SSB and Genetic Instability

Federica Andreoni

Thesis presented for the degree of Doctor of Philosophy
Institute of Cell Biology
The University of Edinburgh
September 2008
Declaration

I hereby declare that this thesis was composed by me, and the research presented is my own, except where otherwise stated.

Federica Andreoni
September 2008
Acknowledgements

I would like to thank David for giving me the opportunity and privilege to work in his lab and for his guidance during the course of my Ph.D.

Thanks to Wilson for founding my Ph.D. and for giving me the opportunity to explore a new field of research.

Thanks to Dr. Garry Blakely (University of Edinburgh) for providing the SOS reporter plasmid pGB150 used in this study. Thanks to Prof. McMacken (Johns Hopkins) for providing the anti-SSB antibody.

Present members of the lab, thanks, you are ace guys! Particular thanks go to Elise for attentive and detailed proofreading and critical reading of this thesis. Thanks to AJ for GM support and for having silly chats with me while I was submerged by endless GM analyses. Thanks to Martin boy both for microscopy tips and for being so efficient in providing the lab with good old Scrumpy cider. Also, many thanks to Ewa for her technical support during my Ph.D. and for the “how to be a good lab manager” tips.

Hi-5 to the past lab members for the banter supplied during my first two years of studies. Thanks to Dr. John E. for all his constructs and strain I used. Thanks to Kebaab-ya-bas for sharing silliness and laughter and for giving me the excuse to wrap up everything that was on her bench in tin foil. Man, that was fun! Many thanks to Miss Michelle Stevense for being such a good laugh and for the precious help given during the “tin foil business” (see above). Thanks to Laura-Pakora, I am really missing the random-noises sessions we were having in the lab! I am sooo glad she became a T-shirt :) Jezbian: FA FA NAPOLI stinky head!!! Spanish Donkey, you’re the best… Today!

Thanks to my new Boss Heidrun for giving me the opportunity to stick around Schottland a while longer and for taking care of my beloved dragon(fly?) tree.
Acknowledgments

Thanks to the sequencing department of the University of Edinburgh for being so punctual and helpful. Special thanks go to Anna-Banana, for her quickies and curries and fun-tennis games.

Thanks to the Bball and Monday-Footie guys and especially to the organisers Kate and The Neel. Thanks to the climbing bunch, Emma, TomAdams and his beard, Jezza, Lucio, Stefano, Maria&Ninja and Andy from Andyland.

I have to say I had some of the best times, parties and nights out with “los Latinos”. Guys, you have been soooooo great! Muchas gracias a los Latinos auténticos Claudia, Paola, Ana Arel y El Profesor Sebastian Lopez Escarcena. Gracias por las clases de Mexicano y Chileno, por la tequila, el Manjar, y la musica! Thanks to the honorary Latinos, il Simo, Nikos, Cecile, Alex and Dave for being such good fun! A special mention goes to Simone’s flatmates Daria, Rupy-Rupy-Rupy-Rupeeeee and Shona for being always so welcoming, for coocking such good dinners and make the flat available for the craziest parties.

Grazie alla compa del Bruco, Dick, Bug, Tony, Frenk, Rosmy, Silvia, Silvietta, Frasci e Ggioggina, mi basta leggere una delle vostre mail per spanciarmi dallerrisa e dimenticarmi di una giornata al catztz. Scibboni scibbelli scitanti fratelli.

Grazie alla Rella e a True per supporto morale e assistenza tecnica specialmente durante il primo anno di dottorato, you are the best. And of course grazie al Dottor Piazza per la consulenza scientifica, bella Ricci!

Grazie alla mia famiglia, Mamma, Papá e Cello, per il supporto datomi durante gli anni di dottorato. Ho sempre voi come punto di riferimento e una costante nostalgia di casa. Un ringraziamento particolare alla nonna Maria che mi ha rattrappato un numero infinito di maglioni e pantaloni strappati. Ovviamente grazie anche a tutto il resto della famiglia, zii, zie e cugini, che mi han fatto sentire a casa ad ogni rientro.

Il ringraziamento piu’ speciale va alla persona piu’ speciale, Astuccio, che mi e’ stata vicina in questi tre-quasi-quattro anni e che mi ha sempre appoggiata, specialmente nei momenti piu’ difficili. E’ stato essenziale averti al mio fianco.
Finally thanks to EPSRC for funding this research at the University of Edinburgh.
ABSTRACT

Genome stability has great importance in maintaining cell viability and optimal functionality of cellular processes. Loss of genome stability can lead to cell death in the simplest organisms and to deregulation of the cell proliferation machinery in higher organisms, potentially causing cancer or morbid states.

The Single Stranded DNA Binding (SSB) protein of *Escherichia coli* is an essential protein that binds and stabilises ssDNA stretches. Its role is particularly crucial during DNA replication, recombination and repair processes and it has therefore been predicted to play a prominent role in the maintenance of genome stability. The role of SSB in genome instability was investigated using an *E. coli* strain in which, the expression of the *ssb* gene was placed under the control of the arabinose promoter. The level of SSB protein present in the cell could therefore be tuned by varying the arabinose concentration in the medium.

A wide characterisation of the behaviour of the strain at low SSB level was carried out. Viability and growth tests showed that a threshold level of protein is required to allow normal growth. Microscopy analyses were carried out to follow cell division, nucleoid morphology and SOS response activation. Cells grown at low SSB level, showed a phenotype consistent with impaired cell division and altered nucleoid morphology. The SOS response was activated at low SSB levels and cell elongation was detected.

Lowering the arabinose concentration in solid medium allowed the selection of suppressor strains that could form colonies under the new conditions. Sequencing of the entire genome of one such suppressor strain was carried out revealing a possible candidate for the phenotype change.

The stability of a 105bp and of a 246bp DNA imperfect palindromes and the stability of CAG·CTG trinucleotide repeat arrays, inserted in the *E. coli* chromosome, were investigated
in correlation to the SSB cellular level. Lowering the SSB level in cells grown on solid medium, increased the instability of the 105bp palindrome presumably by increasing the number of slippage events. On the other hand, SSB overexpression did not have an effect on the stability of the 246bp palindrome. The stability of a leading strand \((\text{CAG})_{75}\) repeat array was highly increased by overexpressing SSB, while the same effect was not observed for a leading strand \((\text{CTG})_{137}\) repeat array. Furthermore, excess SSB caused a change in the deletion size distribution profile for the leading strand \((\text{CAG})_{75}\) strain, lowering the bias towards big deletions. This is consistent with SSB being able to preferentially impede the formation of big DNA hairpins. Also, SbcCD nuclease was shown to have an effect on the deletion size distribution profile of the leading strand \((\text{CTG})_{137}\) strain. The lack of SbcCD led to a slight reduction of the number of big deletions.
Abbreviations

\( A_{260} \) Absorbance at 260nm
\( A_{280} \) Absorbance at 280nm
A, C, G, T Adenosine, cytosine, guanine, thymine
APS Ammonium persulfate
Ara Arabinose
bp Base-pair
ddH\(_2\)O Double distilled water
DMSO Dimethyl Sulfoxide
DNA Deoxyribonucleic acid
dNTPs Deoxyribonucleotides triphosphate
DSB Double-strand break
\( E. \) coli \textit{Escherichia coli}
EM Electron microscopy
Glc Glucose
IPTG Isopropyl \( \beta \)-D-1-thiogalactopyranoside
J Jules
kb Kilobase
l Litre
M Molar
m Metre
ml Millilitre
n Nano
\( \text{OD}_{600} \) Optical density at a wavelength 600nm
PCR Polymerase chain reaction
\( Pfu \) pol \textit{Pyrococcus furiosus} polymerase
pH Power of hydrogen \((-\log_{10}[\text{H}^+]\))
Phage Bacteriophage
RNA Ribonucleic acid
SDS Sodium dodecyl sulphate
<table>
<thead>
<tr>
<th>Abbreviations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sqrt</td>
</tr>
<tr>
<td>SSB</td>
</tr>
<tr>
<td>ssDNA</td>
</tr>
<tr>
<td>T</td>
</tr>
<tr>
<td>\textit{Ta}q pol</td>
</tr>
<tr>
<td>Tm</td>
</tr>
<tr>
<td>Tris</td>
</tr>
<tr>
<td>UV</td>
</tr>
<tr>
<td>V</td>
</tr>
<tr>
<td>v/v</td>
</tr>
<tr>
<td>w/v</td>
</tr>
<tr>
<td>\mu</td>
</tr>
<tr>
<td>X-gal</td>
</tr>
<tr>
<td>°C</td>
</tr>
<tr>
<td>%</td>
</tr>
</tbody>
</table>
# Table of contents

CHAPTER I ......................................................................................................................... 1

INTRODUCTION ................................................................................................................... 1

1.1 GENOME INSTABILITY ............................................................................................... 2

1.1.1 Mutations and evolution ....................................................................................... 3

1.1.2 DNA palindromes ................................................................................................. 3

1.1.2.1 Palindrome-associated diseases ...................................................................... 4

1.1.2.2 SbeCD and Rad50/Mre11 ............................................................................... 5

1.1.2.3 Palindromes instability ..................................................................................... 6

1.1.3 Trinucleotide repeats .............................................................................................. 8

1.1.3.1 TNRs associated diseases .............................................................................. 8

1.1.3.2 Mechanisms of instability ............................................................................. 10

1.1.3.2.1 DNA replication models ........................................................................... 10

1.1.3.2.2 DNA recombination and repair models ..................................................... 14

1.2. THE SINGLE STRANDED DNA BINDING PROTEIN OF *E. coli* ........................ 15

1.2.1 Structure of *E. coli* SSB .................................................................................... 16

1.2.2 *E. coli* SSB ssDNA binding features .................................................................. 18

1.2.3 SSB in DNA replication, recombination and repair ........................................... 20

1.2.3.1 DNA Replication ............................................................................................. 20

1.2.3.2 DNA Recombination ....................................................................................... 21

1.2.3.3 DNA Repair ..................................................................................................... 22

1.2.4 SSB protein and TNR instability ......................................................................... 23

1.2.5 SSB interactions .................................................................................................... 24

1.2.5.1 Recombination and repair: RecO, RecJ, RecQ and RuvAB .............................. 24

1.2.5.2 Stalled replication forks: PriA, RecQ and RecG ............................................. 26

1.2.5.3 DNA polymerase III ε subunit, polymerase V and TopB .............................. 27

1.2.5.4 Uracil Excision Repair: UDG ........................................................................ 27

1.2.6 SSB homologues .................................................................................................. 28

1.3. THE SOS RESPONSE .............................................................................................. 30

1.4. THE ARABINOSE OPERON OF *E. coli* ................................................................. 31
# Table of Contents

CHAPTER II .................................................................................................................. 33

MATERIALS AND METHODS....................................................................................... 33

2.1 Bacterial methods.................................................................................................... 34

2.1.1 Transformation protocol .................................................................................. 34

2.1.2 P1 lysates preparation ...................................................................................... 34

2.1.3 P1 transduction ................................................................................................ 35

2.1.4 pTOF24 gene replacement technique .............................................................. 35

2.1.5 Gentamycin resistance gene insertion .............................................................. 36

2.1.6 ΔsbcDC mutants construction ......................................................................... 36

2.1.7 Spot test ............................................................................................................ 36

2.1.8 Growth curves .................................................................................................. 36

2.1.9 Viability assay .................................................................................................. 37

2.1.10 U.V. sensitivity test ......................................................................................... 37

2.2 DNA methods ........................................................................................................ 37

2.2.1 Genomic DNA extraction ................................................................................ 37

2.2.2 Crossover PCR ................................................................................................ 38

2.2.3 Colony PCR ..................................................................................................... 39

2.2.4 Plasmids ........................................................................................................... 39

2.2.4.1 pTOF24 ........................................................................................................ 39

2.2.4.2 pTOF24/pwr-A-ssbCO2 .............................................................................. 40

2.2.4.3 pAM34/ssb .................................................................................................. 41

2.2.4.4 pLacD2/pal105 ........................................................................................... 42

2.2.4.5 pGB150 ....................................................................................................... 44

2.3 Western blot analysis ............................................................................................ 45

2.3.1 Cell lysate preparation ..................................................................................... 45

2.3.2 Polyacrylamide gel electrophoresis .................................................................. 45

2.4 Microscopy ............................................................................................................. 46

2.4.1 DAPI staining .................................................................................................. 46

2.4.2 SOS response activation assessment ............................................................... 47

2.5 Instability assays ................................................................................................ 47

2.5.1 105bp palindrome deletion assay .................................................................... 48
Table of Contents

2.5.2 246bp palindrome associated viability .................................................. 49
2.5.3 CAG CTG repeat instability assay.......................................................... 49
  2.5.3.1 Ex-test PCR .................................................................................. 50
  2.5.3.2 Gene Mapper analysis .................................................................... 51
2.6 E. coli BACTERIAL STRAINS ................................................................. 53
  2.6.1 Background strains ......................................................................... 53
  2.6.2 Engineered strains ........................................................................... 53
2.7 DNA PRIMERS ...................................................................................... 55
2.8 MEDIA .................................................................................................... 55
  2.8.1 LB broth ......................................................................................... 55
  2.8.2 LB agar ........................................................................................... 55
  2.8.3 LC agar ............................................................................................ 56
  2.8.4 LC top agar ..................................................................................... 56
  2.8.5 BBL agar ........................................................................................ 56
  2.8.6 BBL top agar .................................................................................. 56
  2.8.7 Phage buffer ................................................................................... 56
  2.8.8 TM buffer ....................................................................................... 57
2.9 DRUGS AND SUGARS ......................................................................... 57

CHAPTER III ............................................................................................... 58

EFFECTS OF SSB PROTEIN DEPLETION .............................................. 58

3.1 INTRODUCTION .................................................................................... 59
3.2 RESULTS ............................................................................................... 60
  3.2.1 Viability test on plates ...................................................................... 60
  3.2.2 Growth and viability in liquid medium .......................................... 62
  3.2.3 SSB expression level assessment .................................................. 64
  3.2.4 Cell density and SSB expression coordination ............................ 67
  3.2.5 Growth curves in glucose ............................................................... 68
  3.2.6 U.V. sensitivity test ....................................................................... 72
3.3 DISCUSSION AND CONCLUSIONS ....................................................... 74
  3.3.1 SSB expression and cell growth .................................................... 74
  3.3.2 Cell density and growth ................................................................. 75
Table of Contents

3.3.3 Growth in glucose and arabinose ................................................................. 77
3.3.4 SSB expression and U.V. sensitivity ............................................................... 78

CHAPTER IV ............................................................................................................. 80
CELL DIVISION, NUCLEOID MORPHOLOGY AND SOS RESPONSE ........... 80

4.1 INTRODUCTION ................................................................................................. 81
4.1.1 SSB depletion, cell division and nucleoid morphology ..................................... 81
4.1.2 SSB depletion and the SOS response ............................................................... 81
4.2 RESULTS .............................................................................................................. 82
4.2.1 SSB depletion impairs cell division and changes nucleoid morphology ............. 82
4.2.2 SSB protein depletion causes activation of the SOS response ......................... 84
4.3 DISCUSSION AND CONCLUSIONS ................................................................ 89

CHAPTER V ............................................................................................................. 90
ISOLATION AND ANALYSIS OF SUPPRESSOR STRAINS .............................. 90

5.1 INTRODUCTION ................................................................................................. 91
5.2 RESULTS ............................................................................................................. 91
5.2.1 Suppressor clones .......................................................................................... 91
5.2.1.1 Isolation of suppressor clones .................................................................. 91
5.2.1.2 Analysis of suppressor clones .................................................................. 95
5.2.2 Analysis of the DL3622 suppressor strain ....................................................... 95
5.2.2.1 Growth of the DL3622 suppressor strain .................................................. 95
5.2.2.2 U.V sensitivity of DL3622 suppressor strain ............................................ 98
5.2.2.3 Genome sequencing of the DL3622 suppressor strain ............................... 99
5.3 DISCUSSION AND CONCLUSIONS ................................................................. 101

CHAPTER VI .......................................................................................................... 104
SSB AND PALINDROME INSTABILITY ............................................................ 104

6.1 INTRODUCTION ............................................................................................... 105
Table of Contents

6.1.1 105bp palindrome instability assay ................................................................. 105
6.1.2 246bp palindrome instability assay ..................................................................... 106
6.2 RESULTS ........................................................................................................... 107
  6.2.1 SSB depletion lowers the stability of a 105bp imperfect palindrome flanked by short direct repeats ................................................................. 107
  6.2.2 SSB overexpression does not improve the viability of recA strains carrying a 246bp imperfect palindrome ................................................................. 109
6.3 DISCUSSION AND CONCLUSIONS .................................................................. 112

CHAPTER VII................................................................................................. 114

SSB AND CAG CTG INSTABILITY ................................................................... 114

7.1 INTRODUCTION .............................................................................................. 115
  7.1.1 CAG CTG repeat instability ............................................................................ 115
  7.1.2 SSB and CAG CTG repeat instability ............................................................. 116
  7.1.3 SbcCD and CAG CTG repeat instability ......................................................... 117
  7.1.4 Aim of the work ............................................................................................. 117
7.2 RESULTS ......................................................................................................... 118
  7.2.1 SSB protein level increases in cells carrying the pAM34/ssb plasmid ......... 118
  7.2.2 Overexpression of SSB protein stabilises the (CAG)_{25} repeat array but does not influence the stability of the (CTG)_{37} repeat array both in sbcDC^+ and sbcDC^- backgrounds ........................................ 120
  7.2.3 (CAG)_{25} repeat array deletion size distribution profile is affected by SSB overexpression and by the lack of the sbcCD nuclease .......................................................... 124
  7.2.4 Lack of sbcCD changes the deletion size distribution profile in (CTG)_{37} strains .......................................................... 127
  7.2.5 The proportion of the number of expansions is not influenced by SSB overexpression both in sbcDC^+ and sbcDC^- backgrounds .......................................................... 129
7.3 DISCUSSION .................................................................................................. 131
  7.3.1 SSB overexpression stabilises CAG CTG repeat arrays in an orientation dependent manner in the E. coli chromosome .......................................................... 131
  7.3.2 SSB overexpression may preferentially impede the formation of big CTG hairpins on the lagging strand template by binding onto ssDNA ........................................ 132
  7.3.3 SbcCD Nuclease May Stabilise CAG DNA Extrusions ................................ 133
7.4 CONCLUDING REMARKS .............................................................................. 134
Table of Contents

CHAPTER VIII .................................................................................................................. 136
CONCLUSIONS ................................................................................................................. 136

8.1 SUMMARY .................................................................................................................... 137

REFERENCES .................................................................................................................... 143
List of figures

Figure 1.1-DNA secondary structures 3
Figure 1.2-Model for TNR instability occurring during replication 12
Figure 1.3-Template push model for TNR instability 14
Figure 1.4-Schematic SSB protein structure 17
Figure 1.5-OB-fold structure 18
Figure 1.6-Schematic representation of the *E. coli* arabinose operon 33

Figure 2.1-Crossover PCR technique 39
Figure 2.2-pTOF24 plasmid map 40
Figure 2.3-Crossover PCR strategy 41
Figure 2.4-pTOF24/*urrA*/ssbCO2 plasmid 42
Figure 2.5-pAM34/ssb plasmid 43
Figure 2.6-105bp palindrome construct 49
Figure 2.7-Gene Mapper 4.0 graphs 53

Figure 3.1-DL1582 and DL2844 strains; ssb locus in the chromosome 61
Figure 3.2-DL1582 and DL2844 strains viability on plate 63
Figure 3.3-DL1582 and DL2844 strains; growth curve 64
Figure 3.4-DL2844 strain viability in liquid culture 65
Figure 3.5-Western blot analysis 67
Figure 3.6-Cell density and growth 69
Figure 3.7-DL2844 (SSB<sub>md</sub>) growth curve of cells grown in the presence of glc and ara 70
Figure 3.8-DL2844 (SSB<sub>md</sub>) viability of cells grown in the presence of glc and ara 71
Figure 3.9-Western blot of cells grown in the presence of glc and ara 72
Figure 3.10-U.V sensitivity test 74

Figure 4.1-DAPI staining 84
Figure 4.2-SOS response activation control 86
Figure 4.3-SOS response activation of DL3121 and DL2984 87
List of Figures

Figure 4.4-SOS response activation and cell length of DL3121 and DL2984 88
Figure 4.5-Cell length and fluorescence of DL2984 grown in 10⁻⁴ % ara 89

Figure 5.1-Colony phenotype 93
Figure 5.2-1% arabinose streaks phenotype 93
Figure 5.3-Phenotype of suppressor strains streaked on plates 95
Figure 5.4-DL3622 suppressor strain U.V. sensitivity spot test 97
Figure 5.5-Growth and viability in liquid medium 98
Figure 5.6-DL3622 U.V. sensitivity test 100

Figure 6.1-105 bp imperfect palindrome construct 108
Figure 6.2-Plating of a blue sectored colony 110
Figure 6.3-Palindrome deletion assay results 111
Figure 6.4-Viability of recA strains containing the 246bp palindrome 112
Figure 6.5-Strains DL2075 and DL3750 growth curve 113
Figure 6.6-Viability in liquid culture 113

Figure 7.1-SSB level assessment in overnight cultures 120
Figure 7.2-SSB level assessment during instability assay 122
Figure 7.3-(CAG)₁₅ instability assay graphs 124
Figure 7.4-(CTG)₁₃₇ instability assay graphs 125
Figure 7.5-(CAG)₁₅ deletion size distribution profile curves 128
Figure 7.6-(CTG)₁₃₇ deletion size distribution profiles 130
Figure 7.7-Proportion of expansions 132
Figure 7.8-SbcCD stabilisation of CAG extrusions model 136
List of tables

Table 2.1- *E. coli* background strains 54
Table 2.2- *E. coli* engineered strains 54
Table 2.3-DNA primers 56
Table 2.4-Drugs and sugars 58

Table 5.1-Candidate target genes and promoters 96
Table 5.2-Point mutations 102

Table 7.1-Best fit lines slopes 126
Chapter I

Introduction
1.1 Genome Instability

Genome stability has great importance in maintaining cell viability and optimal functionality of cellular processes. Loss of genome stability can therefore lead to cell death in the simplest organisms and to deregulation of the cell proliferation machinery in higher organisms, potentially causing the development of cancer or morbid states (Cadman and McGlynn 2004). Genome replication, recombination and repair are precise processes which participate in preserving and maintaining the integrity of the genome. However, errors during replication, recombination and repair can lead to genome instability. Sequences that can allow the formation of DNA secondary structures are particularly likely to be a source of instability. Palindromic sequences and trinucleotide repeat (TNRs) sequences can be found in the human genome and can allow the formation of particular DNA structures such as slipped strands, hairpins, cruciform structures and quadruplex DNA (Fig.1.1). The formation of these structures is mediated by the pairing between complementary bases.

![Image](Hairpin structure)

![Image](Slipped strand structure)

![Image](Cruciform structure)

![Image](CGG quadruplex)

Figure 1.1 - DNA secondary structures. These particular structures can be formed by palindromic DNA sequences (A and B) or trinucleotide repeats (B, C and D). In the case of trinucleotide repeats an imperfect pairing is observed with two bases out of three that interact. The formation of quadruplex has been observed in vitro only while the other structures have been detected in vivo (Sinden, Potaman et al. 2002).
1.1.1 Mutations and evolution

Surprisingly, genome instability is not always deleterious. Under certain circumstances it can in fact be used by microorganisms as a survival mechanism. Studies on different bacteria populations show how mutator strains, which present a high mutation rate, are relatively frequently found in nature or in laboratory isolated strains (LeClerc, Li et al. 1996; Matic, Radman et al. 1997; Oliver, Canton et al. 2000; Bjorkholm, Sjolund et al. 2001; Richardson, Yu et al. 2002; Prunier, Malbruny et al. 2003; del Campo, Morosini et al. 2005). Having a mutator phenotype may confer an advantage to a strain enduring a selective pressure. Most of the isolated mutator strains have defects in the mismatch repair system genes, encoding proteins involved in DNA repair and responsible for accurate DNA replication. These cells can adapt faster than normal cells to selective conditions. The mutator phenotype can however become detrimental to cells that have adapted to the new conditions; reversion to a non mutator phenotype before the quantity of deleterious mutations becomes too high is in fact necessary for survival. A balance between genetic stability and genetic instability is therefore required for microorganisms adaptation (Denamur and Matic 2006). Higher organisms do not generally benefit from a mutator phenotype as mutations and genetic instability can give rise to diseases and disorders such as cancer.

1.1.2 DNA palindromes

Perfect DNA palindromes are formed by two adjacent inverted-repeat DNA sequences. Long DNA palindromes have been proven to be very unstable in vivo and can be selectively removed from the genome of E. coli (Leach 1994). Their instability is potentially due to the fact that they can form DNA secondary structures by self base pairing.
Cruciform structures can form within dsDNA by extrusion and intramolecular base pairing. The first step in the formation of cruciform structures is believed to be the melting of the dsDNA followed by annealing within the palindrome sequence; the base pairing can then be extended by branch migration (Sullivan and Lilley 1986). Perfect palindromes can form cruciforms more easily than interrupted ones. This is due to a lack of symmetry that characterises interrupted palindromes in which a short non-palindromic sequence separates the two inverted repeats (Zheng, Koehl et al. 1991).

DNA hairpins can form within ssDNA stretches and their formation is therefore favoured during replication, recombination and repair processes. In particular, during DNA replication, the formation of hairpins is thought to occur mainly on the lagging strand template as ssDNA stretches are present between the Okazaki fragments (Trinh and Sinden 1991; Pinder, Blake et al. 1998).

Deletion mutations in prokaryotic cells are often associated with short direct repeats, and inverted repeats flanked by short direct repeats are associated with specific deletion events in bacteriophages, E. coli and yeast. The presence of DNA secondary structure could lead to such deletions, by inducing replication errors, such as transient pausing of the replication machinery and DNA slippage (Weston-Hafer and Berg 1991).

1.1.2.1 Palindrome-associated diseases

The presence of DNA palindromes in the genomes of S. cerevisiae and E. coli has been associated with increased homologous recombination frequency caused by the formation of DNA DSBs (Nag and Kurst 1997; Nasar, Jankowski et al. 2000; Eijkelenboom, Blackwood et al. 2008). Similarly, chromosomal translocations in humans have been associated with the presence of DNA palindromes near the breakpoint cluster regions in the chromosomes. The
Chapter I - Introduction

t(11;22)(q23;q11) non–Robertsonian translocation is one of the best characterised examples. An AT-rich palindrome present at 22q11 and flanked by short GC-rich regions can form very stable cruciform structures. Such structures can be cleaved by specific enzymes leading to formation of DSBs which are believed to be at the origin of the translocation process (Kurahashi, Inagaki et al. 2004; Kurahashi, Inagaki et al. 2006).

DNA palindromes are also involved in gene amplification whereby a cell can increase the copy number of a gene or a genomic region. This process may associate with neoplasias (e.g. myc gene) and drug resistance (Butler, Yasuda et al. 1996; Tanaka, Tapscott et al. 2002; Rattray, Shafer et al. 2005; Tanaka, Bergstrom et al. 2005).

Many mutations are associated with the presence of repeated sequences in the human genome. For example, imperfect inverted repeats flanked by short direct repeats are the cause of deletions occurring in α-globulin and lactate dehydrogenase genes (Efstratiadis, Posakony et al. 1980; Maekawa, Sudo et al. 1990).

1.1.2.2 SbcCD and Rad50/Mre11

SbcC and SbcD proteins of E. coli and their eukaryotic homologues Rad50 and Mre11 respectively, are implicated in the metabolism of DNA ends and possess DNA binding and nucleolytic activities. In vivo these proteins associate to form a large complex of about 1.2MDa that has ATP-dependent double-strand DNA exonuclease activity and ATP-independent single-strand DNA endonuclease activity (Connelly and Leach 1996; Connelly, de Leau et al. 1997). Interestingly, palindromic sequences cloned in plasmid or bacteriophage vectors are not propagated in wild type strains of E. coli, but can be propagated in sbcC or sbcD deficient strains. The SbcCD complex can in fact cleave DNA hairpins in a sequence
Chapter I - Introduction

independent manner at the 5’ end of the loop, leaving a 5’ phosphate and a 3’ hydroxyl end (Connelly, Kirkham et al. 1998).

SbcC is related to the SMC (Structural Maintenance of Chromosomes) family of proteins which contain Walker A and Walker B nucleotide binding motifs. Additionally, a long α-helical region, predicted to form two coiled coil domains, is placed between the two nucleotide binding motifs. SbcC possesses a sequence similarity to the human, mouse and yeast Rad50 proteins, and to the gp46 protein of T4 bacteriophage. These proteins are involved in genetic recombination and in the processing of DNA double strand breaks.

SbcD shows a sequence similarity with human, mouse and yeast Mre11 proteins and gp47 of T4 bacteriophage. Both SbcD and Mre11 contain the conserved sequence DXH(X)25GDXXD(X)25GNHD/E and belong to the phosphoesterase family (Sharples and Leach 1995). Mre11 is essential to allow a normal cell proliferation in embryonic mouse cell lines (Xiao and Weaver 1997). Rad50 and Mre11 interact in yeast and human cells and their T4 bacteriophage homologues gp46 and gp47 proteins are believed to be essential for DNA recombination and replication (Petrini, Walsh et al. 1995).

1.1.2.3 Palindromes instability

Palindrome instability or host inviability have been observed after the introduction of long DNA palindromes in E. coli. A length of 150-200 bp is enough to confer host inviability or complete or partial loss of the palindrome (Collins 1981; Warren and Green 1985). In vitro evidences showed that sequences prone to give rise to DNA secondary structures can block DNA replication (Sherman and Gefter 1976; Challberg and Englund 1979; Kaguni and Clayton 1982; Weaver and DePamphilis 1982; Kang, Ohshima et al. 1995). Recently, alu sequence inverted repeats have been shown to stall replication forks in vivo in bacteria, yeast
and mammalian cells (Voineagu, Narayanan et al. 2008). The paper proposed DNA hairpins forming on the lagging strand template during replication are responsible for the replication block.

In *E. coli* the SbcCD protein complex can cleave DNA hairpins (Connelly, Kirkham et al. 1998). SbcCD can recognise and cleave cruciform and hairpin structures leaving a DNA double strand break (DSB) that can be repaired by homologous recombination or single-strand annealing followed by ligation (Bzymek and Lovett 2001). A conservative and replication dependent model was been proposed by Leach (Leach, Okely et al. 1997). SbcCD may cleave hairpins forming on the lagging strand template during replication, leading to the creation of a DSB. The break can then be repaired by homologous recombination using the sister chromosome as template resulting in conservation of the palindrome. Evidence for this mechanism was found, demonstrating that SbcCD induction in strains containing a 246bp imperfect palindrome leads to the formation of DSB that can be repaired by homologous recombination. The formation of breaks was shown to be replication dependent (Cromie, Millar et al. 2000; Eykelenboom, Blackwood et al. 2008).

A replication independent model that involves resolution of cruciform structures has also been proposed (Bzymek and Lovett 2001). Following the formation of a cruciform, SbcCD could cleave the hairpins leaving a double strand break. One of the strands on each side of the break would get resectioned, annealed to the remaining strands and finally ligated. This would lead to the formation of a deletion product resulting in palindrome loss (Bzymek and Lovett 2001).

Palindromes can also be deleted in a SbcCD independent manner, via replication slippage; this occurs when the palindrome is flanked by short direct repeats. The formation of DNA hairpins during replication may bring the direct repeats close together creating a barrier that may stop replication. The consequent dissociation of the DNA polymerase from the
replication machinery may facilitate template misalignment. The polymerase can then replicate across the palindrome leading to a deletion (Leach 1994; Pinder, Blake et al. 1998; Bzymek and Lovett 2001).

1.1.3 Trinucleotide repeats

Repeated DNA sequences are present in great number in the human genome, both in expressed and unexpressed DNA tracts and have been demonstrated to be hot spots for recombination and random integration (Jeffreys, Holloway et al. 2004; Yant, Wu et al. 2005; Durkin and Glover 2007). A particular kind of repeat is represented by trinucleotide repeats (TNR), sequences composed of three nucleotides repeated in tandem. The expansion of tracts of TNR sequences underlies several neurological and neurodegenerative diseases that exhibit anticipation: penetration and severity of the disease increases with an observed decrease of the age of onset in successive generations (Sinden, Potaman et al. 2002). In particular, the instability of CTG, CGG and GAA triplet repeats is the molecular cause of at least 14 human diseases amongst which are Huntington disease, fragile-X syndrome, myotonic dystrophy type I, spinocerebellar ataxias and Friedreich’s ataxia (Sutherland and Richards 1995; Rosche, Jaworski et al. 1996; Cummings and Zoghbi 2000).

1.1.3.1 TNRs associated diseases

The cause of TNRs associated diseases can be different depending on the repeats position in the genome. The repeat tract can be found in coding or in non-coding regions.

The only trinucleotide stretches found in gene coding regions are GCG and CAG. GCG expansion is responsible for congenital malformation and/or retardation illnesses classified as polyalanine disorders (Amiel, Trochet et al. 2004). CAG expansion is responsible for nine
neurodegenerative illnesses, classified as polyglutamine disorders. These diseases are all chronic, progressive and present a similar onset and progression rate. They include Huntigton’s disease (HD), Dentatorubralpallidoluysian athrophy (DRPLA), spinocerebellar Ataxias (SCA) 1,2,3,6,7,17 and Kennedy’s disease. CAG codes for the amino acid glutamine and the expansion of the CAG stretch leads to a gain of function of the protein that becomes toxic for the cell. Proteins presenting glutamine tags can in fact form aggregates that often accumulate in the nuclei of affected cells. The primary pathogenic importance of those aggregates is however still not certain. Some studies on model systems and investigations on patients dissociated nuclear inclusions from toxicity (Paulson, Perez et al. 1997; Saudou, Finkbeiner et al. 1998; Fischbeck 2001). The toxicity seems to be bound to various factors. The mutated proteins are believed to increase sensitivity to apoptotic stress and to be able to bind other polyglutamine proteins, sequestering them from their normal function (Preisinger, Jordan et al. 1999). Examples are the Hsp40 and Hsp70 chaperones, involved in the clearance of misfolded proteins (Cummings, Mancini et al. 1998) and the signal transduction mediator CREB binding protein (CBP) that have been demonstrated to co-localise with polyglutamine mutant proteins in cell culture. CBP depletion leads to cell death (McCarmel, Taylor et al. 2000).

Diseases caused by TNR expansion in non-coding regions of the genome are usually associated with downregulation of the gene expression level. CGG and GCC repeats cause Fragile-X Syndrome and Fragile-X Mental Retardation respectively. The CpG dinucleotides can be targeted by DNA methylases and hypermethylation of the promoter region of the gene leads to silencing and subsequent loss of the gene product (Eberhart and Warren 1996). In Friedrich’s Ataxia an expansion of the GAA trinucleotide, located in the first intron of the frataxin gene, can cause changes in the chromatin packaging of the DNA, silencing frataxin expression (Saveliev, Everett et al. 2003).
1.1.3.2 Mechanisms of instability

Instability of TNR array length has been observed both in replicating and non-replicating cells, yet replication seems to be the prime mechanism causing instability. Errors occurring during replication, recombination and repair processes have been proposed to account for the instability of TNRs but controversy remains in the field (Sinden, Potaman et al. 2002; Hashem, Rosche et al. 2004). A bias towards expansions has been observed in germ and somatic cells (Kovtun and McMurray 2001; Kovtun, Liu et al. 2007) while deletions seem to occur more frequently in rapidly dividing organisms such as bacteria and yeast. A recent paper proposed that the length of a TNR stretch in mammalian cells can be controlled during the early embryo stage, in which cells are rapidly dividing, favouring deletion rather than expansion events (Kovtun, Thornhill et al. 2004).

Instability caused by TNRs has been associated to DNA replication, recombination and repair processes. An overview of proposed models, accounting for instability, will be presented in the next section.

1.1.3.2.1 DNA replication models

One of the first hypotheses brought forward to explain contractions and expansions of a TNR tract was based on the possibility of template misalignment and DNA strand slippage during replication (Kunkel 1993). This model is consistent with the evidence that TNR stretches can form DNA secondary structures. During DNA replication, secondary structures may be more prone to form on the lagging strand template, which presents ssDNA stretches that can easily undergo conformational changes. In several organisms, a bias towards instability has been observed when the TNR orientation that is more prone to form stable secondary structures was on the lagging strand template (Freudenreich,
Stavenhagen et al. 1997; Miret, Pessoa-Brandao et al. 1998; Cleary, Nichol et al. 2002; Zahra, Blackwood et al. 2007).

The mechanism leading to TNR instability during replication seems to be triggered by problems encountered by the DNA polymerase at the replication fork. It has been shown that T7 DNA polymerases can stall at replication forks in vitro (Kang, Ohshima et al. 1995) and that T4 and T7 DNA polymerases encounter an impediment while trying to replicate through TNR stretches (Delagouette, Goellner et al. 2008). A replication block stimulated by repeats in vivo in a plasmid based system was also observed (Samadashwily, Raca et al. 1997).

A model for TNR instability occurring at stalled replication forks has recently been proposed by Mirkin (Mirkin 2007). The model accounts for both deletions and expansions and is consistent with hairpin formation, during replication, on the lagging strand template. The TNR tract gets deleted when the replication machinery skips the hairpin formed on the lagging strand template. On the other hand, fork reversal and replication restart would lead to the formation of a hairpin on the nascent leading strand causing expansion after a second round of replication (Fig.1.2).
Figure 1.2 - Model for TNR instability occurring during replication (Mirkin 2007). A) The formation of hairpins either on the leading or on the lagging strand template can lead to expansion (left panel) or contraction (right panel) of the TNR tract. B) Model for repeat instability during replication. Contractions (upper panel) occur when the hairpin forms on the lagging strand template and the replication machinery jumps across the hairpin. Expansions (lower panel) occur when the hairpin forms on the nascent leading strand after replication fork reversal.

Recently, the role of the *E. coli* SSB protein on polymerase progression through CAG-CTG repeat arrays was investigated by an *in vitro* primer extension assay and a new orientation-dependent instability model was proposed (Delagoutte, Goellner et al. 2008). The authors found a slowdown in polymerase progression while copying through CAG repeat tracts but not through CTG repeat tracts. Adding SSB to the reaction seemed to facilitate polymerase progression through CAG repeat tracts. The authors proposed a “template push” model which is relying on the formation of hairpins on the leading strand template rather than on the lagging strand template. During replication of a CAG-CTG repeat array (with CAG repeats present on the leading strand template), a differential speed
between leading and lagging strand polymerase would be generated. To avoid uncoupling the
leading strand polymerase would pull towards the helicase, producing a CAG hairpin on the
leading strand template. This model would explain orientation-dependent instability through
a mechanism which is not relying on hairpin stability issues (Fig.1.3) (Delagoutte, Goellner et
al. 2008).

Figure 1.3 - Template push model for TNR instability (Delagoutte, Goellner et al. 2008). The model
shows a replication fork carrying CAG repeat array on the leading strand template. A) Helicase and polymerases
are tightly coupled during DNA replication. B) When the polymerase enters the CAG tract replication slows
down causing uncoupling. C) The polymerase has to bypass a segment of TNRs to couple again with the
helicase, leaving behind an unreplicated tract. After a second round of replication one of the two strands would
give rise to a deleted product.
1.1.3.2.2 DNA recombination and repair models

As well as DNA replication, DNA recombination may be involved in TNR instability. CAG-CTG tracts can stimulate recombination in bacteria, yeast and mammalian cells (Freudenreich, Kantrow et al. 1998; Jakupciak and Wells 1999; Meservy, Sargent et al. 2003). In mammalian cells, reciprocal expansions and contractions of repeat sequences could be triggered by unequal crossing over between homologous chromosomes.

In *E. coli*, the role of proteins involved in homologous recombination in TNR instability have been controversial. Experiments showed that RecA and RecB mediated DSB repair lowers the instability of CAG-CTG in TNR tracts in the *E. coli* chromosome (Hebert, Spitz et al. 2004). However, other studies showed increased instability of repeat tracts in a plasmid-based system, in RecA and RecB deficient strains (Jakupciak and Wells 2000; Hashem, Rosche et al. 2004). Finally, in a recent study, no effect of *recA recB* and *ref* mutations was observed on CAG-CTG repeat instability in the *E. coli* chromosome (Zahra, Blackwood et al. 2007).

Mismatch repair (MMR) proteins can play an indirect role in promoting instability while DNA damage is fixed (Kennedy and Shelbourne 2000). MMR proteins can recognise base pair mismatches and correct them by cleavage and gap filling. As TNRs can form mismatched hairpins, the prediction is that MMR proteins would help stabilising TNR sequences by eliminating mismatches. Studies carried out in mice demonstrated that MMR defective cells had a decreased frequency of repeat expansion and increased frequency of contractions (Manley, Shirley et al. 1999; Savouret, Brisson et al. 2003). A proposed model, accounting for these findings, suggests that the MSH1/MSH2 dimer (MutS protein homologues 2 and 3) may bind to the TNR hairpin with high affinity and stabilise the hairpin, promoting the permanence of a slipped strand structure. In bacteria, controversial results have been found and a role for the MMR system has not been clearly outlined yet.
Findings support the hypothesis that MMR increases TNR instability in _E. coli_ (Jaworski, Rosche et al. 1995; Schmidt, Abbott et al. 2000) whereas other studies argue against this hypothesis (Schumacher, Fuchs et al. 1998; Schmidt, Abbott et al. 2000). Recently, Zahra and collaborators showed that a mutant _E. coli_ strain lacking MutS protein showed a marginal increase in TNR instability, suggesting a stabilising role for the protein (Zahra, Blackwood et al. 2007).

### 1.2. The Single Stranded DNA Binding Protein of _E. coli_

Every domain of life and viruses are characterised by the presence of proteins that can bind single stranded DNA. This enables the cell to sequester and protect single stranded DNA stretches until the double helix has been reformed (Kowalczykowski, Dixon et al. 1994). Differences in protein structure and binding modalities have however been found between eukaryotes and prokaryotes; only the DNA binding domain seems to be conserved among all characterised single-stranded DNA (SSB) binding proteins (Kerr, Wadsworth et al. 2003). Interestingly, homotetrameric SSB proteins have been found both in prokaryotes and eukaryotic mitochondria, suggesting that they are evolutionary conserved (Tomaska, Nosek et al. 2001).

The single-stranded DNA binding protein of _E. coli_ is an essential protein. SSB plays a central role in processes related to DNA metabolism such as replication, repair and homologous recombination. The roles of SSB are mediated by its ability to bind DNA in a sequence independent manner and with high specificity for single versus double stranded DNA (Reddy, Guhan et al. 2001). SSB can modulate the function of DNA processing enzymes via protein-protein interaction or by controlling DNA accessibility (Meyer and
Laine 1990). Several mutations in the DNA sequence of the *ssb* gene have been described; they are localised either in the first part of the N-terminal domain or towards the very end of the sequence. The phenotypes conferred by these mutations are diverse and include temperature-sensitive DNA replication, increased sensitivity to UV radiation and alkylating agents and a defect in RecA protein activation.

### 1.2.1 Structure of *E. coli* SSB

The *E. coli* SSB protein contains 177 amino acids with an estimated Mₐ of 18,843 and its structure is composed by approximately 20% α-helix, 20% β-sheets and 60% random coil. In solution SSB monomers form very stable homotetramers. The sequence can be divided in two functional blocks, N and C-terminal regions, separated by a spacer (Fig.1.4) (Chou and Fasman 1978; Sancar, Williams et al. 1981; Meyer and Laine 1990).

![Schematic SSB protein structure](image)

**Figure 1.4 - Schematic SSB protein structure.** The DNA binding domain and the subunit-subunit interaction domain are shown in blue, the spacer domain is shown in green and the C-terminal tail in yellow.

The N-terminal region comprises about 120 amino acids and represents two thirds of the entire SSB sequence. The region is rich in charged residues (especially positively charged), and highly structured in α-helices and β-sheets. The subunit-subunit interaction domain and the DNA binding region can be found within the first 115 amino acids of the protein (Raghunathan, Ricard et al. 1997). The subunit-subunit interaction domain was localised...
when studying the ssb-1 mutant isoform, in which the histidine in position 55 is replaced by a tyrosine. This mutant isoform shows a loss of ability to form tetramers in vitro (Williams, Murphy et al. 1984). In vivo, the lack of activity of this mutated form generates a temperature sensitive phenotype (Meyer, Glassberg et al. 1979). The DNA binding domain is located in the same region of the protein and the most important residues involved in DNA binding are Trp40, Trp54, Trp88 and Phe60 (Raghunathan, Kozlov et al. 2000). The DNA binding domain has been crystallised showing that two monomers contribute three β-strands each to form a single six-stranded β-sheet that is present in the dimer. Two dimer-dimer interfaces were observed within the crystal (Raghunathan, Ricard et al. 1997). The DNA binding domain has a particular folding structure called OB-fold (oligonucleotide/oligosaccharide binding). The structure of the OB-fold (Fig.1.5) is characterised by the presence of five strands of β-sheet coiled to form a β-barrel. The β-barrel is capped by an α-helix located between the third and the fourth strands (Murzin 1993).

![Figure 1.5 - OB-fold structure](image)

Figure 1.5 - OB-fold structure. The five β-sheet are represented by the arrows, and the α-helix is represented by the cylinder. Unstructured loops are marked by L (Murzin 1993).

The C-terminal region covers the final third of the SSB sequence and can be divided into two different domains. The first domain is unstructured and composed of uncharged amino acids, prevalently glycine, proline, glutamine and asparagine (Meyer and Laine 1990). It spans
about 50 amino acids in length and is not required for DNA-binding \textit{in vitro} (Curth, Genschel et al. 1996). The C-terminal tail (10 amino acids long) constitutes the final part of the protein. Its sequence is highly conserved among prokaryotic SSB proteins. The presence of four aspartate residues confers to this portion of the protein a negative charge that weakens the binding of SSB to the DNA. The unstructured region preceding this tail is needed as a spacer, keeping the negative charges away from the DNA and allowing SSB binding (Meyer and Laine 1990). Deletions of the C-terminal tail leads to a higher SSB/ssDNA binding efficiency \textit{in vitro}, but results in a complete loss of SSB activity \textit{in vivo}. This led to the hypothesis that the C-terminal tail could be involved in protein-protein interactions. A particular mutation, \textit{sib}-113, (Pro176Ser), has been located in the C-terminal tail. \textit{sib}-113 phenotypes present UV and temperature sensitive phenotype (Meyer and Laine 1990).

\subsection{1.2.2 \textit{E. coli} SSB ssDNA binding features}

The complex that SSB forms with ssDNA is very stable (Romer, Schomburg et al. 1984). The kinetics of the interaction between SSB and ssDNA has demonstrated that it is a two steps process. The first step is the slowest and involves the binding of SSB to the ssDNA stretch, while the second step is very quick and involves the wrapping of the ssDNA around the SSB protein (Kozlov and Lohman 2002).

Several experiments demonstrate that the ssDNA is wound around the SSB tetramers giving rise to histone-like nucleosome structures (Krauss, Sindermann et al. 1981; Meyer and Laine 1990). Depending on temperature, pH, protein-binding density and monovalent and divalent ion concentrations, several binding modes referred to as (SSB)$_n$ (where “$n$” is the number of nucleotides bound per each SSB tetramer) were observed (Bujalowski, Overman et al. 1988). The two most common binding modes are (SSB)$_{35}$ in which only two of the four
SSB subunits interact with the ssDNA, and (SSB)$_{65}$ in which the interaction is extended to the four monomers included in the tetramer. The in vitro role of the two binding modes observed in vitro has not been totally explained yet. It has been proposed that the (SSB)$_{35}$ binding mode is likely to work during DNA replication (Lohman, Bujalowski et al. 1988), whereas the (SSB)$_{65}$ one might be used during DNA recombination where it may stimulate RecA mediated strand exchange (Griffith, Harris et al. 1984).

The binding of SSB tetramers to ssDNA is characterised by a high level of positive cooperativity, implying that a SSB tetramer has higher affinity for a ssDNA stretch to which another SSB tetramer is already bound. Notably, two different kind of cooperativity have been observed for the two binding modes, (SSB)$_{35}$ and (SSB)$_{65}$. An unlimited cooperativity is observed while SSB is wrapped by 35nt. This binding mode allows the formation of long clusters of SSB along the DNA. On the other hand, a limited cooperativity is observed for the (SSB)$_{65}$ binding mode. In that situation the formation of long clusters of SSB bound to ssDNA is not permitted and SSB clusters appear to be limited to the formation of octamers (dimers of SSB tetramers)(Kozlov and Lohman 2002). Recent studies have focused on SSB-DNA binding at the molecular level. Apparently the wrapping of ssDNA around SSB tetramers involves a three steps mechanism. The speed of wrapping/unwrapping has been demonstrated to be much faster than the DNA replication machinery speed, consistent with SSB being dynamically present at the replication fork (Kuznetsov, Kozlov et al. 2006). A deletion of the C terminal portion of SSB, has been shown to shift the equilibrium towards the (SSB)$_{35}$ binding mode, suggesting its involvement in SSB subunit-subunit interaction (Roy, Kozlov et al. 2007).
1.2.3 SSB in DNA replication, recombination and repair

1.2.3.1 DNA Replication

In *E. coli* DNA replication initiation occurs at a specific AT-rich region called the origin of replication. This region is characterised by the presence of binding sites for DnaA, a protein which is involved in the first step of DNA replication. After binding, DnaA causes distortions on DNA that result in a local unwinding of the double helix. The association of DnaA to ssDNA recruits the DnaB/DnaC complex. DnaC is a loader that allows the binding of the DnaB helicase on the origin of replication and, after its release from the complex, promotes helicase activation. DnaB mediates the recruitment of the DnaG primase that leads to the synthesis of the RNA primer necessary for the action of DNA polymerase III holoenzyme and for replication to start (Kaguni 2006).

The importance of SSB in DNA replication has been widely demonstrated (Meyer, Glassberg et al. 1979). SSB is involved in the organisation and stabilization of DNA replication origins, primosome assembly, priming specificity and correct helix unwinding. Furthermore, SSB enhances the processivity of chain elongation and the loading of the β clamp onto DNA by a direct interaction with the χ subunit of the DNA polymerase III holoenzyme. This last interaction is also responsible for lowering the polymerase error rate (Meyer and Laine 1990; Kelman, Yuzhakov et al. 1998). Finally, the binding of SSB to ssDNA during replication can prevent the formation of secondary structures, which may cause the replication machinery to stop, potentially leading to misalignment of the newly synthesised strand with the template and appearance of mutations (Rosche, Jaworski et al. 1996).
1.2.3.2 DNA Recombination

Homologous recombination refers to the exchange of DNA sequences that are identical or nearly identical. In bacteria, this process is essential for DSBs repair and important to promote genetic material exchange.

After a DNA break occurs, the initiation step of homologuos recombination triggers the alignment of homologous sequences present on the broken and unbroken chromosomes. The alignment is followed by strand invasion that sees the pairing of short regions of the two different DNA molecules with the formation of a Holliday junction. Branch migration allows the effective exchange of DNA between the two molecules and finally, the resolution of the Holliday junction drives the separation of the two DNA molecules.

In *E. coli* homologous recombination following a DSB is mediated by various proteins. Firstly the RecBCD complex binds unwinds and degrades the DNA at a DSB generating a 3' ssDNA tail. Secondly the ssDNA tail is coated by RecA, a strand exchange protein that catalyses the pairing of homologous DNA sequences promoting strand invasion and strand exchange. Finally RuvA, RuvB and RuvC proteins mediate branch migration and Holliday junction resolution. After cleavage of the Holliday junction, the nicks left in the two DNA molecules are joined by the DNA ligase.

The role of SSB in homologous recombination is carried out by its ability to bind ssDNA thus melting secondary structures and facilitating the formation of RecA-ssDNA nucleoprotein filaments (Meyer and Laine 1990). SSB can also stabilise joint molecules by binding to the ssDNA displaced during the strand exchange step (Lavery and Kowalczykowski 1992). However, *in vitro*, SSB can inhibit the RecA mediated homologous pairing reaction if pre-incubated with the ssDNA. The reaction can be restored by the joint action of RecO and RecR proteins (Umez, Chi et al. 1993). RecO and RecR are involved in DNA gap repair and can bind to SSB coated ssDNA promoting RecA nucleation (Umez
and Kolodner 1994). Interestingly, *recA4441* and *recA4803* mutated versions of RecA protein, exhibit enhanced SSB displacement activity from ssDNA and do not require the action of RecO and RecR for the loading step (Wang, Chang et al. 1993).

1.2.3.3 DNA Repair

SSB mutants may present increased sensitivity to UV-radiation, suggesting an involvement of SSB in DNA repair processes (Glassberg, Meyer et al. 1979; Lieberman and Witkin 1983). SSB is involved in two kinds of DNA repair: the methyl-direct mismatch repair and the SOS response. Methyl-direct mismatch repair occurs on newly synthesised DNA molecules and can remove base pair mismatches. A MutS dimer can recognise and bind the mismatch recruiting MutL and the MutH nuclease. MutH creates a nick in the newly synthesised strand. This is followed by UvrD mediated DNA unwinding, digestion by Exonuclease IV, RecJ or Exonuclease I and re-synthesis by polymerase I. This results in the correction of the mismatch soon after its formation.

The SOS response pathway, as more extensively described in Chapter 1.3, is activated when DNA damage occurs and is initiated by the proteolytic cleavage of LexA, a repressor protein that binds SOS boxes present in the control region of about 40 genes in *E. coli*. The auto-proteolysis of LexA is mediated by ssDNA/RecA binding and results in the expression of genes involved in translesion DNA synthesis. This allows the cell replicate through a lesion present in the DNA double helix when the lesion can not be fixed. Though this mechanism leads to a high nucleotide incorporation error rate, it may enable the cell to survive.

During methyl-direct mismatch repair, the interaction between the SSB terminal tail and Exonuclease I, can stimulate the degradation of ssDNA (Sandigursky and Franklin 1994).
Moreover, SSB can stimulate in vitro RecA mediated λ-repressor cleavage (Weinstock and McEntee 1981; Resnick and Sussman 1982), showing its possible involvement in the activation of the SOS response.

1.2.4 SSB protein and TNR instability

The characteristics of the SSB protein suggest that it might play a stabilising role in TNR repeat instability. As the most accredited model accounting for instability is based on the formation of DNA secondary structures on ssDNA during DNA replication, it is logical to think of SSB as a major player involved in melting secondary structures, decreasing the chances ssDNA has to fold and cause instability. It was shown that replication slippage can take place in vivo in E. coli and that several E. coli and phage polymerases can slip during in vitro reactions (Canceill, Viguera et al. 1999). However SSB plays a role in DNA slippage which depends on its ability to increase the strand displacement activity of the polymerase rather than having a role in melting DNA secondary structures before polymerase passage (Canceill, Viguera et al. 1999) (Viguera, Canceill et al. 2001).

A stabilising role for SSB on TNR dynamics has been reported. The instability of different CAG-CTG repeat lengths, with CTG array present on the leading strand template, was studied in plasmids in E. coli. The temperature sensitive ssb-1 mutant form of SSB was used for the experiments; a temperature of 43°C completely eradicates the activity of the protein. When grown at high temperature, the strain bearing the plasmid displayed a significant increase of instability. The proposed mechanism involves the ability of SSB to prevent the formation of DNA secondary structures (Rosche, Jaworski et al. 1996).
1.2.5 SSB interactions

*E. coli* SSB has been shown to interact with numerous proteins involved in DNA processing. This highlights the fact that SSB can not only be seen as a mere ssDNA binding protein able to iron out secondary structures and protect ssDNA, but that it may hold more specific and extensive roles in the cell, as targeting DNA processing machinery components towards their substrates. The C-terminal domain of SSB can interact with a variety of proteins, suggesting that it is the gene expression profile present in the cell at a set time that makes a specific set of substrates available for binding to the ssDNA-SSB complex.

1.2.5.1 Recombination and repair: RecO, RecJ, RecQ and RuvAB

RecO protein is involved in the catalysis of complementary oligonucleotides annealing and ssDNA invasion of dsDNA (*Luisi-DeLuca and Kolodner 1994*). RecO can bind ssDNA via an OB fold structure and can enhance the loading of RecA onto ssDNA in the presence of RecR. RecA loading can be favoured by the binding of RecO/RecR onto ssDNA or by an interaction with the RecO/RecR complex that changes the ability of RecA to displace SSB (*Eggler, Lusetti et al. 2003*). Physical interaction of SSB with RecO was demonstrated (*Umezu and Kolodner 1994*). SSB is believed to interact with RecO via its C-term tail allowing RecO/RecR complex recruitment to the SSB coated ssDNA, followed by the binding of RecO to the ssDNA that leads to RecA protein nucleation (*Hobbs, Sakai et al. 2007*).

RecJ is a ssDNA exonuclease involved in recombination and in the excision step of methyl-direct mismatch repair (*Lovett and Clark 1984; Viswanathan and Lovett 1998; Burdett, Baitinger et al. 2001; Viswanathan, Burdett et al. 2001*). Recently, a role of SSB in RecJ loading has been proposed as *in vitro* addition of SSB to a RecJ substrate before incubation with RecJ leads to a more efficient substrate processing by the exonuclease (*Han,
Chapter I - Introduction

Cooper et al. 2006). SSB may have a dual role including secondary structure melting and interaction with RecJ leading to the recruitment of RecJ to its substrate. Interestingly there is evidence for physical interaction between SSB C-terminal domain and an opposite polarity *E. coli* exonuclease, exonuclease I (ExoI or SbcB) suggesting that a similar mechanism could be acting for RecJ (Sandigursky, Mendez et al. 1996).

RecQ is a helicase involved in the repair of U.V. damaged DNA structures and in DSB repair in *RecBC* strains in which the normal DSB repair pathway is not functional (Nakayama, Irino et al. 1985). Under these conditions, RecQ is believed to unwind aberrant DNA secondary structures, contribute to the exonuclease action of RecJ and favour RecFOR-mediated RecA loading onto ssDNA. Additionally, RecQ seems to be involved in the suppression of the SOS response and of the illegitimate recombination due to stalled replication forks (Hanada, Ukita et al. 1997; Hishida, Han et al. 2004). *In vitro* studies demonstrated that SSB can interact with RecQ via the C-terminal domain and that this interaction stimulates the RecQ helicase activity (Shereda, Bernstein et al. 2007). *B. subtilis* SSB and RecQ also interact through the C-terminal portion of SSB$_{bs}$ (Lecointe, Serena et al. 2007). A similar interaction was seen between RPA (Replication Protein A) and WRN, respectively SSB and RecQ homologues in eukaryotes (Doherty, Sommers et al. 2005). SSB can stimulate RecQ activity by preventing the reannealing of the unwound DNA and by preventing the binding of RecQ to ssDNA as the RecQ-ssDNA complex formation inhibits the RecQ activity. RecQ/SSB complex might also have a role at stalled replication forks where RecQ is believed to unwind the nascent lagging strand and coordinate fork restart (Shereda, Bernstein et al. 2007). Finally, SSB could also mediate RecQ interaction with RecJ and ExoI.

RuvA and RuvB proteins can branch migrate Holliday junctions formed during homologous recombination and have a role in RecA filament dissociation from duplex DNA
Chapter I - Introduction

(Muller, Tsaneva et al. 1993; Adams, Tsaneva et al. 1994). No evidence for physical interactions between SSB and RuvA and RuvB have been found. However, SSB has been shown to catalyse the branch migration reaction carried out by RuvA and RuvB (Parsons, Stasiak et al. 1995).

1.2.5.2 Stalled replication forks: PriA, RecQ and RecG

Replication fork stalling occurs when DNA replication encounters an obstacle to its progression. A stalled replication fork can be rescued by bypassing or removing the lesion, or by DNA breakage. This last pathway leads to homologous recombination between the free duplex and the intact sister and to the reloading of the replication machinery on the D-loop formed by the recombination intermediate (Horiuchi and Fujimura 1995; Xu and Marians 2003). A major role in fork restart is provided by PriA which is needed for DnaB helicase reloading back onto the chromosome. SSB protein interacts with PriA via the last 15 amino acids of the C-terminal domain stimulating the catalysis of PriA mediated unwinding of branched DNA substrates (Cadman and McGlynn 2004). On the other hand, SSB does not interact with the Rep helicase, involved in a different and less efficient pathway involved in stalled replication fork restart (Sandler 2000). Similarly, there is an interaction between B. subtilis SSB and PriA proteins. Microscopy studies on protein localisation demonstrated that PriA\textsubscript{bs} co-localises with the replication machinery and that this localisation is lost when cells express a C-term truncated version of SSB (\textit{sbdA35}). Additionally, SSB\textsubscript{bs} interacts in a similar manner with the RecG\textsubscript{bs} and the RecQ\textsubscript{bs} helicase. Localisation studies of both proteins gave the same patterns as PriA suggesting a role for SSB in targeting repair proteins to the replication fork in \textit{B. subtilis}. (Lecointe, Serena et al. 2007).
1.2.5.3 DNA polymerase III \(\chi\) subunit, polymerase V and TopB

The C-terminal tail of SSB protein has been shown to interact directly with the \(\chi\) subunit of the DNA polymerase III. This interaction would shield the SSB C-terminal tail from interactions with ssDNA as the C-terminal tail is negatively charged and, left unprotected, it would destabilise SSB binding to ssDNA. A mechanism by which the processivity of the polymerase would increase by preventing premature dissociation of SSB from ssDNA, was proposed (Witte, Urbanke et al. 2003). In vitro evidence for increased processivity of polymerase III bound to SSB was found by two other groups in 1998 (Glover and McHenry 1998; Kelman, Yuzhakov et al. 1998).

Polymerase V (PolV) of \(E.\ coli\) is responsible for translesion DNA synthesis and is formed by two subunits, UmuC and UmuD. When a lesion causes replication to stop, PolV can overcome the problem by replicating across the lesion. Its activity is stimulated by the RecA/ssDNA complex that activates the UmuD subunit and targets it to the lesion. SSB targets PolV to the lesion in \textit{in vitro} reactions (Arad, Hendel et al. 2008). SSB C-terminus interacts with the UmuC subunit and facilitate the interaction with polV substrate by remodelling the RecA filament on ssDNA (Arad, Hendel et al. 2008).

An extensive study on \(E.\ coli\) protein complexes, also demonstrated that SSB co-purifies with the topoisomerase III (TopB) (Butland, Zhang et al. 2006).

1.2.5.4 Uracil Excision Repair: UDG

Uracil residues incorporated by mistake in the DNA by the DNA polymerase or created by spontaneous deamination of cytosine, need to be removed. UDG (Uracil DNA glycosilase) is the first enzyme involved in the removal reaction. SSB is able to enhance the activity of the enzyme in \textit{in vitro} reactions involving structured substrates, thanks to its ability to melt DNA
Chapter I - Introduction

secondary structures (Kumar and Varshney 1997). The *E. coli* and *M. tuberculosis* SSB protein can interact with *E. coli* and *M. tuberculosis* UDG respectively, *in vitro* (Handa, Acharya et al. 2001). Studies in humans have also demonstrated the interaction between RPA (replication protein A, a SSB homologue in eukaryotes) and the human UDG (Nagelhus, Haug et al. 1997).

1.2.6 SSB homologues

Single stranded DNA binding protein studies started with the characterisation of the bacteriophage T4 gene 32 that encodes a protein (gp32) that can bind ssDNA tightly, cooperatively and with great affinity (Alberts and Frey 1970). This protein gave the first insights into ssDNA binding modes and allowed the characterisation of the residues involved in this binding. *In vitro*, gp32 has been shown to be essential for T4 DNA replication, recombination and repair (Alberts 1970; Chase and Williams 1986).

Replication protein A (RPA) is the SSB homologue in eukaryotes. It is formed by three subunits, RPA70, RPA32 and RPA14, which are conserved among eukaryotes and form a stable complex. It contains six OB folds by which it can bind to ssDNA and is essential for replication, recombination and repair processes as well as for DNA damage signalling and firing frequency of replication origins (Zou and Elledge 2003; Shechter, Costanzo et al. 2004; Fanning, Klimovich et al. 2006). As for SSB, RPA can interact with proteins involved in DNA metabolism. RPA30 interacts with Rad52 (RecO analogue) and Uracil DNA glycosylase 2 (Mer, Bochkarev et al. 2000) while RPA70 interacts with the tumour suppressor p53 (Li and Botchan 1993), Rad51 (RecA homologue) (Stauffer and Chazin 2004), BRCA1 and BRCA22 (MutS homologue) (Wong, Ionescu et al. 2003) and Mre11-Rad50-Nbs1 complex (SbcCD homologue) (Robison, Elliott et al. 2004). RPA is hyperphosphorylated in
response to DNA stress or damage (Binz, Sheehan et al. 2004); phosphorylation can modulate the interaction with other proteins and finely direct DNA damage responses. Interestingly, autoantibodies against the RPA32 subunit have been found in sera of patients with breast cancer. This may suggest the presence of a mutation in the *rpa* gene or RPA ability to sequestrate p53 tumour suppressor. A lower availability of p53 would prevent the activation of repair protein genes’ transcription, favouring cells with damaged DNA to progress through the cell cycle (Tomkiel, Alansari et al. 2002). Overexpression of RPA70 and RPA30 has been associated with poor overall survival and prognosticators of adverse outcome in patients with colon carcinomas. A precise mechanism by which RPA could favour cancer growth has not yet been found. However, it has been shown that a reduction of DNA replication rates correlates with a parallel reduction of RPA level in the cell; high replication speed observed in cancer cells would then require RPA overexpression (Givalos, Gakiopoulou et al. 2007).

A second human SSB protein, hSSB1, was recently characterised (Richard, Bolderson et al. 2008). hSSB1 is a conserved SSB homologue primarily involved in activating cellular response to DSBs. It has been shown to accumulate in cells in response to DNA damage. The accumulation is due to ATM kinase-mediated phosphorylation that stabilises the protein and shields it from proteosome degradation. hSSB1 localises at the site of DNA damage and is essential for the activation of the ATM kinase. hSSB1 deficient cells display hypersensitivity to ionising radiations and higher frequency of chromosomal aberrations after exposure to ionising radiations. hSSB1 is also involved in the stimulation of homologous recombination and interacts with Rad51, enhancing D-loop formation in an ATM kinase independent way.
1.3. The SOS response

DNA lesions can occur in bacterial cells during their lifespan and threaten their survival. Bacteria have developed a sophisticated system to sense and overcome DNA lesions that could stop DNA replication and gene transcription. The response to DNA damage and the subsequent cascade of events that follow is called the SOS response and causes deep changes in the gene expression profile of the cell.

Transcription of DNA damage induced genes is normally repressed by the LexA protein that binds to a 20bp sequence present in the gene’s operator region. The SOS response is triggered by the binding of RecA to ssDNA, present where a lesion occurred. Binding to ssDNA causes a conformational change in RecA, which becomes active as a co-protease for the cleavage and consequent deactivation of the LexA repressor. Different SOS genes can be induced at different times and at different levels depending on the affinity of LexA for the promoter and on the promoter strength (Radman 1975; Brent and Ptashne 1981; Little, Mount et al. 1981). So far, more than 40 genes that can be activated by DNA damage have been identified. Their products are primarily involved in DNA repair (UvrA, UvrB and UvrD involved in nucleotide excision repair) and recombination (RecA), processes that allow the cell to repair the damage and pursue replication. SfiA is a protein that can inhibit cellular division; sfiA gene expression is up-regulated during the SOS response to give more time to the cell to repair the damage. Finally transcription of genes like polB, dinB, umuC and umuD (encoding for PolIII, PolIV and PolV) is activated during the SOS response. PolIV and PolV are part of the Y-family polymerases and are involved in the resolution of replication forks stalling at a DNA lesion that PolIII can not overcome. They can insert a nucleotide opposite the lesion resulting in replication restart but low fidelity, which can cause appearance of mutations (Friedberg, Wagner et al. 2002). Despite the fact that the sib gene’s promoter contains a LexA binding site, shared with the uvrA gene, it has been demonstrated that its
transcription is not affected by the SOS response activation (Courcelle, Khodursky et al. 2001).

1.4. The arabinose operon of *E. coli*

The *E. coli* arabinose operon is regulated by the presence of arabinose through the AraC protein, which can act both as an activator or a repressor. The eight genes involved in the control of the arabinose metabolism are scattered around the chromosome, grouped in three different clusters: *araBAD* and *araC, araFGH* and finally *araE, araB, araA* and *araD* gene expression is controlled by the *araBAD* promoter (P_{BAD}) and their products are involved in arabinose metabolism. *araB* codes for the enzyme ribulokinase, *araA* for the L-arabinose-isomerase and *araD* for the L-ribulosephosphate-4-isomerase. *araC* codes for a repressor protein; its transcription direction diverge from that of *araBAD, araG* and *araH* both code for high affinity arabinose transporters and *araF* for an arabinose binding protein. Finally, *araE* codes for a low affinity arabinose transporter which works as a proton symport pump.

In the presence of arabinose, transcription from the P_{BAD} promoter is turned on whereas the presence of glucose completely represses the promoter by reducing the cellular levels of cAMP (Miyada, Stoltzfus et al. 1984; Lobell and Schleif 1990; Carra and Schleif 1993). The induction and repression rate of P_{BAD} is quick and efficient and can be modulated by the addition of different quantities of arabinose (Guzman, Belin et al. 1995). The repression/induction of AraC mediated transcription, depends on the presence of two operator sites and one inducer site in the P_{BAD} sequence. The organisation of the locus and the activation/repression mechanism are described in Figure 1.6.
Figure 1.6 - Schematic representation of the *E. coli* arabinose operon. A) $P_{BAD}$ and $P_C$ represent the promoters of *araBAD* and *araC* genes respectively. Two operator sites ($O_1$ and $O_2$) and one inducer site (*araI*) are present in front of the arabinose operon. CRP represents the binding site for cAMP repressor protein which promotes the rearrangement of AraC in the presence of arabinose. B) In absence of arabinose, AraC (green circle) binds to $O_2$ and *araI* simultaneously creating a DNA loop that renders $P_{BAD}$ inaccessible for the polymerase. The binding to $O_2$ prevents AraC transcription as well. C) In the presence of arabinose AraC changes conformation (green triangles) and binds preferentially to *araI*, favouring the disruption of the loop. If glucose is also absent, the levels of CRP protein (orange squares) are high and its binding to CRP enhances the breakage of the loop that results in transcription start from $P_{BAD}$. 
Chapter II

Materials and Methods
2.1 Bacterial methods

2.1.1 Transformation protocol

All strains were transformed according to the following protocol. After overnight growth in LB cells were diluted 50-fold in fresh medium and grown for 2 hours to log phase. 1ml aliquots were used for each transformation. Following a 30 minutes incubation on ice in 500μl of CaCl₂ (0.1M), cells were incubated on ice for 30 minutes in 100μl CaCl₂ (0.1M) in the presence of the transforming DNA and subsequently heat-shocked at 37°C for 5 minutes. After the addition of 400μl of LB medium cells were grown for 1 hour under agitation and plated on the suitable solid medium.

2.1.2 PI lysates preparation

PI lysates were made as follows. An overnight culture of the strain of interest was diluted 10-fold in 2.5mM CaCl₂ supplemented LB and grown for 2 hours at 37°C. 0.1ml of various dilutions of a PI lysate were incubated at 37°C for 30 minutes with 0.2ml of the culture to which 2.5ml of 5mM CaCl₂ supplemented-LC top-agar were added. The mixture was then poured on 5mM CaCl₂ supplemented LC-agar plates. After overnight incubation the top-agar containing the phage was harvested in 5 ml of phage buffer and 0.5ml of chloroform added. The lysate was incubated at 4°C in the dark for 30 minutes and then centrifuged for 5 minutes at 6,000rpm, the supernatant collected and stored at 4°C in the dark after addition of 0.1ml chloroform.
2.1.3 P1 transduction

P1 transductions were performed as follows. An overnight culture of the strain to transduce was grown in 2.5mM CaCl$_2$ supplemented LB. For each transduction 1ml aliquot was spun down at 13,000rpm for 1 minute and resuspended in 100µl of 2.5mM CaCl$_2$ supplemented LB. 100µl of the appropriate P1 lysate dilution were added to the cells followed by 20 minutes incubation at 37°C, under gentle agitation. Previous incubation at 37°C for 1 hour, 0.8ml of 2mM sodium-citrate supplemented LB were added to the cells. Finally 200µl of the culture were centrifuged, resuspended in 100µl of phage buffer and spread on plates containing gentamycin (10µg/ml). Selected clones were re-streaked twice on gentamycin plates to purify from residual virus.

2.1.4 pTOF24 gene replacement technique

pTOF24 plasmid derivatives were used for gene replacements. The strain of interest was transformed with an engineered pTOF24 plasmid containing two 450bp homology arms flanking the target gene in the chromosome. Transformants were streaked on chloramphenicol supplemented plates and incubated at 42°C. The process was repeated twice to allow selection of clones containing the plasmid inserted into the chromosome. Colonies were picked and grown overnight at 30°C in LB medium, then dilutions were plated on 5% sucrose plates. Clones were selected for Cm sensitivity and the gene replacement was checked by PCR. Positive clones were re-streaked twice to purify them and rechecked by PCR and sequencing.
2.1.5 Gentamycin resistance gene insertion

The gentamycin resistance gene was inserted in \textit{gyrX} using pTOF24 plasmid mediated gene replacement (Link, Phillips et al. 1997; Merlin, McAteer et al. 2002). pDL2812 plasmid was used for this purpose as previously described (Eykelenboom, Blackwood et al. 2008). Mutants were selected on gentamycin supplemented LB agar and confirmed by PCR.

2.1.6 \textit{Δ}sbcDC mutants construction

\textit{Δ}sbcDC mutants were constructed using plasmid mediated gene replacement as previously described (Darmon, Lopez-Vernaza et al. 2007).

2.1.7 Spot test

Spot tests were performed, unless otherwise stated, by adjusting the OD\textsubscript{600} of a cell culture to a value of 0.6. 5\textmu l drops of \(10^9\) to \(10^{-5}\) 10 fold dilutions of the culture were spotted on LB agar plates supplemented with suitable markers and sugars and incubated overnight at 37°C.

2.1.8 Growth curves

After overnight growth in 1% arabinose supplemented LB, DL2844 and DL3622 cultures were diluted 50-fold and grown in 1% arabinose supplemented LB to log phase (OD\textsubscript{600} =\textasciitilde 0.6). The cells were then washed twice in LB to remove the excess arabinose in the medium. The OD\textsubscript{600} was adjusted to a value of 0.05 and cells were grown in LB supplemented with different arabinose concentrations spanning from 1 to \(10^{-4}\)% in LB and in 0.5% glucose supplemented LB. DL1582 was used as a control strain and grown in LB and LB supplemented with 1% arabinose. The OD\textsubscript{660} of the liquid culture was kept between 0.1
and 0.9 by diluting it with fresh medium every hour. To follow cell growth, the OD₆₀₀ of each sample was measured every 30 minutes for a total of 7 hours.

2.1.9 Viability assay

Along with the growth curve experiment, viability tests were performed. Unless otherwise stated, a sample from each growing culture was taken every hour. The OD₆₀₀, adjusted to a value of 0.6 and 5µl drops of 10⁰ to 10⁻³ 10 fold dilutions were spotted on 1% arabinose plates and incubated overnight at 37°C.

2.1.10 U.V. sensitivity test

Overnight cultures of the strains of interest were grown from a single colony in LB (SSB_wt strains) and 1% arabinose supplemented LB (SSB_mae strains) at 37°C. The OD₆₀₀ of the overnights was adjusted to a value of 0.6 and 5µl drops of 10⁰ to 10⁻³ dilutions of the cultures were spotted on LB supplemented with different concentrations of arabinose (1%, 10⁻¹%, 10⁻²%, 10⁻³%, 10⁻⁴%) and on LB plates. The plates were then irradiated with U.V. light (0, 5,000, 10,000, 15,000 and 20,000µJ) and incubated at 37°C overnight.

2.2 DNA methods

2.2.1 Genomic DNA extraction

Genomic DNA was extracted starting from 1ml of overnight culture using the Wizard genomic DNA Purification Kit (Promega) and finally resuspended in 100µl of deionised
water and quantified by spectrometric analysis; the rate \( \Delta_{260}/\Delta_{280} \) was measured as index of DNA purity.

2.2.2 Crossover PCR

Crossover PCR was used to fuse \( P_{\text{BAD}} \) to the \( sib \) gene. The technique is based on two rounds of PCR. The first round allows amplification of the two fusion fragments using F1-R1 and F2-R2 primer sets. The second round consists of an additional PCR reaction, using an equal ratio of the products of the first two PCRs as templates and F1-R2 primer set for the amplification. The two primers sets, F1-R1 and F2-R2 were designed to insert a complementary tail on one end of the amplification product and a restriction site on the other end, allowing the fusion of the two fragments and cloning of the product (Fig.2.1).

![Crossover PCR technique](image)

**Figure 2.1 - Crossover PCR technique.** Step 1: first round of amplification using F1-R1 and F2-R2 primers sets. Steps 2 and 3: second round of amplification using F1-R2 primers set. Step 2 schematically represents the annealing of the two fragments after denaturation (single stranded DNA represented in only one orientation) and step 3 represents the final product of the crossover PCR. A and D: restriction sites inserted in F1 and R2 primers. B and C: complementary 12bp long tails.
2.2.3 Colony PCR

Colony PCR s were performed by picking a single colony and by resuspending it in 50μl H₂O. Samples were then boiled at 99.9°C for 10 minutes and spun down for 1 minute. 1μl of the supernatant was used as a substrate for the PCR reaction.

2.2.4 Plasmids

2.2.4.1 pTOF24

The pKO3 derived vector, pTOF24 (Fig.2.2), was used to replace the sib promoter, knock out or insert genes, insert the 105p palindrome and CAG-CTG arrays. The features of the plasmid include chloramphenicol and kanamycin resistance genes, a temperature sensitive replication gene (repA) and the sacB gene whose product is detrimental for cells if grown in the presence of sucrose and allows the selection of cells that have lost the plasmid.

Figure 2.2 - pTOF24 plasmid map. Chloramphenicol (Cm<sup>R</sup>) and kanamycin (Km<sup>R</sup>) resistance genes are indicated in blue, sacB gene in orange and repA temperature sensitive replication gene in red. PstI and SalI restriction sites are normally used for fragments cloning, replacing Km<sup>R</sup> gene.
2.2.4.2 pTOF24/uvrA-ssbCO2

pTOF24/uvrA-ssbCO2 was constructed to allow the replacement of the sib promoter with \( P_{BAD} \). A crossover PCR was performed to fuse \( P_{BAD} \) to the sib gene allowing the construction of strain DL2844. UvrACO_For and UvrACO2_Rev primers were used to amplify a 419bp region including the uvrA gene promoter and part of the uvrA gene (Fig.2.3A) generating the uvrA' fragment. MG1655 genomic DNA was used as a template for the PCR reaction. pAraB/ssbCO2_For and pAraB/ssbCO_Rev primers were used to amplify the \( P_{BAD}-sib \) gene fusion, using pTOF25araB/ssb2 as a template (Fig.2.3B, Dr. John Eykelenboom and Rachel Salz, University of Edinburgh). The amplification product \( P_{BAD}-sib' \) contains \( P_{BAD} \) (356bp) and the first 452bp of sib.

![Diagram](image)

**Figure 2.3 - Crossover PCR strategy.** A) Location of UvrACO_For and UvrACO2_Rev, used for the amplification of a 419bp long region, containing the uvrA gene promoter and part of the uvrA gene (uvrA-PuvrA). B) Location of pAraB/ssbCO2_For and pAraB/ssbCO_Rev, used for the amplification of \( P_{BAD}-sib' \) fragment.

UvrACO_For and pAraB/ssbCO_Rev primers contains respectively a \( PstI \) and a \( SalI \) restriction sites in 5’, while pAraB/ssbCO2_For and UvrACO2_Rev have 24bp homology tail that allowed the fusion of the two PCR product obtained in the first round of amplification. To enable the replacement of the sib promoter with the arabinose promoter at
genomic level, the fragment obtained by crossover PCR (uvrA'/P_{BAD}-ssb') was cloned in pTOF24 plasmid at PstI and SalI restriction sites (Fig.2.4). pTOF24uvrA'/ssbCO2 plasmid was used to replace the ssb promoter with P_{BAD} at genomic level in a BW27784 background.

![Diagram of the pTOF24UvrA/ssbCO2 plasmid](image)

**Figure 2.4 - pTOF24uvrA/ ssbCO2 plasmid.** The uvrA'-P_{uvrA}/ParaBAD-ssb' gene fusion was cloned in pTOF24 plasmid at PstI and SalI restriction sites. The ssb' and P_{uvrA-uvrA'} regions provide homology arms necessary for the integration of the plasmid at the ssb locus in the *E. coli* chromosome.

### 2.2.4.3 pAM34/ ssb

SSB protein was overexpressed in the strains DL3788, DL3800, DL3714 and DL3715 (Tab.2.2) using the pAM34/ ssb plasmid (Fig.2.5), a pAM34 derived vector (Gil and Bouche 1991). The ssb gene and promoter were amplified using ssb-F1 and ssb-R1 primers (Tab.2.3 pag 55) and the plasmid was constructed by inserting the ssb gene and promoter at EcoRI and PstI restriction sites in pAM34. pAM34/ ssb carries an ampicillin resistance gene and an IPTG dependent origin of replication. Cells carrying the plasmid were selected in IPTG and ampicillin containing medium.
Figure 2.5 - pAM34/ssb plasmid. pAM34/ssb plasmid is derived from pAM34 plasmid (Gil and Bouche 1991). The ssb gene and promoter were amplified from MG1655 strain DNA and cloned at PstI/EcoRI restriction sites in pAM34 plasmid.

2.2.4.4 pLacD2/pal105

The pLacD2 plasmid (Zahra, Blackwood et al. 2007), a pTOF24 derived vector containing homology arms for the beginning of lacZ, has been used as a tool to insert the 105bp palindrome in the lacZ gene. Primers LacZ-Mut_For and LacZ-Mut_Rev were used to introduce a double mutation in lacZ second homology arm in pLacD2 to create an 8bp sequence identical to an 8bp sequence present at the beginning of the 105bp palindrome. A site direct mutagenesis kit (Stratagene) was used to introduce the mutation. The 105bp palindrome (Pinder, Blake et al. 1998) was amplified from DRL242 phage λ strain (this lab) lysate using pal-CF2 and pal-CR1 primers. The DL51 strain was grown overnight in LB at 37°C, dilute 10-fold and grown for 3 hours in LB supplemented with 5mM MgSO₄ 0.2% maltose and 0.01g/l vitamin B1. 1ml of the culture was mixed with 1ml of TM buffer (10mM Tris-hydrochloride pH 7.6 containing 10mM CaCl₂ and 0.5mM MgCl₂). 250µl of the diluted
culture were mixed with 2ml of BBL top agar (LB agar supplemented with 5mM MgSO₄, 0.2% maltose and 0.01g/l vitamin B1) and laid on BBL agar plates. A premade DRI.242 phage λ strain lysate was diluted to 10⁻³ (ten fold dilutions) in phage buffer and 10μl drops were spotted on DL51 layer. The plate was incubated overnight at 37°C. An isolated plaque formed by the virus on the bacteria layer was picked, resuspended in 30μl dH₂O and boiled for 10 minutes at 99.9°C. 1μl of the boiled plaque was used as template for the amplification reaction, expecting a fragment of 533bp in length. Pfu polymerase (Promega) was used, to ensure high fidelity, according to the following protocol:

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Stock concentration</th>
<th>Volume (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pfu Buffer</td>
<td>10X</td>
<td>5</td>
</tr>
<tr>
<td>dNTPs</td>
<td>2.5mM</td>
<td>5</td>
</tr>
<tr>
<td>Fwd primer</td>
<td>10μM</td>
<td>2</td>
</tr>
<tr>
<td>Rev primer</td>
<td>10μM</td>
<td>2</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>3U/μl</td>
<td>31</td>
</tr>
<tr>
<td>Pfu</td>
<td>3U/μl</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>T(°C) and time</th>
</tr>
</thead>
<tbody>
<tr>
<td>95°C for 2’</td>
</tr>
<tr>
<td>95°C for 30”</td>
</tr>
</tbody>
</table>
| 58°C for 15”  | 25 cycles  
| 73°C for 2’   |  
| 73°C for 5’   |  
| 4°C hold      |  

The PCR fragment was digested with *Eco*RI (Roche) and pLacD2 modified plasmid with *Mfe*I (NEB) enzymes respectively according to the following protocols:

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume(μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffer H</td>
<td>3</td>
</tr>
<tr>
<td><em>Eco</em>RI</td>
<td>2</td>
</tr>
<tr>
<td>PCR product</td>
<td>20</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>5</td>
</tr>
</tbody>
</table>
### Chapter II – Materials and Methods

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffer 4</td>
<td>2.5</td>
</tr>
<tr>
<td>MfeI</td>
<td>1.5</td>
</tr>
<tr>
<td>pLacD2 modified</td>
<td>20</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>1</td>
</tr>
</tbody>
</table>

The 105bp palindrome was extracted from an agarose gel using qIagen protocol and eluted in 30ul ddH₂O. Digested pLacD2 modified plasmid was treated with CIP (NEB) before ligation (see table below), purified using Qiagen protocol and eluted in 30μl ddH₂O.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffer 3</td>
<td>2.5</td>
</tr>
<tr>
<td>CIP</td>
<td>2</td>
</tr>
<tr>
<td>Digested pLacD2 modified</td>
<td>20</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Quick ligase enzyme (Promega) was used to perform the ligation reaction to insert the 105bp palindrome into pLacD2 modified plasmid:

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quick Buffer</td>
<td>10</td>
</tr>
<tr>
<td>Quick Ligase</td>
<td>1</td>
</tr>
<tr>
<td>Digested pLacD2 modified</td>
<td>1</td>
</tr>
<tr>
<td>105bp palindrome</td>
<td>6</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>3</td>
</tr>
</tbody>
</table>

#### 2.2.4.5 pGB150

The pGB150 plasmid (Dr. Garry Blakely, University of Edinburgh) was used to detect the activation of the SOS response in the strains DL3121 and DL2984. The plasmid carries the chloramphenicol resistance gene as a selection marker and the gfp gene under the control of the sfiA gene promoter. The sfiA promoter is induced when the SOS response is activated
leading to \( \beta \) gene transcription. GFP protein expression level can be detected by fluorescence microscopy and used as a marker for induction of the SOS response.

### 2.3 Western blot analysis

#### 2.3.1 Cell lysate preparation

Cell samples were taken from growing liquid cultures; the OD<sub>600</sub> was measured and adjusted to a value of 0.6 prior to lysis. 1ml of OD<sub>600</sub>=0.6 cell culture was centrifuged at 13,000rpm for 1 minute, cells were washed once in PBS buffer and resuspended in 100μl of ddH<sub>2</sub>O and 100μl of 2X lysis buffer (2mM Tris-HCl pH6.8, 200mM DTT, 4% SDS, 0.2% bromophenol blue, 20% glycerol). Cell lysates were then boiled for 10 minutes and stored at -20°C.

#### 2.3.2 Polyacrylamide gel electrophoresis

Western blot analysis was carried out using α-SSB antibody (Prof. Roger McMacken, Johns Hopkins Bloomberg School of Public Health, Baltimore, Md). Concentrations of 15% and 7% of polyacrylamide were used to make stacking and separating gels respectively and prepared as follows:

<table>
<thead>
<tr>
<th>Reagent</th>
<th>7%</th>
<th>15%</th>
</tr>
</thead>
<tbody>
<tr>
<td>30% ( w/v ) acrylamide (ratio 37.5:1 bis acrylamide, Severn Biotech l.t.d.)*</td>
<td>2.3ml</td>
<td>5ml</td>
</tr>
<tr>
<td>Buffer A (1.5M Tris-HCl pH 8.8)</td>
<td>-</td>
<td>2.5ml</td>
</tr>
<tr>
<td>Buffer B (0.5M Tris-HCl pH 6.8)</td>
<td>2.5ml</td>
<td>-</td>
</tr>
<tr>
<td>SDS 10% solution</td>
<td>100μl</td>
<td>100μl</td>
</tr>
<tr>
<td>ddH&lt;sub&gt;2&lt;/sub&gt;O</td>
<td>5.1ml</td>
<td>2.4ml</td>
</tr>
<tr>
<td>APS</td>
<td>5μl</td>
<td>10μl</td>
</tr>
<tr>
<td>Temed</td>
<td>50μl</td>
<td>50μl</td>
</tr>
</tbody>
</table>
5µl of cell lysate were loaded on gel and the run was performed at 100V for about 3 hours. SeeBlue®Plus2 Prestained Standard (Invitrogen) was used as a marker. Proteins were transferred to a nitrocellulose membrane for 1.5 hours at 80V (4°C) using a wet system (transfer buffer: 250mM Tris, 2M glycine) and the membranes were blocked overnight in 5% milk (dissolved in 0.001% Tween 20 supplemented PBS). Membranes were then incubated for 30 minutes in 3% milk (dissolved in 0.001% Tween 20 supplemented PBS) containing the α-SSB antibody (1:40,000 dilution of the PBS diluted stock), washed for 30 minutes in 0.001% Tween 20 supplemented PBS, incubated for 45 minutes in 1% milk solution (dissolved in 0.001% Tween 20 supplemented PBS) containing α-rabbit antibody (1:5,000 dilution of the stock) and washed for an hour in 0.001% Tween 20 supplemented PBS. The signal was detected on a radiographic film using ECL reagent (Pierce) after 10-20 seconds exposure.

2.4 Microscopy

A Zeiss Axiovert 200 microscope with a Semrock dichroic filter set was used to obtain both bright-field and fluorescent images. Metamorph v 6.3r2 program was used to acquire, analyse and manipulate images. Microscopy slides were prepared by pouring 300µl of 1% w/v agarose dissolved in ddH₂O. 5µl of the cell culture of interest were spotted on the agarose layer and covered with a coverslip prior to analysis.

2.4.1 DAPI staining

DAPI staining was performed on DL1582 and DL2844. Overnight cultures of the strains were respectively grown in LB and 1% arabinose supplemented LB, diluted 50-fold in the
same fresh medium and grown for 1.5 hours. DL2844 culture was washed twice in LB, split and grown at different arabinose concentrations (0, 10^{-2}, 10^{-3}, 10^{-4}) for 240 minutes. DL1582 was grown for 240 minutes in LB. Prior to the microscopy analysis, samples were taken from the growing culture and incubated at 37°C for 10 minutes in the presence of 1μg/ml of DAPI and 250μg/ml of chloramphenicol.

2.4.2 SOS response activation assessment

Strains DL3121 and DL2984 were used for microscopy analysis to measure the SOS response activation on SSB depletion over time. Cells were grown overnight from a single colony in 5ml of LB supplemented with chloramphenicol or 1% arabinose and chloramphenicol respectively, diluted 50-fold in fresh medium keeping the same conditions and grown for 1.5 hours. DL2984 was then washed twice in LB and the culture split in different conditions (LB supplemented with chloramphenicol and arabinose concentrations of 1% 10^{-2}%, 10^{-3} % and 10^{-4}%). The strains were then grown for 240 minutes and the OD_{600} was kept between 0.1 and 0.9 by diluting the culture at regular intervals. Samples for microscopy analysis were taken at time 0 and then every 60 minutes. 5μl of the cultures were spotted on 1% agarose slides and analysed under the microscope.

2.5 Instability Assays

Instability assays were performed on strains containing either palindromes or CAG-CTG repeat arrays to test the stability of those sequences as a function of the SSB protein level.
2.5.1 105bp palindrome deletion assay

A palindrome deletion assay was designed to assess the ability of a 105bp imperfect palindrome to form hairpins during replication in the presence of normal or depleted SSB levels. The palindrome was cloned in a modified pLacD2 plasmid as previously described. The lacZ arm downstream of the palindrome was engineered to create a sequence identical to an 8bp sequence at the beginning of the palindrome (Fig.2.6). Strain DL2944 was created by inserting the palindrome in the lacZ gene of strain DL2551 by using the plasmid-based replacement technique. A P1 lysate from DL2944 was used to insert the palindrome in strains DL1582 and DL2844 to create DL2959 and DL2957 respectively. DL2959 and DL2957 were grown overnight in LB and 1% arabinose supplemented LB respectively, washed twice in LB to remove the arabinose in excess and plated on different arabinose concentrations (1%, 0.5% 10⁻¹ %, 5*10⁻¹ %). Analysis of the colonies was performed after 3 days incubation at 37°C.

**Figure 2.6 - 105bp palindrome construct.** A) Schematic representation of the lacZ locus. lacZ gene is represented in black, the 105bp palindrome in blue and the homologous 8bp sequences in orange. B) Details of the construct at the base pair level. Only the beginning and the very end of the palindrome sequence are represented in the figure.
2.5.2 246bp palindrome associated viability

To generate the strains DL3749 and DL3750, DL2006 and DL2075 were transformed with pAM34/sb plasmid allowing SSB overexpression. Overnight cultures of these strains were grown in LB supplemented with 0.5% glucose to repress SbcCD transcription. DL3749 and DL3750 strains were grown in ampicillin, IPTG and 0.5% glucose supplemented medium.

A viability test was performed on plates. Cells were washed twice in LB and 5µl of 10^6 to 10^5 10-fold dilutions were plated on 0.2% arabinose and 0.5% glucose supplemented LB-agar plates, in the presence or absence of IPTG and ampicillin. Before diluting, the OD_{600} of the cultures was adjusted to a value of 0.6.

A viability test in liquid culture was performed by growing overnight cultures of DL2006 in LB supplemented with 0.5% glucose and DL3750 in ampicillin, IPTG and 0.5% glucose supplemented LB. The cultures were diluted 50-fold and grown to log phase in the same medium. Each culture was then washed twice in LB and split in two aliquots; one of them was grown in LB supplemented with 0.5% glucose and the second one in 0.2% arabinose supplemented LB. Ampicillin and IPTG were added to DL3750 cultures. The OD_{600} was checked every 30 minutes and the viability of the strain every 120 minutes via spot test. The OD_{600} of the culture was adjusted to a value of 0.6 before carrying out the spot test. 5µl of 10^0 to 10^{-5} 10 fold dilutions of the cultures were spotted on LB supplemented with 0.5% glucose and then incubated overnight at 37°C.

2.5.3 CAG-CTG repeat instability assay

The instability assay was carried out over time to follow the variation of the instability level of a growing culture. A single colony was picked from a streak of the desired strain and
inoculated in 5ml LB or IPTG and ampicillin supplemented LB. Cells were grown overnight, the OD<sub>600</sub> of the culture was assessed and cells were diluted in fresh media to and OD<sub>600</sub> value of 10<sup>-5</sup>. Every 20 generations (approximatively 8 hours) the process was repeated to allow a continuum growth. Samples were taken for analysis after the first overnight growth (generation 0) and then every 20 generations (generations 20, 40, 60, 80 and 100). The OD<sub>600</sub> of the culture was adjusted to a value of 0.6 and 100µl of a 10<sup>-5</sup> ten fold dilution of it were plated on LB-agar or IPTG and ampicillin supplemented LB-agar plates. 96 colonies per plate were picked and analysed by colony PCR. The experiment was repeated twice for each strain.

2.5.3.1 Ex-test PCR

Ex-test PCRs were performed as colony PCRs using FAM-Ex-test-F and Ex-test-R primers (Tab.2.3). The forward primer was labelled using 6-fam fluorescent tag (Meion) to allow detection of the PCR fragment for Gene Mapper analysis (see below). The protocol used to amplify TNR arrays is listed below; the volumes used refer to a 25µl total volume reaction. Taq polymerase from Roche was used for the amplification and DMSO as it can facilitate the amplification reaction on GC-reach templates.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Stock concentration</th>
<th>Volume (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR Buffer</td>
<td>10X</td>
<td>2.5</td>
</tr>
<tr>
<td>dNTPs</td>
<td>2.5mM</td>
<td>2</td>
</tr>
<tr>
<td>DMSO</td>
<td>100%</td>
<td>1.25</td>
</tr>
<tr>
<td>Fwd primer</td>
<td>5µM</td>
<td>0.5</td>
</tr>
<tr>
<td>Rev primer</td>
<td>5µM</td>
<td>0.5</td>
</tr>
<tr>
<td>Taq</td>
<td>5U/µl</td>
<td>0.125</td>
</tr>
</tbody>
</table>
### T(°C) and time

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>95°C</td>
<td>5’</td>
</tr>
<tr>
<td>95°C</td>
<td>30”</td>
</tr>
<tr>
<td>53°C</td>
<td>15”</td>
</tr>
<tr>
<td>72°C</td>
<td>1’15”</td>
</tr>
<tr>
<td>72°C</td>
<td>10’</td>
</tr>
<tr>
<td>4°C</td>
<td>hold</td>
</tr>
<tr>
<td></td>
<td>30 cycles</td>
</tr>
</tbody>
</table>

### 2.5.3.2 Gene Mapper analysis

After amplification, each of the PCR reactions was diluted 5-fold with ddH₂O. 1.5μl of a fluorescent size standard (Gene Scan-500LIZ and Gene Scan-1200LIZ for (CAG)₇₅ and (CTG)₁₀⁷ respectively) was added to 1ml of HiDi reagent (ABI) and 1μl of the diluted PCR product was added to 9μl of the HiDi mixture. The samples were then run in a capillary electrophoresis apparatus to detect the size of the fragments (ABI3730 analyser). The output data of the electrophoretic separation are graphs displaying peaks corresponding to the length of the PCR product. The graphs were analysed using the Gene Mapper program (ABI) (Fig. 2.7). Only parental length variations happening in the liquid culture were counted as instability events (Fig.2.7b and 2.7c). Therefore deletion or expansion peaks present in graphs together with parental length peaks, were not counted as instability events as they derive from deletions or expansions which happened on the plate (Fig.2.7d). For the same experiment, sibling clones (two samples presenting the same expanded or deleted length of the array and belonging from the same cell culture) were not counted more than once as instability events, it being likely that they were generated by the same deletion or expansion event occurred during growth in liquid culture. The proportion of instability was calculated as the percentage of the number of events (deletions + expansions) over the total number of significant clones analysed. Standard error was used to draw error bars.
Figure 2.7 - Gene Mapper 4.0 graphs. The graphs show the output of Gene Mapper analysis for PCR products derived from strains containing (CAG)$_{75}$ repeat array. The same considerations apply for (CTG)$_{137}$ containing strains. The peak on the far right represents the size of the amplified PCR fragment while the stutter peaks building up right before it are to be considered PCR artifacts. a) (CAG)$_{75}$ parental size. b) Deletion peak. c) Expansion peak. d) In this graph both expansion and deletion peaks are present together with parental peak.
Chapter II – Materials and Methods

2.6 *E. coli* bacterial strains

2.6.1 Background strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>BW27784</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
rmB3  
ΔlacZ4787 
hsaR514 
DE(araBAD)567  
DE(araBAD)  
68 DE(araFGH) 
Φ(DnaE ep P_C::araE) | (Kholebnikov, Datsenko et al. 2001) |
| CSH100   | F' lacpro A+ B+ (lacIq lacPL8)            | Cold Spring Harbour Laboratory                   |
| MG1655   | F- lambda- iscG- rfb-50 rfb-1             | -                                               |
| AB1157   | ΔrecB21 ΔrecC22 ΔshcCB13                  | -                                               |
| XL-1 Blue| recA1 endA1 gyrA496 thi-1 bsdR17 supF44 relA1 lac | Stratagene                                     |

Table 2.1 - *E. coli* background strains

2.6.2 Engineered strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Background</th>
<th>Genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>DL51</td>
<td>AB1157</td>
<td>AB1157</td>
<td>-</td>
</tr>
<tr>
<td>DL844</td>
<td>CHS100</td>
<td>CHS100</td>
<td>Cold Spring Harbour Laboratory</td>
</tr>
<tr>
<td>DL928</td>
<td>CHS100</td>
<td>DL844 mutλ::Tn10</td>
<td>Laboratory</td>
</tr>
<tr>
<td>DL1582</td>
<td>BW27784</td>
<td>BW27784</td>
<td>(Kholebnikov, Datsenko et al. 2001)</td>
</tr>
<tr>
<td>DL1786</td>
<td>MG1655</td>
<td>lacZ+ lacI+ Zeo998 X+</td>
<td>Dr. John Eykelenboom</td>
</tr>
<tr>
<td>DL1995</td>
<td>MG1655</td>
<td>lacZ+ lacI+ Zeo998 X+ lacZ::(CAG)75</td>
<td>Dr. Rabaab Zahra</td>
</tr>
<tr>
<td>DL2006</td>
<td>BW27784</td>
<td>P_RAD iscCD gyrX::Gm8 lacZ+ lacI+ lacZ::pal246bp</td>
<td>Dr. John Eykelenboom</td>
</tr>
<tr>
<td>DL2009</td>
<td>MG1655</td>
<td>DL1786 lacZ::(CTG)95</td>
<td>Dr. Rabaab Zahra</td>
</tr>
<tr>
<td>DL2075</td>
<td>BW27784</td>
<td>DL2006 recA::Cm8</td>
<td>Dr. John Eykelenboom</td>
</tr>
<tr>
<td>DL2305</td>
<td>MG1655</td>
<td>lacZ+ lacI+ Zeo9998 X+ lacZ::(CTG)140</td>
<td>Dr. Rabaab Zahra</td>
</tr>
<tr>
<td>DL2551</td>
<td>MG1655</td>
<td>lacZ+ lacI+ gyrX::Gm8</td>
<td>Dr. John Eykelenboom</td>
</tr>
<tr>
<td>DL2751</td>
<td>MG1655</td>
<td>DL2009 marA</td>
<td>Dr. Rabaab Zahra</td>
</tr>
<tr>
<td>DL2844</td>
<td>BW27784</td>
<td>DL1582 P_RAD isc</td>
<td>This work</td>
</tr>
<tr>
<td>DL2944</td>
<td>MG1655</td>
<td>DL2551 lacZ::pal105</td>
<td>This work</td>
</tr>
<tr>
<td>DL2957</td>
<td>BW27784</td>
<td>DL2844 lacZ::pal105 gyrX::Gm8 lacZ+ lacI+</td>
<td>This work</td>
</tr>
<tr>
<td>DL2959</td>
<td>BW27784</td>
<td>DL1582 lacZ::pal105 gyrX::Gm8 lacZ+ lacI+</td>
<td>This work</td>
</tr>
</tbody>
</table>
### Table 2.2 - *E. coli* engineered strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Background</th>
<th>Genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>DL2984</td>
<td>BW27784</td>
<td>DL2844 + pGB150</td>
<td>This work</td>
</tr>
<tr>
<td>DL3121</td>
<td>BW27784</td>
<td>DL1582 + pGB150</td>
<td>This work</td>
</tr>
<tr>
<td>DL3297</td>
<td>MG1655</td>
<td>DL1995 gyrX::Gm&lt;sup&gt;+&lt;/sup&gt;</td>
<td>This work</td>
</tr>
<tr>
<td>DL3298</td>
<td>MG1655</td>
<td>DL2305 gyrX::Gm&lt;sup&gt;+&lt;/sup&gt;</td>
<td>This work</td>
</tr>
<tr>
<td>DL3311</td>
<td>BW27784</td>
<td>DL1582 lacZ&lt;sup&gt;+&lt;/sup&gt; lacI&lt;sup&gt;+&lt;/sup&gt; Zeo&lt;sup&gt;+&lt;/sup&gt; lacZ::(CAG)&lt;sub&gt;15&lt;/sub&gt;</td>
<td>This work</td>
</tr>
<tr>
<td>DL3315</td>
<td>BW27784</td>
<td>DL2844 lacZ&lt;sup&gt;+&lt;/sup&gt; lacI&lt;sup&gt;+&lt;/sup&gt; Zeo&lt;sup&gt;+&lt;/sup&gt; lacZ::(CAG)&lt;sub&gt;15&lt;/sub&gt;</td>
<td>This work</td>
</tr>
<tr>
<td>DL3548</td>
<td>BW27784</td>
<td>DL3311 ΔsbeDC</td>
<td>This work</td>
</tr>
<tr>
<td>DL3622</td>
<td>BW27784</td>
<td>DL2844 repressor strain selected on 10&lt;sup&gt;-3&lt;/sup&gt;% ara</td>
<td>This work</td>
</tr>
<tr>
<td>DL3640</td>
<td>BW27784</td>
<td>DL1582 recA</td>
<td>This work</td>
</tr>
<tr>
<td>DL3714</td>
<td>BW27784</td>
<td>DL3311 + pAM34/psb</td>
<td>This work</td>
</tr>
<tr>
<td>DL3715</td>
<td>BW27784</td>
<td>DL3548 + pAM34/psb</td>
<td>This work</td>
</tr>
<tr>
<td>DL3749</td>
<td>BW27784</td>
<td>DL2006 + pAM34/psb</td>
<td>This work</td>
</tr>
<tr>
<td>DL3750</td>
<td>BW27784</td>
<td>DL2075 + pAM34/psb</td>
<td>This work</td>
</tr>
<tr>
<td>DL3772</td>
<td>BW27784</td>
<td>DL2844 repressor strain selected on 10&lt;sup&gt;-3&lt;/sup&gt;% ara</td>
<td>This work</td>
</tr>
<tr>
<td>DL3773</td>
<td>BW27784</td>
<td>DL2844 repressor strain selected on 10&lt;sup&gt;-3&lt;/sup&gt;% ara</td>
<td>This work</td>
</tr>
<tr>
<td>DL3774</td>
<td>BW27784</td>
<td>DL2844 repressor strain selected on 10&lt;sup&gt;-3&lt;/sup&gt;% ara</td>
<td>This work</td>
</tr>
<tr>
<td>DL3785</td>
<td>BW27784</td>
<td>DL1582 lacZ&lt;sup&gt;+&lt;/sup&gt; lacI&lt;sup&gt;+&lt;/sup&gt; Zeo&lt;sup&gt;+&lt;/sup&gt; lacZ::(CTG)&lt;sub&gt;140&lt;/sub&gt;</td>
<td>This work</td>
</tr>
<tr>
<td>DL3788</td>
<td>BW27784</td>
<td>DL3785 + pAM34/psb</td>
<td>This work</td>
</tr>
<tr>
<td>DL3799</td>
<td>BW27784</td>
<td>DL3785 ΔsbeDC</td>
<td>This work</td>
</tr>
<tr>
<td>DL3800</td>
<td>BW27784</td>
<td>DL3799 + pAM34/psb</td>
<td>This work</td>
</tr>
</tbody>
</table>
2.7 DNA Primers

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ex-test-R</td>
<td>GGCGATTAAGTTGGGTAACG</td>
</tr>
<tr>
<td>FAM-Ex-test-F</td>
<td>TTATGCTTCCGGGCTCGGTATG</td>
</tr>
<tr>
<td>LacZ-Mut_For</td>
<td>GCGTCTTCGTGTTGTTGTAACACGTGACTGGG</td>
</tr>
<tr>
<td>LacZ-Mut_Rev</td>
<td>CCCAGTCACGACGTTGTACACCGCAAGACGC</td>
</tr>
<tr>
<td>Pal-CF2</td>
<td>ATACCCAGATTGCAGACACC</td>
</tr>
<tr>
<td>Pal-CR1</td>
<td>ACAACCTGACCCCAGAAAAA</td>
</tr>
<tr>
<td>ssb-F1</td>
<td>AAAAAACGAGCTACGGCCGATCACAAAC</td>
</tr>
<tr>
<td>ssb-R1</td>
<td>AAAAAAGAGTTCGCTGGCAGATGCCCTTAAATC</td>
</tr>
<tr>
<td>UvrACO_For</td>
<td>AAAAAACGAGCTGCTGATTTCGGGATTGTG</td>
</tr>
<tr>
<td>UvrACO2_Rev</td>
<td>AGTATGAAAGTACTGTATATTCCATTTGTC</td>
</tr>
<tr>
<td>pAraB/ssbCO2_For</td>
<td>TGAATATACAGTTTTCATACACTCCACCATT</td>
</tr>
<tr>
<td>pAraB/ssbCO_Rev</td>
<td>AAAAAAGTCGACCCCGCCGCTGAACGTTGC</td>
</tr>
</tbody>
</table>

Table 2.3 - DNA primers

2.8 Media

2.8.1 LB broth

LB broth was used as a standard growth medium in liquid. 10g bacto-tryptone (Difco), 5g yeast extract (Difco) and 10g NaCl made up to 1 litre with distilled water and autoclaved.

2.8.2 LB agar

10g Bacto-tryptone (Difco), 5g yeast extract (Difco), 10g NaCl and 15 g Bacto-agar (Difco) made up to 1 litre with distilled water and adjusted to pH 7.5 using NaOH and autoclaved.
2.8.3 LC agar

LC agar was used to grow P1 phage. 10g tryptone, 5g yeast extract, 5g NaCl and 10g Difco-agar made up to 1 litre with distilled water and adjusted to pH 7.2 with NaOH and autoclaved.

2.8.4 LC top agar

10g tryptone, 5g yeast extract, 5g NaCl and 7g Difco-agar made up to 1 litre with distilled water and adjusted to pH 7.2 with NaOH and autoclaved.

2.8.5 BBL agar

10g Trypticase-peptone, 5g NaCl, 10g agar (Difco) made up to 1 litre with distilled water and autoclaved.

2.8.6 BBL top agar

10 g Trypticase-peptone, 5 g NaCl, 6.5 g agar (Difco) made up to 1 litre with distilled water and autoclaved.

2.8.7 Phage buffer

3g KH₂PO₄, 7g Na₂HPO₄, 5g NaCl, 1mM MgSO₄, 1mM CaCl₂ and 1% (w/v) gelatine made up to 1 litre with distilled water and autoclaved.
2.8.8 TM buffer

Made up to 10mM Tris (using 1M Tris pH7.5) and 10mM MgSO₄ (using 1M MgSO₄) using Milli-Q water. Sterilised using a 0.45µm syringe filter and tored at room temperature.

2.9 Drugs and Sugars

For the different experiments purposes, the following chemicals were added to media at the indicated working concentrations:

<table>
<thead>
<tr>
<th>Chemicals</th>
<th>Brand</th>
<th>Working Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chloramphenicol</td>
<td>Calbiochem</td>
<td>50 µg/ml</td>
</tr>
<tr>
<td>Ampicillin</td>
<td>GlaxoSmithKline</td>
<td>100 µg/ml</td>
</tr>
<tr>
<td>Kanamicin</td>
<td>Calbiochem</td>
<td>50 µg/ml</td>
</tr>
<tr>
<td>Gentamycin</td>
<td>Sigma</td>
<td>10 µg/ml</td>
</tr>
<tr>
<td>IPTG</td>
<td>Melford labs</td>
<td>1 mM</td>
</tr>
<tr>
<td>X-gal</td>
<td>Melford labs</td>
<td>40 µg/ml</td>
</tr>
<tr>
<td>Glucose</td>
<td>Sigma</td>
<td>0.5%</td>
</tr>
<tr>
<td>Sucrose</td>
<td>Sigma</td>
<td>5.0%</td>
</tr>
</tbody>
</table>

Table 2.4 - Drugs and sugars
Chapter III

Effects of SSB Protein Depletion
3.1 Introduction

The single-stranded binding protein (SSB) of *E. coli* is an essential protein that specifically binds single stranded DNA and is believed to have a prominent role in the maintenance of genome stability. The manipulation of the SSB level in cells could therefore lead to the possible creation of an unstable cellular system. This would not only allow a deeper characterisation of the roles of the SSB protein in the maintenance of the genome stability, but also the study and characterisation of biological and evolutionary aspects triggered by a loss of DNA stability.

For this purpose, the *E. coli* strain BW27784 (Khlebnikov, Datsenko et al. 2001), DL1582 in this work, was engineered permitting a fine control of the expression of the *ssb* gene, placed under the control of the arabinose-inducible *araBAD* genes promoter (P_{BAD}) at chromosomal level. DL1582 is a K12-derived strain that was used as a background strain to ensure a homogeneous intake of arabinose in the cell population overall. DL1582 can not utilise arabinose as an energy source and the intake of arabinose is controlled by the AraE transporter only, the expression of which is under the control of a constitutive promoter and therefore equal in every cell (Khlebnikov, Datsenko et al. 2001).

The natural *ssb* promoter was replaced by P_{BAD} (Fig.3.1) using pTOF24 replacement vector to give strain DL2844. Care was taken not to disrupt the functionality of the adjacent *uwrA* gene and promoter. In this work we will refer to a strain with a native *ssb* gene control as “SSB_{wt}” (SSB wild type). On the other hand we will refer to a strain in which the *ssb* gene expression is controlled by the arabinose promoter as “SSB_{ind}” (SSB inducible).
Figure 3.1 - Strains DL1582 and DL2844: ssb locus in the chromosome. A) DL1582 strain B) DL2844 strain. In DL2844 the ssb promoter was replaced by the arabinose inducible promoter (P_{BAD}) so the ssb gene expression could be controlled by the level of arabinose in the medium.

Preliminary experiments were carried out to characterise the newly created strain. Viability tests and growth in solid and liquid media were performed together with an assessment of the SSB level present in cells grown at different arabinose concentrations. This revealed differences between growth in liquid and solid medium when the arabinose concentration was depleted in the growth medium. U.V. sensitivity tests were also carried out.

3.2 Results

3.2.1 Viability test on plates

*ssb* gene transcription, controlled by P_{BAD}, can be inhibited by the presence of glucose in the growth medium and can be induced and tuned by exposing cells to different concentrations of arabinose. SSB being an essential protein, we were expecting the cells not to be viable if grown in the presence of glucose and to present different growth rates if exposed to different arabinose concentrations. The ability of DL2844 strain to grow at different arabinose concentrations on solid medium was assessed by performing spot tests.
Chapter III – Effects of SSB Protein Depletion

Cells were grown and spotted as described in the material and methods. DL1582 was used as a control, its growth being independent of glucose or arabinose.

A pronounced decrease in cell viability correlated with the decrease in arabinose concentration in the medium (Fig.3.2). In particular, different levels of viability could be associated with different concentrations of arabinose, confirming that the ssb gene transcription could be modulated. Strikingly a 2-3 log loss in viability was observed when cells were shifted from 1 to 0.1% arabinose supplemented plates. The complete absence of arabinose almost completely inhibited cells growth. The presence of glucose exacerbated this effect.

**Figure 3.2 - DL1582 and DL2844 strains viability on plate.** DL1582 and DL2844 strains were grown overnight in LB and 1% arabinose supplemented LB respectively. The OD<sub>600</sub> of the overnight culture was adjusted to a value of 0.6 and 5μl drops of 10<sup>0</sup> to 10<sup>-3</sup> 10 fold dilutions were spotted on plates supplemented with different arabinose concentrations (as indicated in the figure), LB and 0.5% glucose supplemented plates. DL1582 was used as a control, its growth being independent of arabinose or glucose. The numbers on the left side of the picture refer to the arabinose concentration (expressed as a percentage) present in the LB-agar, 0.5 refers to the glucose percentage present in the medium.
3.2.2 Growth and viability in liquid medium

DL2844 was assessed for growth trend and viability in liquid culture in LB supplemented with different arabinose concentrations. The strain was grown as described in materials and methods: the culture was kept in log phase by diluting it every hour with fresh medium. Diluting the culture during growth allowed the arabinose concentration and consequently the SSB concentration, to set on the required level without being influenced by the starting conditions.

The OD$_{600}$ of the culture was measured every half hour to follow the growth trend of the strain. The OD$_{600}$ of each sample was plotted against time (Fig.3.3) in a semi-logarithmic graph. DL2844 grown in the presence of 0.5% glucose stopped growing after 180 minutes. The growth of DL2844 in the presence of 10% arabinose stopped at 210 minutes. An arabinose concentration as little as 0.0005% proved to be sufficient for the strain to grow normally in liquid medium. These results differ strongly from the ones obtained from growth on plate. Cells growing on solid medium needed an arabinose concentration of about 2-3 logs higher than cells growing in liquid medium. Moreover, growth in liquid medium did not allow the appreciation of a set of different behaviours, suggesting that profound differences exist between growth in liquid and solid medium.
Figure 3.3 - DL1582 and DL2844 strains: growth curve. DL2844 strain was grown in LB supplemented with different arabinose concentrations. Growth was followed for 420 minutes and the OD\textsubscript{600} measured every 30 minutes. DL1582 strain was used as control and grown in 1% arabinose supplemented LB. The numbers in the legend refer to the sugar percentage added to the medium. Glc stands for glucose. DL2844 grown in LB behaved like DL2844 grown in LB supplemented with 0.5% glucose (data not shown).

During growth in liquid medium at different arabinose concentrations, cell viability was assessed every hour. A spot test was performed by spotting $10^0$ to $10^{-4}$ 10 fold dilutions of the culture on 1% arabinose plates to allow full recovery of viable cells. The OD\textsubscript{600} of the culture was adjusted to a value of 0.2 before dilution (Fig.3.4). As it can be observed, the cell viability dropped about an hour before the cells stopped increasing their mass, consistent with inhibition of cell division followed by cell elongation.
Chapter III – Effects of SSB Protein Depletion

Figure 3.4 - DL2844 strain viability in liquid culture. DL2844 strain was grown in 1 and 10^{-6}% arabinose and 0.5% glucose supplemented LB and 10^0 to 10^{-4} 10 fold dilutions were spotted on 1% arabinose plates at each time point. The OD_{600} of the cultures was adjusted to a value of 0.2 before dilution. DL2844 strain grown in 1%, 10^{-1}%, 10^{-2}% and 5*10^{-4}% arabinose behaved like DL1582 (data not shown).

3.2.3 SSB expression level assessment

Western blot analyses were performed to assess the SSB protein level in DL2844 grown in LB supplemented with different arabinose concentrations. The SSB level was measured during post exponential and exponential phase growth. Overnight growth was performed in LB for DL1582 and in the presence of 1% arabinose for DL2844. Exponential phase growth was performed in LB for DL1582 and in media containing different arabinose concentrations, spanning from 1% to 10^{-6}%, for DL2844. On each polyacrylamide gel DL1582 samples were loaded as a control together with DL2844 samples grown in different
Chapter III – Effects of SSB Protein Depletion

conditions. The bands appearing on the film were quantified using ImageJ program and the band intensities were plotted against time.

Figure 3.5 shows the SSB expression level in DL2844 grown in the presence of different arabinose concentrations. It can be observed how SSB expression was not strongly affected by the arabinose level when 1% to $10^{-1}$ % concentrations were used. A remarkable drop of the SSB level was found when the strain was grown in $10^{-4}$ % arabinose (Fig.3.5). This is in accordance with the data published by Khlebnikov: a remarkable decrease of the expression of $P_{\text{BAD}}$ controlled genes was observed for arabinose concentrations ranging in the order of $10^{-4}$ % (Khlebnikov, Datsenko et al. 2001). However, a fine modulation of the SSB expression could not be observed, suggesting that, SSB being an essential protein, a slight drop of its level below a certain threshold may lead to strong consequences on DNA replication and cell viability.
Figure 3.5 - Western blot analysis. SSB expression level was assessed in DL1582 and DL2844 strains by western blot. Both strains were grown in liquid culture for 210 minutes and samples were taken after overnight growth (marked as minute 0) and during exponential phase growth after 90, 150 and 210 minutes of growth. DL1582 was grown in LB and DL2844 in LB supplemented with different arabinose concentrations as specified below. A) 1% arabinose B) 10⁻¹ % arabinose C) 10⁻² % arabinose D) 10⁻³ % arabinose E) 10⁻⁴ % arabinose F) Western blot pictures. The number on the left side refers to the arabinose concentration used during DL2844 growth, 0, 90, 150 and 210 is the time (minutes).
Chapter III – Effects of SSB Protein Depletion

3.2.4 Cell density and SSB expression coordination

During growth the ssb gene is expressed in exponential phase and the expression is shut down during stationary phase (Moreau 1987; Meyer and Laine 1990). Putting the ssb gene under the control of a different promoter may disrupt this kind of regulation, possibly leading to the creation of obstacles to the cell cycle progression.

Experiments on strains DL1582 and DL2844 were performed to test this hypothesis. Overnight cultures of the two strains were grown in LB and 1% arabinose supplemented medium respectively, diluted to $10^2$, $10^4$ and $10^6$ OD$_{600}$ values and grown to post exponential phase. For the strains used, a culture with an OD$_{600}$ value of $10^6$ contains approximately 300 cells/ml. DL1582 strain was grown in LB while DL2844 was grown in 1% and $10^1$ % arabinose supplemented medium.

DL2844 growth was strongly influenced by the starting value of the cultures’ optical density. The arabinose concentration in the medium did not influence the growth of DL2844 when growth started from an OD$_{600}$ value of $10^2$. An arabinose concentration of $10^1$ % caused impaired growth for both $10^4$ and $10^6$ OD$_{600}$ starting conditions while 1% arabinose allowed normal growth for the $10^4$ OD$_{600}$ starting condition but not when growth was started from an OD$_{600}$ value of $10^6$. (Fig.3.6). This may suggest that a wild type regulation of the ssb promoter is important to ensure a good functionality of the replication machinery.
Figure 3.6 - Cell density and growth. Strain DL1582 and DL2844 were grown overnight in LB and 1% arabinose supplemented LB respectively. Cultures were diluted to an OD\textsubscript{600} value of 10\textsuperscript{2}, 10\textsuperscript{4} and 10\textsuperscript{6} and grown in LB (DL1582), 1% and 10\textsuperscript{-1} % arabinose supplemented LB (DL2844). The OD\textsubscript{600} of the cultures was measured every 30 minutes and plotted against time. A Starting OD\textsubscript{600} 10\textsuperscript{2}. B Starting OD\textsubscript{600} 10\textsuperscript{4}. C Starting OD\textsubscript{600} 10\textsuperscript{6}.

3.2.5 Growth curves in glucose

To be able to control the SSB level more precisely in liquid culture and to create a condition in which cells would be able to survive at slightly lower-than-normal SSB levels, DL2844 (SSB\textsubscript{mut}) was grown in media containing a combination of arabinose and glucose. The presence of both sugars allowed the creation of an unstable system in which the effect of SSB depletion was not as dramatic and threatening for cell survival. A wide range of arabinose concentrations were tested at a fixed glucose concentration of 1%. Cell growth and viability were assessed together with the SSB protein level in the cell (Fig.3.7, 3.8 and 3.9).

DL2844 grown in the presence of 1% glucose and 1%, 0.5% and 0.25% arabinose did not show impaired growth, while dropping the arabinose concentration to 10\textsuperscript{-1} %, 10\textsuperscript{-2} % and
$10^{-3}$ % arabinose caused a growth arrest at 450, 360, 270 and 210 minutes respectively (Fig.3.7). However, even if cell growth was not influenced, a drop of the cell viability was observed in the presence of 0.5% and 0.25% arabinose (Fig. 3.8). Cells grown in the presence of 1% arabinose did not show a viability drop. Cells grown in LB and LB supplemented with $10^{-1}$%, $10^{-2}$%, $10^{-3}$% arabinose supplemented LB and LB displayed a viability drop that was more severe.

Figure 3.7 - DL2844 (SSB<sub>mad</sub>) growth curve of cells grown in the presence of glucose and arabinose. The effect of a mixture of glucose and arabinose was tested on DL2844 strain growth and viability. A fixed concentration of 1% glucose was used in combination with different arabinose concentrations, as specified in the graph. A control sample was grown in LB. The data present in the graph are pooled from 2 independent experiments. The first experiment was followed for 420 minutes (1%, $10^{-2}$%, $10^{-3}$% arabinose) the second experiment was followed for 480 minutes (0.5%, 0.25%, 0.1% arabinose and LB).
Figure 3.8 - DL2844 (SSB
del) viability of cells grown in glucose and arabinose. Samples for viability test were taken every hour. 5μl drops of 10⁰ to 10⁻⁵ 10 fold dilutions were spotted on 1% arabinose plates. The OD₆₀₀ was adjusted to a value of 0.6 before diluting the culture. Numbers on the left hand side refer to growth time (minutes) while numbers on the top refer to arabinose percentage used during growth.
Chapter III – Effects of SSB Protein Depletion

To monitor the SSB level during growth in 1% glucose and 10⁻¹ % arabinose, samples were taken every hour and prepared for western blot analysis. DL1582 and DL2844 strains containing (CAG)₅₅ repeat stretch, DL3311 and DL3315 respectively, were used for this analysis. The experiment was performed on those strains as the aim of the experiment was to assess TNR instability when the SSB level was depleted. 5µl of each sample was loaded on a gel and band intensity measured and plotted against time (Fig.3.9). Strikingly, in the presence of 1% glucose and 0.1% arabinose, DL3315 strain cell growth was only slightly impaired but the SSB protein level was lower that in the DL3311 strain that displays a wild type level of SSB.

![Figure 3.9 - Western blot of cells grown in the presence of glucose and arabinose. Strain DL3311 (SSBwt) and DL3315 (SSBmut) were grown in 0.1% ara and 1% glucose supplemented LB. The behaviours of these strains in liquid culture followed those of DL1582 and DL2844 respectively (data not shown). 5µl of cell lysates were loaded for each time point for each strain analysed. The first 5 time points (0-240) and the last four time points (300-480) were loaded in two different gels and are therefore not directly comparable. Strain number and time points (minutes) are reported on top of the gel pictures. Time point 0 refers to the SSB level present in the overnight culture.](image-url)
3.2.6 U.V. sensitivity test

As the construction of DL2844 involved the manipulation of the promoter of *uvrA* and *uvrA* being involved in nucleotide excision repair pathway, the U.V. sensitivity of the strain was tested to check if the functionality of *uvrA* was affected. Furthermore, as *ssb* mutants show U.V sensitivity (Glassberg, Meyer et al. 1979; Whittier and Chase 1981; Lieberman and Witkin 1983), the viability of cells expressing low SSB levels was studied after U.V exposure. A U.V test was carried out as described in materials and methods. DL1582 (SSB<sub>add</sub>), DL2751 (*uvrA*) and DL3640 (*recA*) strains were used as controls. Dilutions of the overnight cultures were plated in the presence of 0%, 1%, 10<sup>-1</sup> %, 10<sup>-2</sup> %, 10<sup>-3</sup> %, 10<sup>-4</sup> % arabinose concentrations and irradiated with 0, 5,000 and 10,000 micro Jules (μJ) of U.V. light. After U.V exposure, no viability problems were observed for DL1582 at all conditions tested while a drop in viability was observed for DL2751 and DL3640. DL2844 showed a slightly decreased U.V resistance when plated in the presence of 1% arabinose and irradiated with 10,000μJ. Surprisingly, it showed an increased U.V. resistance when plated on lower arabinose concentrations. In the presence of 10<sup>-4</sup> % arabinose a slight decrease of viability was observed (Fig.3.10).
Figure 3.10 - U.V sensitivity test. DL2844 (SSBind) U.V. sensitivity was tested on LB-agar supplemented with different arabinose concentrations on plates, as indicated on the left hand side of the picture. DL1582 (SSB<sub>wt</sub>), DL3640 (∆recA<sub>1</sub>) and DL2751 (∆uvrA<sub>1</sub>) were used as control strains. Cells were spotted after overnight growth and irradiated with 0, 5,000 and 10,000 µJ power as indicate at the top of the figure. DL1582, DL3640 and DL2751 showed the same behaviour for all the arabinose concentrations tested (data not shown).
3.3 Discussion and conclusions

3.3.1 SSB expression and cell growth

SSB is an essential protein and varying its level may be lethal. In this work we regulated the SSB level by putting the ssb gene under the control of the arabinose inducible promoter, $P_{BAD}$. Experiments on cell growth and viability were performed on strains in which the SSB level was lowered and led to some interesting findings.

DL2844 viability on plates was strictly dependent on the arabinose concentration added to the LB-agar suggesting a modulation of the expression level mediated by arabinose. Colonies resulting from plating at low arabinose concentrations were smaller than normal after overnight growth (see chapter V, Fig.5.1), suggesting that lowering the SSB level may interfere with replication and cell cycle progression. However, results obtained from growth in liquid medium led to completely different results. Strikingly, the arabinose concentration necessary for DL2844 to grow properly in liquid medium was about three orders of magnitude lower that the one required for normal growth on plate. Moreover, in liquid culture, gradually lowering the arabinose concentration to a threshold level led either to cell survival or to cell death with no gradual change of phenotype, as observed on plates.

These behaviour differences could be explained by the ability of the strain to uptake arabinose from the growing medium. Arabinose diffusion and uptake may in fact be easier in liquid medium than on solid medium. Growth on plates may therefore require the presence of a higher arabinose concentration in the medium to reach a similar level of arabinose, and therefore of SSB expression, in the cells. To test this hypothesis, viability experiments were carried out on plates in which the concentration of agar was lowered to 0.5%, being the standard concentration used 1.5%. Different arabinose concentrations were tested showing that the viability levels were independent of the agar concentration in the plate (data not
Chapter III – Effects of SSB Protein Depletion

shown). This might be due to the percentage of agar being still too high to allow a more efficient uptake of the sugar.

Another aspect to consider is that cells undergo stress response activation while transferred from liquid culture onto solid medium. In particular, heat shock and oxidative-stress regulons are induced (Cuny, Lesbats et al. 2007). After plating, cells go through a lag phase during which no growth can be observed (Cuny, Lesbats et al. 2007). Moreover SSB overexpression has been detected in immobilised cold shocked cells (Perrot, Hebraud et al. 2001). Maintaining an accurate SSB level may therefore be important during the first time window after cells get transferred from liquid to solid medium. The required SSB level may be higher than normal to allow the cells to go over the lag phase. An accurate and timed SSB expression might also be crucial to guarantee the expression of genes involved in stress responses.

It has been proposed that easily culturable bacterial cells that undergo stresses may enter in a state in which they do not grow but they can still be considered viable, their membrane integrity and metabolic activity being preserved (Roszak and Colwell 1987). This could be what might happen to SSB strains grown on solid medium and could explain why in liquid culture cells are either dead or alive as the arabinose concentration is reduced, while on plate they gradually lose culturability. Cells shocked by the transfer on solid medium may still be viable but need higher SSB levels to divide and form colonies.

3.3.2 Cell density and growth

When starting DL2844 growth in liquid culture from an OD$_{600}$ value of 0.05, the presence of either 1% or 10$^{-1}$% arabinose does not influence growth and viability. Both conditions behaved as the control strain (Fig. 3.3 and 3.4). However, adjusting the starting OD$_{600}$ of the
liquid culture to lower values ($10^4$ and $10^5$) proved to be deleterious for cell growth and
growth rate of DL2844 (Fig.3.6). Strikingly, the quantity of arabinose added to the growing
medium proved to be important when cell growth started at low cell density as the behaviour
of DL2844 grown in the presence of either 1% or $10^4$ % arabinose was different. These data
indicate that the SSB level is important for driving cell growth at low cell density and that a
marginal variation in the SSB level could be detrimental.

$sab$ gene expression regulation may play an important part in the process. The natural $sab$
promoter regulation may be very tight and constantly changing during different stages of
growth. SSB expression is in fact higher during exponential phase growth (Moreau 1987;
Meyer and Laine 1990). Having the $sab$ gene under the control of a different promoter may
have disrupted this fine regulation and therefore disrupted the capacity of cells to behave like
SSB$_a$ cells.

Quorum sensing is described as the phenomenon whereby the accumulation of specific
molecules in the environment, allows bacteria to sense the presence of other bacteria and
register the cell density in the medium. This system enables cells to behave as part of a
population and not as an isolated cell and to elicit concerted responses to variation in cell
number (Bassler and Losick 2006; Diggle, Crusz et al. 2007). Quorum sensing controls
several processes such as bioluminescence, competence, virulence, sporulation, biofilm
formation and antibiotics production. It has been found in both gram-positive and gram-
negative bacteria and it is mediated by specific signalling molecules that accumulate in the
extracellular environment (Diggle, Crusz et al. 2007).

The fact that strain DL2844, in which the $sab$ gene expression is not regulated by the native
promoter, faces growth problems when cell growth started at low density, suggests that
sensing the presence of other cells could be crucial for the cells to start replicating. Since the
only difference between strains DL1582 and DL2844 is $sab$ gene expression regulation, a

76
Chapter III – Effects of SSB Protein Depletion

certain cell density may be necessary to fully induce the \( P_{\text{BAD}} \) promoter. A similar situation may pertain when cells are plated in the presence of low arabinose concentrations. Cell density largely decreases during the plating process and this may influence the response cells have to different arabinose concentrations.

In these experiments growth was started from an overnight culture. An interesting test would be to grow the cells to exponential phase prior to dilution. This would indicate if exiting stationary phase is the problem cells encounter when growth starts at a very low cell density or if replication would still slow down. However, cells grown in the presence of \( 10^1 \) % arabinose from an initial \( \text{OD}_{600} \) value of \( 10^4 \), showed a slow-growth phenotype meaning that they exited stationary phase but still present impaired growth abilities. Cells grow in the presence 1% arabinose from an initial \( \text{OD}_{600} \) value of \( 10^6 \) behaved in the same way. These findings argue against the need for quorum sensing.

To further investigate this point, viability test on slow growing cultures should be performed together with a western blot analysis. Following growth of those cultures and changing the growth conditions by increasing the arabinose concentration in the medium it would be interesting to check if the slow-growth phenotype could be reverted.

3.3.3 Growth in glucose and arabinose

Growth in the presence of glucose and arabinose was carried out in order to be able to perform long term instability experiments in liquid culture. The aim was to slightly lower the SSB level possibly perturbing genome instability without causing cell death. The presence of both sugars allowed the creation of a system in which the SSB level could be kept at a lower-than-\textit{wild type} level, but where cells were still growing at normal rates for a time window of about 7 hours. However, cell viability was affected to a certain extent. SSB is an essential
protein and even a small variation of its level may lead to cell death, as demonstrated in previous experiments. To be able to survive the stringent conditions, cells may have accumulated mutations in the arabinose or glucose transport system or in the arabinose promoter regulation. As a consequence it is also possible that the distribution of the two sugars would not be even in the cell population overall, causing some cells to die and some others to survive.

We found that the growth of cells grown in the presence of 0.5%, 0.25% and 0.1% arabinose was comparable to the growth rate of cell grown in 1% arabinose supplemented medium, but we found that the mortality rate was higher. We can therefore speculate that the SOS response may be active in some of these cells, causing cell elongation and therefore a normal OD_{600} increase rate but decreased viability. This suggests dis-homogeneity in the cell population, regarding sugar uptake and SSB expression level.

More investigations on cell viability, accurate measurement of the SSB level during growth and SOS response studies will be necessary before the system could be used for long term experiments on SSB depletion. It is however interesting to highlight the fact that, in the presence of certain quantities of arabinose, the presence of glucose, is not able to completely switch off the activity of P_{BAD}.

3.3.4 SSB expression and U.V sensitivity

A few ssb mutations have been associated with increased U.V. radiation sensitivity (Glassberg, Meyer et al. 1979; Johnson 1984; Meyer and Laine 1990). Given that the system used here was aimed at lowering the SSB protein level, we were expecting an increased U.V. sensitivity for DL2844 grown at low arabinose concentration.
Chapter III – Effects of SSB Protein Depletion

However, the results obtained from the spot test clearly show that U.V. sensitivity did not increase but slightly decreased for most of the arabinose concentrations tested and that only arabinose concentrations limiting growth both on plate and in liquid culture led to a similar level of survival between irradiated and not irradiated cells. We can therefore conclude that, after U.V. exposure, the sensitivity to arabinose depletion of this strain clearly decreased. The irradiated cells can survive at lower arabinose concentrations, on plates. An explanation for these data could lie in the capacity of U.V. irradiation to generate suppressor clones that could survive at lower arabinose concentrations on plates. As thoroughly explained in chapter V, the selection of suppressors able to grow on plates at lower arabinose concentration was possible. Irradiating the cells with U.V. may have led to the generation of such suppressors that would have a growth advantage at the conditions present on plates.
Chapter IV

Cell Division, Nucleoid Morphology and SOS Response


4.1 Introduction

4.1.1 SSB depletion, cell division and nucleoid morphology

An interesting question to ask was how SSB depletion would affect cell division and nucleoid morphology in strain DL2844. One would expect SSB to be essential to avoid replication arrest thanks to its ability to bind and protect ssDNA and to target other protein to the replication fork. DAPI staining was used to follow cell division and nucleoid morphology during cell growth in the presence of different SSB levels. DAPI (4',6-diamidino-2-phenylindole) is a fluorescent dye that can enter live cells and strongly bind to DNA. It can be detected as it gets excited by ultraviolet light and emits in the blue light spectrum.

4.1.2 SSB depletion and the SOS response

A second interesting point to investigate was the effect that SSB depletion would have had on the SOS response activation. As described in the introduction, RecA binding to ssDNA leads to the activation of RecA proteolytic activity towards the LexA repressor. This is the step that triggers induction of the SOS response induction. SSB may be able to enhance the SOS response induction as it can facilitate RecA binding to ssDNA by avoiding the formation of DNA secondary structures. On the other hand, the presence of SSB bound to ssDNA may compete with RecA binding and create an obstacle for triggering the induction of the SOS response.

The plasmid pGB150, containing the green fluorescent protein (gfp) gene under the control of the sfiA gene’s promoter, was used for the quantification of the SOS response activation. The SOS response induction could therefore be followed by fluorescence microscopy,
assessing the fluorescence level in a cell population carrying the plasmid. The SOS response is involved in the activation of \textit{slrA} gene transcription. Expression of \textit{slrA} leads to inhibition of cell division which causes cell elongation. Cell elongation was used as a second marker to assess the SOS response activation.

4.2 Results

4.2.1 SSB depletion impairs cell division and changes nucleoid morphology

After overnight growth in 1% arabinose, a DL2844 culture was split and grown respectively in LB and LB supplemented with 1 % arabinose for 240 minutes. A DAPI staining analysis was performed every 60 minutes on the growing strain. DL1582 was used as control. The strains were treated with DAPI and chloramphenicol before microscope analysis, as described in materials and methods. Chloramphenicol induces condensation of the nucleoid into a toroid shape thus facilitating their localisation (Zimmerman 2002).

After 180 minutes growth in LB, DL2844 showed cell elongation and impaired chromosome condensation. After 240 minutes growth, the DAPI signal was completely diffused in the whole cell and the formation of condensed nucleoids could not be observed anymore (Fig.4.1). Experiments carried out on strain DL2844 showed that SSB depletion led to impaired cell division and to a change in nucleoid morphology.
**Figure 4.1 - DAPI staining.** Brightfield and DAPI images were merged; DAPI staining is shown in blue. A) DL1582 after 60 minutes growth in LB. B) DL2844 after 60 minutes growth in 1% arabinose supplemented LB. C, D, E, F) DL2844 grown in LB for 60, 120, 180, and 240 minutes respectively. DL1582 grown in LB and DL2844 grown in 1% arabinose supplemented LB maintained the same cell length and DNA pattern distribution in the cell for the whole duration of the experiment (time points 120, 180 and 240 not shown). DL2844 grown in LB showed a variation in the DNA distribution pattern and cell length starting from minute 180. At minute 240 a dramatic cell filamentation and DNA diffusion in the cell were observed.
4.2.2 SSB protein depletion causes activation of the SOS response

The effect of SSB depletion on the SOS response activation, was measured as a function of fluorescence intensity in strains DL3121 and DL2984, respectively DL1582 and DL2844 transformed using the plasmid pGB150. 300 cells per sample were analysed and fluorescence intensity was plotted in increasing order against the number of each individual cell (1-300). A second type of analysis was carried out by plotting fluorescence intensity against cell length. Several conditions were tested. Arabinose concentrations of 1%, 10⁻² %, 10⁻³ % and 10⁻⁴ % were used to grow DL2984 while control strain DL3121 was grown in LB. Cell growth was followed for 240 minutes and samples analysed for SOS response induction every 60 minutes.

The SOS response induction was first measured in DL3121 and DL2984 overnight cultures grown in LB and 1% arabinose supplemented LB respectively; DL2844 was grown in 1% arabinose supplemented LB and used as control strain to assess the fluorescence background level in the absence of pGB150 (Fig.4.2A). DL3121 and DL2984 overnights were then diluted and grown for 90 minutes, respectively in LB and 1% arabinose supplemented LB, to get the cell out of stationary phase. The SOS response activation was reassessed after this period of time (Fig.4.2B). DL2984 was then split and grown under different conditions. DL3121 grown in LB and DL2984 grown in the presence of 1%, 10⁻² %, 10⁻³ % arabinose did not show activation of the SOS response (data not shown) while DL2984 grown in 10⁻⁴ % arabinose supplemented LB showed activation of the SOS response starting after about 180 minutes growth (Fig.4.3 and 4.4). Pictures of cells grown in 10⁻⁴ % arabinose, show how SOS response activation level and cell length increased when the SSB protein was depleted (Fig.4.5).
Figure 4.2 - SOS response activation control. Fluorescence intensity was plotted against cell length to assess the SOS response activation. A) SOS response activation in strains DL3121 (SSB<sub>ex</sub>) and DL2984 (SSB<sub>mut</sub>) after overnight growth in LB and 1% arabinose supplemented LB respectively. DL2844 was used as a control for fluorescence background. B) SOS response activation in strains DL3121 and DL2984 grown respectively in LB and 1% arabinose supplemented LB, after the initial growth period of 90 minutes.
Figure 4.3 - SOS response activation of DL3121 (SSB\textsubscript{om}) and DL2984 (SSB\textsubscript{sd}). To assess the SOS response activation as function of cell fluorescence only, the fluorescence intensity of 300 cells was measured and plotted in growing order against the individual cell number (1-300). A and B) 60 and 120 minutes growth. No changes in fluorescence intensity were registered for the first two time points analysed. C) 180 minutes growth. The SOS response activation started to be visible with a slight increase of cell fluorescence for DL2984 grown in 10^{-4}% arabinose supplemented LB. D) 240 minutes growth. The SOS response induction became more evident with a net increase of cell fluorescence for DL2984 grown in 10^{-4}% arabinose supplemented LB.
Figure 4.4 - SOS response activation and cell length of strains DL3121 (SSB<sub>red</sub>) and DL2984 (SSB<sub>blue</sub>). Fluorescence intensity was plotted against cell length to assess the SOS response activation. A and B) 60 and 120 minutes growth. No changes in cell length and fluorescence intensity were registered. C) 180 minutes growth. The SOS response activation started to be visible with a slight increase of cell fluorescence and elongation for DL2984 grown in 10^{-4}% arabinose supplemented LB. D) 240 minutes growth. The SOS response induction became more evident with a net increase of fluorescence intensity and cell elongation for DL2984 grown in 10^{-4}% arabinose supplemented LB.
**DL2984 (SSB\textsubscript{ind} + pGB150)**

![Images of DL2984 cells at different times: A 60', B 120', C 180', D 240']

**Figure 4.5 - Cell length and fluorescence of DL2984 (SSB\textsubscript{ind}) grown in 10\(^{-4}\)% arabinose.** The pictures represent the look of DL2984 cells grown in 10\(^{-4}\)% arabinose over time. **A, B and C**) Time 60, 120, 180 minutes cells are small and occasionally some fluorescent ones are found. **B**) Time 240 minutes, cells start elongating dramatically and the fluorescence level is high in pretty much all of them. DL3121 was used as a control strain and grown in LB for 240 minutes. Conditions A, B and C presented the same phenotype as the control strain (data not shown).
4.3 Discussion and conclusions

By visualising the DNA distribution pattern in the cell we demonstrated that SSB protein is essential for the cells to go through a normal cell division and for chromosome segregation.

Experiments carried out on cephalalexin treated cells, in which cell division but not replication was inhibited, showed how the nucleoids kept on organising in a tidy way along the cell (Zimmerman 2003). The results obtained here from the DAPI staining show that the regular pattern described in the paper is not maintained when SSB is depleted and cells start elongating, consistent with impaired nucleoid condensation and missegregation. The diffusion of the DAPI stain signal within the whole cell may suggest that DNA degradation occurs. To make this point, quantification of the DAPI signal would be necessary together with more microscopy experiments involving tagging of loci in the chromosomes and counting the number of foci present in each cell.

We also demonstrated that SSB depletion led to the SOS response cascade activation. Increased cell fluorescence intensity and cell elongation were observed in DL2984 grown in 10^-4 % arabinose supplemented LB. As previously discussed, SSB may play an essential role in the activation of the SOS response by facilitating the binding of the RecA protein to ssDNA. However, as SSB and RecA compete towards ssDNA binding, a lower level of SSB may create a situation in which RecA would have an easier access to ssDNA. Moreover, a lower level of SSB may lead to a higher level of damage occurring to DNA. These considerations suggest that lowering the SSB level may lead to induction of the SOS response.
Chapter V

Isolation and Analysis of Suppressor Strains
Chapter V – Suppressor Strains Isolation and Analysis

5.1 Introduction

As discussed in chapter 3, the behaviour of strain DL2844 exposed to varied arabinose concentrations and therefore expressing different levels of SSB, was different in liquid and solid medium. The quantity of arabinose required for growth on plate was about three orders of magnitude higher than the one required for growth in liquid culture. Lowering the SSB level creates a situation of stress for the strain and suppressor clones, able to cope with the new conditions, could be selected for their capacity to grow at arabinose concentrations that would not allow growth of the original strain.

To investigate the difference between growth in liquid and solid medium and to elucidate how DL2844 strain could evolve to survive at lower arabinose concentrations on plates, cells that could survive at low arabinose concentrations were isolated and their phenotype and genotype analysed, leading to some interesting findings.

5.2 Results

5.2.1 Suppressor clones

5.2.1.1 Isolation of suppressor clones

Suppressor clones were isolated starting from a culture of DL2844 grown overnight in LB supplemented with 1% arabinose and by plating an appropriate dilution on 0.1% arabinose supplemented plates. DL2844 cells deriving from the same overnight culture were plated in the presence of 1% arabinose and used as a control. The phenotype of the resulting colonies was different in the two cases, after overnight incubation at 37°C. Cells plated on 1% arabinose showed a healthy phenotype and uniform colony size while cells plated on 0.1% arabinose gave rise to a mixture of healthy-looking and small colonies (Fig.5.1).
Figure 5.1 - Colony phenotype. A) DL2844 plated in the presence of 1% arabinose. B) DL2844 plated in the presence of 0.1% arabinose. Plating in the presence of 1% arabinose resulted in healthy and evenly sized colonies while plating on 0.1% arabinose resulted in a mixture of small and big colonies. 100μl of a ten fold 10^{-3} dilution were plated in the presence of 1% arabinose while 100μl ten fold 10^{-2} dilution were plate in the presence of 0.1% arabinose.

2 colonies were picked from the 1% arabinose plate, 2 big and 2 small colonies were picked from the 0.1% arabinose plate and streaked on 1% arabinose plates with the aim of separate suppressor cells from nonsuppressor cells in case colonies were a mixture of both. Colonies grown from the streaks presented a healthy phenotype, as expected. No difference in size was registered between streaks deriving from big or small colonies picked from 0.1% arabinose plates (Fig.5.2).

Figure 5.2 - 1% arabinose streaks phenotype. Example of the phenotypes the colonies assumed after streak on 1% arabinose plates. A) Streak of a colony selected on 1% arabinose plate. B) Streak of a colony selected on 0.1% arabinose plate, small colony. C) Streak of a colony selected on 0.1% arabinose plate, big colony. No difference in size between small and big colonies was noticed while plated on 1% arabinose.
After the purification step, five colonies from each streak were picked and streaked on 1%, 0.5% and 0.1% arabinose plates to check the phenotype of the resulting colonies. A total of 10 colonies selected on 1% arabinose plates and 20 colonies (10 from big colonies and 10 from small colonies) selected 0.1% arabinose plates were analysed.

Colonies selected on 1% arabinose plates streaked poorly on 0.5% and 0.1% arabinose but not on 1% arabinose, retaining the original phenotype.

Every colony selected from 0.1% arabinose plates was able to grow healthily on 1% and 0.5% arabinose plates. Streaking these colonies on 0.1% arabinose gave rise to three different phenotypes. Some of the streaks looked healthy with a colony size comparable to the one of the non-suppressor strain, some showed smaller colonies and some showed impaired growth and mixed colony sizes (Fig.5.3). In particular it was observed that streaks on 0.1% arabinose, obtained from big colonies, gave rise either to healthy streaks or to poor streaks containing a mixture of big and small colonies. Streaks coming from small colonies gave rise either to healthy streaks or streaks containing smaller colonies with no mixed sizes (data not shown).
Figure 5.3 – Phenotype of suppressor strains streaked on plates. The pictures represent streaks of colonies selected on 0.1% arabinose plates in the first step, purified on 1% arabinose and streaked on 1 and 0.1% arabinose plates. A) The colony size does not change if the arabinose concentration on plate is dropped (suppressor phenotype). B) The colony size slightly decreases when the arabinose concentration is dropped (suppressor phenotype). C) Lowering the arabinose concentration completely inhibits growth (non-suppressor phenotype). The arabinose concentration present on plate is indicated at the top of the figure. Colonies selected on 1% arabinose plate in the first step all behaved like case C when streaked at lower arabinose concentration.
5.2.1.2 Analysis of suppressor clones

All the clones selected on 1% and 0.1% arabinose plates were analysed to search for mutations that could be responsible for the changed behaviour of some of them and ability to grow healthily at lower arabinose concentrations on plates.

Five possible candidate genes and promoters were selected for the analysis, being involved in $P_{BAD}$ regulation, arabinose intake, SSB protein function or ssDNA binding (Tab.5.1). Sequencing was performed on the 10 colonies isolated from the 1% arabinose plate (control) and on the 20 colonies isolated from the 0.1% arabinose plate. All the genes and promoter listed in table 5.1 were fully sequenced but no mutations were found.

<table>
<thead>
<tr>
<th>Target</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{BAD}$</td>
<td>Regulates $sib$ expression</td>
</tr>
<tr>
<td>$araC$ gene and promoter</td>
<td>Involved in negative and positive regulation of $P_{BAD}$</td>
</tr>
<tr>
<td>$araE$ gene and promoter</td>
<td>Regulates arabinose intake</td>
</tr>
<tr>
<td>$sib$ gene</td>
<td>Essential gene</td>
</tr>
<tr>
<td>$recA$ gene and promoter</td>
<td>Involved in ssDNA binding and SOS response activation</td>
</tr>
</tbody>
</table>

Tab 5.1 - Candidate target genes and promoters. In the table are listed all the obvious candidates that could have been mutated and affected growth on plates.

5.2.2 Analysis of the DL3622 suppressor strain

5.2.2.1 Growth of the DL3622 suppressor strain

One of the suppressor clones isolated from the plate containing 0.1% arabinose was purified and saved for further analysis (strain DL3622). Growth of DL3622 on plate at different arabinose concentrations was assayed by spot test. Unlike DL2844, DL3622 suppressor strain retained viability when plated at low arabinose concentrations. A slight loss in viability could be observed at $10^3 \%$ arabinose while $10^1 \%$ and $10^2 \%$ concentrations did
not affect growth (Fig.5.4). The presence of 10^{-4} \% arabinose or plain LB equally affected growth of strains DL2844 and DL3622.

\[
\begin{array}{c|c}
\text{DL2844} & \text{DL3622} \\
1 & \\
10^{-1} & \\
10^{-2} & \\
10^{-3} & \\
10^{-4} & \\
LB & \\
\end{array}
\]

\textbf{Figure 5.4 - DL3622 spot test.} A spot test was performed on plates containing different arabinose concentrations. DL3622 showed the ability to grow at lower arabinose concentrations when compared to the non-suppressor strain DL2844.

A growth curve was performed to assess the arabinose sensitivity of DL3622 in liquid medium. The strain was grown as described in materials and methods, the OD_{600} of the culture was assessed every 30 minutes and the viability every 120 minutes by spot test. As the graphs show, the behaviour of the suppressor strain in liquid culture was comparable to the one of the non-suppressor strain (Fig.5.5). Growth was not affected at 10^{-2} \% and 10^{-3} \% arabinose concentrations. When grown in 10^{-4} \% arabinose, the OD_{600} stopped increasing after about 210 minutes growth for both strains. The viability test confirmed the fact that DL2844 and DL3622 strains behaved similarly in liquid culture. A loss of viability was observed after about 180 minutes growth for both strains (Fig.3.4 and 5.5).
Figure 5.5 - Growth and viability in liquid medium. DL2844 and DL3622 were grown in liquid culture in the presence of different arabinose concentrations as specified in the legend. DL1582 was used as a control. The OD_{600} was measured every 30 minutes and viability tested at time 60, 180 and 300 minutes. The behaviour of the suppressor strain in liquid culture followed exactly the one of the non-suppressor strain. During growth, the viability of DL3622 (above) was comparable to the one of DL2844 (see Figure 3.4).

During growth in liquid culture in the presence of different arabinose concentrations, the SSB level of DL3622 was assessed by western blotting to compare it with the SSB level present in DL2844 strain. The behaviour observed was the same as for DL2844 (data not shown). At 10^{-4} % arabinose the SSB level decreased dramatically after about 60 minutes growth, while the SSB level was higher than wild type for higher arabinose concentrations.
5.2.2.2 U.V sensitivity of DL3622 suppressor strain

The U.V sensitivity of DL3622 was assessed by spot test on plates containing different arabinose concentrations. U.V. dosages of 5,000μJ, 10,000μJ, 15,000μJ and 20,000μJ were tested. DL3622 did not display increased U.V. sensitivity when plated at lower arabinose concentrations and U.V. irradiated (Fig.5.6).

Strains DL1582 and DL2844 were used as controls. DL1582 showed between one and two order of magnitude viability loss when irradiated with 20,000μJ power. No effect on viability was observed for lower irradiation intensities. To cause the same effect on DL3622 and DL2844 strain, respectively 15,000μJ and 10,000μJ were sufficient consistent with the fact that SSB expression regulation is crucial for correct cell growth.
Chapter V – Suppresser Strains Isolation and Analysis

Figure 5.6 - DL3622 suppressor strain U.V. sensitivity test. U.V. sensitivity of strains DL1582, DL3622 and DL2844 was assessed at different arabinose concentrations. 5,000µJ, 10,000µJ, 15,000µJ and 20,000µJ powers, were used. DL3622 did not show increased U.V. sensitivity correlated to low arabinose concentrations present on plates but the strains showed increased U.V. sensitivity when compared to DL1582 and decreased U.V. sensitivity when compared to DL2844. DL1582 strain behaviour was constant for all the arabinose concentrations tested (data not shown).

5.2.2.3 Genome sequencing of the DL3622 suppressor strain

As the partial sequencing analysis performed on the clones isolated from 1 and 0.1% arabinose plates did not give any indication about the origin of the suppressor mutation, the
whole genome of strain DL3622 was sequenced to search for point mutations or insertions and deletions that could have caused the appearance of the suppressor phenotype.

Solexa® Technology was used to sequence the genome. DNA samples of strains DL1582, DL2844 and DL3622 were analysed and the sequences obtained aligned against *E. coli* K-12 genome sequence. DL3622 sequence was compared to DL1582 and DL2844 sequences and six point mutations, present in DL3622 only, were found (Tab.5.2).

The most promising of those candidates is the *ydcN* gene, predicted to be a DNA-binding transcriptional regulator. The mutation causes an amino acid change from glycine to glutamic acid, which present very different structural and chemical characteristics.

Unfortunately no time was left to characterise this mutation and to prove it is involved in the phenotype change. The first and easiest approach to follow would probably be to amplify and sequence *ydcN* region in strains DL2844 and DL3622 to confirm the mutation is present in DL3622 only. A plasmid containing the wild type sequence of *ydcN* could be created and the un-mutated gene inserted in DL3622 strain at the place of the mutated one. A reverse approach could be followed too, cloning the mutated gene and inserting it in DL2844 strain to check its growth on plates.

We are still waiting for the results relative to insertion and deletions that may be present in the genome and that may be the real cause of the phenotype change.
Chapter V – Suppressor Strains Isolation and Analysis

<table>
<thead>
<tr>
<th>Position</th>
<th>Mutation</th>
<th>Target</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>568383</td>
<td>G-T</td>
<td>ybcK</td>
<td>Predicted recombinase</td>
</tr>
<tr>
<td>568384</td>
<td>A-C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1309854</td>
<td>A-G</td>
<td>tonB promoter</td>
<td>tonB gene: energy transducer, involved in iron and cyanocobalamin uptake, phage sensitivity</td>
</tr>
<tr>
<td>1309855</td>
<td>C-T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1504684</td>
<td>G-A</td>
<td>ydcN</td>
<td>Predicted DNA-binding transcriptional regulator</td>
</tr>
<tr>
<td>4547955</td>
<td>T-A</td>
<td>gntP promoter</td>
<td>gntP gene: Gluconate Gnt transporter</td>
</tr>
</tbody>
</table>

Table 5.2 - Point mutations. DL3622 genome sequence was aligned to DL1582 and DL2844 genome sequences. Five point mutations were found, the characteristics are listed in the table above.

5.3 Discussion and conclusions

Suppressor strains may appear and be selected due to a change of the environmental conditions the bacterial population has to cope with. SSB is an essential protein and cells that are put through a drop of the protein level may have the chance to overcome the problem and survive if a favourable mutation appears in the genome.

Interestingly we managed to select a suppressor strain in which growth on plate and in liquid culture required a comparable amount of arabinose. Growth on plate was improved at low arabinose concentration when compared to the non-suppressor strain. This may have more than one explanation.

The suppressor’s ability to uptake arabinose from the growth medium may be increased or mutation in the $P_{BAD}$ regulation system may have arisen, allowing a higher expression of SSB even at low arabinose concentrations. When discussing the differences between growth on plate and liquid culture in chapter three, it was mentioned that cells that get transferred from a growing culture to plates, go through stress and may lose the ability to form colonies. It was shown that such “injured cells” can re-gain the ability of forming colonies if reactive oxygen...
Chapter V – Suppressor Strains Isolation and Analysis

species that accumulate in the growth medium get removed (McDonald, Hackney et al. 1983; Mackey and Seymour 1987; Wang, Chang et al. 1993; Dukan and Nystrom 1999; Cuny, Lesbats et al. 2007). A mutation in the regulation of those genes’ transcription may lead to a more efficient transcription or mutations in the genes themselves may make the enzymes work in a more efficient way. A first investigation of the suppressors we isolated was done by sequencing candidate genes involved either in $P_{BAD}$ and arabinose transport regulation or involved in ssDNA binding. Those candidates were the most probable and expected options for mutations to appear. A mutation in the promoter regulating the transcription of the arabinose transporter AraE would have improved arabinose uptake leading to increased SSB expression. A mutation in the araC repressor gene or $P_{BAD}$ would have led to a partial loss of the arabinose dependent regulation of $P_{BAD}$ and therefore resulted in increased SSB expression level. Mutations in the $ssb$ gene were more interesting candidates with the possibility of finding a new isoform of SSB protein able to maintain its role with the same effectiveness even when present at a lower level. RecA is involved in ssDNA binding and SOS response induction. $recA$ mutants have the ability to bind ssDNA with higher affinity and to displace SSB protein in a more efficient way (Wang, Chang et al. 1993). A prompt SOS response activation may be crucial for cells to overcome the stress they are put through when transferred onto plates. A $recA$ mutation may elicit a more efficient SOS response cascade overcoming the growth problems the cells may encounter. None of those prime candidates was however found responsible for the phenotype change of the suppressor strains and a deeper analysis was necessary to solve the problem.

One of the suppressor strains isolated, DL3622, was purified and fully characterised. Roughly the same quantity of arabinose was required by DL3622 to grow in liquid and solid culture while the non-suppressor strain DL2844 required a much higher concentration of arabinose to grow on solid medium. Both strains died when plated on LB, demonstrating
Chapter V – Suppressor Strains Isolation and Analysis

that the suppressor strain still requires arabinose for growth. The SSB expression level of DL3622 was not different from the SSB expression level of DL2844 at the same arabinose concentrations in liquid culture. Growth curves and viability tests were also comparable for the two strains. This proved that the phenotype change of the suppressor strain was only related to growth on solid medium and that the quantity of arabinose present in liquid medium was not influencing the growth pattern if compared to the non-suppressor strain.

U.V. sensitivity tests revealed differences between strains DL3622 and DL2844. While U.V. irradiation stimulated the development of a partially mutated phenotype in DL2844 for arabinose concentrations higher than 10^{-4} \% and lower that 1\%, the same effect was not found for DL3622. Interestingly, DL3622 displayed higher U.V. resistance than DL2844 when exposed to high radiation powers (15,000\mu J and 20,000\mu J) even at high arabinose concentrations, highlighting the difference between the two strains and their different ability of surviving on plate. DL3622 requires lower arabinose concentrations that DL2844 to survive on plates. If plated at the same arabinose concentration, DL3622 may therefore be able to produce a higher quantity of SSB than DL2844. This would provide the strain with the ability of repairing U.V. damage more efficiently.

The whole genome sequence of DL3622 suppressor strain led to the location of a mutation present in ydcN gene. ydcN is predicted to code for a DNA-binding transcriptional regulator which would be an interesting candidate responsible for phenotype change of DL3622. Unfortunately no time was left to analyse the role of ydcN mutation. The protein may be involved in the regulation of P_{BAD}, responsible for ssb expression level control or regulation of the transcription of either araC or araE, responsible respectively P_{BAD} regulation and arabinose uptake.
Chapter VI

SSB and Palindrome Instability
Chapter VI – SSB and Palindrome Instability

6.1 Introduction

The potential for palindromes to cause problems during replication is thought to be bound to their ability to form stable DNA hairpins; such hairpins can be attacked and cleaved by the SbcCD nuclease (Connelly, Kirkham et al. 1998). The cleavage leads to the formation of a DSB that can be repaired by homologous recombination, allowing the palindrome to be kept (Eykelenboom, Blackwood et al. 2008). Hairpins are likely to form, during replication, on the lagging strand template where ssDNA stretches are present and would allow the folding of the palindrome. SSB can bind ssDNA and may therefore be involved in impeding the formation of palindrome hairpins. In this work we investigated both the effects of SSB depletion and SSB overexpression on the stability of a 105bp and a 246bp imperfect palindromes. Both palindromes were inserted at the beginning of the lacZ gene and their presence disrupts LacZ functionality. SSB overexpression was obtained by transforming cells with pAM34/sib in which the SSB gene is controlled by an IPTG inducible promoter. The change of the SSB expression level is showed in figure 7.1

6.1.1 105bp palindrome instability assay

The 105bp palindrome derives from a 246bp imperfect palindrome as a deletion product (Pinder, Blake et al. 1998) and was used to assess the potential of palindromes to form hairpins when the SSB level was lower than normal. The lacZ tract downstream the palindrome was engineered to create a sequence identical to an 8bp sequence located in the initial tract of the palindrome (Fig.6.1) as thoroughly described in materials and methods. Palindrome deletions may occur, when a hairpin is formed, by polymerase slippage and replication going across the hairpin. Having direct repeats flanking the palindrome has been shown to enhance polymerase slippage and precise deletion events (Pinder, Blake et al. 1998).
The construct was used as a tool to increase the hairpin-detection sensitivity of the assay. LacZ functionality being disrupted by the presence of the palindrome, cells carrying the palindrome could be distinguished from cells that had lost the palindrome by plating them on IPTG supplemented medium and by checking the colour of the resulting colonies.

**Figure 6.1 - 105 bp imperfect palindrome construct.** The 105bp imperfect palindrome was inserted in lacZ gene as shown in figure. A) black: lacZ gene; blue: palindrome; orange: complementary sequences. B) Sequence of the construct, same colour code. The central portion of the palindrome sequence is not represented.

### 6.1.2 246bp palindrome instability assay

SSB overexpression experiments were performed on recA strains carrying a 246bp imperfect palindrome and in which the sbcDC genes were under the control of P<sub>BAD</sub> sbcDC genes’ expression could therefore be repressed by growing cells in the presence of glucose and activated by adding arabinose to the growth medium.

SSB overexpression was obtained by transforming one of the strains with pAM34/sbr and by growing the transformants in the presence of IPTG and ampicillin. As it has been previously shown, the induction of SbcCD expression leads to a severe loss of cell viability, in recA strains containing a 246bp imperfect palindrome (Eykelenboom, Blackwood et al. 2008). Palindrome hairpins can in fact be attacked by the SbcCD nuclease resulting in the creation of a DSB that can only be repaired via homologous recombination.
Chapter VI – SSB and Palindrome Instability

In this work, the viability of \( ncoA \) strains containing the palindrome was assessed, after SbcCD induction, at normal and overexpressed SSB levels. SSB overexpression was expected to decrease the number of hairpins formed by the palindrome leading to increased viability of the \( ncoA \) strains containing the palindrome after SbcCD induction.

6.2 Results

6.2.1 SSB depletion lowers the stability of a 105bp imperfect palindrome flanked by short direct repeats

A 105bp imperfect palindrome was introduced at the beginning of the \( lacZ \) gene as mentioned before. The instability of the palindrome was measured as a function of the arabinose concentration in a plate-based assay. Strains DL2959 and DL2957, DL1582 and DL2844 containing the 105p palindrome inserted in \( lacZ \) respectively, were used for the analysis. The strains were grown overnight in 1% arabinose and LB, grown maintaining the same conditions for 1.5 hours to make cells entering the exponential phase and then plated at different arabinose concentrations (1%, 0.5%, 0.1% and 0.05%).

Interestingly, the appearance of blue colonies could not be observed after a first overnight growth at 37°C, but the presence of blue sectors or blue papillae in the colonies was observed after about 3 days of growth. Blue sectored colonies were streaked on IPTG containing plates and gave rise to a mixture of blue and white colonies (Fig.6.2). A number of such colonies were analysed by PCR to detect the presence or absence of the palindrome. All the blue colonies analysed had lost the palindrome while it was still present in the white colonies, as expected (data not shown).
The total number of blue-sectored colonies appearing on plates was counted and the percentage of blue-sectored colonies calculated over the total number of colonies present on plates. Strikingly, SSB depletion had an effect on palindrome instability which greatly increased. Blue-sectored colonies started appearing at 0.5% arabinose concentration and their number increased as soon as the arabinose concentration was lowered (Fig.6.3).

Figure 6.2 - Plating of a blue sectored colony. A colony presenting a blue sector was re-suspended in 200µl LB medium. Dilutions of it were plated on 1% arabinose and IPTG supplemented LB-agar and grown overnight at 37°C. A mixture of blue and white colonies resulted indicating the presence or absence of the 105bp palindrome.
Figure 6.3 - Palindrome deletion assay results. The percentage of blue sectored colonies found on plates was plotted against the arabinose percentage used. As shown in the graph, the palindrome deletion frequency increased while the arabinose percentage, and therefore the SSB level, was lowered. The error bars represent the standard error (standard deviation/sqrt(n)) where n=number of experiments) of three separate experiments.

6.2.2 SSB overexpression does not improve the viability of recA strains carrying a 246bp imperfect palindrome

The viability of recA strains containing a 246bp palindrome was measured, after induction of the SbcCD nuclease expression, both at normal (DL2075) and overexpressed (DL3750) SSB level. Strains DL2006 and DL3749 do not contain the palindrome and were used as control for normal and overexpressed SSB respectively.

To assess viability on plates, the strains were grown overnight in 0.5% glucose to repress sbeDC transcription. After washing cells twice in LB, a spot test was then performed on both on 0.5% glucose and 0.2% arabinose supplemented plates. SSB overexpression did not
improve the viability of recA strains containing the palindrome when SbcCD was expressed (Fig.6.4).

**Figure 6.4 – Viability of recA strains containing the 246bp palindrome.** The viability of strains DL2006, DL2075, DL3749 and DL3750 was assessed in absence and presence of SbcCD and normal and overexpressed SSB level. The conditions used on plate are listed on the side of the pictures, strains number are on top. SSB overexpression did not improve the viability of recA- strains carrying the palindrome and in which SbcCD expression was induced.

Viability was assessed in liquid medium too. Overnight cultures were grown in 0.5% glucose, each culture was split and 0.2% arabinose was added to one sample. Cells were grown for 240 minutes, the OD$_{600}$ measured every 30 minutes and viability assessed every 120 minutes by spot test. DL3750 strain stopped increasing OD$_{600}$ after about 120 minutes growth while DL2075 OD$_{600}$ stopped increasing a little later, at time 150 minutes. Interestingly, strains in which SSB was overexpressed managed to exit stationary phase before strains in which SSB level was normal (Fig.6.5). The viability of the strains grown in 0.2% arabinose dropped earlier than the OD$_{600}$ dropped, as expected (Fig.6.6). In liquid culture SSB overexpression seemed not to have an effect on the viability of recA strains carrying the palindrome and in which the expression of sbcDC was induced.
Figure 6.5 – Strains DL2075 and DL3750 growth curve. DL2075 and DL3750 strains were grown in glucose and arabinose supplemented media. DL3750, in which SSB was overexpressed, exited stationary phase in a faster way. DL2075 and DL3750 grown in arabinose stopped growing after 150 and 120 minutes respectively.

Figure 6.6 - Viability in liquid culture. The viability of strains DL2075 and DL3750 was measured in liquid culture in the presence or absence of SbeCD protein induction. Overexpression of SSB did not improve the viability of DL3750 when SbeCD was induced.
6.3 Discussion and Conclusions

The SSB protein is thought to be able to impede the formation of DNA secondary structures by binding and protecting ssDNA. In this work we investigated the ability of SSB to impede the formation of palindrome DNA hairpins using two different approaches and substrates.

The instability of a 105bp imperfect palindrome was investigated at low SSB levels on plates. Lowering the SSB level resulted in increased palindrome deletion frequency. The deletion frequency registered was however very low and it was measured as the percentage of blue sectored colonies appearing on plate. Totally-blue colonies could not be detected suggesting no deletion events happened at the early stages of colony growth. Blue sectors could only be detected after 3 days of growth, stressing the fact that deletions happen with very low frequency. From the data obtained we can state that a lower level of SSB protein does have an effect on hairpins formation, even if the effect we could detect with our system was small. The importance that SSB depletion has on palindrome instability may be affected by two factors. SSB depletion may increase the rate of palindrome hairpins formation. In addition SSB has been reported to have an effect on polIII slippage in vitro. SSB can increase the processivity of polIII thus increasing its slippage rate across a palindrome substrate (Cancell and Ehrlich 1996). The assay we were using is based on slippage and these two effects elicited by SSB depletion act in opposite ways on palindrome instability detection and this may explain the very low frequency of palindrome deletion we observed. Also, lowering the level of SSB is predicted to slow replication down. As polymerase slippage is replication dependent, this could be another explanation of why the frequency of palindrome deletion events observed was so low.

The potential for a 246bp palindrome to form hairpin structures was investigated using a different approach. recA strains containing the 246bp imperfect palindrome lose viability
when the SbcCD nuclease expression is induced. SbcCD is believed to cleave palindrome hairpins generating a double strand break that has to be repaired by homologous recombination (Eykelenboom, Blackwood et al. 2008). We were expecting that SSB overexpression would rescue such strains viability, by reducing the formation of hairpins formed during DNA replication and therefore reducing the frequency of hairpin processing. The results obtained showed however no positive effect of SSB overexpression on viability. This may be consistent with the fact that palindromes can form very stable hairpins and that the dynamics of hairpin formation is really fast. SSB excess would not be able to impede the formation of such strong hairpins.

Interestingly, SSB overexpression stimulated the exit from stationary phase of the DL3750 recA strain. After 60 minutes of growth, DL3750 growth rate was comparable to the one of the control strain DL2075.
Chapter VII

SSB and CAG·CTG Instability
7.1 Introduction

7.1.1 CAG·CTG repeat instability

As previously described in the introductory section of this thesis, the expansion of trinucleotide repeat tracts underlies several neurological and neurodegenerative diseases. Expansions of CAG·CTG trinucleotide repeat arrays are in particular responsible for nine illnesses, known as polyglutamine disorders.

Instability of trinucleotide repeat (TNR) sequences has been extensively studied in bacteria, yeast, mouse models and cultured cells. It has been shown to be dependent on factors bound to the nature of the repeat array such as length and purity, orientation and distance from the origin of replication. External factors like proteins involved in replication, recombination and repair processes have also been proven to be involved in modulating stability (Maurer, O'Callaghan et al. 1996; Petruska, Arnheim et al. 1996; Freudenreich, Stavenhagen et al. 1997; Cleary, Nichol et al. 2002; Zahra, Blackwood et al. 2007).

DNA replication instability models are based on the fact that CAG·CTG repeat arrays can form DNA secondary structures that, if not erased, may lead to polymerase slippage. Contractions may arise if the secondary structures form on the template strand, expansions if structures form on the nascent strand. During replication, secondary structures are more likely to form on the lagging strand template, where ssDNA stretches are created after unwinding of the double helix. This may explain why, in rapidly dividing organisms, the expansion rate is much lower than the contraction rate. CAG·CTG repeat array instability has also been proven to be orientation dependent. When CAG repeats are on the leading strand template, the instability of the sequence is much greater than it is when CTG repeats are on the leading strand template. The explanation for this may be bound to the relative stability of
CAG and CTG hairpins. *In vitro* studies showed in fact that CTG sequences form more stable hairpins than CAG sequences, due to the higher bulkiness created by mis-paired adenines.

### 7.1.2 SSB and CAG-CTG repeat instability

The role of *E. coli* SSB on CAG-CTG repeat arrays instability was investigated in a plasmid based system using the *ssb*·1 temperature sensitive mutant form of the protein (Rosche, Jaworski et al. 1996). It was demonstrated that affecting the functionality of SSB greatly increased the instability of (CTG)\(_{100}\) and (CTG)\(_{180}\) repeats when the CTG sequence was on the leading strand template.

Recently, the role of the *E. coli* SSB protein on polymerase progression through CAG-CTG repeat arrays was also investigated by primer extension *in vitro* assay and a new orientation-dependent instability model was proposed. The authors found a slowdown in the polymerase progression while copying through CAG repeat tracts but not through CTG tracts. Adding SSB to the reaction seemed to facilitate polymerase progression through CAG repeat tracts. The authors proposed a “template push” model which relies on the formation of hairpins on the leading stand template rather then on the lagging strand template. During replication of the CAG-CTG repeat array (with CAG repeats present on the leading strand template) and at normal SSB levels, a differential speed between leading and lagging strand polymerase would be generated. To avoid uncoupling of the two polymerases, the leading strand polymerase would pull towards the helicase, leading to the creation of a CAG hairpin on the leading strand template (Delagoutte, Goellner et al. 2008). This model would explain orientation-dependent instability through a mechanism which does not rely on hairpin stability.
7.1.3 SbcCD and CAG-CTG repeat instability

SbcCD nuclease (Rad50/Mre11 homologue) is also believed to have a role in TNR instability as it can attack hairpin structures \textit{in vitro} (Connelly, de Leau et al. 1999). SbcCD nuclease has also been demonstrated to cleave palindrome hairpins \textit{in vivo}. These hairpins are believed to form on the lagging strand template during replication, their processing leading to the formation of double strand breaks (Eykelenboom, Blackwood et al. 2008). A destabilising effect of SbcCD on leading strand CAG repeat arrays was observed in \textit{E. coli}. In the same work, the absence of SbcCD, in a proofreading deficient (\textit{dnaQ}) strain, has been shown to change the deletion size distribution profile of leading strand CTG repeat arrays. These findings support the hypothesis that DNA secondary structures and their processing are involved in TNR instability (Zahra, Blackwood et al. 2007).

7.1.4 Aim of the work

In this work we investigated the role of SSB protein overexpression and SbcCD nuclease on CAG-CTG repeat instability in the \textit{E. coli} chromosome. The instability of a leading strand (CAG)\textsubscript{75} and a leading strand (CTG)\textsubscript{137} repeat array were studied giving insights into the role and mechanism by which SSB can act on CAG-CTG repeat instability. The instability of the TNR lengths in the chromosome of a \textit{wild type} \textit{E. coli} strain was previously quantified (Zahra, Blackwood et al. 2007) and they were chosen because they present a similar instability level which allowed direct comparison between the two orientations.
7.2 Results

7.2.1 SSB protein level increases in cells carrying the pAM34/ssb plasmid

A system that allowed overexpressing SSB protein without influencing the viability of the strains examined was set up. The plasmid pAM34/ssb, in which the ssb gene is under the control of its natural promoter, was used to increase the SSB expression level. Western blotting was used to assess the SSB protein level. Preliminary experiments were performed on overnight cultures of control strains and strains carrying pAM34/ssb. We found an increase in SSB protein expression level of about 30-40 times in strains containing the plasmid and grown in the presence of IPTG and ampicillin (Fig.7.1). IPTG is needed for the replication of pAM34 plasmid.

![Figure 7.1 - SSB level assessment in overnight cultures](image)

**Figure 7.1 - SSB level assessment in overnight cultures.** SSB protein level was measured after overnight growth in the presence or absence of pAM34/ssb plasmid. The first two columns refer to DL3311 and DL3785 strains (normal SSB level) while the second two columns refer to DL3714 and DL3788 strains (overexpressed SSB via pAM34/ssb plasmid). 5µl of a 1:10 dilution of the cell lysate sample were loaded for DL3714 and DL3788 strains to be able to compare the SSB level in the presence of pAM34/ssb plasmid to the SSB
physiological level. The SSB level appears to be 30-40 times higher in strains carrying the plasmid grown in the presence of IPTG and ampicillin.

To assess CAG-CTG instability variation over time, an instability assay was designed as described in materials and methods. The instability assay was performed on both (CAG)$_{35}$ and (CTG)$_{135}$ repeat containing strains, at normal and high SSB level. Cells were grown for 100 generations; instability and SSB level were monitored during growth at intervals of 20 generations. Control strains (DL3311 and DL3785) maintained a constant SSB expression level while strains grown in IPTG and ampicillin for the first overnight (DL3714 and DL3788) showed a drop of the SSB level as soon as the drugs were depleted from the growth medium. Strains grown in IPTG and ampicillin for the whole duration of the experiment (DL3714 IA and DL3788 IA, IA standing for IPTG and ampicillin) maintained a uniform SSB expression level which was higher than the physiological one (Fig.7.2).
Figure 7.2 - SSB level assessment during instability assay. To assess the SSB protein level over the duration of the experiment, samples were taken at generation 0, 20, 40, 60, 80 and 100. Generation 0 refers to samples taken after a first overnight growth. DL3311 (CAG\textsubscript{75}) and DL3785 (CTG\textsubscript{140}) were grown in LB, DL3714 (CAG\textsubscript{75} pAM34/SSB) and DL3788 (CTG\textsubscript{140} pAM34/SSB) were grown in IPTG and ampicillin supplemented medium for the first overnight and then grown in LB. DL3714 IA and DL3788 IA were grown in IPTG and ampicillin supplemented medium for the whole duration of the experiment. Western blot analysis was carried out using α-SSB antibody and band intensity was measured using ImageJ program. Panels C/D and E/F refer to the same experiment but show two different gels. Band intensity values obtained from the two gels were plotted in the same graph (A and B) even if not directly comparable.

7.2.2 Overexpression of SSB protein stabilises the (CAG)\textsubscript{75} repeat array but does not influence the stability of the (CTG)\textsubscript{137} repeat array both in \textit{sbcDC}\textsuperscript{+} and \textit{sbcDC}\textsuperscript{-} backgrounds

To assess the instability level of (CAG)\textsubscript{75} and (CTG)\textsubscript{137} repeat arrays over time, an instability assay was carried out as described in materials and methods. Strains were grown
for 100 generations and instability was assessed at intervals of 20 generations by plating cells and measuring the length of the TNR arrays of the resulting colonies. Only parental length variations happening in liquid culture were counted as instability events.

Overexpression of SSB protein had a dramatic effect on (CAG)$_{75}$ repeat instability in a $sbeDC^+$ background (Fig.7.3A). When a normal level of SSB was maintained (DL3311), the instability level steadily rose as a function of the generation number. Increasing the SSB level in the first overnight, by inducing pAM34/sib transcription, led to a lower instability level in the overnight culture (DL3714). After the SSB level was brought down to normal by depleting IPTG and ampicillin in the growth medium, the instability increased following the trend of the control strain (Tab.7.1). Finally, SSB overexpression (DL3714 IA) caused a remarkable decrease of the (CAG)$_{75}$ repeat array instability. The instability level of (CTG)$_{137}$ repeat tract was on the other hand completely unaffected by overexpressing SSB protein (Fig.7.4A).

The same experiment was repeated on $sbeDC$ strains to investigate the role of SbcCD nuclease on TNR repeat instability upon SSB overexpression. The role of SbcCD nuclease on CAG-CTG instability was investigated by Zahra et al. (2007) The lack of SbcCD had a 1.8 fold stabilising effect on a (CAG)$_{25}$ repeat array while no effect was detected for a (CTG)$_{95}$ repeat array (Zahra, Blackwood et al. 2007). In our case and in the conditions tested, the absence of SbcCD nuclease did not show a marked effect on the (CAG)$_{75}$ repeat array instability (Fig.7.3B). This may be due to the different nature and sensitivity of the assay carried out. The slopes of the best fit lines relative to DL3714-IA and DL3715-IA strains were comparable suggesting that stabilisation mediated by excess SSB occurs in both $sbeDC^+$ and $sbeDC$ backgrounds. No marked SbcCD effect on instability accumulation was observed in (CTG)$_{137}$ strains, either in absence or presence of overexpressed SSB (Fig.7.4B). In the absence of SSB overexpression, a steeper instability increase was observed for (CAG)$_{75}$
strains when compared to (CTG)\textsubscript{137} strains, despite the remarkable difference in length. This supports the finding that instability is strongly dependent on repeat array orientation.

![Graph A](image1)

**Figure 7.3 - (CAG)\textsubscript{75} Instability assay graphs.** A) (CAG)\textsubscript{75} repeat instability in *sbcDC*\textsuperscript{+} (DL3311 and DL3714) background. B) (CAG)\textsubscript{75} repeat instability in *sbcDC* (DL3548 and DL3715) background. Overexpression of SSB (DL3714 IA and DL3715 1A) stabilises the repeat array. Strains DL3311 and DL3548 were grown in LB. Strains DL3714 and DL3715 were grown in IPTG and ampicillin supplemented medium for the first overnight growth (generation 0) and then grown in LB. Strains DL3714 IA and DL3715 IA were grown in IPTG and ampicillin supplemented medium for the whole duration of the experiment. The lines represent the best fit lines interpolating the experimental points. Each experiment was repeated twice and the standard error was used to draw error bars. DL3311 (CAG\textsubscript{75}); DL3714 (CAG\textsubscript{75} pAM34/ssl); DL3548 (CAG\textsubscript{75} *sbcCD*); DL3715 (CAG\textsubscript{75} *sbcCD* pAM34/ssl).
Figure 7.4 – \((\text{CTG})_{137}\) Instability assay graphs

A) \((\text{CTG})_{137}\) repeat instability in \(sbcDC^+\) (DL3785 and DL3788) background. B) \((\text{CTG})_{137}\) repeat instability in \(sbcDC\) (DL3799 and DL3800) background. Overexpression of SSB (DL3788 IA and DL3800 IA) does not have an effect on the repeat array stability. Strains DL3785 and DL3799 were grown in LB. Strains DL3788 and DL3800 were grown in IPTG and ampicillin supplemented medium for the first overnight growth (generation 0) and then grown in LB. Strains DL3788 IA and DL3800 IA strains were grown in IPTG and ampicillin supplemented medium for the whole duration of the experiment. The lines represent the best fit lines interpolating the experimental points. Each experiment was repeated twice and the standard error was used to draw error bars.
<table>
<thead>
<tr>
<th>DL3311</th>
<th>DL3714</th>
<th>DL3714 IA</th>
<th>DL3548</th>
<th>DL3715</th>
<th>DL3715 IA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0325</td>
<td>0.0371</td>
<td>0.0075</td>
<td>0.0325</td>
<td>0.0469</td>
<td>0.0016</td>
</tr>
<tr>
<td>DL3785</td>
<td>DL3788</td>
<td>DL3788 IA</td>
<td>DL3799</td>
<td>DL3800</td>
<td>DL3800 IA</td>
</tr>
<tr>
<td>0.0203</td>
<td>0.0215</td>
<td>0.0181</td>
<td>0.0247</td>
<td>0.0186</td>
<td>0.0181</td>
</tr>
</tbody>
</table>

Table 7.1 - Best fit lines slopes. In table are shown the values relative to the slopes of the best fit lines for each of the strains analysed.

7.2.3 (CAG)₇₅ repeat array deletion size distribution profile is affected by SSB overexpression and by the lack of the sbcCD nuclease

In rapidly dividing organisms, contractions occur with substantially higher frequency then expansions and a bias towards big deletions in the *E. coli* chromosome has also been observed (Zahra, Blackwood et al. 2007). That may be consistent with contractions caused by the formation of big DNA hairpins or by multiple small deletions. To analyse the deletion size distribution profile of the strains used in this work all the data available on deletions, which occurred both on plate and in liquid culture, were pooled. For events that occurred on plate, single deletion peaks accompanied by a parental length peak were considered for the analysis. The percentage of the deleted sizes relative to the parental length was plotted against the frequency of the deletion events, expressed as percentage of the total number of deletion events. The deletion distribution profile of strains DL3714, DL3715, DL3788 and DL3800 was created omitting data deriving from the first overnight growth, carried out at high SSB levels.

In strains carrying the (CAG)₇₅ repeat array at a normal SSB level, the profiles of the curves were skewed towards big deletions both in *sbcDC*⁺ and *sbcDC* backgrounds (Fig.7.5A, B, D, E) When SSB was overexpressed the profile of the curve dramatically changed presenting a flat shape with no preferred deletion sizes (Fig.7.5C, F).
Chapter VII – SSB and CAG-CTG Instability

At a normal SSB level, even though the absence of SbcCD nuclease did not have a noticeable effect on (CAG)$_{35}$ instability under the conditions tested, the analysis of the deletion size distribution profile revealed a slight shift of the curve towards smaller deletions in the $sbeDC$ mutant strain (Fig. 7.5). This may suggest the ability of SbcCD to target CTG hairpins forming on the lagging strand template during replication, showing size selectivity on substrates choice.
Figure 7.5 - (CAG)$_{75}$ deletion size distribution profile curves. Deletion size distribution graphs were drawn by plotting the percentage of the deleted array length relative to the parental length (percentage of parental length) against the percentage of the event frequency (percentage of event frequency) calculated on the total number of deletion events. The best fit curve modelling the distribution of the points was approximated using a 4th order polynomial using Origin program. A, B, C) (CAG)$_{75}$. D, E, F) (CAG)$_{75}$ sbcDC. C, F) overexpressed SSB protein.
7.2.4 Lack of sbcCD changes the deletion size distribution profile in (CTG)$_{137}$ strains

It has been previously shown that in a sbcDC dnaQ double mutant, (CTG)$_{95}$ repeat array instability increased and that the deletion size distribution bias towards big deletions was disrupted (Zahra, Blackwood et al. 2007). In this work we show that SbcCD nuclease has, on its own, an effect on the deletion size distribution profile of a (CTG)$_{137}$ repeat arrays and that the effect is independent of SSB overexpression. In the absence of SbcCD nuclease, the bias towards big deletions was partially lost (Fig.7.5) suggesting a role for SbcCD in recognising big CAG hairpins possibly forming on the lagging strand template.
Figure 7.6 - (CTG)$_{127}$ deletion size distribution profiles. Deletion size distribution graphs were drawn by plotting the percentage of the deleted array length relative to the parental length (percentage of parental length) against the percentage of the event frequency (percentage of event frequency) calculated on the total number of deletion events. The best fit curve modelling the distribution of the points was approximated using a 4th order polynomial. A, B, C) (CTG)$_{127}$. D, E, F) (CTG)$_{127}$ sbcDC. C, F) overexpressed SSB protein.
7.2.5 The proportion of the number of expansions is not influenced by SSB overexpression both in \(sbcDC^+\) and \(sbcDC\) backgrounds

The mechanism involved in the formation of expansion is yet not well understood and models of possible pathways have been proposed, as described in the introductory section of this work. We calculated the proportion of expansion events out of the total number of instability events and found that SSB overexpression did not influence the proportion of the number of expansions for both (CAG)\(_{75}\) and (CTG)\(_{137}\) repeat arrays (Fig.7). For the analysis, both expansions and deletions occurred in liquid culture and on plate were counted as instability events, as previously described for the deletion distribution profile.
Figure 7.7 - Proportion of expansions. The proportion of expansions was calculated on the total number of instability events (expansions + deletions). A) SbcDC+ expansions proportion. B) SbcDC expansions proportion. The conditions specified in the graph refer to plating conditions (LB or IPTG+ampicillin). Error bars were calculated using the standard error between two different experiments. DL3311 (CAG3); DL3785 (CTG180); DL3714 (CAG75 pAM34/xib); DL3788 (CTG140 pAM34/xib); DL3548 (CAG75 sbcCD); DL3715 (CAG75 sbcCD pAM34/xib); DL3788 (CAG75 sbcCD pAM34/xib); DL3799 (CAG75 sbcCD pAM34/xib); DL3800 (CAG140 sbcCD pAM34/xib).
7.3 Discussion

7.3.1 SSB overexpression stabilises CAG-CTG repeat arrays in an orientation dependent manner in the *E. coli* chromosome

As mentioned in the introductory session of this chapter, according to a widely accepted model, orientation dependent instability would be caused by the different ability of CAG and CTG repeats to form hairpins. It has in fact been shown that CTG repeats can form more stable hairpins than CAG repeats *in vitro* (Miret, Pessoa-Brandao et al. 1997). DNA hairpins may have the chance to form on the lagging strand template during replication. The potential to fold is given by the fact that stretches of ssDNA form on the lagging strand template after unwinding of the double helix. CTG hairpins would form more easily than CAG hairpins thus causing greater instability problems. Therefore the CAG orientation, where CTG repeats lie on the lagging strand template, would be less stable than the CTG orientation.

The results obtained for the (CAG)\textsubscript{15} repeat array are consistent with the fact that, SSB overexpression may impede the formation of CTG hairpins potentially forming on the lagging strand template, by more efficiently binding ssDNA stretches forming at the replication fork. This would dramatically decrease the chances the CTG arrays have to fold and to cause stability issues.

The stabilising effect given by SSB overexpression was not observed for the (CTG)\textsubscript{137} repeat array. In this case CAG hairpins may have the chance to form on the lagging strand template. If the mechanism leading to instability were the same for the two orientations, we would have in both cases expected to see increased stability when SSB was overexpressed. The fact that SSB overexpression did not influence the (CTG)\textsubscript{137} array stability suggests that, in this orientation, hairpin structure formation is not the predominant cause for instability. Expansions and contractions may be caused by the formation of unstructured CAG
extrusions that would cause polymerase slippage. This implies the fact that SSB may impede the formation of DNA hairpins but would be incapable of impeding the formation of unstructured DNA loop-outs. The proposed hypothesis is based on the fact that CTG repeat arrays are prone to form hairpins and that on the contrary, CAG repeat arrays mainly form unstructured DNA extrusions. In vitro data obtained from the study of slipped strand structures support this hypothesis (Miret, Pessoa-Brandao et al. 1997). Electron microscopy (EM) analysis of CAG-CTG slipped strand DNA structures showed that CAG extrusions are covered in SSB tetramers, suggesting that SSB does not influence the formation of such structures (Miret, Pessoa-Brandao et al. 1997).

### 7.3.2 SSB overexpression may preferentially impede the formation of big CTG hairpins on the lagging strand template by binding onto ssDNA

As we showed in section 7.2.2, SSB overexpression changed the deletion size distribution profile of (CAG)$_35$ repeat arrays by removing most of the bias towards big deletions. This supports the hypothesis that SSB excess can selectively impede big DNA hairpin formation in vivo. A decrease in the number of such structures, would in fact lead to a decrease in the number of big deletions. A basal level of big deletions is however still present, even in the presence of overexpressed SSB. Excess SSB levelled the percentage of big and small deletions but it did not completely eliminate the big ones. This suggests that CTG repeats may occasionally form extrusions that do not have the chance organise into hairpins and whose formation is independent of SSB overexpression. Deletion of such extrusions would be possible by polymerase slippage.
7.3.3 SbcCD Nuclease May Stabilise CAG DNA Extrusions

As we previously showed, the absence of SbcCD nuclease slightly reduced the bias of the (CTG)$_{137}$ repeat array towards big deletions.

This may suggest the existence of a SbcCD dependent and of a SbcCD independent pathway that generate big deletions for this orientation. Here we propose a model based on CAG hairpin dynamics to underline the role of the nuclease (Fig.7.8). CAG repeat arrays may be able to form DNA extrusions whose size could change dynamically. Occasionally some of the big loopouts may form stable hairpins that could get recognised and cleaved by SbcCD nuclease, giving rise to a slightly higher number of big deletions. On the other hand SbcCD may be able to recognise and stabilise big CAG extrusions without cleaving them but increasing the chances of polymerase slippage. The deletion size distribution profile was not influenced by SSB overexpression, suggesting that SSB sensitive hairpin formation is not predominant in this orientation.
Figure 7.8 - SbcCD stabilisation of CAG extrusions model. A) When the CTG repeat array lies on the leading strand template, CAG loopouts may have the chance to form on the lagging strand template during replication. B) The extrusion could organise to form a hairpin structure. C) The equilibrium between the two forms is likely to be shifted towards the unstructured extrusion. A big extrusion could also dynamically convert into a small extrusion. SbcCD could either attack the hairpin structures that occasionally form, giving rise to a big deletion, or bind and stabilise the big unstructured extrusion. This would give rise to deletions occurring by polymerase slippage. SSB (green dots), newly synthesized strand (blue), repeat array (red).

7.4 Concluding Remarks

SSB overexpression has been shown to influence CAG-CTG repeat array instability in an orientation dependent manner in the *E. coli* chromosome. SSB being able to bind and stabilise ssDNA, we conclude that DNA hairpins formation on the lagging strand template during replication stimulates instability and that SSB has a specific role in impeding the formation of CTG hairpins.
We moreover hypothesize that CAG repeat arrays mainly form non-structured DNA extrusions whose formation is SSB overexpression independent; this would explain SSB influence on CAG leading orientation only.

The role of SSB in impeding CTG hairpins formation is also reflected in the change of the deletion size distribution profile, where SSB overexpression has been shown to lower the number of big deletions.

SbcCD nuclease has been shown to influence the deletion distribution profile of the (CTG)$_{137}$ repeat array. The absence of the nuclease partially reduced the bias towards big deletions. This suggests either a role in cutting CAG hairpins occasionally forming on the lagging strand template or a stabilising effect on CAG loopouts that can cause polymerase slippage. As, in the absence of SbcCD, a slight effect on the deletion size distribution profile of (CAG)$_{38}$ repeat array strains was observed, the nuclease may also be involved in cleaving CTG hairpins giving rise to big deletions.

Finally, SSB overexpression did not seem to influence the proportion of expansion suggesting a ssDNA dependent pathway leading to both expansions and deletions.
Chapter VIII

Conclusions
8.1 Summary

The aim of this project was to study and assess genome instability in *Escherichia coli*, following perturbation of the system. *E. coli* is a widely characterised gram negative bacterium that has been successfully used as a model organism for the investigation of basic principles of biology. *E. coli* genetics is well known and the manipulation of its genome is easy and controllable. Moreover, strains used in laboratories as *E. coli* K12 and derivatives are safe to handle and easy to grow. *E. coli* is a great tool to study genetics as proteins found in *E. coli* often have homologues in higher organisms and provide a good platform for basic research investigations. In this work the single stranded DNA binding protein was used as a tool to perturb and investigate genome instability in *E. coli*.

Single stranded DNA binding proteins have been widely characterised and their roles have been studied in prokaryotic and eukaryotic cells. Particular attention was at first given to *E. coli* SSB mutant forms and to the phenotypes associated with their expression. SSB was demonstrated to be essential, a lack of the protein causing cell death. *In vitro* experiments showed that SSB is required for DNA replication, recombination and repair processes (Meyer and Laine 1990). Recently the study of physical interactions between SSB and other proteins involved in DNA metabolism has been of high interest and new insights on SSB possible roles have been highlighted. The interaction with SSB is believed to provide a targeting system for proteins involved in replication, recombination and repair processes. More attention has also been given to eukaryotic SSB proteins (RPAs). Interestingly, RPA overexpression has been associated to breast cancer and colon carcinomas (Tomkiel, Alansari et al. 2002, Givalos, Gakiopoulou et al. 2007). A new RPA isoform, hSSB1, has been recently characterised and demonstrated to have a main role in the activation of the response to DSBs.
Chapter VIII - Conclusions

_E. coli_ was used as model organism to study the consequences of SSB depletion and overexpression on genome stability. In this work, the creation of a novel system in which the SSB level could either be lowered or increased, allowed the investigation of the consequences of SSB depletion on cell growth and viability, nucleoid morphology and cell division, SOS response activation and of mechanisms that may lead to enhanced instability of DNA palindromes and CAG-CTG trinucleotide repeats.

Previously, SSB mutants were used to investigate and characterise phenotype changes and properties in _E. coli_. The _ssb^-1_ temperature sensitive mutant was used to measure the stability of CAG-CTG trinucleotide repeats in response to a variation of the levels of active SSB. These mutant forms do not however allow controllable variations and quantification of the active SSB protein quantities present in cells. The system employed in this work, made use of the arabinose inducible promoter, P_BAD, that allowed fine variations of the SSB levels. Overexpression of SSB was on the other hand obtained using an IPTG inducible plasmid, pAM43/ ssb, containing the _ssb_ gene under the control of its natural promoter. SSB levels could then be assessed and quantified by western blot allowing the tracking of variations occurring in response to the quantity of arabinose added to the growth medium.

Experiments performed on the SSB\textsubscript{ind} strain (DL2844), in which the expression of _ssb_ was controlled by P_BAD, showed a striking difference between growth in liquid and solid medium. To obtain normal growth of DL2844 on solid medium, the arabinose concentration required was about three orders of magnitude higher than the one required for growth in liquid culture. Given the impossibility to measure the SSB level present in cells grown on solid medium, a direct comparison between the two conditions was not possible. The sensitivity to arabinose may have changed for cells grown on plates. Either a higher arabinose concentration may be required for the production of the same amount of SSB, or the
Chapter VIII - Conclusions

arabinose intake may be impaired in cells growing on plates. The latter would result in a lower expression of ssb and therefore in a lower survival rate.

Another interesting difference between growth on solid and liquid medium was the ability of cells growing on plates to respond differently to different arabinose concentrations. DL2844 presented various grades of survival which were bound to the arabinose concentrations present on plates. On the contrary, cells grown in liquid medium abruptly died when the arabinose concentration drop reached a certain threshold. This characteristic allowed the selection of suppressor strains that showed similar growth behaviour on plates and in liquid culture, by losing the capacity to respond to different arabinose concentrations on solid medium. Suppressor strains could grow at lower levels of arabinose on plates; death occurred abruptly at the same arabinose concentration threshold both in liquid and solid medium. The characterisation of one such suppressor was carried out and its whole genome was sequenced. Six point mutations were found but unfortunately no time was left for a deeper analysis of the effect of a reversion of the mutations found. A good candidate is represented by ydcN, predicted to be a transcriptional regulator that may be involved in the regulation of the transcription of ssb, controlled by P_{BAD}.

Interestingly, microscope analysis showed that SSB depletion leads to impaired cell division and influences the morphology of nucleoids. Cells deprived of SSB presented an elongated shape and diffusion of the nucleoids in the cytosol. This finding may suggest DNA replication arrest and DNA degradation caused by a lack of SSB but further investigations are required to prove this point. The SOS response was also active in SSB deprived cells consistent with SSB protecting ssDNA from damage. As discussed in chapter IV, SSB may have a dual role in the activation of the SOS response. It helps RecA loading by entangling ssDNA stretches but at the same time it competes with RecA for binding ssDNA. The SOS
response activation may therefore happen when a certain threshold of protein level is reached, allowing a balance to be created between the two situations.

The stability of DNA palindromes has been extensively studied, giving their involvement in human diseases by enhancing chromosomal translocations, gene’s duplications and genes’ deletions. One of the questions we addressed in this work was how the stability of widely characterised DNA palindromes inserted in the *E. coli* chromosome would vary in response to a variation of the SSB level in cells. DNA palindromes have been shown to form DNA hairpins during replication. Those hairpins are believed to form on the lagging strand template, where stretches of ssDNA are present. The hypothesis that was tested in this work relates to the ability of SSB to impede the formation of DNA hairpins by binding ssDNA and therefore contribute to the maintenance of the stability of the genome. The stability of a 105bp palindrome was tested at lower-than-*wild type* SSB levels while the stability of a 246bp palindrome, from which the 105bp palindrome derived by partial deletion, was tested at overexpressed SSB levels. Interestingly lowering the arabinose concentration on plates increased the 105bp palindrome deletion frequency supporting the hypothesis that a lower SSB level may increase the frequency of hairpins formation and therefore the frequency of palindrome deletions occurring by replication slippage. The stability of a 246bp palindrome was tested as a function of the viability of a *recA* strain containing the palindrome and in which the SbcCD nuclease was induced. SbcCD can cut palindrome hairpins causing the formation of a DSB that can only be repaired by homologous recombination in *E. coli*. Unexpectedly, overexpression of SSB did not improve the viability of such a strain, suggesting that hairpin formation is not influenced by an increased level of the protein. The dynamics of hairpin formation may involve a fast and efficient step at the beginning of the process and having excess SSB may not be enough to avoid the folding. Alternatively,
lowering the level of SSB may have a drastic effect on the capacity of the palindrome to form hairpins as longer stretches of ssDNA would remain uncovered and prone to fold.

The stability of CAG-CTG TNR sequences was also assessed in relation to the SSB level present in cells. CAG-CTG repeats expansions have been associated with several neurological and neurodegenerative diseases and their stability has been extensively studied both in prokaryotic and eukaryotic systems. The potential of these sequences to fold into secondary structures has been thought to be the major cause of instability. The stability of a leading strand (CAG)$_{75}$ and of a leading strand (CTG)$_{137}$ repeat arrays was measured as a function of SSB overexpression. Striking differences were found between the two orientations suggesting the existence of two different mechanisms leading to instability for the two different orientations. The results found indicate that SSB overexpression leads to a strong reduction of the instability of the (CAG)$_{75}$ repeat array while no effect on instability was detected for the (CTG)$_{137}$ repeat array. This supports the hypothesis that CTG hairpins forming on the lagging strand during (CAG)$_{75}$ repeat array replication may be the source of instability and that SSB overexpression may impede their formation lowering the chances the repeat array has to expand or contract. It can be hypothesised that, in the opposite orientation, instability can still be related to CAG random coil structures forming on the lagging strand template. It was shown in vitro that CAG-CTG repeats can form slipped strand structures. Electron microscopy analysis showed that CTG extrusions form hairpins while CAG extrusions tend to form random coil structures, often covered by SSB tetramers. This suggests that the formation of these random coil extrusions may be independent of SSB. Interestingly, SSB overexpression influenced the deletion size distribution profile of (CAG)$_{75}$ repeat arrays by attenuating the bias towards big deletion. This may suggest that SSB can preferentially impede the formation of big DNA hairpins. The ratio between expansion and deletion events was also measured for both repeats orientations, leading to the conclusion that SSB
overexpression influences both expansions and deletions to the same extent. This may suggest the existence of a common mechanism that triggers both expansions and deletions.

These last experiments highlighted the difference existing between the formation of palindrome hairpins and TNR hairpins. According to the results obtained and according to the hypothesis formulated on the mechanisms involved in palindrome and \((\text{CAG})_{15}\) repeat array instability, SSB overexpression could not impede the formation of palindrome hairpins but very effectively impeded the formation of CTG hairpins, possibly forming on the lagging strand during the replication of the \((\text{CAG})_{15}\) repeat array. CTG hairpins contain mispairs and are therefore less stable than palindrome hairpins. It can be hypothesised that SSB overexpression could influence the formation of those hairpins because of this stability difference.

The depletion of the SSB level has highlighted differences between growth on liquid and solid medium and interestingly stimulated the selection of suppressor clones, able to cope with the new conditions. Discovering the mechanism by which these suppressors overcome the variation of arabinose present on plates, would be of great interest. The modulation of the SSB level has proved to be a powerful tool that, according to the hypothesis formulated, allowed to find differences between instability generated by palindromes and CAG-CTG repeats. Moreover, SSB overexpression allowed the distinction between the mechanisms which trigger instability of CAG-CTG repeat tract in two different orientations. It would be interesting to be able to test how SSB depletion would act on the CAG-CTG repeat tract instability. The system that has been created during this work, in which SSB expression is controlled by adding both arabinose and glucose to the growth medium, still has to be optimised but is suitable for the purpose. This system would also enable the study of evolutionary aspects triggered by a loss of genome stability.
References


References


References


References


References


References


References


