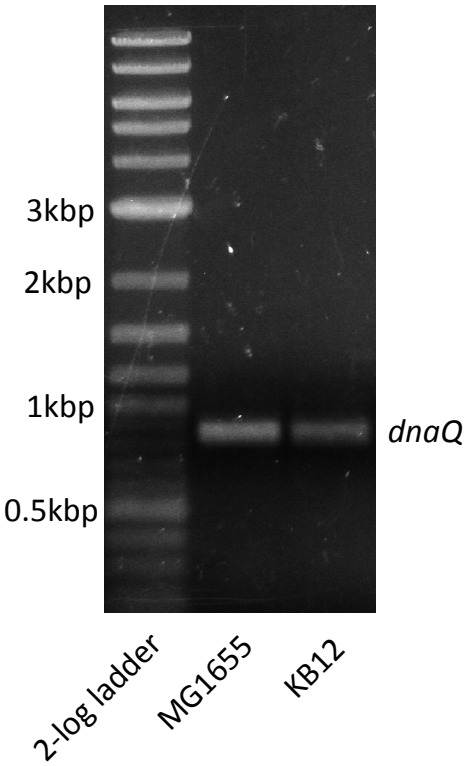


Figure 39: *The electrophoresis gel showing the DNA fragment from the PCR with primers 30 bp upstream of* dnaQ *to 30 downstream of the gene.*

The PCR fragment was sent to sequencing at LGC. The sequences obtained from LGC sequencing were compared using BLAST, the result can be found in C.1. The

BLAST analysis revealed four point mutations, where three of them were silent, giving no alteration in the protein sequence. The last point mutation was located in the amino acid number 15 on the N-terminal end of the sequence. The point mutation was a missense, non-conservative mutation giving an alteration in the protein sequence. The alteration provided a change from the hydrophilic amino acid threonine to the hydrophobic isoleucine, which is the correct point mutation in the *mutD*5 mutation (Taft-Benz and Schaaper [1998]).

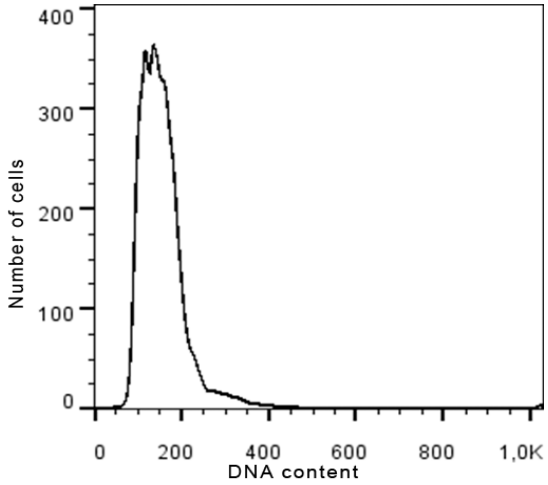
3.3.3. Measuring Doubling Time and Analysis by Flow Cytometry of SF149, KB09 and KB12

Before we start measuring distances between foci in the different strains by microscopy, investigation of the cell growth and doubling time, DNA content and the cell cycle must be conducted. It is important to have knowledge about the cell cycle for each strain when analyzing foci, because the number of foci per cell depends on where the cycle the cell is situated and in which generation initiation occurs. Flow samples were therefore prepared for the three strains SF149, KB09 and KB12. The samples were grown exponentially in ABB_1 Glu Caa medium at 28℃ to early exponential phase (OD reached 0.15) at which time the samples were split into three. Two were used to prepare flow samples and one for microscopy samples. The OD was measured periodically and the doubling time was found for the three strains. The doubling time for SF149 was measured to 62 min, the doubling time of the two other strains KB09 and KB12 was decreased to 58 min and increased 78 min, respectively. The preparation of flow samples and measuring of doubling time was performed three times for the strains KB09 and KB12 with no change in the doubling times (see section C.2 in supplementary).

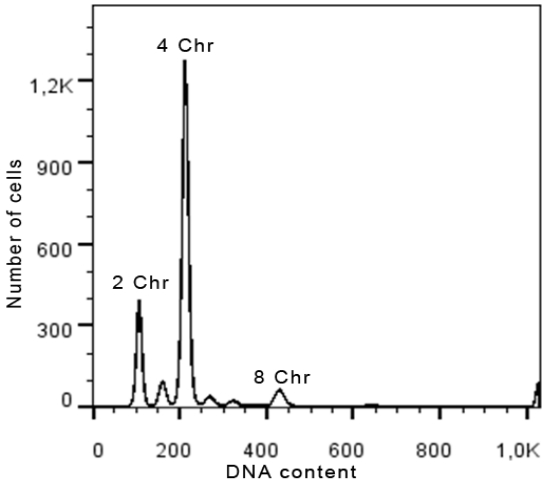
Table 5: *The doubling time for the three strains SF149, KB09 and KB12 found by measuring OD in exponentially grown cells.*

	Doubling time
SF149	62 min
KB09	58 min
KB12	78 min

A part of the exponentially grown cells were treated with rifampicin (150 µg/ml) and cephalexin (10 µg/ml) as described in 2.11. These samples were called run out samples, and had only fully replicated chromosomes. The samples were fixed in 77 % ethanol and stained with FITC and Hoechst 33258 before analysis was conducted by the flow cytometer. Hoechst is an equilibrium dye which binds DNA and gives the DNA distribution of the culture.

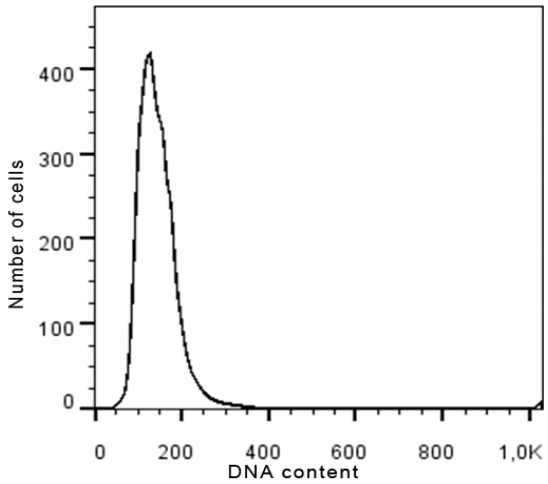


(a) *The DNA content of the exponentially grown cells*

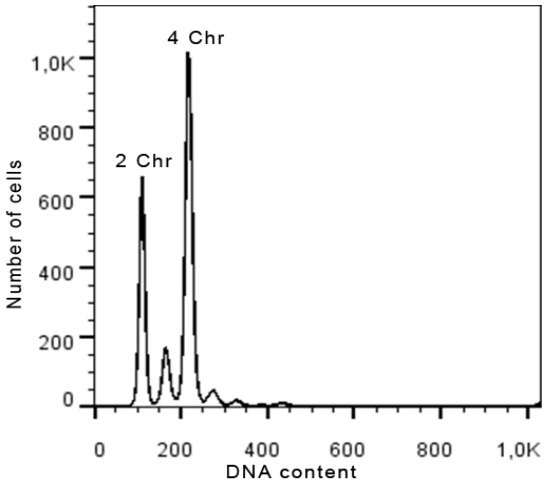


(b) *The DNA content from the run out sample*

Figure 40: *Flow samples from SF149*

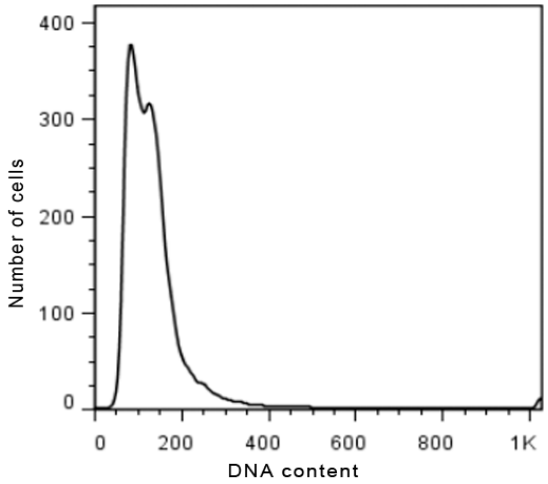


(a) *The DNA content of the exponentially grown cells*

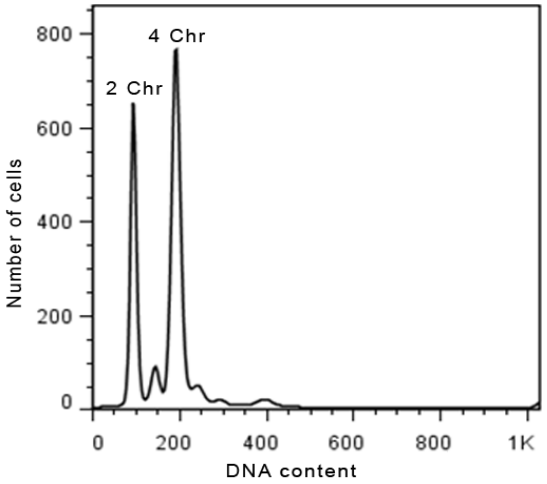


(b) *The DNA content from the run out sample*

Figure 41: *The DNA histogram of the KB09 cell*



(a) *The DNA content of the exponentially grown cells*



(b) *The DNA content from the run out sample*

Figure 42: *The DNA histogram of the KB12 cell*

The cells form the different strains generally ended up fully replicated chromosomes, with DNA contents representing the correct integral number of chromosomes (2^n).

This indicates that the sequestration was not compromised by the YFP-tagged SeqA protein. KB09 and KB12 both initiate replication at two origins in the mother cell, as SF149 (as we can see from the two and four chromosome peaks in the run out histograms). The two chromosome peak and half of the cells in the three chromosome peak correspond to the cells that had not initiated at the time when rifampicin and cephalexin were added. The four chromosome peak including the other half of the three chromosome peak represents the cells that had initiated replication when the drugs were added. The DNA histograms have not change notably form the SF149 to the KB09 strain, except for a somewhat higher two chromosome peak in the KB09 (figure 40b) . The change from the SF149 to the KB12 strain however is slightly larger. The size of the 2 chromosome peak in the KB12 run out sample is about the same height as the four chromosome peak (see figure 42b), indicating that the cells initiate replication at a later stage in the division cycle.

3.3.4. Simulating Cell Cycles for SF149, KB09 and KB12

The knowledge obtained from measuring of the doubling times and analysis of flow samples was used to simulate the cell cycles. The cell cycle simulating program introduced in the paper by Caroline Stokke (Stokke et al. [2012]) was used in the simulating process (explained in section 2.12). The figures 43, 44 and 45 show the DNA histograms obtained from the cell cycle simulating program and the simulated cell cycles. The histograms show the theoretical distribution in pink compared to the experimental distribution in black. For SF149 and KB09 the theoretical distributions fit quite well. For proper cell cycle analysis all cells must go through the same cycle pattern, generation after generation. This criterion might not be completely fulfilled for KB12 cultures. The experimental distribution is slightly too wide and it was not possible to obtain a good fit.

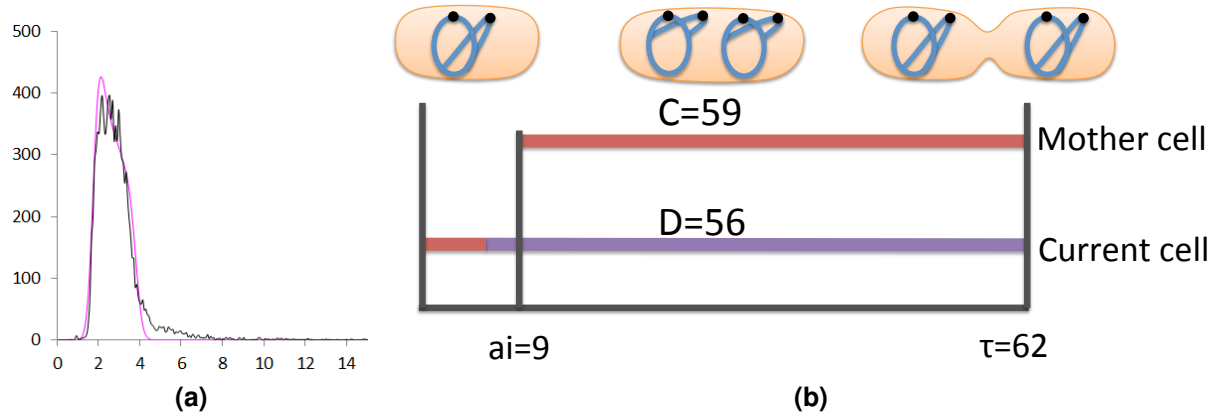


Figure 43: *The cell cycle achieved form the simulation of SF149: C-period= 59 min, D-period = 56 min, initiation time a_i = 9 min and the generation time τ=62 min*

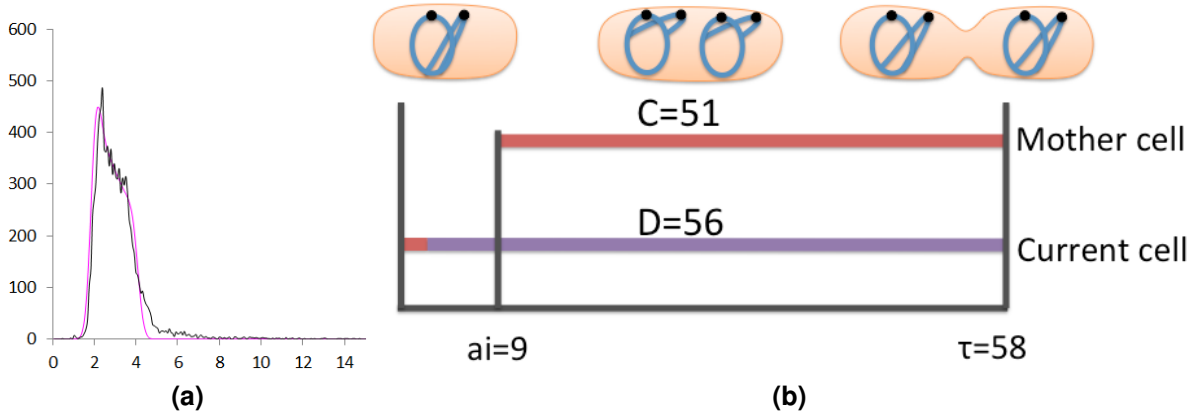


Figure 44: *The cell cycle achieved form the simulation of KB09: C-period= 51 min, D-period = 56 min, initiation time (a_i) = 9 min and the generation time τ=58 min*

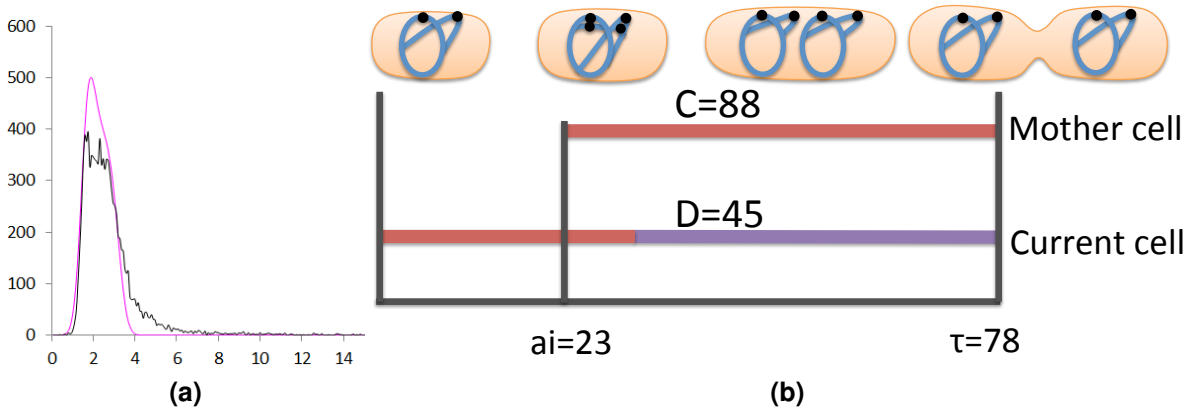


Figure 45: *The cell cycle achieved form the simulation of KB12: C-period= 88 min, D-period = 45 min, initiation time (a_i) = 23 min and the generation time τ=78 min*

The cell cycles for KB09 and KB12 were simulated for the three flow samples made (standard deviation for τ, a_i and C refer to section C.2). The simulated cell cycles show differences between the strains. It was demonstrated in section 3.3.3 that the doubling time (see table 5) slightly decreased in the KB09 strain, and increased for the KB12 strain. The same tendency can be seen in the C-periods in the simulated cell cycle for both the two strains. It appears as if the DNA replication is slightly affected by the mismatch repair system, because when SF149 has a C-period of about the same length as the generation time, KB09 had a C-period shorter than the generation time. This can indicate that DNA replication occurs faster in the absence of mismatch repair. The opposite result was found for KB12, where the C-period was increased. When the cell must cope with a high a amount of mismatch repair, the speed of the DNA replication forks seems to be affected.

3.3.5. Preparation of Microscopy and Image Analysis of SF149, KB09 and KB12

The exponentially grown cells (at OD=0.15) were prepared for microscopy. The samples were centrifuged and the pellet was resuspended in 1 x PBS. The cells were immobilized

by an 1 % agarose pad (B.16). The images were made in the microscope Leica DM6000, and the sample was illuminated by the two filters complementary to YFP and CFP in 700ms each (see table 2). Image analysis was carried out and the measurements were performed from the brightest pixel in middle of the SeqA focus, to the brightest pixel in the middle of the SSB focus. The chosen colors were red for SeqA and green for SSB and the light thresholds were adjusted to optimize each focus.

The strain SF149 made by Solveig Fossum-Raunehaug was the base for both the KB09 and the KB12. The measurement between the two foci (SeqA and SSB) was conducted in all three strains and the distance was measured in nm. The distances were measured between 150 foci, at three different occasions for KB09 and KB12. The measurements performed on SF149 on the other hand were only conducted to verify earlier results and the measurements were therefore carried out between 300 foci in one sample.

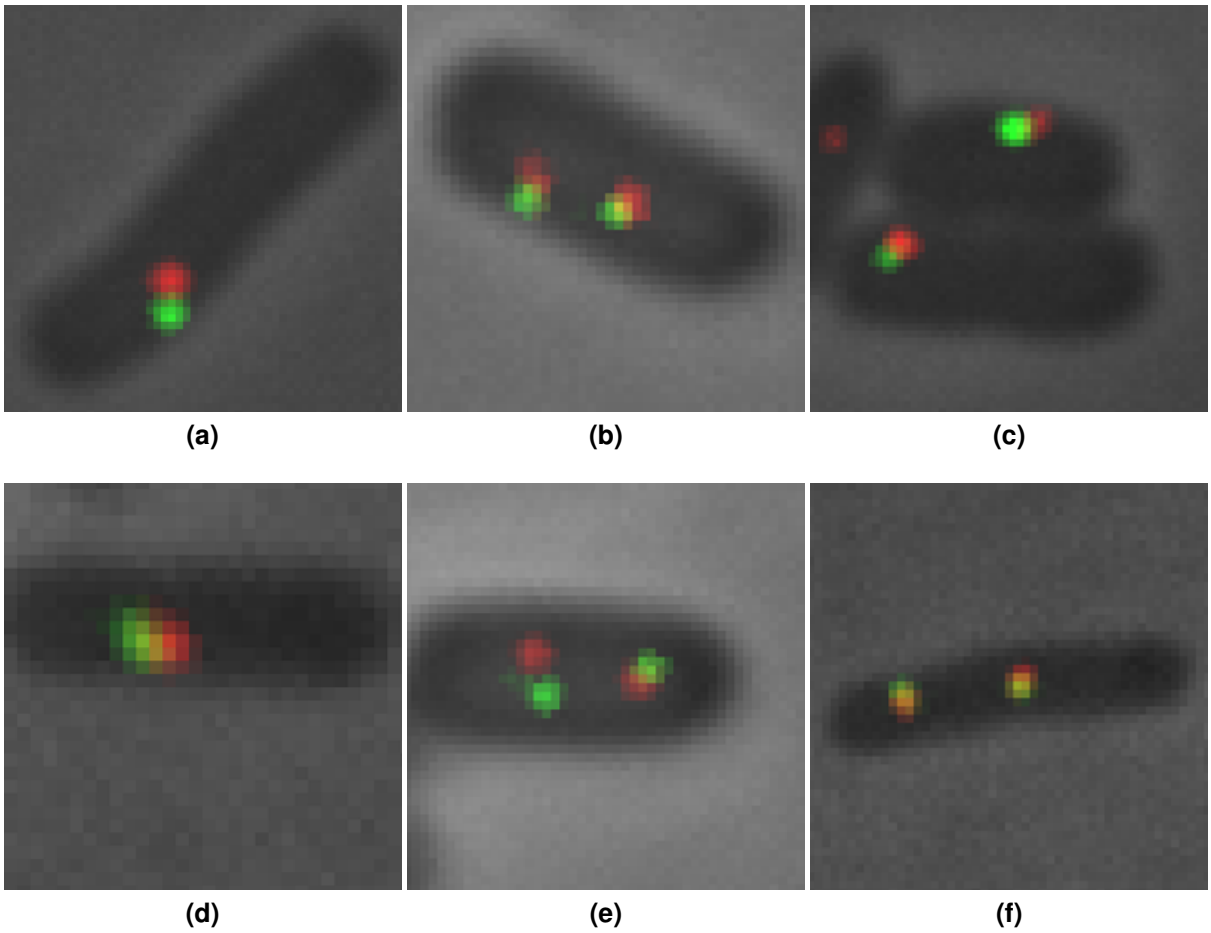


Figure 46: *Microscopy images of the SF149 strain with SeqA (red) and SSB (green).*

After measuring the average distance between 300 foci in SF149, the mean was calculated to 197 nm with a standard deviation of 103 nm. The foci from the two target protein complexes were very close to each other, but only occasionally had the cells colocalized foci (yellow foci). Most of the cells had partly overlapping foci and some of the cells had non overlapping foci.

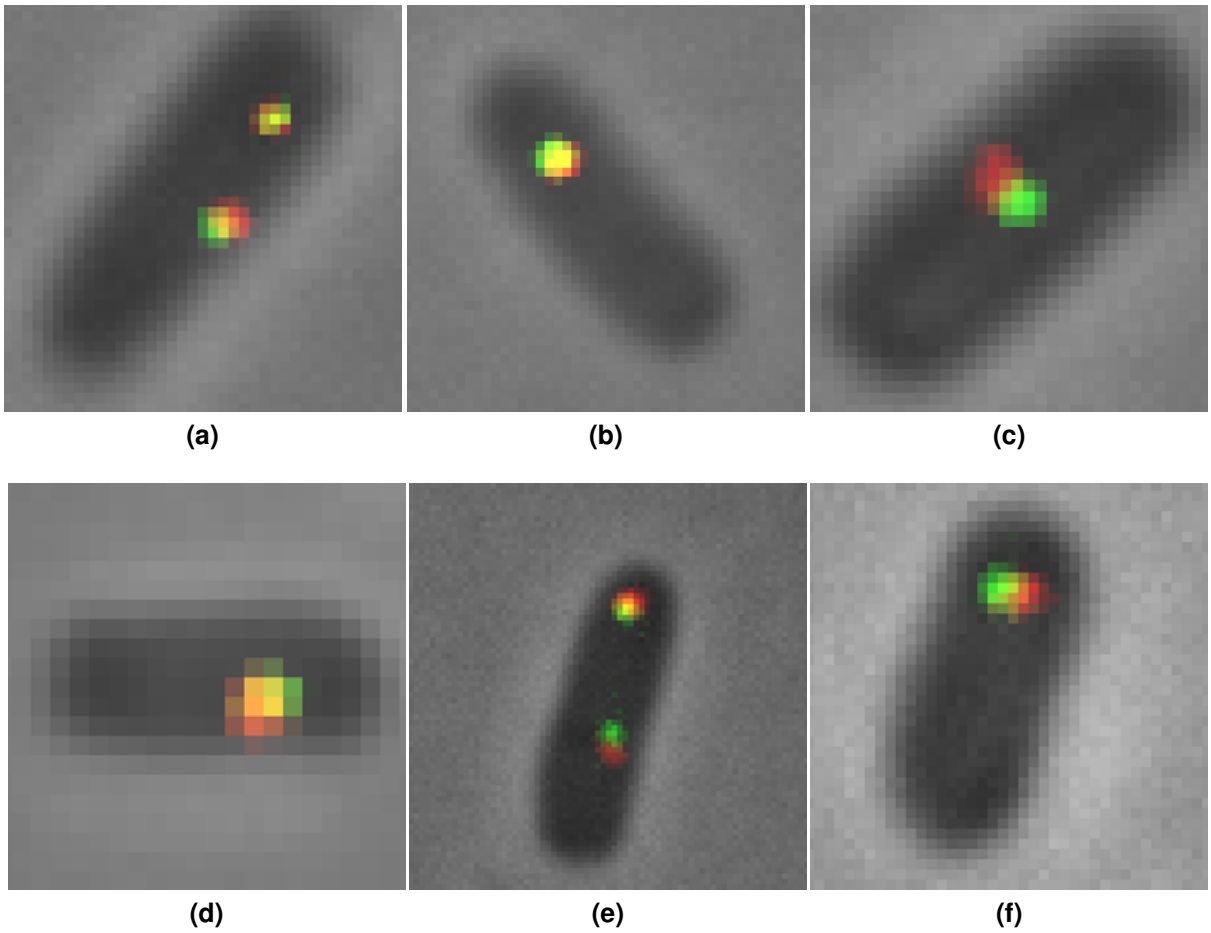


Figure 47: *Microscopy images of the KB09 strain with SeqA (red) and SSB (green).*

The average distance in the strain KB09 was calculated to 150 nm with a standard deviation of 68 nm after the distances between 450 foci were measured. The foci from the two target protein complexes were very close to each other, closer than what we observed for SF149. The cells had an equal distribution between colocalized and partly overlapping foci, and only very few of the cells had non overlapping foci.

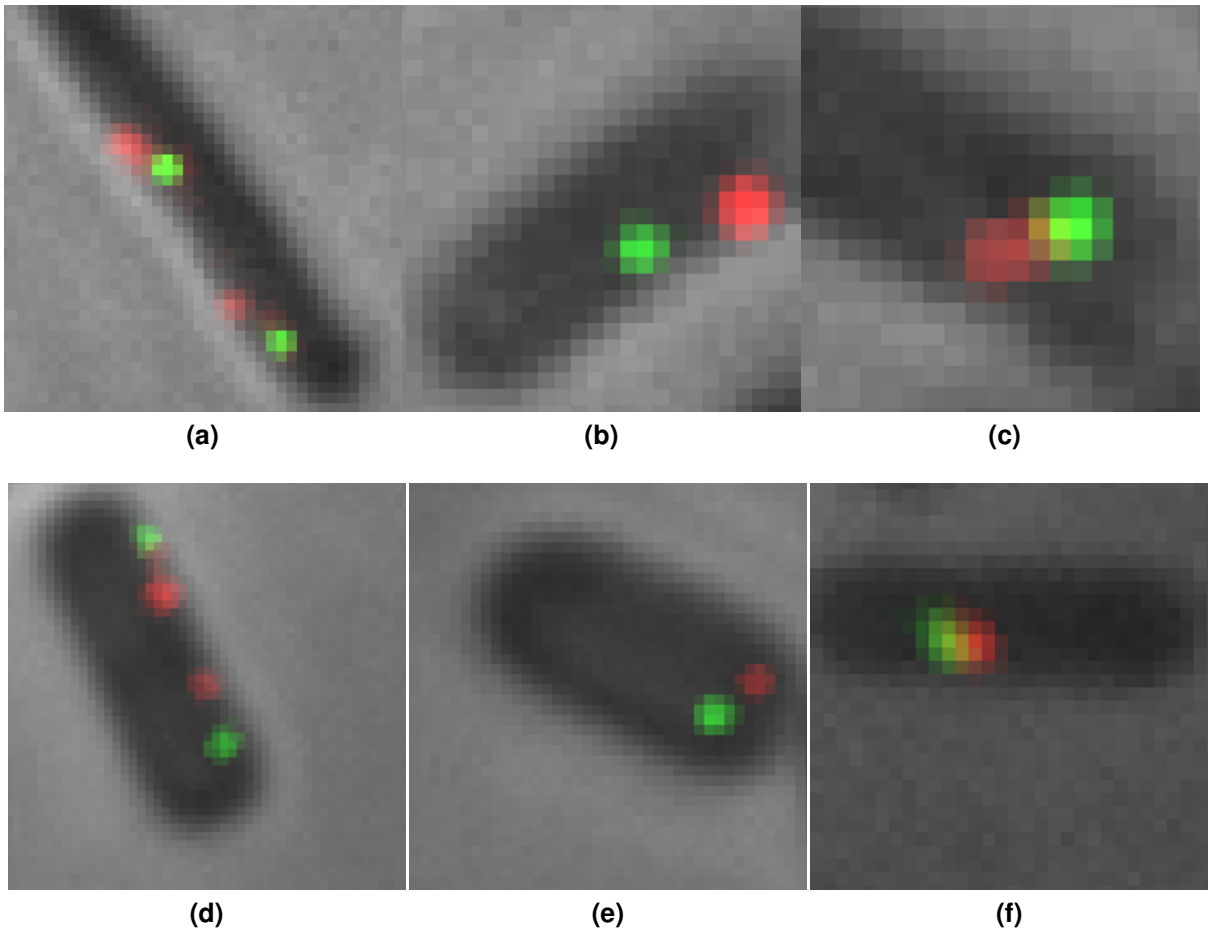


Figure 48: *Microscopy images of the KB12 strain with SeqA (red) and SSB (green).*

The mean distance was also measured for the KB12 strain. The average was calculated to 239 nm after 450 foci were counted. The distances between the two protein complexes varied more than what was observed in the two other strains and standard deviation was calculated to 140 nm. Some of the cells had colocalized foci, but most of the cells had either partly overlapping or non overlapping foci.

Table 6: *A table showing the mean, variance and the number of samples counted of SF149, KB09 and KB12.*

Strain	Mean	St. dev	Number of foci counted
SF149	197	103	300
KB09	150	68	450
KB12	239	140	450

The average distance between the two foci appeared to have been changed from the strain SF149 to KB09 and from SF149 to KB12, showing a decreased or increased average of about 50 nm. However we also observed relatively a high standard deviation in the three strains. To be certain that the three strains belong to different populations, statistical hypothesis tests were conducted, see section C.3 and C.4. It was here demonstrated with 1% level of significance that the means of the three experiments were from different populations.

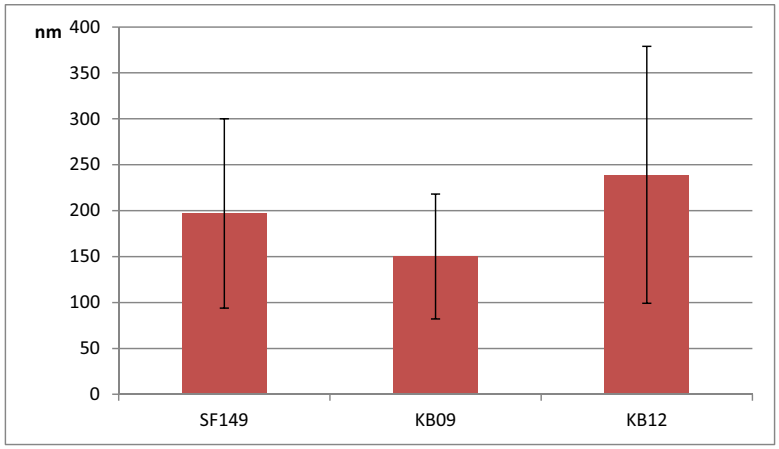


Figure 49: *A bar chart showing the average distance and standard deviation in nm between SeqA and SSB in the three strains SF149, KB09 and KB12*

3.3.6. Measuring the Mutation Ratio

The mutation ratio was measured in the strains AB1157, MG1655, KS9921, SF149, KB09 and KB12, mainly to verify the lack of mismatch repair in KB09 and the lack

of proofreading in KB12. The strain KS9921 with Δ SeqA was included to find if the presence of SeqA is important in mismatch repair. AB1157 and MG1655 wild type strains were included as controls. The mutation ratio was calculated by counting the number of colonies grown on a rifampicin agar (100 µg/ml) (see B.17). Dilution series were made from the overnight cultures. The undiluted overnight culture and the 10^{-1} dilution were plated on the rifampicin agar. The 10^{-6} dilution was plated on a LB agar plate to calculate the number of cells plated on the rifampicin agars. The mutation ratio was calculated as explained in 2.13. Measurements of mutation ratio were performed in three different experiments. An overview of the number of cells on each of the plates and the mutation ratio of all the three experiments are presented in table 23 in section C.5. The table 7 shows the average value obtained in the three experiments.

Table 7: *The mutation ratio of the strains AB1157, MG1655, KS9921, SF149, KB09 and KB12*

Strain	Mutation ratio
AB1157	$9.75 \cdot 10^{-8}$
MG1655	$1.49 \cdot 10^{-8}$
KS9921	$2.20 \cdot 10^{-8}$
SF149	$1.23 \cdot 10^{-8}$
KB09	$5.30 \cdot 10^{-6}$
KB12	$2.25 \cdot 10^{-5}$

The three strains AB1157, MG1655 and SF149 had as expected a low mutation ratio. The same applies for KS9921, indicating that SeqA is not important in the mismatch repair process. KB09 and KB12 had a much higher mutation ratio than the wild type AB1157 strain and SF149. KB09 had $5 \cdot 10^2$ times a higher mutation ratio than SF149, while KB12 had $2 \cdot 10^3$ times a higher mutation ratio than SF149. This means that KB12 (the strain lacking the DNA polymerase proofreading center) had over 4 times higher mutation ratio than KB09 (the strain lacking mismatch repair).

4. Discussion

4.1. Loss of the YgdQ Protein

In the attempt to find out where MutH is localized on hemimethylated DNA compared to SeqA, an experiment was performed to try to tag *mutH* on the chromosome with the sequence of the fluorescent protein mKate2. The *E. coli* chromosome has a high density of genes, so the gens lie relatively close to each other. During chromosomal insertion with Red/ET cloning, 30 of the chromosomal base pairs were removed downstream of the *mutH* gene to make it easier to insert mKate2 followed by the chloramphenicol resistance marker. These happened to remove the promoter of the *ygdQ* gene, and the gene was probably no longer expressed.

The cells that were expected to lack the YgdQ protein (KB04) were large and round (see figure 33), and had a doubling time of 83 min in LB medium. The DNA histogram from the flow cytometry analysis demonstrated a DNA content equivalent to around 10 *E. coli* chromosomes inside every cell.

Little is known about the YgdQ protein, except that YgdQ is a seven α helical transmembrane protein in the inner membrane of the *E. coli* cell wall (Rapp et al. [2004]). Membrane damage studies performed in *E. coli* in the presence of TiO2 and UV light demonstrated that the cells become round primarily when the damage affects the peptidoglycan layer (Liu et al. [2010]). This corresponds with other studies performed on the peptidoglycan layer and cell shape (Furchtgott et al. [2011]), making us believe that the YgdQ protein possibly could affect the peptidoglycan layer. Other protein deletions that have shown similar phenotype are the penicillin binding protein (Pbp2) and RodA, where both deletions give spherically shaped cells, and both have important roles in peptidoglycan elongation (Spratt [1975], Spratt [1980], Ishino et al. [1986]Iwaya et al. [1978]).

A paper written by Moon and Gottesman explains that *E. coli* cells in the presence

of low concentrations of Mg^{2+} activate the membrane kinase PhoQ which in turn phosphorylates and activates the transcription factor PhoP. The PhoP is a transcription factor for a number of genes including the sRNA MgrR. The sRNA MgrR negatively regulate two proteins, one of them YgdQ. The YgdQ membrane protein is downregulated 9-folds when the concentration of Mg^{2+} is low. (Moon and Gottesman [2009]). A medium low on Mg^{2+} is LB (Nikaido [2009]). Wild type cells grown in LB tend to be more round than in other media. Could this perhaps be explained by the downregulation of YgdQ?

The spherical phenotype of the KB04 cell indicates that the YgdQ membrane protein is important in maintaining the characteristic rod shape structure of the *E. coli* membrane, possibly by affecting the peptidoglycan layer. The cells struggle to divide when they lack this protein, which also might be explained by a nonfunctional peptidoglycan layer. Since the protein seems to be affected by low Mg^{2+} concentration, leading to the idea that the protein YgdQ is important in magnesium regulated control of the cell shape. These are possible explanations of the observations and results obtained in this thesis and in previous studies. To be able to say with certainty that it is the lack of YgdQ that provides the strain with this phenotype we must once again introduce the protein back into the strain to observe if the cells regenerate.

4.2. Does Binding of MutH Affect Replication Fork Dynamics?

The tagging of MutH turned out unsuccessfully in the time periode available, and we had to consider a new strategy. We constructed two new strains based on SF149, one with a *mutH* deletion (KB09) and one with a mutation in the *dnaQ* gene. The KB09 cells have presumably no mismatch repair. Even though the strain still has the two other MMR proteins MutS and MutL, these cells lack the endonuclease protein MutH which discriminates the daughter strand from the parental strand. The KB12 cells have a higher mismatch frequency due to the *mutD5* mutation which is located

in the *dnaQ* gene, which codes for the ε subunit in the polymerase (illustrated in figure 4) (Scheuermann et al. [1983], Scheuermann and Echols [1984]). The strain therefore has a nonfunctional ε subunit and lacks the proofreading property, causing a very strong mutator phenotype (Cox and Horner [1983], Degnen and Cox [1974], Scheuermann et al. [1983],Taft-Benz and Schaaper [1998]). The idea behind the new approach was to examine if MutH affects the binding of SeqA to newly synthesized hemimethylated DNA downstream of the replisome, by investigating if the distances between SeqA-YFP and SSB-CFP are different in the new strains (as explained in 3.3).

4.2.1. The Validity of the Foci Results

The SF149 and the strains prepared from SF149 (KB09 and KB12) have both SSB-CFP and SeqA-YFP markers. The SSB proteins bind and protect single stranded DNA mainly at the lagging strand by the replication fork. The SSB-CFP was used here to represent where in the cell the replisome is localized. The SSB is tagged with CFP on its C-terminal end. Unfortunately, both the N- and C-terminal end of the SSB are essential for its function, thus the CFP marked SSB does not function properly (Reyes-Lamothe et al. [2012], Reyes-Lamothe [2012], Antony et al. [2013], Naue et al. [2013]). The cells therefore require the wild type SSB additionally. This means that the CFP foci observed in the microscopy images were solely from the population of non-functional SSB proteins. But the natural tetrameric structure of the SSB (Greipel et al. [1987], Overman et al. [1988], Weiner et al. [1975], Williams et al. [1983], Williams et al. [1984]) are believed to keep the functional wt SSB proteins and the unfunctional SSB-CFP at the same and correct localization. This is confirmed in the paper by Rodrigo Reyes-Lamothe where he explains that over 90% of the SSB foci colocalized with the replisome markers DnaQ, HolC, DnaE and HolD (Reyes-Lamothe et al. [2008]).

We do not encounter the same issue with the SeqA-YFP, since all the SeqA proteins

are tagged. However, earlier attempts on tagging SeqA with the fluorescent proteins YPet and mKate2 have shown asynchronous replication. SeqA-YFP does not have this problem, as we also saw in the flow run out sample (see section 3.3.3). The reason for this is not known, but it seems as the YFP tag does not interfere with SeqA's function in the same way as for the two other fluorescent proteins.

It is important to be aware of the problems connected to *in vivo* tagging of proteins, in particular the SeqA-YFP and the SSB-CFP, but we believe that tagged proteins represent the protein localization in the cell accurately. Nevertheless, we believe they provide qualitatively valid results concerning the change in distance between the two proteins in the three strains.

4.2.2. The Raised Mutation Ratio in KB09 and KB12

The mutation ratio was measured for the strains AB1157, MG1655, KS9921, SF149, KB09 and KB12 (see table 7), mainly to verify the raised mutation frequency in the two newly constructed strains. The low mutation ratio in the KS9921 strain, which lacks the SeqA protein, indicated that SeqA does not assist MutH or any of the other MMR proteins in the repair process. The obtained results in the mutation ratio study also revealed a higher mutation ratio for KB09 and KB12, as expected. This is because the ε proofreading center increases the replication fidelity roughly from 10^5 to 10^7, while the MMR system further increases the fidelity a thousand fold to 10^{10} (explained in section 1.4 and 1.7). Intuitively, the KB09 should have a higher mutation ratio than KB12, since MMR decreases the mutation frequency almost 1000 folds, while the proofreading center only decreases the mutation frequency 100 folds. However, the result from the experiment turned out to be that KB12 had a four times higher mutation ratio than KB09. The raised mutation ratio in the KB12 most likely comes from a mismatch saturation in the mismatch repair system. This is backed up by previous studies on the subject which show that a raised mismatch insertion by a mutated ε subunit will

temporarily overwhelm the mismatch repair system giving cells with an even higher mutation frequency than the expected 100 fold increase (Krishnaswamy et al. [1993], Radman [1981], Schaaper [1988], Schaaper and Radman [1989], Damagnez V and M [1989]) The MMR saturation occurs only in rapidly growing cells, and the mismatch repair process will return to full efficiency if the cells are supplied with the MMR proteins MutL and MutH or by transfer to minimal media (Schaaper and Radman [1989]).

4.2.3. Measuring Distances Under the Diffraction Limit

To see if the distances between SeqA-YFP and SSB-CFP in the newly constructed cells had changed relative to those of SF149, an analysis using fluorescence microscopy was conducted (see section 3.3.5). Some of the measurements in this section refer to distances less than the diffraction limit described in section 2.10. However, since the two foci are emitting light with two different wavelengths and the pictures were taken separately before they were laid on top of each other, the problem of overlapping point spread functions was avoided. Nonetheless, the light from the point sources is still spread out according to the point spread function. It is common to approximate the PSF by a Gaussian (bell) curve. Using this knowledge about the PSF, the positions of the pixel sources are approximated by measuring from the point with the highest intensity (explained in section 2.10.2). This allows for measurement of distances under the diffraction limit, but gives some uncertainty to the measurements. However, by doing a large number of measurements this uncertainty is decreased.

4.2.4. Verifying the Distance Between SeqA and SSB in SF149

Analysis of the distances between foci from microscope images of SF149 was first done by Solveig Fossum-Raunehaug. She found that SeqA and SSB foci normally are only partly overlapping and thereby discarded the theory that SeqA binds right behind the replisome. In her studies she observed an average distance between SeqA and

SSB of 234 nm. The distances were measured by another principle than what was used in this thesis, though. Measurements of distances between the two proteins in SF149 have also been performed in this thesis to verify the findings (see section 3.3.5), and to use the result in comparison studies with KB09 and KB12. After measuring the distances between 300 foci in SF149, an average of 197 nm (st. dev 103 nm) was obtained. The studies on overlapping foci in this study showed that some of the foci were colocalized, but most of them were partly or non-overlapping. The difference between the average distances obtained here and the results obtained by Solveig Fossum-Raunehaug in her experiments is relatively large. Due to the differences between the analyzing methods we will discard her results in the comparison studies, using only the results obtained in this study.

4.2.5. The Decreased Distance Between the Foci Pairs in KB09

The same analysis was conducted for KB09 and KB12 (see section 3.3.5). The KB09 strain was found to have significantly shorter distances between SeqA and SSB than what was measured for the SF149 strain (see section C.3 for a statistical hypothesis test). The average was measured to 150 nm (st. dev 68 nm). The microscopy analysis also shows a distinctly higher number of colocolized foci. We see from these results that the cells lacking MutH have a shorter distance between SeqA and SSB than cells with MutH.

Even though the MMR system has been well studied, there are no papers on the localization of the MMR proteins relative to the replisome or SeqA in *E. coli* known to the author. Studies of the gram positive bacteria *Bacillus subtilis* show that almost 50 % of the MutS foci colocalize with the replisome (Smith et al. [2001]). Studies performed on other model organisms (including Bacillus) have demonstrated interactions between MutS and the β-clamp (Kleczkowska et al. [2001], Flores-Rozas et al. [2000], Lopez de Saro et al. [2006], Lopez de Saro and O'Donnell [2001], Noirot-Gros et al.

[2002], Simmons et al. [2008], Shell et al. [2007]), and a new study has shown that the β-clamp works as a platform for MMR and is involved in over 90 % of the mismatch repair (Lenhart et al. [2013]). Even if these results are connected to cells with methyl-independent mismatch repair (cells lacking both SeqA and MutH), the mismatch repair system is highly conserved between species and the results may be used for comparison with results obtained in *E. coli*. The mentioned studies indicate that MMR occurs right downstream of the replisome and are in accordance with the results obtained in this thesis.

There are two potential binding patterns for MutH behind the replisome: MutH binds either indepentently or dependently of a mismatch. If MutH binds independently, the MutH binding occurs in a small window right behind the replisome, disallowing the binding of SeqA until it dissociates from the binding site. If the binding is mismatch dependent, MutH competes with SeqA for the hemimethylated sites near a mismatch. In the data obtained here, the observed shift in the distribution of distances (between SeqA and SSB) in SF149 and KB09 is not compatible with the last hypothesis. The fact that a mismatch on average only occurs every second round of replication, does not fit with the relatively large difference between the two strains. Most of the foci in KB09 were either colocalized or partly overlapping, but some of the foci pairs were not overlapping. If the first theory is true, i.e. that MutH binds mismatch independently, disallowing the SeqA binding; why are not all the foci pairs colocalized in the KB09 strain?

4.2.6. The Increased Distance Between the Foci Pairs in KB12

The average distance between the SeqA and SSB foci in the KB12 strain was measured to 239 nm (st.dev 140 nm), a higher average than SF149. A statistical hypothesis test was performed (see section C.4), confirming a difference between the strains with a 1 % significant level. The study of foci colocalization showed a tendency towards

longer distances between the foci compared to SF149 (see table 3.3.5). Thus we have observed that the average distance between the two proteins decrease in cells with a nonfunctional MMR and increase in cells with a functional MMR. The results here in the context of earlier studies indicate that MMR takes place almost right after replication occurs, affecting the SeqA binding slightly. It is hard to discuss the results achieved with the KB12 strain, since we do not know what causes the increased distance. The MMR saturation explained in section 4.2.2 makes it difficult to tell if the increased distance is caused by a mismatch repair process or by accumulation of MMR proteins at an unrepaired mismatch site which prevents a rapid SeqA binding. The KB12 cells were also found to have an increased C-period. This could either be due to each replication fork traveling more slowly or that the replication forks pause more frequently. The low rate of replication may either be due to the mutation in the ε subunit or the the MMR proteins' increased activity.

Earlier studies performed on MMR have indicated that mismatch repair happens so fast that MMR foci do not occur (Elez et al. [2010]). Only those mismatches that remain unrepaired have enough time to form distinct MMR foci, which means that the increased distances between the proteins not necessarily come from more or longer MutH binding associated with MMR, but could likely be caused by the MutL polymerization (see figure 9) suggested in Elez et al. [2012] (and backed up by the articles of Grilley et al. [1989], Damagnez V and M [1989], Feng et al. [1996], Negishi et al. [2002], Ivan Matic and Radman [2003], Selmane et al. [2003]). This theory could also explain why not all the foci in the KB09 strain are colocalized.

4.3. Further Studies

We believe from the results obtained here and earlier results on the mismatch repair proteins that MMR takes place right downstream of the replication fork, however we cannot say if it is the lack of MutH or the lack of a functional MMR that allows SeqA

to bind closer to the replisome. Some indications has emerged in the work with this thesis, but further investigation is needed to find the answer to the question posed in this thesis, namely; how does the MutH binding affect the binding of SeqA.

We want to try tag MutH on the N-terminal end with respect to the promoter using Red/ET cloning. We will then see if, and possibly where, MutH forms foci relative to SeqA or the replisome. We could then also determine if MutH binds mismatch dependently or independently by counting the number of foci in the population. We suspect that the tagging of MutH may be hard since no results of fluorescently tagged MutH have yet been published, despite several articles published on fluorescently tagged MutS and MutL. An alternative may be to introduce a plasmid with a fluorescently marked MutH.

If the window for mismatch repair and MutH binding is too short for foci formation as suggested by Elez et al. [2010] other methods shall be used for investigation. The mutant MutHE56A has a mutation leading to a change from the amino acid glutamine acid to alanine at position 56. This mutation does not affect the binding ability of MutH to hemimethylated DNA, but the protein loses its endonuclease property. By introducing this protein into the strain KB09 we can observe if the distance between SeqA and SSB increases.

The binding pattern of SeqA to hemimethylated GATCs was investigated with ChIP-on-chip technique in the paper by Waldminghaus et al. [2012]. This method could also be used to find the binding pattern of MutH and compare with the SeqA results.

We would also like to delete *mutL* in both the KB09 and the KB12 strain to investigate if this changes the average distance between SeqA and SSB, to confirm or discard the theory of mutL polymerization.

References

Bacterial stain jm109, 2013. URL `http://no.promega.com/products/cloning-and-dna-markers/cloning-tools-and-competent-cells/bacterial-strains-and-competent-cells/bacterial-strain-jm109/`.

Mortimer Abramowitz. Optical microscopy primer anatomy of the microscope, 2004. URL `http://micro.magnet.fsu.edu/primer/anatomy/numaperture.html`.

S. Acharya, P. L. Foster, P. Brooks, and R. Fishel. The coordinated functions of the E. coli MutS and MutL proteins in mismatch repair. *Mol. Cell*, 12(1):233–246, Jul 2003.

S. Adachi, M. Kohiyama, T. Onogi, and S. Hiraga. Localization of replication forks in wild-type and mukB mutant cells of Escherichia coli. *Mol. Genet. Genomics*, 274(3): 264–271, Oct 2005.

D. J. Allen, A. Makhov, M. Grilley, J. Taylor, R. Thresher, P. Modrich, and J. D. Griffith. MutS mediates heteroduplex loop formation by a translocation mechanism. *EMBO J.*, 16(14):4467–4476, Jul 1997.

E. Antony, E. Weiland, Q. Yuan, C. M. Manhart, B. Nguyen, A. G. Kozlov, C. S. McHenry, and T. M. Lohman. Multiple C-Terminal Tails within a Single E. coli SSB Homotetramer Coordinate DNA Replication and Repair. *J. Mol. Biol.*, 425(23): 4802–4819, Nov 2013.

T. Asai, M. Takanami, and M. Imai. The AT richness and gid transcription determine the left border of the replication origin of the E. coli chromosome. *EMBO J.*, 9(12): 4065–4072, Dec 1990.

K. G. Au, K. Welsh, and P. Modrich. Initiation of methyl-directed mismatch repair. *J. Biol. Chem.*, 267(17):12142–12148, Jun 1992.

T. Bach, M. A. Krekling, and K. Skarstad. Excess SeqA prolongs sequestration of oriC and delays nucleoid segregation and cell division. *EMBO J.*, 22(2):315–323, Jan 2003.

C. Ban and W. Yang. Crystal structure and ATPase activity of MutL: implications for DNA repair and mutagenesis. *Cell*, 95(4):541–552, Nov 1998.

F. R. Blattner, G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao. The complete genome sequence of Escherichia coli K-12. *Science*, 277(5331):1453–1462, Sep 1997.

E. Boye and A. Løbner-Olesen. Bacterial growth control studied by flow cytometry. *Res. Microbiol.*, 142(2-3):131–135, 1991.

E. Boye, T. Stokke, N. Kleckner, and K. Skarstad. Coordinating DNA replication initiation with cell growth: differential roles for DnaA and SeqA proteins. *Proc. Natl. Acad. Sci. U.S.A.*, 93(22):12206–12211, Oct 1996.

E. Boye, A. Løbner-Olesen, and K. Skarstad. Limiting DNA replication to once and only once. *EMBO Rep.*, 1(6):479–483, Dec 2000.

D. Bramhill and A. Kornberg. A model for initiation at origins of DNA replication. *Cell*, 54(7):915–918, Sep 1988.

J. L. Campbell and N. Kleckner. E. coli oriC and the dnaA gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork. *Cell*, 62(5):967–979, Sep 1990.

K. M. Carr and J. M. Kaguni. Stoichiometry of DnaA and DnaB protein in initiation

at the Escherichia coli chromosomal origin. *J. Biol. Chem.*, 276(48):44919–44925, Nov 2001.

M. R. Cassler, J. E. Grimwade, and A. C. Leonard. Cell cycle-specific changes in nucleoprotein complexes at a chromosomal replication origin. *EMBO J.*, 14(23): 5833–5841, Dec 1995.

S. Cooper and C. E. Helmstetter. Chromosome replication and the division cycle of Escherichia coli B/r. *J. Mol. Biol.*, 31(3):519–540, Feb 1968.

Chroma Technology Corporation. Chroma spectra viewer, 2012. URL `http://www.chroma.com/spectra-viewer?fluorochromes=845`.

E. C. Cox and D. L. Horner. Structure and coding properties of a dominant Escherichia coli mutator gene, mutD. *Proc. Natl. Acad. Sci. U.S.A.*, 80(8):2295–2299, Apr 1983.

A. B. Cubitt, R. Heim, S. R. Adams, A. E. Boyd, L. A. Gross, and R. Y. Tsien. Understanding, improving and using green fluorescent proteins. *Trends Biochem. Sci.*, 20(11):448–455, Nov 1995.

Doutriaux M P Damagnez V and Radman M. Saturation of mismatch repair in the mutd5 mutator strain of escherichia coli. *Journal of Bacteriology*, 171:4494–4497, 1989.

K. A. Datsenko and B. L. Wanner. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.*, 97(12): 6640–6645, Jun 2000.

G. E. Degnen and E. C. Cox. Conditional mutator gene in Escherichia coli: isolation, mapping, and effector studies. *J. Bacteriol.*, 117(2):477–487, Feb 1974.

N. E. Dixon and A. Kornberg. Protein HU in the enzymatic replication of the chromosomal origin of Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.*, 81(2): 424–428, Jan 1984.

Duke-University. Spectra of common fluorophores, 2012. URL `http://microscopy.duke.edu/spectra.html`.

M. Elez, A. W. Murray, L. J. Bi, X. E. Zhang, I. Matic, and M. Radman. Seeing mutations in living cells. *Curr. Biol.*, 20(16):1432–1437, Aug 2010.

M. Elez, M. Radman, and I. Matic. Stoichiometry of MutS and MutL at unrepaired mismatches in vivo suggests a mechanism of repair. *Nucleic Acids Res.*, 40(9): 3929–3938, May 2012.

L. Fang, M. J. Davey, and M. O'Donnell. Replisome assembly at oriC, the replication origin of E. coli, reveals an explanation for initiation sites outside an origin. *Mol. Cell*, 4(4):541–553, Oct 1999.

G. Feng, H. C. Tsui, and M. E. Winkler. Depletion of the cellular amounts of the MutS and MutH methyl-directed mismatch repair proteins in stationary-phase Escherichia coli K-12 cells. *J. Bacteriol.*, 178(8):2388–2396, Apr 1996.

R. Fishel. Mismatch repair, molecular switches, and signal transduction. *Genes Dev.*, 12(14):2096–2101, Jul 1998.

H. Flores-Rozas, D. Clark, and R. D. Kolodner. Proliferating cell nuclear antigen and Msh2p-Msh6p interact to form an active mispair recognition complex. *Nat. Genet.*, 26(3):375–378, Nov 2000.

S. Fossum, S. Søreide, and K. Skarstad. Lack of SeqA focus formation, specific DNA binding and proper protein multimerization in the Escherichia coli sequestration mutant seqA2. *Mol. Microbiol.*, 47(3):619–632, Feb 2003.

S. Fossum, E. Crooke, and K. Skarstad. Organization of sister origins and replisomes during multifork DNA replication in Escherichia coli. *EMBO J.*, 26(21):4514–4522, Oct 2007.

E. C. Friedberg. Out of the shadows and into the light: the emergence of DNA repair. *Trends Biochem. Sci.*, 20(10):381, Oct 1995.

L. Furchtgott, N. S. Wingreen, and K. C. Huang. Mechanisms for maintaining cell shape in rod-shaped Gram-negative bacteria. *Mol. Microbiol.*, 81(2):340–353, Jul 2011.

GeneBridges. How red/et works, 2013. URL `http://www.genebridges.com/gb/red_et_principles.php`.

Christine L. Case Gerard J. Tortora. Berdell R. Funke. *Mastering Microbiology*. Benjamin Cummings, 2012.

S. Gradia, S. Acharya, and R. Fishel. The human mismatch recognition complex hMSH2-hMSH6 functions as a novel molecular switch. *Cell*, 91(7):995–1005, Dec 1997.

J. Greipel, G. Maass, and F. Mayer. Complexes of the single-stranded DNA-binding protein from Escherichia coli (Eco SSB) with poly(dT). An investigation of their structure and internal dynamics by means of electron microscopy and NMR. *Biophys. Chem.*, 26(2-3):149–161, May 1987.

M. Grilley, K. M. Welsh, S. S. Su, and P. Modrich. Isolation and characterization of the Escherichia coli mutL gene product. *J. Biol. Chem.*, 264(2):1000–1004, Jan 1989.

A. Guarne, Q. Zhao, R. Ghirlando, and W. Yang. Insights into negative modulation of E. coli replication initiation from the structure of SeqA-hemimethylated DNA complex. *Nat. Struct. Biol.*, 9(11):839–843, Nov 2002.

A. Guarne, S. Ramon-Maiques, E. M. Wolff, R. Ghirlando, X. Hu, J. H. Miller, and W. Yang. Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair. *EMBO J.*, 23(21):4134–4145, Oct 2004.

A. Guarne, T. Brendler, Q. Zhao, R. Ghirlando, S. Austin, and W. Yang. Crystal structure of a SeqA-N filament: implications for DNA replication and chromosome organization. *EMBO J.*, 24(8):1502–1511, Apr 2005.

F. G. Hansen and K. von Meyenburg. Characterization of the dnaA, gyrB and other genes in the dnaA region of the Escherichia coli chromosome on specialized transducing phages lambda tna. *Mol. Gen. Genet.*, 175(2):135–144, Sep 1979.

R. Heim, D. C. Prasher, and R. Y. Tsien. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A.*, 91(26): 12501–12504, Dec 1994.

Morten Helbæk. *Statistikk for kjemikere*. Tapir Akademisk Forlag, 2001.

S. Hiraga, C. Ichinose, H. Niki, and M. Yamazoe. Cell cycle-dependent duplication and bidirectional migration of SeqA-associated DNA-protein complexes in E. coli. *Mol. Cell*, 1(3):381–387, Feb 1998.

F. Ishino, W. Park, S. Tomioka, S. Tamaki, I. Takase, K. Kunugita, H. Matsuzawa, S. Asoh, T. Ohta, and B. G. Spratt. Peptidoglycan synthetic activities in membranes of Escherichia coli caused by overproduction of penicillin-binding protein 2 and rodA protein. *J. Biol. Chem.*, 261(15):7024–7031, May 1986.

Ana Babic Ivan Matic and Miroslav Radman. 2-aminopurine allows interspecies recombination by a reversible inactivation of the escherichia coli mismatch repair system. *Journal of Bacteriology*, 185:1459–1461, 2003.

M. Iwaya, C. W. Jones, J. Khorana, and J. L. Strominger. Mapping of the mecillinam-resistant, round morphological mutants of Escherichia coli. *J. Bacteriol.*, 133(1):196–202, Jan 1978.

Stephen P Bell Alexander Gann Michael Levine Richard Losick James D Watson, Tania A Baker. *Molecular Biology of the Gene*. Pearson Education, 2008.

J. Jiang, L. Bai, J. A. Surtees, Z. Gemici, M. D. Wang, and E. Alani. Detection of high-affinity and sliding clamp modes for MSH2-MSH6 by single-molecule unzipping force analysis. *Mol. Cell*, 20(5):771–781, Dec 2005.

David William Russell Joseph Sambrook. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, 2001.

M. S. Junop, G. Obmolova, K. Rausch, P. Hsieh, and W. Yang. Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair. *Mol. Cell*, 7(1):1–12, Jan 2001.

K. Kasho and T. Katayama. DnaA binding locus datA promotes DnaA-ATP hydrolysis to enable cell cycle-coordinated replication initiation. *Proc. Natl. Acad. Sci. U.S.A.*, 110(3):936–941, Jan 2013.

T. Katayama, T. Kubota, K. Kurokawa, E. Crooke, and K. Sekimizu. The initiator function of DnaA protein is negatively regulated by the sliding clamp of the E. coli chromosomal replicase. *Cell*, 94(1):61–71, Jul 1998.

T. Katayama, S. Ozaki, K. Keyamura, and K. Fujimitsu. Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. *Nat. Rev. Microbiol.*, 8(3):163–170, Mar 2010.

J. Kato and T. Katayama. Hda, a novel DnaA-related protein, regulates the replication cycle in Escherichia coli. *EMBO J.*, 20(15):4253–4262, Aug 2001.

R. Kitagawa, T. Ozaki, S. Moriya, and T. Ogawa. Negative control of replication initiation by a novel chromosomal locus exhibiting exceptional affinity for Escherichia coli DnaA protein. *Genes Dev.*, 12(19):3032–3043, Oct 1998.

H. E. Kleczkowska, G. Marra, T. Lettieri, and J. Jiricny. hMSH3 and hMSH6 interact with PCNA and colocalize with it to replication foci. *Genes Dev.*, 15(6):724–736, Mar 2001.

H. K. Klungsoyr and K. Skarstad. Positive supercoiling is generated in the presence of Escherichia coli SeqA protein. *Mol. Microbiol.*, 54(1):123–131, Oct 2004.

D. Kowalski and M. J. Eddy. The DNA unwinding element: a novel, cis-acting component that facilitates opening of the Escherichia coli replication origin. *EMBO J.*, 8(13):4335–4344, Dec 1989.

S. Krishnaswamy, J. A. Rogers, R. J. Isbell, and R. G. Fowler. The high mutator activity of the dnaQ49 allele of Escherichia coli is medium-dependent and results from both defective 3'–>5' proofreading and methyl-directed mismatch repair. *Mutat. Res.*, 288 (2):311–319, Aug 1993.

T. A. Kunkel and K. Bebenek. Recent studies of the fidelity of DNA synthesis. *Biochim. Biophys. Acta*, 951(1):1–15, Nov 1988.

R. S. Lahue, K. G. Au, and P. Modrich. DNA mismatch correction in a defined system. *Science*, 245(4914):160–164, Jul 1989.

M. H. Lamers, A. Perrakis, J. H. Enzlin, H. H. Winterwerp, N. de Wind, and T. K. Sixma. The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature*, 407(6805):711–717, Oct 2000.

H. Lee, S. Kang, S. H. Bae, B. S. Choi, and D. S. Hwang. SeqA protein aggregation is necessary for SeqA function. *J. Biol. Chem.*, 276(37):34600–34606, Sep 2001.

J. S. Lenhart, A. Sharma, M. M. Hingorani, and L. A. Simmons. DnaN clamp zones provide a platform for spatiotemporal coupling of mismatch detection to DNA replication. *Mol. Microbiol.*, 87(3):553–568, Feb 2013.

LifeTechnologiesCorporation. Tetraspeck (tm) fluorescent microspheres size kit, 2013. URL `http://www.lifetechnologies.com/order/catalog/product/T14792`.

P. Liu, W. Duan, Q. Wang, and X. Li. The damage of outer membrane of Escherichia coli in the presence of TiO2 combined with UV light. *Colloids Surf B Biointerfaces*, 78(2):171–176, Jul 2010.

A. Løbner-Olesen, M. G. Marinus, and F. G. Hansen. Role of SeqA and Dam in Escherichia coli gene expression: a global/microarray analysis. *Proc. Natl. Acad. Sci. U.S.A.*, 100(8):4672–4677, Apr 2003.

T. Loh, K. C. Murphy, and M. G. Marinus. Mutational analysis of the MutH protein from Escherichia coli. *J. Biol. Chem.*, 276(15):12113–12119, Apr 2001.

F. J. Lopez de Saro and M. O'Donnell. Interaction of the beta sliding clamp with MutS, ligase, and DNA polymerase I. *Proc. Natl. Acad. Sci. U.S.A.*, 98(15):8376–8380, Jul 2001.

F. J. Lopez de Saro, M. G. Marinus, P. Modrich, and M. O'Donnell. The beta sliding clamp binds to multiple sites within MutL and MutS. *J. Biol. Chem.*, 281(20): 14340–14349, May 2006.

M. Lu, J. L. Campbell, E. Boye, and N. Kleckner. SeqA: a negative modulator of replication initiation in E. coli. *Cell*, 77(3):413–426, May 1994.

Kraft CA Ray K Cranfill PJC Markwardt ML, Kremers G-J. An improved cerulean fluorescent protein with enhanced brightness and reduced reversible

photoswitching, 2011. URL http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0017896.

S. W. Matson. Escherichia coli helicase II (urvD gene product) translocates unidirectionally in a 3' to 5' direction. *J. Biol. Chem.*, 261(22):10169–10175, Aug 1986.

M. V. Matz, A. F. Fradkov, Y. A. Labas, A. P. Savitsky, A. G. Zaraisky, M. L. Markelov, and S. A. Lukyanov. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat. Biotechnol.*, 17(10):969–973, Oct 1999.

M. V. Matz, K. A. Lukyanov, and S. A. Lukyanov. Family of the green fluorescent protein: journey to the end of the rainbow. *Bioessays*, 24(10):953–959, Oct 2002.

M. L. Mendillo, C. D. Putnam, and R. D. Kolodner. Escherichia coli MutS tetramerization domain structure reveals that stable dimers but not tetramers are essential for DNA mismatch repair in vivo. *J. Biol. Chem.*, 282(22):16345–16354, Jun 2007.

W. Messer, U. Bellekes, and H. Lother. Effect of dam methylation on the activity of the E. coli replication origin, oriC. *EMBO J.*, 4(5):1327–1332, May 1985.

W. Messer, F. Blaesing, D. Jakimowicz, M. Krause, J. Majka, J. Nardmann, S. Schaper, H. Seitz, C. Speck, C. Weigel, G. Wegrzyn, M. Welzeck, and J. Zakrzewska-Czerwinska. Bacterial replication initiator DnaA. Rules for DnaA binding and roles of DnaA in origin unwinding and helicase loading. *Biochimie*, 83(1):5–12, Jan 2001.

Halvor Rollag Miklos Degre, Berit Hovig. *Medisinsk mikrobiologi*. Gyldendal, 2008.

P. Modrich. DNA mismatch correction. *Annu. Rev. Biochem.*, 56:435–466, 1987.

P. Modrich and R. Lahue. Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.*, 65:101–133, 1996.

F. Molina and K. Skarstad. Replication fork and SeqA focus distributions in Escherichia coli suggest a replication hyperstructure dependent on nucleotide metabolism. *Mol. Microbiol.*, 52(6):1597–1612, Jun 2004.

K. Moon and S. Gottesman. A PhoQ/P-regulated small RNA regulates sensitivity of Escherichia coli to antimicrobial peptides. *Mol. Microbiol.*, 74(6):1314–1330, Dec 2009.

M. L. Mott, J. P. Erzberger, M. M. Coons, and J. M. Berger. Structural synergy and molecular crosstalk between bacterial helicase loaders and replication initiators. *Cell*, 135(4):623–634, Nov 2008.

G. Natrajan, M. H. Lamers, J. H. Enzlin, H. H. Winterwerp, A. Perrakis, and T. K. Sixma. Structures of Escherichia coli DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates. *Nucleic Acids Res.*, 31(16):4814–4821, Aug 2003.

N. Naue, M. Beerbaum, A. Bogutzki, P. Schmieder, and U. Curth. The helicase-binding domain of Escherichia coli DnaG primase interacts with the highly conserved C-terminal region of single-stranded DNA-binding protein. *Nucleic Acids Res.*, 41 (8):4507–4517, Apr 2013.

K. Negishi, D. Loakes, and R. M. Schaaper. Saturation of DNA mismatch repair and error catastrophe by a base analogue in Escherichia coli. *Genetics*, 161(4): 1363–1371, Aug 2002.

Hiroshi Nikaido. The limitations of lb medium, November 2009. URL `http://schaechter.asmblog.org/schaechter/2009/11/the-limitations-of-lb-medium.html`.

M. F. Noirot-Gros, E. Dervyn, L. J. Wu, P. Mervelet, J. Errington, S. D. Ehrlich, and P. Noirot. An expanded view of bacterial DNA replication. *Proc. Natl. Acad. Sci. U.S.A.*, 99(12):8342–8347, Jun 2002.

G. Obmolova, C. Ban, P. Hsieh, and W. Yang. Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature*, 407(6805):703–710, Oct 2000.

I. Odsbu, H. K. Klungsøyr, S. Fossum, and K. Skarstad. Specific N-terminal interactions of the Escherichia coli SeqA protein are required to form multimers that restrain negative supercoils and form foci. *Genes Cells*, 10(11):1039–1049, Nov 2005.

A. Oka, K. Sugimoto, M. Takanami, and Y. Hirota. Replication origin of the Escherichia coli K-12 chromosome: the size and structure of the minimum DNA segment carrying the information for autonomous replication. *Mol. Gen. Genet.*, 178(1):9–20, Apr 1980.

D. L. Ollis, P. Brick, R. Hamlin, N. G. Xuong, and T. A. Steitz. Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP. *Nature*, 313 (6005):762–766, 1985.

M. Ormo, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, and S. J. Remington. Crystal structure of the Aequorea victoria green fluorescent protein. *Science*, 273(5280): 1392–1395, Sep 1996.

L. B. Overman, W. Bujalowski, and T. M. Lohman. Equilibrium binding of Escherichia coli single-strand binding protein to single-stranded nucleic acids in the (SSB)65 binding mode. Cation and anion effects and polynucleotide specificity. *Biochemistry*, 27(1):456–471, Jan 1988.

B. O. Parker and M. G. Marinus. Repair of DNA heteroduplexes containing small heterologous sequences in Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.*, 89(5): 1730–1734, Mar 1992.

P. Polaczek and A. Wright. Regulation of expression of the dnaA gene in Escherichia coli: role of the two promoters and the DnaA box. *New Biol.*, 2(6):574–582, Jun 1990.

D. C. Prasher, V. K. Eckenrode, W. W. Ward, F. G. Prendergast, and M. J. Cormier. Primary structure of the Aequorea victoria green-fluorescent protein. *Gene*, 111(2): 229–233, Feb 1992.

Promega. Dna purification, 2013a. URL `http://no.promega.com/resources/product-guides-and-selectors/protocols-and-applications-guide/dna-purification/`.

Promega. pgem®-t easy vector systems, 11 2013b. URL `http://no.promega.com/products/pcr/pcr-cloning/pgem_t-easy-vector-systems/?origUrl=http%3a%2f%2fwww.promega.com%2fproducts%2fpcr%2fpcr-cloning%2fpgem_t-easy-vector-systems%2f`.

QIAGEN. Qiaprep miniprep handbook, 05 2012.

R. Qiu, V. C. DeRocco, C. Harris, A. Sharma, M. M. Hingorani, D. A. Erie, and K. R. Weninger. Large conformational changes in MutS during DNA scanning, mismatch recognition and repair signalling. *EMBO J.*, 31(11):2528–2540, May 2012.

Dohet C. Bourguignon M F. Doubleday O P. Lecomte P. Radman, M. Chromosome damage and repair. *Plenum Press*, pages 431–445, 1981.

M. Rapp, D. Drew, D. O. Daley, J. Nilsson, T. Carvalho, K. Melen, J. W. De Gier, and

G. Von Heijne. Experimentally based topology models for E. coli inner membrane proteins. *Protein Sci.*, 13(4):937–945, Apr 2004.

R. Reyes-Lamothe. Use of fluorescently tagged SSB proteins in in vivo localization experiments. *Methods Mol. Biol.*, 922:245–253, 2012.

R. Reyes-Lamothe, X. Wang, and D. Sherratt. Escherichia coli and its chromosome. *Trends Microbiol.*, 16(5):238–245, May 2008.

R. Reyes-Lamothe, E. Nicolas, and D. J. Sherratt. Chromosome replication and segregation in bacteria. *Annu. Rev. Genet.*, 46:121–143, 2012.

A. Roth and W. Messer. The DNA binding domain of the initiator protein DnaA. *EMBO J.*, 14(9):2106–2111, May 1995.

R. Rottenfusser, E. Wilson, E, and V. Davidson, M. Education in microscopy and digital imaging. URL `http://zeiss-campus.magnet.fsu.edu/articles/basics/resolution.html`.

C. E. Samitt, F. G. Hansen, J. F. Miller, and M. Schaechter. In vivo studies of DnaA binding to the origin of replication of Escherichia coli. *EMBO J.*, 8(3):989–993, Mar 1989.

R. M. Schaaper. Mechanisms of mutagenesis in the Escherichia coli mutator mutD5: role of DNA mismatch repair. *Proc. Natl. Acad. Sci. U.S.A.*, 85(21):8126–8130, Nov 1988.

R. M. Schaaper. Base selection, proofreading, and mismatch repair during DNA replication in Escherichia coli. *J. Biol. Chem.*, 268(32):23762–23765, Nov 1993.

R. M. Schaaper and M. Radman. The extreme mutator effect of Escherichia coli mutD5 results from saturation of mismatch repair by excessive DNA replication errors. *EMBO J.*, 8(11):3511–3516, Nov 1989.

R. Scheuermann, S. Tam, P. M. Burgers, C. Lu, and H. Echols. Identification of the epsilon-subunit of Escherichia coli DNA polymerase III holoenzyme as the dnaQ gene product: a fidelity subunit for DNA replication. *Proc. Natl. Acad. Sci. U.S.A.*, 80 (23):7085–7089, Dec 1983.

R. H. Scheuermann and H. Echols. A separate editing exonuclease for DNA replication: the epsilon subunit of Escherichia coli DNA polymerase III holoenzyme. *Proc. Natl. Acad. Sci. U.S.A.*, 81(24):7747–7751, Dec 1984.

T. Selmane, M. J. Schofield, S. Nayak, C. Du, and P. Hsieh. Formation of a DNA mismatch repair complex mediated by ATP. *J. Mol. Biol.*, 334(5):949–965, Dec 2003.

S. S. Shell, C. D. Putnam, and R. D. Kolodner. The N terminus of Saccharomyces cerevisiae Msh6 is an unstructured tether to PCNA. *Mol. Cell*, 26(4):565–578, May 2007.

A. Shimada, Y. Kawasoe, Y. Hata, T. S. Takahashi, R. Masui, S. Kuramitsu, and K. Fukui. MutS stimulates the endonuclease activity of MutL in an ATP-hydrolysis-dependent manner. *FEBS J.*, 280(14):3467–3479, Jul 2013.

Johnson F. H. Saiga Y Shimomura, O. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, aequorea. *J. Cell Comp.*, 59:223–239, 1962.

L. A. Simmons, B. W. Davies, A. D. Grossman, and G. C. Walker. Beta clamp directs localization of mismatch repair in Bacillus subtilis. *Mol. Cell*, 29(3):291–301, Feb 2008.

K. Skarstad, H. B. Steen, and E. Boye. Escherichia coli DNA distributions measured by flow cytometry and compared with theoretical computer simulations. *J. Bacteriol.*, 163(2):661–668, Aug 1985.

K. Skarstad, E. Boye, and H. B. Steen. Timing of initiation of chromosome replication in individual Escherichia coli cells. *EMBO J.*, 5(7):1711–1717, Jul 1986.

K. Skarstad, T. A. Baker, and A. Kornberg. Strand separation required for initiation of replication at the chromosomal origin of E.coli is facilitated by a distant RNA–DNA hybrid. *EMBO J.*, 9(7):2341–2348, Jul 1990.

K. Skarstad, R. Bernander, and E. Boye. Analysis of DNA replication in vivo by flow cytometry. *Meth. Enzymol.*, 262:604–613, 1995.

S. Slater, S. Wold, M. Lu, E. Boye, K. Skarstad, and N. Kleckner. 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(6):927–936, Sep 1995.

B. T. Smith, A. D. Grossman, and G. C. Walker. Visualization of mismatch repair in bacterial cells. *Mol. Cell*, 8(6):1197–1206, Dec 2001.

C. Speck and W. Messer. Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA. *EMBO J.*, 20(6):1469–1476, Mar 2001.

B. G. Spratt. Distinct penicillin binding proteins involved in the division, elongation, and shape of Escherichia coli K12. *Proc. Natl. Acad. Sci. U.S.A.*, 72(8):2999–3003, Aug 1975.

B. G. Spratt. Deletion of the penicillin-binding protein 5 gene of Escherichia coli. *J. Bacteriol.*, 144(3):1190–1192, Dec 1980.

C. Stokke, I. Flatten, and K. Skarstad. An easy-to-use simulation program demonstrates variations in bacterial cell cycle parameters depending on medium and temperature. *PLoS ONE*, 7(2):e30981, 2012.

S. S. Su and P. Modrich. Escherichia coli mutS-encoded protein binds to mismatched DNA base pairs. *Proc. Natl. Acad. Sci. U.S.A.*, 83(14):5057–5061, Jul 1986.

S. A. Taft-Benz and R. M. Schaaper. Mutational analysis of the 3'–>5' proofreading exonuclease of escherichia coli dna polymerase iii. *Nucleic Acid Research*, 26: 4005–4011, 1998.

S. A. Taft-Benz and R. M. Schaaper. The theta subunit of Escherichia coli DNA polymerase III: a role in stabilizing the epsilon proofreading subunit. *J. Bacteriol.*, 186(9):2774–2780, May 2004.

Profiling the E.coli Chromosome. Gene report : muth, 2013. URL `http://www.shigen.nig.ac.jp/ecoli/pec/genes.List.DetailAction.do?fromListFlag=true&featureType=1&orfId=2781`.

N. K. Torheim and K. Skarstad. Escherichia coli SeqA protein affects DNA topology and inhibits open complex formation at oriC. *EMBO J.*, 18(17):4882–4888, Sep 1999.

R. Y. Tsien. The green fluorescent protein. *Annu. Rev. Biochem.*, 67:509–544, 1998.

M. Viswanathan and S. T. Lovett. Single-strand DNA-specific exonucleases in Escherichia coli. Roles in repair and mutation avoidance. *Genetics*, 149(1):7–16, May 1998.

U. von Freiesleben, K. V. Rasmussen, and M. Schaechter. SeqA limits DnaA activity in replication from oriC in Escherichia coli. *Mol. Microbiol.*, 14(4):763–772, Nov 1994.

U. von Freiesleben, M. A. Krekling, F. G. Hansen, and A. Løbner-Olesen. The eclipse period of Escherichia coli. *EMBO J.*, 19(22):6240–6248, Nov 2000.

T. Waldminghaus and K. Skarstad. The Escherichia coli SeqA protein. *Plasmid*, 61(3): 141–150, May 2009.

T. Waldminghaus, C. Weigel, and K. Skarstad. Replication fork movement and methylation govern SeqA binding to the Escherichia coli chromosome. *Nucleic Acids Res.*, 40(12):5465–5476, Jul 2012.

J. H. Weiner, L. L. Bertsch, and A. Kornberg. The deoxyribonucleic acid unwinding protein of Escherichia coli. Properties and functions in replication. *J. Biol. Chem.*, 250(6):1972–1980, Mar 1975.

K. M. Welsh, A. L. Lu, S. Clark, and P. Modrich. Isolation and characterization of the Escherichia coli mutH gene product. *J. Biol. Chem.*, 262(32):15624–15629, Nov 1987.

K. R. Williams, E. K. Spicer, M. B. LoPresti, R. A. Guggenheimer, and J. W. Chase. Limited proteolysis studies on the Escherichia coli single-stranded DNA binding protein. Evidence for a functionally homologous domain in both the Escherichia coli and T4 DNA binding proteins. *J. Biol. Chem.*, 258(5):3346–3355, Mar 1983.

K. R. Williams, J. B. Murphy, and J. W. Chase. Characterization of the structural and functional defect in the Escherichia coli single-stranded DNA binding protein encoded by the ssb-1 mutant gene. Expression of the ssb-1 gene under lambda pL regulation. *J. Biol. Chem.*, 259(19):11804–11811, Oct 1984.

F. Yang, L. G. Moss, and G. N. Phillips. The molecular structure of green fluorescent protein. *Nat. Biotechnol.*, 14(10):1246–1251, Oct 1996.

W. Yang. Structure and function of mismatch repair proteins. *Mutat. Res.*, 460(3-4): 245–256, Aug 2000.

T. C. Zahrt, G. C. Mora, and S. Maloy. Inactivation of mismatch repair overcomes the barrier to transduction between Salmonella typhimurium and Salmonella typhi. *J. Bacteriol.*, 176(5):1527–1529, Mar 1994.

J. W. Zyskind and D. W. Smith. The bacterial origin of replication, oriC. *Cell*, 46(4): 489–490, Aug 1986.

A. Materials

A.1. Bacterial Strains

Table 8: *Bacterial strains used in this thesis*

Strain name	Genotype/relevant features	Selections	Source
AB1157	*F-, thr*-1, *ara*-14, *leuB6*, Δ(*gpt-proA*) 62, *lacY*1, *tsx*-33, *supE44*, *galK*2, *rac*-, *his*G4(Oc), *rfb*D1, *mgl*-51, *kdgK51*, *xyl*-5, *mtl*-1, *qsr*,-*rfb*C1, *mgl*-51, *rpo*S396(Am), *rps*L31(strR), *xyl*A5, mtl-1, *arg*E3(Oc), *thi*-1	-	Howard-Flanders 1964
CM735	*met*E46, *trp*.3, *his-4*, *thi*-1, *gal*K2, *lac*Y1 or *lac*Z4, *mtl*-1, *ara*-9, *tsx*-3, *ton*-1, *rps*-9, *sup*E44, λ–	-	Hansen and von Meyenburg [1979]
MG1655	*F-*, *lambda-*, *rph-1*	-	Guyer, et al 1981
DH10B	F– *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lacZ*ΔM15Δ*lacX*74 *rec*A1 *end*A1 *ara*D139 Δ(*ara leu*) 7697*galU galK rpsL nupG* λ–	-	Invitrogen
One Shot (R) Top10	*mcr*A, Δ(mrr-hsdRMS-mcrBC), Phi80*lacZ*(del)M15, Δ*lacX*74, *deoR*, *recA*1, *ara*D139, Δ(*ara-leu*)7697, *galU*, *galK*, *rpsL*(SmR), *end*A1, *nupG*	-	Invitrogen
GL224	MG1655 with *ssb*-CFP::cmR in *lamB*	cam	Grace Leung
ME125	MG1655 Δ*mutH*::kan, *mutL*218::tn10, Plac-YFP-*mutL*::cam, pBad24*mutH*E56A	kan, amp, cam, tet	Elez et al. [2012]
FR680	MG1655 *mutD*5zae::tn10	tet	Elez et al. [2010]
AB1157/pRed/ET	pRed/ET	tet 3	Gene bridges
EH04	Top10/pEH04 (pSF36 with mKate2 insert)	cam, amp	L
SF95	Top10/pSF36	cam, amp	L
SF149	AB1157 *seqA*-YFP, *ssb*-CFP::cam in *lamB*	cam	Solveig Fossum-Raunehaug (unpuplished)
KB01	JM109 pKB01	amp	This study
KB02	DH10B pKB02	cam, amp	This study
KB03	DH10B pKB03	cam, amp	This study
KB04	*AB1157*Δ *ygdQ* promoter	cam	This study
KB05	AB1157	cam	This study
KB09	SF149 Δ*mutH*::kan	cam, kan	This study
KB12	SF149 *mutD*5::tet	cam, tet	This study
KS9921	MG1655 *seqA*::tn10		Løbner-Olesen et al. [2003]

A.2. Plasmids

Table 9: *Plasmids used in this thesis*

Plasmids	Relevant features	Selection	Source
pGEM-T easy	T+ sticky ends	amp	Promega
pKB01	pGEM-T easy with mCerulean3 inseted	amp	Study creation
pSF36	pUC19 + cam-FRT-HindIII (1014 bp)	amp, cam	Solveig Fossum Raunhaug
pKB02	pSF36 with mCerulean3 inserted reverse upstream of the chloramphenicol resistance gene	amp, cam	Study creation
pKB03	pSF36 with mCerulean3 inserted inserted the same direction and upstream of the chloramphenicol resistance gene	amp, cam	Study creation
pRed/ET	(Redα/Redβ) important in homologous recombination	tet	Gene bridges

A.3. Primers

Table 10: *Primers used in this thesis*

Primer	Nucleotide sequence	Tm	Producer
Left_mCerulean3	5'GGCGGCATCGATGGATCCGTGA GCAAGGGCGAGGAGC3'	57.6℃	BioNordika
Right_mCerulean3	5'GGCGGCATCGATTTAGCTCGTC CATGCCGAGAGTG3'	57.9℃	BioNordika
Left_mKate2_MutH	5'TTGAAGAAGAATTTCACCAGTGC ACTACTGGCCCGTCATTTTCTGAT CCAGTCTGGTTCTGGTTCTGGTTC TGGTTCTGGTTCTGGTGTGAGCGA GCTGATTAAGGAGAA3'	57.9℃	BioNordika
Right_mKate2_30bpMutH	5'ATGCAAATAACATCTAAAAAGC CCTGCCAAAAAAATGAAGCGGTA ATTATACCATGATTACGCCAAGCT TGTGTA3'	55.7℃	BioNordika
Right_mKate2_6bpMutH	5'CCAAAAAAATGAAGCGGTAATT ATATGCCCGGAAAGCGGCAGGTC AAAGCCCATGATTACGCCAAGCT TGTGTA3'	57.4℃	BioNordika
Left_Insert_behind_MutH	5'CCCGGGGCGAACGGATTCTGAC3'	62.3℃	BioNordika
Right_Insert_behind_MutH	5'AAGAACGATCTCCAGCAGCGTC AGCTG3'	62.8℃	BioNordika
Left_deltaMutH	5'CGCATGTGGAAAAAGTTACACT GC3'	55.7℃	BioNordika
Right_deltaMutH	5'CCATGCAAATAACATCTAAAAA GCCC3'	54.8℃	BioNordika
Left_MutD5	GGTAGACTTCCTGTAATTGAATC GA	54,.4℃	BioNordika
Right_MutD5	5'CTCATCTTTTGAATCGTTTGCT GC	54.0℃	BioNordika

A.4. Buffer Solutions

Table 11: *Buffer solutions*

Buffer	Contains	Relevant features
50X TAE buffer	24.2 % Trisbase, 0.05M EDTA, Acetic acid to pH 8, dH20	Agarose gel
0.1 M PB buffer	0.1 M KH2PO4 / K2H2PO4 pH 9	Flow samples
1 X TE buffer	10 mM Tris-HCl pH 7.5, 1mM EDTA	Flow samples
0.002 M TBS buffer	100 mM TrisHCl pH7.5, 29.2 % NaCl	Flow samples
0.02 M PBS buffer	0.02 M KH2PO4 /K2HPO4, 130 mM NaCl	Microscopy

,

Table 12: *Commercial buffer solutions*

Buffer	Relevant features
Pfu Ultra II buffer	Agilent Technologies
Pfu Ultra buffer	Agilent Technologies
EcoRI buffer	New England Biolabs® Inc.
Antarctic phosphatase buffer	New England Biolabs® Inc.
Electro ligase buffer	New England Biolabs® Inc.
Thermopol buffer	New England Biolabs® Inc.
T4 DNA ligase buffer	Promega
NEBuffer 3	New England Biolabs® Inc.
BSA	New England Biolabs® Inc.

,

A.5. Media

Table 13: *The content of ABB_1 Caa Glu*

ABB_1 CAA Glu				
5xA	B	B1	Glu	CAA
$(NH_4)_2SO_4$ 1 g	$CaCl_2$ 0.1 mM	Thiamin 0.01 mg/ml	Glucose 0.2%	CAA 0.5%
Na_2HPO_4 3 g	$MgCl_2$ 1mM			
KH_2PO_4 1.5 g	$FeCl_3$ 0.003mM			
NaCl 1.5 g				
dH_2O 100 ml				

,

Table 14: *The content in LB media*

	LB	**LB lowsalt**	**LB Agar**
Tryptone	10 g	10 g	10 g
Yeast extract	5 g	5 g	5 g
NaCl	10 g	5 g	10 g
Bacto-agar	-	-	20 g
dH_2O	1000 ml	1000 ml	1000 ml

A.6. Antibiotics

Table 15: *Stock solutions of the antibiotics used in this thises*

Antibiotic	Stock solution	Producer
Ampicillin	100 mg/ml	Bristol-Meyers Squibb
Tetracycline	5 mg/ml	Arcopharma AS
Chloramphenicol	30 mg/ml	Sigma-Aldrich ®
Kanamycin	50 mg/ml	Sigma-Aldrich ®
Rifampicin	30 mg/ml	Fluka
Cephalexin	1.0 mg/ml	Eli Lilly

A.7. Reaction Kit

Table 16: *Reaction kits used in this thesis*

Kit	Producer
JETstar (MIDI-prep): The novel plasmid purification system	Promega
QIAprep Spin Miniprep Kit	QIAGEN
Wizared ® Genomick DNA Purification Kit	Genomed
Wizard ® SV Gel and PCR clean-up system	Genomed

A.8. Enzymes

Table 17: *Reaction enzymes used in this thesis*

Enzyme	Concentration	Buffer	Relevant features	Source
EcoR1 HF	20 000 U/ml	EcoR1 buffer	G|AATTC (cutting sequence)	New England Biolabs® Inc.
BamH1	10 000 U/ml	NEBuffer3, BSA	G|GATCC (cutting sequence)	New England Biolabs® Inc.
T4 DNA ligase	3U/µl	T4 DNA ligase buffer	Ligation	Promega
Electroligase		Electroligase buffer	Ligation	New England Biolabs® Inc.
Taq Polymerase	5000 U/ml	Thermopol buffer	Poly-A-tailing	New England Biolabs® Inc.
Pfu Ultra II HS		Pfu Ultra II buffer	PCR	Agilent Technologies
Antarctic phosphatase	5000 U/ml	Antarctic phosphatase buffer	Dephosphorylation	New England Biolabs® Inc.

A.9. Chemicals

Table 18: *Chemicals and reagents used in this thesis*

Chemicals	Concentration	Relevant features	Producer
Agarose, Type 1		DNA separation	Sigma-Aldrich ®
GelRed®	1:10 000	Nucleic Acid stain	Life Technologies
Gel Loading Dye Blue	6x	DNA separation	New England Biolabs ®
2 Log DNA-Ladder	10x	Reference	New England Biolabs ®
Hoechst 33258	1 mg/ml	DNA dye	Sigma-Aldrich ®
FITC	0.3 mg/ml	Protein dye	Sigma-Aldrich ®
IPTG	100 mM	Induses transcription of lac operon	Sigma-Aldrich ®
X-gal	50 mg/ml	Lactose analog	Sigma-Aldrich ®
Nuclease-Free Water	-	Nuclease free water	QIAGEN
MgCl2	1 M	Growth medium	Sigma-Aldrich ®
FeCl2	0.003 M	Growth medium	Sigma-Aldrich ®
$CaCl_2$	0.1 M / 1 M	Growth medium	Sigma-Aldrich ®
Ethanol	77 %	Fixation	Kemetyl Norge AS
Glycerol	87 %	Cell storage	Sigma-Aldrich ®
Isopropanol	100 %	Presipitation	Sigma-Aldrich ®
dNTP	40 mM	PCR	Agilent Technologies
dATP	10 mM	Poly-A tailing	Agilent Technologies
dH_2O	-	-	MilliQ
CAA	10 %	Growth medium	Sigma-Aldrich ®
Glucose	20 %	Growth medium	Sigma-Aldrich ®
Thiamin	10 mg/ml	Growth medium	Fluka

A.10. Equipment and Apparatus

Table 19: *Equipment used in this thesis*

Equipment	Model	Producer
NanoDrop (tm)	ND-1000 Spectrophotometer	Thermo Science
Qubit ®	2.0	Invitrogen
UV Spectrophotometer	UV1800	Shimadzu
Spectrophotometer cuvette	Semi-micro	BrandTech Scientific ®
PCR instrument	C1000 (TM) Thermal cycler	BIO-RAD
Gel Electrophoresis Apparatus	Power PAC 300	BIO-RAD
Electroporator	Gene pulser ® II	BIO-RAD
Electroporation cuvette	1 mm 2 mm	BIO-RAD
Refrigerated Microcentrifuge	5417 R	Eppendorf
Microsentrifuge	Mikro 200R	Hettich
Refrigerated Centrifuge	5930	KUBOTA
Microwave	Wavedom	LG
Shake incubator	Thermomixer	Eppendorf
Waterbath	Innova (TM) 3100	New Brunswick Scientific
Heat Block	QBD2	Grant Instrument
Pipette	0.5-5 µl 2-20 µl 10-100 µl 100-1000 µl	Thermo Scientific
Pipette tip	0.1-10 µl 10-100 µl 100-1000 µl	Thermo Scientific
Large pipette filler	PIPETBOY comfort	IBS-INTEGRA BIOSCIENCES
Needle	Mod 3	BD microlance
Vacuum suction	Broen	Broen
Vortex	MS3 digital	IKA Works, Inc,
Weight	Extend	Sartorius
Gas burner	FLAMEBOY	IBS-INTEGRA BIOSCIENCES
Deep Freezer	ULT Freezer, -76 C	Thermo Scientific
Liquid nitrogen	-	Laboratory stock
Parafilm	Parafilm M	Bemis
Petridish	90-15 mm	Heger AS
Weightdish	35 x 15 mm Style	BD Falcon
Object glass	76 x 26 mm	Thermo Scentific
Coverslip	13 mm	Thermo Scentific
Flowcytometer	LSRII	Becton Dickinson
Fluorescent Microscope	DM6000 B	Leica

B. Method Protocol

B.1. Cell Growth

B.1.1. Overnight Culture

1. Puncture holes in the lid of a 2 ml growth tube to provide air to the overnight culture.
2. Add 1.5 ml of LB medium or ABB_1 GluCAA medium to the tube.
3. Add antibiotics if necessary.
4. Add cells from -80℃ glycerol stock or a colony from an LB agar plate.
5. Incubate at 37℃ with shaking (1100 rpm) overnight.

B.1.2. Preparation of Cells for Deep Freeze Storage

1. Mix overnight cell culture with 87% glycerol in a 1:1 ratio 2 ml screw capped eppendorf tube.
2. Turn the tube upside-down gently 5-6 times.
3. Place the tube in deep freezer at -80℃.

B.1.3. Colony Growth

1. Melt LB agar and cool down to until 50℃
2. Supply with antibiotics if necessary.
3. Fill a Petri dish with 20 ml of the LB agar.
4. Solidify at room temperature for approximately 15 min.

5. Apply and distribute ~50-100ul cell culture on the top of the agar and incubate at 37℃ overnight.

B.1.4. Control of the Growth Rate of Cell Cultures in the Exponential Phase

1. Dilute overnight 1:1000 in either LB or ABB_1 GluCAA medium.
2. Apply antibiotics if necessary. Incubate with shaking at 37℃.
3. Measure OD every half hour.
   LB is measured at wavelength 600, and grown until OD reached 0.3.
   ABB_1 Glu Caa is measured at wavelength 450, and grown until OD reached 0.15
   Cells grown longer than the mentioned ODs will no longer have an exponential growth.

B.2. Agarose gel electrophoresis

B.2.1. Agarose Gel

100 ml 0,8% Agarose gel: 0,8 g Agarose 100ml 1xTAE buffer

1. Heat up solution in microwave, boil until the solution becomes homogenous.
2. Transfer solution over to an agarose gel container, and add a comb. The comb makes wells with the approximate size of 40ul, but the size depends on gel dept and comb size.
   a) If bigger wells are needed, tape together two or more wells

B.2.2. DNA Separation

1. Move the stiffen agarose gel into a agarose gel container with 1xTAE.

2. Add gel loading buffer to DNA samples and apply to the wells.

3. Run the gel at 100 volt for 90 minutes

B.2.3. DNA Staining

1. Stain the gel in GelRed for 20-30 min.

B.3. PCR Amplification

PCR reaction 50 ul:	Volume	Final amount
DNA template (10 ng/µl - 30 ng/µl)	~1 ul	-
10 x PfuUltra II reaction Buffer	5 µl	1x
10 mM dNTP	1.25 µl	250 µM
10 µM Left primer	1 µl	0,2 µM
10 µM Right primer	1 µl	0,2 µM
Pfu Ultra II fusion HS Polymerase	1 µl	1 U
dH_2O	39,75 µl	-
Total	50 µl	

Cycles:	Temperature	Time
1.	95℃	2 min
2.	94℃	0.20 min
3.	Primer Tm - 5℃	0.30 min
4.	72℃	0.45 min
5.	72℃	3 min
6.	8℃	Forever
Stage 2-4	30 cycles	

B.4. Purification of Plasmid from Bacterial Cells

B.4.1. QIAprep Spin Miniprep Kit from Qiagen

1. Centrifuge a 1.5 ml overnight culture 5 min with 15.000rpm, and discard the supernatant.
2. Resuspend the bacterial pellet in 250 µl Buffer P1 and transfer sample to a microcentrifuge tube.
3. Add 250 µl Buffer P2, and mix by inverting the tube approximately 10 times.
4. Add 350 µl Buffer N3 and immediately mix the sample by inverting the tube approximately 10 times.
5. Keep in room temperature for 5 min before centrifugation for 10 min at 4℃ ~13,000rpm.
6. Apply the formed supernatant to a QIAprep spin column and centrifuge for 60s ~13,000rpm.
7. Discard the flow-through.
8. Wash the column by adding 500 µl Buffer PB and centrifuge for 60s ~13,000rpm before discarding the flow-through.
9. Wash column by adding 750 µl Buffer PE and centrifuge for 60s ~13,000rpm.
10. Discard flow-through and repeat centrifugation
11. Place the column in a clean 1.5 ml microcentrifuge tube and add 50 µl nuclease free water. Incubate for 1 min before centrifugation for 1 min ~13,000rpm.

B.4.2. JetStart 2.0 Midi Kit from Genomed

1. The JetStar columns need to be equilibrated before use by adding 10 ml of solution E4. Allow columns to drain by gravity flow.
2. Centrifuge 25 ml of overnight culture in the KUBATOR 5930 centrifuge at 8500 rpm, 4℃ for 10 min and discard supernatant.
3. Apply 4 ml of solution E1 to the pellet and resuspend until the suspension is homogeneous.
4. Add 4 ml solution E2 and mix carefully until the suspension appear homogeneous.
5. Incubate 5 min at room temperature.
6. Apply 4 ml of solution E3 and mix immediately by inverting the tube multiple times until homogeneous suspension is obtained.
7. Centrifuge the sample at 8500 rpm, 4℃ for 15 min.
8. Apply the achieved supernatant to the equilibrated JetStar column and allow it to drain by gravity flow.
9. Wash the column with 10 ml of solution E5 twice, and let the column empty by gravity flow.
10. Elute the DNA from the column by applying 5 ml of solution E6.
11. Precipitate the DNA by adding 3.5 ml isopropanol and centrifuge at 4℃ and 8500 rpm for 30 min.
12. Wash the plasmid with 3 ml 70% cold ethanol and recentrifuge.
13. Air dry the pellet for 10 min and dissolve the DNA pellet in 50 µl nuclease free water.

B.5. Purification of Genomic DNA

1. Centrifuge 1 ml of overnight culture for 3 min at ~15.000 rpm, 4℃ and discard supernatant.
2. Apply 600 µl Nuclei Lysis Solution to the pellet cells and pipet gently to mix.
3. Incubate the sample at 80℃ for 5 min and cool down to room temperature.
4. Add 3 µl of RNase Solution and mix carefully.
5. Incubate 30 min at 37℃, and cool down to room temperature.
6. Apply 200 µl of Protein Precipitation Solution and vortex 1 min.
7. Incubate the sample on ice for 5 min and centrifuge at ~15.000 rpm for 3 min.
8. Transfer the supernatant to a clean tube and add 600 µl isopropanol, and mix by inverting the tube multiple times.
9. Place the sample in the freezer at -20℃, preferably overnight.
10. Centrifuge sample 20 min at ~15.000 rpm and discard the supernatant.
11. Wash pellet with 600 µl of 70 % cold ethanol.
12. Centrifuge for 3 min at ~15.000 rpm and discard the supernatant.
13. Air dry the DNA pellet for 15 min and rehydrate in 100 µl of Rehydration Solution overnight at 4℃.

B.6. Purification of DNA from Agarose Gel and PCR Reaction

1. If DNA in agarose gel;

a) Cut out the DNA from the agarose gel after electrophoresis, and place it in a 15 ml tube.

b) Add 1µl Membrane Binding solution for every 1 mg of gel.

c) Vortex and incubate at 50-65℃ until the gel has dissolved.

2. If PCR product;

a) Apply equal amount of Membrane Binding solution as the PCR product, and mix thoroughly.

3. Insert a SV Minicolumn into a collection tube.

4. Transfer the PCR product to the minicolumn, and incubate at room temperature for 1 min.

5. Centrifuge at ~15.000 rpm for 1 min and discard the flow through. (If the volume of the prepared solution volume is >700 µl the repeat last step)

6. Wash the column by adding 700 µl of Membrane Wash Solution (ethanol added), and centrifuge at ~15.000 rpm for 1 min and discard the flow through.

7. Repeat the last step with 500 µl Membrane Wash Solution and centrifuge for 5 min.

8. Empty the collection tube and centrifuge one more time for 1 min with the centrifuge lid off to allow evaporation of ethanol residuals.

9. Elute the DNA from the column by carefully transferring it to a clean 1,5 eppendorf tube.

10. Add 50 µl of nuclease free water and incubate in room temperature for 1 min, before centrifugation at ~15.000 rpm for 1 min.

B.7. Cutting of Plasmids with Restriction Enzymes

Cutting reaction	Volume	Final amount
0.5 - 1 µg DNA	-	
10000 U/ml Enzyme	1 µl	10 U
10x Buffer	2 µl	1x
10 x BSA	2 µl	1x
dH_2O	-	
Total	20 µl	

1. Set up restriction enzyme digestion as shown above
2. Incubate reaction for 1 hour at 37℃
3. Inactivate enzyme at 65℃ for 20 min

B.8. Dephosphorylation

Dephosphorylation of vector	Volume	Final amount
Cutting rection	-	-
10 x Antarctic Phosphatase Reaction Buffer	3 µl	1 x
5000 U/ml Antarctic Phosphatase	0.6 µl	3 U
dH_2O	-	-
Total	30 µl	

1. Add the Antarctic phosphatase, Antarctic Phosphatase Reaction Buffer and dH_2O to the inactivated cutting reaction as shown above.
2. Incubate at 37℃ for 1 hour
3. Heat inactivate enzyme at 70℃ for 5 min

B.9. Ligation

Ligation of insertion into vector

$$\frac{\text{X ng vector} \cdot \text{bp insert} \cdot 3}{\text{bp vector}} = \text{X ng insert}$$

Ligation reaction	Volume	Amount
2x Ligation buffer	10 µl	1 x
Vector	1 µl	50 ng
Insert	-	-
DNA ligase	1 µl	1 U
Nuclease free water	-	-
Total	20 µl	

1. Calculate the amount of insert and add all components shown above.
2. Incubate 1 hour in 25℃.
3. Heat inactivate enzyme at 65℃ in 15 min.

B.10. A-tailing with Taq DNA Polymerase

A-tailing	Volume	Amount
100 ng/µl PCR product	7 µl	700 ng
10 mM dATP	1 µl	1 mM
10 x NEB PCR buffer	1 µl	1x
5000 U/ml Taq Polymerase	1 µl	5 U
Nuclease free water	2 µl	
Total	10 µl	

B.11. DNA Precipitation

1. Add 1:10 of 3M Na-acetate to the sample.
2. Add 2-3 times the final volume of cold 100 % ethanol, mix by inverting 10 times. Place the sample at -20℃ overnight.
3. The next day, centrifuge the sample at ~15.000 rpm at 4℃ for 30 min.
4. Discard the supernatant and wash the pellet twice in ice cold 70 % ethanol and centrifugation for 10 min ~15.000 rpm in between each round.
5. Air dry the pellet for 20 min, and resuspended in 1/5 of the start volume with nuclease free water.

B.12. Transformation

B.12.1. Transformation of Plasmid into Chemically Competent Cells

1. Transformation into JM109 and Top10 follow the same procedure
2. Thaw cells on ice and prechill a tube with punched holes on ice.
3. Add 50 µl of chemical competent cells to a prechilled tube.
4. Add 2 µl of desired DNA and incubate on ice for 30 min
5. Move sample over to a water bath with exactly 42℃, and keep it there for 30 sec before transferring sample back on ice.
6. Incubate at 37℃ with shaking (300 rpm) for 1 hour
7. Plate on a LB agar plate with antibiotics.

B.12.2. Transformation of Plasmid into Electrocompetent Cells

Preparation Electrocompetent Cells

1. Day 1: Make an overnight culture of DH10B in low-salt LB medium at 37℃
2. Day 2: Dilute overnight culture 1:1000 in 500 ml LB low-salt medium, transfer to a 2 L growth bottle. Grow cells at 37℃ until OD reaches 0.3.
3. Place the 500 ml bottle on ice for 15 min and transfer the culture to a pre chilled centrifuge bottle. Centrifuge at 4000 rpm for 15 min at 4℃.
4. Resuspended by shaking the cells in 15 ml milliQ H_2O and centrifuge again at 4000 rpm for 15 min at 4℃. Repeat step twice.
5. Resuspend pellet 1:4 in chilled milliQ water. Divide the solution into 50 µl per eppendorf tube, quick freeze by throwing the tubes in a bottle of liquid nitrogen.

Transforming into Electrocompetent Cells

1. Thew electrocompetent cell on ice.
2. Add 1 µl of electroligation mix. Transfer the solution to a pre chilled 2 mm cuvette.
3. Insert the cuvette in the electroporator, and turn it on 2500 V, 400 ohm and 25 uF. Hold the two buttons until the sound of a beep.
4. Apply 700 µl of LB to the cuvette, and transfer the whole amount to an eppendorf tube with holes in the lid.
5. Shake the sample in 300 rpm at 37℃ for 1 hour. Plate out on a LB agar with antibiotics.

B.13. P1 Transduction

B.13.1. Making P1 Phage Lysate

1. Day 1: Make overnight culture of donor strain in 3 ml LB medium with the necessary antibiotics.
2. Day 2: Add 15 µM $CaCl_2$ to the overnight culture.
3. Dilute P1 stock 10^{-2} to 10^{-5} in LB medium also containing 15 mM $CaCl_2$ to the final volume of 100 µl. Keep the lid of the tube of for 10 min to remove chloroform.
4. Add 200 µl of donor strain to each of the dilutions and incubate for 20 min in 37℃ (without shaking).
5. Make top agar by warming 30 µl LB and add 10 of premelted LB agar, cool down and keep it at approximately 50-55℃. Add $CaCl_2$ to the final concentration of 15 mM.
6. Add 6 ml top agar per dilution and quickly plate in on an LB agar plate. Let the agar solidify. Keep the plate with the agar bottom down and incubate overnight at 37℃.
7. Day 3: Choose the plate where the plaques grows almost overlapping, and scrape together the top agar with a glass stick and transfer it to a 50 ml tube.
8. Wash the plate with 1 ml LB with 15 mM $CaCl_2$ and add to 50 ml tube. Add 50 µl of chloroform to kill the bacteria cells and centrifuge at 8500 rpm for 30 min.
9. Transfer the supernatant to an eppendorf tube and add 10 µl of chloroform and 20 mM MgCl2. Incubate for 60 min at 20℃ before use.

B.13.2. Transduction to Recipient with P1 Lysate

1. Day 2: Make overnight culture of recipient strain in LB with necessary antibiotics.
2. Day 3: Add 15 mM $CaCl_2$.
3. Dilute P1 lysate 1, 10^{-1} and 10^{-2} in LB containing 15 mM $CaCl_2$ to the final volume of 100 µl. Incubate 5 min in 37℃.
4. Add 200 µl of recipient culture to each of the dilutions and incubate for 20 min at 37℃.
5. Add the dilutions on to LB agar plates with chosen antibiotic selection and incubate at 37℃ overnight.
6. Day 4: Pick colonies and restreak to verify antibiotic resistance.

B.14. ET Cloning with RED/ET

1. Day 1: Make an overnight culture of AB1157 with pRED/ET in low-salt LB with 3 µg/ml tet.
2. Day 2:Take two eppendorf tubes and make 3-4 holes in the lid. Fill the two tubes with 1.4 ml low salt LB and 3 tet.
3. Add 30 µl of the overnight culture to each of the tubes and incubate at 30℃ with shaking for approximately 2 hours until OD reaches ~0.3.
4. Add 50 µl of 10 % arabinose to one of the two tubes and incubate at 37℃ with shaking for 1 hour.
5. Prepare cells for electroporation by centrifugation at 11000 rpm 30 sec in a cooled benchtop centrifuge. Discard supernatant by tipping it quickly, and place the pellet on ice.

6. Apply 1 ml chilled dH_2O and resuspend the pellet. Repeat the centrifugation and resuspension step again.
7. Centrifuge and remove the supernatant, then add 20 µl dH_2O and resuspend.
8. Add 400 to 800 ng of purified PCR product. Pipette the suspension over to a chilled electroporation cuvette (1 mm).
9. Set the electroporator at 1. 35 kV, 10 µF, 600 Ohms, and press the two buttons until the sound of a beep.
10. Add 1 ml LB medium to the cuvette and mix carefully by pipetting up and down.
11. Transfer the content in the cuvette to an eppendorf tube with holes in the lid and shake at 37℃ for 3 hours (recombination occur).
12. Centrifuge for 30 sec and remove 900 µl of the supernatant. Resuspend the pellet in the remaining medium and spread on a LB agar plate with antibiotic and incubate at 37℃ overnight.

B.15. Flow Cytometry

1. Day 1: Prepare overnight culture in 1.5 ml in ABB_1 Glu CAA medium with antibiotics and shake at 28℃ for 16 hours.
2. Day 2: Dilute the overnight culture 1:1000 in 25 ml ABB_1 Glu CAA medium and measure the OD.
3. Keep the culture in shaking device at 28℃. Grow cells until OD 0.15.
4. At OD 0.15 take out 2 x 1.5 ml of culture and add rifampicin (300 µg/ml) and cephalexin (10 µg/ml) is added to the rif/chx sample, incubation is continued for 3-4 hours with shaking at 28℃.

5. Harvest cells by centrifugation at 15000 rpm for 5 min at 4℃.
6. Wash and resuspend pellet in 1 ml filtered cold TE.
7. Centrifuge and remove 900 µl of the supernatant.
8. Resuspend cells in the 100 µl TE left.
9. Fixate cells by adding 1 ml filtered cold 77 % ethanol.

B.15.1. Staining Samples with FITC/Hoechst

FITC solution: dissolve 3 mg FITC in 10 ml 0.1M PB (0,3 mg/ml stock solution). 100 ul of this stock is dissolved in 10 ml PB giving a final concentartion 3 µg/ml.

Hoechst 33258: Make a solution of 3 µg/ml in 0.002M TBS from the 1 mg/ml stock. The diluted Hoechst 33258 can be stored at 4℃.

1. Spin down the fixed samples including one tube of 5x CM735 standard cells, and wash the pellet in 1 ml of PB.
2. Resuspend samples in 250 µl PB. Splitt the 5x standard in two, 1/10 is treated like the samples and called 1x standard. The second 9/10 standard is not stained with FITC and called FITC negative standard cells.
3. Add 250 µl freshly made FITC (3 µg/ml) to samples and 1 x standard and store in 4℃ overnight.
4. The next day spin down the cells and wash in 1 ml 0.02 M TBS including the 5 x standard.
5. Resuspend samples and 1xstd in 250 µl TBS, and the 5 x standard in 50 TBS.

6. Add 250 µl of 3 µg/ml Hoechst 33258 and 3 µl 5 x standard to samples and 1 x standard.

7. Let cells stain in dark on ice for at least 30 min.

B.16. Microscope Samples

1. Prepare ONC of the samples.

2. The next day, dilute the ONC 1:500 in ABB_1 Glu Caa, and grow until OD 0.15.

3. Make agarose pads: 1% agarose (1 x PBS and 0.5g agarose). Heat up until boiling and let it chill down for 5-10 min. Prepare a frame on every glass slide and add 500 µl of the agarose mix, put a plastic cover on plus lens paper and a heavy block to press it down.

4. Harvest cells by spinning at 15000 rpm for 5 min at room temperature. Resuspend pellet in 30 µl 1x PBS.

5. Place 30 µl of cell suspension on the agarose pad and and wait 10 min to let it dry before adding a coverslip.

B.17. Measuring Mutation Frequency

1. Prepare ONC of the strains in LB with antibiotics

2. Make a dilution series with 1, 10^{-1} , 10^{-3} , 10^{-5} and 10^{-6} in LB.

3. Plate out the 10^{-5} and 10^{-6} dilution samples on a LB agar plate, and the 1 and 10^{-1} dilution samples on a 100 µg/ml rifampicin plate. Incubate overnight in 37℃.

4. Count the cells on the LB agar with a reasonable number of cells and multiply the number by the dilution to obtain the number of cells plated on the rifampicin agar.

Count the number of cells on the rifampicin agar and divide it by the number of cells plated.

C. Supplementary

C.1. Sequencing KB12

The sequence of the *dnaQ* gene of the two strains MG1655 and KB12. The query sequence MG1655, subject KB12:

```
Score    Expect   Identities      Gaps    Strand
1330 bits(720)  0.0     728/732(99%)    0/732(0%)       Plus/Plus
Query  1    ATGAGCACTGCAATTACACGCCAGATCGTTCTCGATACCGAAACCACCGGTATGAACCAG  60
            |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Sbjct  39   ATGAGCACTGCAATTACACGCCAGATCGTTCTCGATACCGAAATCACCGGTATGAACCAG  98

Query  61   ATTGGTGCGCACTATGAAGGCCACAAGATCATTGAGATTGGTGCCGTTGAAGTGGTGAAC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  99   ATTGGTGCGCACTATGAAGGCCACAAGATCATTGAGATTGGTGCCGTTGAAGTGGTGAAC  158

Query  121  CGTCGCCTGACGGGCAATAACTTCCATGTTTATCTCAAACCCGATCGGCTGGTGGATCCG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  159  CGTCGCCTGACGGGCAATAACTTCCATGTTTATCTCAAACCCGATCGGCTGGTGGATCCG  218

Query  181  GAAGCCTTTGGCGTACATGGTATTGCCGATGAATTTTTGCTCGATAAGCCCACGTTTGCC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  219  GAAGCCTTTGGCGTACATGGTATTGCCGATGAATTTTTGCTCGATAAGCCCACGTTTGCC  278

Query  241  GAAGTAGCCGATGAGTTCATGGACTATATTCGCGGCGCGGAGTTGGTGATCCATAACGCA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  279  GAAGTAGCCGATGAGTTCATGGACTATATTCGCGGCGCGGAGTTGGTGATCCATAACGCA  338

Query  301  GCGTTCGATATCGGCTTTATGGACTACGAGTTTTCGTTGCTTAAGCGCGATATTCCGAAG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  339  GCGTTCGATATCGGCTTTATGGACTACGAGTTTTCGTTGCTTAAGCGCGATATTCCGAAG  398
```

```
Query  361  ACCAATACTTTCTGTAAGGTCACCGATAGCCTTGCGGTGGCGAGGAAAATGTTTCCCGGT  420
            |||||||| ||||||||||||||||||||||||||||||||||||||||||||||| |||
Sbjct  399  ACCAATACCTTCTGTAAGGTCACCGATAGCCTTGCGGTGGCGAGGAAAATGTTTCCTGGT  458

Query  421  AAGCGCAACAGCCTCGATGCGTTATGTGCTCGCTACGAAATAGATAACAGTAAACGAACG  480
            |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Sbjct  459  AAGCGCAACAGCCTCGATGCGTTATGTGCCCGCTACGAAATAGATAACAGTAAACGAACG  518

Query  481  CTGCACGGGGCATTACTCGATGCCCAGATCCTTGCGGAAGTTTATCTGGCGATGACCGGT  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  519  CTGCACGGGGCATTACTCGATGCCCAGATCCTTGCGGAAGTTTATCTGGCGATGACCGGT  578

Query  541  GGTCAAACGTCGATGGCTTTTGCGATGGAAGGAGAGACACAACAGCAACAAGGTGAAGCA  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  579  GGTCAAACGTCGATGGCTTTTGCGATGGAAGGAGAGACACAACAGCAACAAGGTGAAGCA  638

Query  601  ACAATTCAGCGCATTGTACGTCAGGCAAGTAAGTTACGCGTTGTTTTTGCGACAGATGAA  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  639  ACAATTCAGCGCATTGTACGTCAGGCAAGTAAGTTACGCGTTGTTTTTGCGACAGATGAA  698

Query  661  GAGATTGCAGCTCATGAAGCCCGTCTCGATCTGGTGCAGAAGAAAGGCGGAAGTTGCCTC  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  699  GAGATTGCAGCTCATGAAGCCCGTCTCGATCTGGTGCAGAAGAAAGGCGGAAGTTGCCTC  758

Query  721  TGGCGAGCATAA  732
            ||||||||||||
Sbjct  759  TGGCGAGCATAA  770
```

C.2. Cell Cycle Parameters

Table 20: *Flow and cell cycle analysis were performed three times, the table show τ, C-period and a_i with standard deviation for KB09 and KB12.*

Strain	Genotype	τ	C-period	a_i
SF149	*seqA*-YFP, *ssb*-CFP	62 min	59 min	9 min
KB09	*seqA*-YFP, *ssb*-CFP, Δ *mutH*	58 ± 0 min	51 ± 4 min	9 ± 2 min
KB12	*seqA*-YFP, *ssb*-CFP, *mutD5*	78 ± 0 min	88 ± 2 min	23 ± 1 min

C.3. Statistical Hypthesis Test Between SF149 and KB09

Table 21: *A table showing the mean, variance and the number of samples counted in both SF149 and KB09.*

	Number	Mean	Variances
SF149	300	197 nm	103
KB09	450	150 nm	68

To investigate whether the two samples were significantly different, a statistical hypthesis test was conducted. The procedure from Helbæk [2001] was followed. First the variances were investigated to determine which assumptions had to be made for the mean test:

$$H_0 \quad : \quad \sigma^2_{SF149} = \sigma^2_{KB09}$$
$$H_1 \quad : \quad \sigma^2_{SF149} \neq \sigma^2_{KB09}$$

The test statistic used was

$$F_{obs} = \frac{S^2_{\mathrm{SF149}}}{S^2_{\mathrm{KB09}}}$$

Where

$$S_{SF149} > S_{KB09}$$

Where S is the sample standard deviations. Insertion of the sample standard deviations yields

$$F_{obs} = \frac{103^2}{68^2} = 2.31$$

The number of degrees of freedom is the pair (m_{SF149}, m_{KB09}) . With a 1% significance level

$$\begin{aligned} F_{crit} &= F_{0.005}(300, 450) = 1.309 \\ F_{obs} &> F_{crit} \end{aligned}$$

Thus H_0 is rejected with a 1% significance level, and conclude that the population variances are unequal.

A Welch's t-test, a t-test used for two samples with unequal variance was performed to investigate if the difference of the two populations means was statistically significant.

$$\begin{aligned} H_0 &: \mu_{SF149} = \mu_{KB09} \\ H_1 &: \mu_{SF149} \neq \mu_{KB09} \end{aligned}$$

The t-test

$$T_{obs} = \frac{|\bar{X}_{\mathrm{SF}} - \bar{X}_{\mathrm{KB}}|}{\sqrt{(S^2_{\mathrm{SF}}/n_{\mathrm{SF}}) + (S^2_{\mathrm{KB}}/n_{\mathrm{KB}})}}$$

Insertion of the sample values yields

$$T_{obs} = \frac{197 - 150}{\sqrt{(103^2/300) + (68^2/450)}} = 6,94$$

The number of degrees of freedom can be calculated using

$$m = \frac{(S^2_{\mathrm{SF149}}/n_{\mathrm{SF09}}) + (S^2_{\mathrm{KB149}}/n_{\mathrm{KB09}})}{((S^2_{\mathrm{SF149}}/n_{\mathrm{SF09}})/n_{\mathrm{SF149}}+1)) + ((S^2_{\mathrm{KB09}}/n_{\mathrm{KB09}})/n_{\mathrm{KB09}}+1))} - 2$$

In our case:

$$m = \frac{(103^2/300) + (68^2/450)}{((103^2/300)/300+1)) + ((68^2/450)/450+1))} - 2 = 1004$$

The number of degrees of freedom is 1004. With a 1% significance level

$$\begin{aligned} T_{crit} &= T_{0.005}(1004) = 2.813 \\ T_{obs} &> T_{crit} \end{aligned}$$

Thus H_0 is rejected with a 1% significance level, and we can conclude that the population means are 99% likely to be unequal.

C.4. Statistical Hypthesis Test Between SF149 and KB12

Table 22: *A table showing the mean, variance and the number of samples counted in both SF149 and KB12.*

	Number	Mean	Variances
SF149	300	197 nm	103
KB12	450	238.8 nm	140

To investigate whether the two samples were significantly different, a statistical hypthesis test was conducted. The procedure from Helbæk [2001] was followed. First the variances were investigated to determine which assumptions had to be made for the mean test:

$$\begin{aligned} H_0 &: \sigma^2_{SF149} = \sigma^2_{KB12} \\ H_1 &: \sigma^2_{SF149} \neq \sigma^2_{KB12} \end{aligned}$$

The test statistic used was

$$F_{obs} = \frac{S^2_{\text{KB12}}}{S^2_{\text{SF149}}}$$

Where

$$S_{KB12} > S_{SF149}$$

Where S is the sample standard deviations. Insertion of the sample standard deviations yields

$$F_{obs} = \frac{140^2}{103^2} = 1.85$$

The number of degrees of freedom is the pair (m_{SF145}, m_{KB12}) . With a 1% significance level

$$\begin{aligned} F_{crit} &= F_{0.005}(300, 450) = 1.309 \\ F_{obs} &> F_{crit} \end{aligned}$$

Thus H_0 is rejected with a 1% significance level, and conclude that the population variances are unequal.

A Welch's t-test, a t-test used for two samples with unequal variance was performed to investigate if the difference of the two populations means was statistically significant.

$$\begin{aligned} H_0 &: \mu_{SF149} = \mu_{KB12} \\ H_1 &: \mu_{SF149} \neq \mu_{KB12} \end{aligned}$$

The t-test

$$T_{obs} = \frac{|\bar{X}_{\mathrm{KB12}} - \bar{X}_{\mathrm{SF149}}|}{\sqrt{(S^2_{\mathrm{KB12}}/n_{\mathrm{KB12}}) + (S^2_{\mathrm{SF149}}/n_{\mathrm{SF149}})}}$$

Insertion of the sample values yields

$$T_{obs} = \frac{239 - 197}{\sqrt{(140^2/450) + (103^2/300)}} = 4.72$$

The number of degrees of freedom can be calculated using

$$m = \frac{(S^2_{\mathrm{KB12}}/n_{\mathrm{KB12}}) + (S^2_{\mathrm{SF149}}/n_{\mathrm{SF149}})}{((S^2_{\mathrm{KB12}}/n_{\mathrm{KB12}})/n_{\mathrm{KB12}} + 1)) + ((S^2_{\mathrm{SF149}}/n_{\mathrm{SF149}})/n_{\mathrm{SF149}} + 1))} - 2$$

In our case:

$$m = \frac{(140^2/450) + (103^2/300)}{((140^2/450)/450 + 1)) + ((103^2/300)/300 + 1))} - 2 = 365$$

The number of degrees of freedom is 365. With a 1% significance level

$$\begin{aligned} T_{crit} &= T_{0.005}(365) = 2.824 \\ T_{obs} &> T_{crit} \end{aligned}$$

Thus H_0 is rejected with a 1% significance level, and we can conclude that the population means are 99% likely to be unequal.

C.5. Mutation Ratio

Table 23: *The mutation ratio of the strains AB1157, MG1655, KS9921, SF149, KB09 and KB12*

		LB plate	Rifampicin	Rifampicin	Mutation ratio
	Dilution of ONC	10^{-6}	1	10^{-1}	
1	AB1157	192	13	0	$6.77 \cdot 10^{-8}$
2	AB1157	367	4	0	$1.09 \cdot 10^{-8}$
3	AB1157	103	22	0	$2.14 \cdot 10^{-7}$
1	MG1655	496	4	2	$8.06 \cdot 10^{-9}$
2	MG1655	754	15	1	$1.98 \cdot 10^{-8}$
3	MG1655	1557	26	4	$1.67 \cdot 10^{-8}$
1	KS9921	543	6	0	$1.10 \cdot 10^{-8}$
2	KS9921	286	11	0	$3,84 \cdot 10^{-8}$
3	KS9921	419	7	1	$1.67 \cdot 10^{-8}$
1	SF149	52	1	0	$1.92 \cdot 10^{-8}$
2	SF149	668	4	0	$5.98 \cdot 10^{-9}$
3	SF149	22	0	0	-
1	KB09	732	-	872	$1.18 \cdot 10^{-5}$
2	KB09	783	-	155	$1.97 \cdot 10^{-6}$
3	KB09	744	-	159	$2.14 \cdot 10^{-6}$
1	KB12	836	-	938	$1.12 \cdot 10^{-5}$
2	KB12	331	-	1337	$4.03 \cdot 10^{-5}$
3	KB12	481	-	766	$1.59 \cdot 10^{-5}$