INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Detection and Characterization of
Spontaneous and N-2-acetylaminofluorene-induced
Deletion Mutations

by

Yen H. Ly, B.Sc.

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for the degree of

Master of Science

Department of Biology

Carleton University
Ottawa, Ontario
September 1, 1998

copyright

Yen Ly

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-36935-8

The undersigned recommend to the Faculty of Graduate Studies
and Research acceptance of the thesis

DETECTION AND CHARACTERIZATION OF
SPONTANEOUS AND N-2-ACETYLAMINOFLUORENE-INDUCED
DELETION MUTATIONS.

submitted by Yen H. Ly, B.Sc.

in partial fulfillment of the requirements for
the degree Master of Science

Thesis Supervisor

Chair, Department

Carleton University

September, 1998

ABSTRACT

Deletion are a class of mutation that is not very well characterized due to their relatively rare occurrence in comparison to frameshift and base substitution mutations. However, recent studies have found that deletion mutations may be related to cancers and genetic diseases, causing interest in this class of mutations to intensify. N-2-acetylaminofluorene (AAF) is a potent mutagen capable of inducing frameshifts and base substitutions. This thesis describes the development of three plasmid systems, pNS, plac and ptac, for the detection of deletion mutations. The effect of AAF-modification, SOS induction and sequence context of the insert sequence, on deletion frequency in these plasmid-based reversion assays were examined.

The pNS system revealed that asymmetric palindromic inserts were deleted with a 185-fold or 22-fold greater frequency than non-palindromic inserts, depending on the orientation of the *Ori*; these results suggest that deletion mutations occur preferentially during lagging strand synthesis. The plac/ptac plasmid systems did not show a marked difference in deletion frequency between palindromic, non-palindromic and triplet repeat inserts. However, deletion frequency in plasmids with inserts situated close (~400 bp) to the origin was significantly increased relative to plasmids with inserts situated ~2500 bp from the origin. AAF-modification caused a small increase in the deletion frequency of inserts in several of the plac/ptac plasmids, but not the pNS plasmids. The effect of SOS induction on spontaneous and AAF-induced deletion frequencies of inserts was found to be inconsistent in the plasmid systems examined.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my research supervisor, Dr. Iain B. Lambert, for his support and guidance throughout the course of my project. I would like to acknowledge my advisors, Dr. Smith and Dr. Bonen for their suggestions and advice. I am grateful to Dr. Carmody for his helpful advice on statistical analysis. I would like to thank Dr. Bichara and Nathalie Samuel, who helped construct and also modified my plasmids.

I am very grateful to Reza Nokhbeh for the many valuable discussions, technical advice and support. I am also grateful to Dr. Suzanne Paterson for her advice in the laboratory and for reading the first draft of this thesis. I would like to acknowledge Craig Carrol, Cristina Micali, Jacqui Whiteway, Katalin Bertenyi and Sheryl Hubbard for all their help in the laboratory and for making the lab an enjoyable place to study.

I would like to thank my parents, Hung Ly and Phuong Luu, and my brother, Chan, for their support and encouragement. I am especially grateful to my great-uncle, Ly Hy, who has encouraged me throughout my studies. Finally, but not least, I would like to thank my fiancé Sebastian for being so supportive and understanding.

Financial support was provided by a grant to Dr. Lambert from the Natural Sciences and Engineering Research Council.

TABLE OF CONTENTS

ACCEPTANCE SHEET	ii
ABSTRACT	iii
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS	v
LIST OF TABLES	x
LIST OF FIGURES	xii
LIST OF ABBREVIATIONS	xiv
1. INTRODUCTION	
1.1 Mutagenesis	1
1.1.1 Spontaneous mutagenesis	1
1.1.2 Deletion mutations	3
1.1.2.1 Mechanisms of deletion mutation	8
1.1.3 Induced mutagenesis	11
1.1.3.1 DNA repair	12
1.1.3.2 DNA damage tolerance	15
1.1.3.3 Types of induced mutations	16
1.2 SOS Response	16
1.2.1 SOS regulatory network	16
1.2.2 Model for SOS regulation	17
1.2.3 The role of SOS response in mutagenesis	20

1.3	N-2-Acetylaminofluorene	22
1.3.1	Formation of AAF and AF adducts	24
1.3.2	DNA adduct conformation	25
1.3.3	Mutagenicity of AAF and AF	28
1.3.4	Mechanisms by which AAF adducts induce mutations	32
1.3.4.1	Effect of genetic control and replication hindrance on mutagenesis	32
1.3.4.2	Induction of -1 frameshift mutations in repeated G sequences	36
1.3.4.3	Induction of -2 frameshift mutations in alternating GC sequences	40
1.4	Objectives	44
2.	MATERIALS AND METHODS	
2.1	Bacterial strains, plasmids and primers	45
2.2	Growth of bacterial cultures	46
2.3	Preparation of electrocompetent cells	47
2.4	Preparation of SOS induced electrocompetent cells	47
2.5	Transformation by electroporation	48
2.6	Extraction of plasmid DNA	48
2.6.1	Boiling miniprep method for the extraction of plasmid DNA	48
2.6.1.1	Purification of plasmid DNA by Gene Clean	49
2.6.2	Extraction of plasmid DNA using Wizard™ Miniprep Kit	49

2.7	Large scale extraction of plasmid DNA	50
2.7.1	Large scale plasmid DNA extraction by CsCl-EtBr gradient	50
2.7.2	Wizard Midiprep System for extraction of plasmids DNA	52
2.8	Double-stranded DNA sequencing	52
2.8.1	Annealing using double-stranded templates	52
2.8.1.1	Preparation and annealing of single-stranded templates	53
2.8.2	Sequencing reactions	53
2.8.3	Sequencing polyacrylamide gel electrophoresis	54
2.9	Plasmid construction	55
2.9.1	Digestion of parent plasmids: pLC, pLCR, pTC and pTCR	55
2.9.2	Preparation of inserts	55
2.9.3	Ligation of digested vector and insert	56
2.9.4	Screening for clones containing insert in the correct orientation	56
2.9.5	Calculation of ΔG values	57
2.10	Modification of plasmid DNA with N-acetoxy-N-2-acetylaminofluorene	57
2.11	Statistical analysis of deletion frequency data	58
3.	RESULTS	
3.1	Plasmid constructs	62
3.1.1	The pNS plasmids	62

3.1.1.1	Optimization of reversion assay conditions for pNS plasmids	63
3.1.2	The plac plasmids	65
3.1.2.1	Optimization of reversion assay conditions for the plac system	68
3.1.3	Construction of the ptac plasmids	68
3.1.3.1	Optimization of conditions for ptac plasmids	71
3.1.4	Cloning of 20p, 20np and 20T inserts into plac and ptac plasmids	71
3.2	Spontaneous deletion frequencies	73
3.2.1	Origin of replication	73
3.2.2	Spontaneous deletion frequencies in pNS plasmids	74
3.2.3	Spontaneous deletion frequencies in plac plasmids	77
3.2.4	Spontaneous deletion frequencies in ptac plasmids	79
3.3	Molecular nature of the deletion mutants	82
3.3.1	The pNS mutants	82
3.3.2	The plac mutants	85
3.3.3	The ptac mutants	88
3.4	Induced deletion frequency in AAF-modified plasmids	92
3.4.1	Deletion frequencies in AAF-modified pNS plasmids	92
3.4.2	Survival of N-acetoxy-N-2-acetylaminofluorene-modified plasmids	93

3.4.3	Deletion frequencies in AAF-modified plac plasmids	96
3.4.4	Deletion frequencies in AAF-modified ptac plasmids	97
3.5	Molecular nature of the AAF-induced deletion mutants	99
3.6	Dimer formation in deletion mutants	101
4.	DISCUSSION	103
4.1	Spontaneous deletions	104
4.1.1	The pNS plasmid system	104
4.1.2	The plac and ptac plasmid systems	110
4.1.3	Molecular nature of spontaneous deletion mutants	114
4.2	Induced deletions in AAF-modified plasmids	115
4.2.1	Survival of N-Aco-AAF-modified ptac plasmids	115
4.2.2	AAF-induced deletion mutations	116
4.3	Dimer formation in deletion mutants	121
5.	CONCLUSIONS	124
6.	REFERENCES	125

LIST OF TABLES

TABLE	DESCRIPTION	PAGE
2.1	<i>E. coli</i> stains used in this study	45
2.2	Plasmids used in this study	45
2.3	Primers and oligomers used in this study	46
3.1	Mean spontaneous deletion frequencies of 17 bp palindromic and non-palindromic inserts in pNS plasmids transformed into <i>E. coli</i> TK603 cells and the statistical comparisons of the deletion frequencies between the SOS ⁻ and SOS ⁺ cells	74
3.2	Statistical analysis of the deletion frequencies of the pNS plasmids in TK603 cells	75
3.3	Mean spontaneous deletion frequencies of 20 bp palindromic, non-palindromic and triplet repeats inserts in plac plasmids transformed into <i>E. coli</i> TG1uvrA-F' cells and the statistical analyses of the deletion frequencies between SOS ⁻ and SOS ⁺ cells	77
3.4	Statistical results from comparisons of the deletion frequencies of inserts in plac plasmids in SOS ⁻ and SOS ⁺ TG1uvrA-F' cells	79
3.5	Mean spontaneous deletion frequencies of 20 bp palindromic, non-palindromic and triplet repeat inserts in ptac plasmids transformed into <i>E. coli</i> TG1uvrA-F' cells	80
3.6	Statistical results from comparisons of the deletion frequencies of inserts in ptac plasmids in SOS ⁻ and SOS ⁺ TG1uvrA-F' cells	82
3.7	Types of deletion mutations found in pNS plasmids introduced into both SOS ⁻ and SOS ⁺ TK603 cells (sequencing data)	83
3.8	Types of deletion mutations found in plac plasmids transformed into SOS ⁻ and SOS ⁺ TG1uvrA-F' cells (from sequencing data)	88
3.9	Types of deletion mutations found in ptac plasmids transformed into SOS ⁻ and SOS ⁺ TG1uvrA-F' cells (from sequencing data)	89

3.10	Mean deletion frequencies of plasmid pNSR17p containing different levels of AAF modification, transformed into SOS ⁻ and SOS ⁺ TK603 cells	93
3.11	Mean deletion frequencies of unmodified and AAF-modified plac plasmids transformed into SOS ⁻ and SOS ⁺ <i>E. coli</i> TG1uvrA-F' cells and statistical analyses of the deletion frequencies between the unmodified and AAF-modified plasmids	97
3.12	Mean deletion frequencies of unmodified and AAF-modified ptac plasmids transformed into SOS ⁻ and SOS ⁺ <i>E. coli</i> TG1uvrA-F' cells and statistical analyses of the deletion frequencies between the unmodified and AAF-modified plasmids	98
3.13	Types of deletion mutations found in pNS, plac and ptac plasmids transformed into SOS ⁻ and SOS ⁺ cells (from sequencing data)	101

LIST OF FIGURES

FIGURE	DESCRIPTION	PAGE
1.1	Schematic diagram of a model of the SOS system	18
1.2	Structures of aromatic amines	23
1.3	Formation of the major adduct, dG-C8-AAF	26
1.4	Schematic diagram of damage avoidance and translesion synthesis mechanisms	33
1.5	Slippage model for -1 frameshift mutagenesis induced by an AAF adduct	38
3.1	Plasmid maps of parental pNS plasmids	64
3.2	Schematic diagram of the construction of the <i>lacZ</i> -Kan gene and of the parental pLC and pTC plasmids	66
3.3	Plasmid maps of parental pLC and pLCR plasmids	67
3.4	Plasmid maps of parental pTC and pTCR plasmids	69
3.5	Sequence of 342 bp fragment containing the ptac promoter	70
3.6	Schematic diagram of the cloning of the 20 bp palindromic insert into a pLC plasmid	72
3.7	Schematic diagram of the parental, insert and deletion mutant plasmid sequences for pNS17p	84
3.8	Schematic diagram of the parental, insert and deletion mutant plasmid sequences for pLC20p	86
3.9	Photograph showing the sequence difference between plasmid pLC20np and the deletion mutant	87
3.10	Photograph showing the sequence difference between plasmid pTC20p and the deletion mutant	90

3.11	Photograph showing the sequence difference between plasmid pTC20T and the deletion mutant	91
3.12	Graph of survival rate of AAF-modified ptac plasmids	95
4.1a	A model for the generation of a deletion mutant in pNS17p via a slippage mechanism (using top strand as template)	106
4.1b	A model for the generation of a deletion mutant in pNS17p via a slippage mechanism (using bottom strand as template)	107
4.2	A model for the generation of a deletion mutant in pLC20p or pTC20p, via a slippage mechanism (using top strand as template)	112

LIST OF ABBREVIATIONS

20p	20 base pair palindromic insert
20np	20 base pair non-palindromic insert
20T	20 base pair triplet repeat insert
2-AF	2-aminofluorene
2-AAF	N-2-acetylaminofluorene
A	Adenine
Amp	Ampicillin
APS	Ammonium persulfate
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
C	Cytosine
°C	degrees Celcius
cfu	colony forming units
CI	chloroform:isoamyl alcohol
cm	centimetre(s)
Cm	Chloramphenicol
CsCl	Cesium chloride
dG-C8-AF	N-(Deoxyguanosin-8-yl)-2-aminofluorene
dG-C8-AAF	N-(Deoxyguanosin-8-yl)-2-acetylaminofluorene
dG-N ² -AAF	3-(Deoxyguanosin- N ² -yl)-2-acetylaminofluorene

DNA	deoxyribonucleic acid
ddH ₂ O	distilled deionized water
dNTP	deoxynucleoside triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EtOH/NaAc	ethanol acetate
G	Guanine
Kan	Kanamycin
KCl	Potassium chloride
KV	kilovolt(s)
LB	Luria-Bertania
M	molar
MgSO ₄	Magnesium sulfate
min.	minute(s)
mL	millilitre(s)
mM	millimolar
N-Aco-AAF	N-acetoxy-N-2-acetylaminofluorene
N-OH-AAF	N-hydroxy-N-2-acetylaminofluorene
NaAc	Sodium acetate
NaCl	Sodium chloride
NaI	Sodium iodide
NaOH	Sodium hydroxide

OD	optical density
PCI	Phenol:chloroform:isoamyl alcohol
plac	<i>lacI</i> promoter
pmol	picomole(s)
ptac	tac promoter
STET	Sucrose Tris EDTA Triton-X
T	Thymine
TBE	Tris borate EDTA electrophoresis buffer
TE	Tris EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
Tn	transposon
TS	Tris-Sucrose
μCi	microcurri(es)
μg	microgram(s)
μL	microlitre(s)
μM	micromolar
U	units (of enzyme)
UV	ultraviolet
v/v	volume per volume
w/v	weight per volume
W/cm ²	Watts/centimetre square

1. INTRODUCTION

1.1 Mutagenesis

Mutations are the net result of all mutagenic and antimutagenic cellular processes (Smith, 1992). Several types of mutations (base substitutions, frameshifts, deletions and insertions) can be produced spontaneously, and genetic factors can influence the type and proportion of mutations that arise. Spontaneous mutations are suspected to play a major role in evolution, aging and carcinogenesis. However, much attention has been focused on the study of induced mutations due to concerns about chemical and physical mutagens that exist in the environment (Smith, 1992).

1.1.1 Spontaneous mutagenesis

The rate of spontaneous mutation per genome varies by many orders of magnitude between different groups of organisms, but is very similar within broad groups of organisms. The rate of spontaneous mutation per genome per replication for *Escherichia coli* (*E. coli*), with a genome size of 4.6×10^6 base pairs, is 5.4×10^{-10} (Drake *et al.*, 1998). The production of these spontaneous mutations may be caused by errors in nucleotide incorporation by DNA polymerase during replication or repair. Other factors involved in spontaneous mutagenesis are the cells' growth conditions, the intrinsic instability of DNA and other cellular metabolic mechanisms (Smith, 1992).

During DNA synthesis, base substitution mutations may result from the misincorporation of a nucleotide. However, this is rare because of the different activities of DNA polymerases that enhance fidelity of DNA replication. *In vitro* studies indicated that DNA polymerases are more efficient at incorporating correct nucleotides than incorrect nucleotides due to tighter binding of the correct deoxynucleoside triphosphate (dNTP) to the polymerase and a more rapid rate of phosphodiester bond formation (Minnick and Kunkel, 1996). There is also a 3' to 5' proofreading mechanism, in which the polymerase-associated 3' to 5' exonuclease increases fidelity by excising misincorporated nucleotides at their points of origin (Goodman and Fyngenson, 1998). It has been found that *E. coli* PolII favours excising mispaired nucleotides over correctly paired nucleotides (Goodman and Fyngenson, 1998). Another important activity which helps increase DNA synthesis fidelity is postreplicational mismatch repair, which can decrease error rates by up to two orders of magnitude (Minnick and Kunkel, 1996).

Depurination of DNA is the most frequently occurring spontaneous alteration in the chemical structure of DNA under physiological conditions, and has the potential to give rise to a mutation (Loeb and Cheng, 1990). In this case, the N-glycosidic bond that attaches the purine base to the deoxyribose sugar is broken, leaving the sugar backbone intact, but lacking a base at this site. DNA synthesis on templates containing abasic sites requires that DNA polymerase incorporate nucleotides opposite a non-instructional lesion. Many studies have suggested that deoxyadenosine is preferentially incorporated into the daughter strand opposite such lesions, sometimes leading to the formation of base substitution mutations (Loeb and Cheng, 1990).

The slipped misalignment model proposed by Streisinger *et al.* (1966) has been used to explain the formation of frameshift mutations during DNA replication. For example, when a misalignment occurs in a repetitive sequence, a bulge of one or more extrahelical nucleotides will be produced, but the flanking repetitive nucleotides will still base pair correctly, resulting in a stabilizing effect on the misaligned intermediate. When DNA synthesis continues from this misaligned intermediate, a deletion will occur if the bulged nucleotide(s) is in the template strand, and an addition mutation will result if the unpaired base(s) is in the primer strand (Owen and Streisinger, 1985).

1.1.2 Deletion mutations

Deletions are an important class of mutation whose existence has been known for well over fifty years. Interest in deletion mutations has intensified recently because of their relation to cancers and other genetic diseases (Balbinder, 1993). These mutations can occur spontaneously, or can be induced by chemical and physical mutagenic agents. In this thesis, deletions will be defined as mutations in which three or more nucleotides have been deleted; losses of one or two bases will be referred to as frameshift mutations.

Although the existence of deletion mutations has been known for several decades, the mechanisms through which they arise remain poorly understood. Deletions constitute a relatively small proportion of mutations (in comparison to frameshifts and base substitutions); therefore, unless an experimental system is in place to specifically characterize deletions, a very large number of mutations have to be analyzed before significant information about deletions can be gathered. In addition, deletion mutations

are difficult to isolate, using techniques such as *in vitro* packaging of phage shuttle vectors or *in vivo* recombination, due to their large size. The size of deletions also poses problems because they extend beyond the target genes and into flanking genes, some of which may be essential. Therefore, the cells or vectors which contain deletions are frequently non-viable (DeMarini *et al.*, 1989).

Studies of the DNA sequences of both spontaneous and chemically-induced mutations have revealed that deletion mutations occur most frequently at DNA sequences which contain reiterated bases, palindromic (inverted repeats) or quasi-palindromic sequences, and direct repeats, leading to the suggestion that these types of sequences somehow promote the formation of deletions (Albertini *et al.*, 1982; Glickman and Ripley, 1984). Not only does the presence of direct repeats affect deletion frequency, but the lengths of the direct repeats are a factor as well. To study the effects of palindromes and direct repeat length on deletion frequency, Pierce *et al.* (1991) cloned 76 bp palindromic and non-palindromic inserts, flanked by direct repeats, into gene 1.3 of bacteriophage T7. They found that when the direct repeats were less than 6 bp long, the deletion frequencies of the inserts were too low to be measured. However, the deletion frequencies of the inserts increased exponentially as the lengths of the direct repeats increased from 8 to 20 bp. Also, the deletion frequency of the palindromic insert was not significantly different from that of the non-palindromic insert when they were flanked by 10 bp direct repeats. In contrast, when the inserts were flanked by 5 bp direct repeats, the deletion frequency of the palindromic insert was two orders of magnitude greater than that of the non-palindromic insert. These findings suggested that in this system, the

homology at the endpoints (*e.g.*, the length of direct repeats) had a greater effect on deletion frequency than the presence of palindromes.

Using a similar system, with 93 bp long inserts flanked by varying lengths of direct repeats (10, 13, 16 or 20 bp), Kong and Masker (1994) found that between 13 and 20 bp, each additional 3 bp in the length of the direct repeat resulted in an order of magnitude increase in deletion frequency. They also demonstrated that sequence context surrounding the deletion site has a great impact on deletion frequency. Results of Balbinder *et al.* (1989) suggest that the stability of the intermediate formed is also a factor in determining deletion frequency.

A plasmid-based reversion assay has been developed which facilitates the detection of deletions between reiterated sequences. Plasmid pBR325, a ColE1-derived plasmid, contains an ampicillin resistance gene, a chloramphenicol resistance gene (with the transcribed strand as the leading template strand), and a unidirectional origin of replication (*Ori*) (Trinh and Sinden, 1991). Derivatives of pBR325 were prepared by inserting a 17 bp palindromic (or 17 bp non-palindromic) sequence flanked by 6 bp direct repeats into the *EcoRI* site of the chloramphenicol resistance gene. The four resultant plasmids used in this assay contained the chloramphenicol resistance gene in the same orientation as the *Ori* or in the opposite orientation, interrupted by a 17 bp palindromic or 17 bp non-palindromic insert. This system was used to examine the effect of reversing the chloramphenicol resistance gene on deletion frequency. Reversing the chloramphenicol resistance gene resulted in a reduction in deletion frequency by a factor

of 20, suggesting that replication-dependent deletion between direct repeats may occur preferentially in the lagging strand (Trinh and Sinden, 1991).

One other type of repeated sequence that may be linked to deletions is the triplet repeat. Long triplet repeats have been associated with several human hereditary diseases, including myotonic dystrophy, Huntington's disease and fragile X syndrome (Rosche *et al.*, 1996). The genes associated with these diseases contain heritable, unstable triplet repeat sequences which increase or decrease in number with each successive generation (Jaworski *et al.*, 1995). The cause of the expansion or deletion in the number of triplet repeats is not yet known, but it has been proposed that slippage during replication may be involved (Rosche *et al.*, 1996).

When a (CTG)₁₈₀ insert was cloned into plasmids near the ColE1 origin, the deletion frequency of long CTG repeats was very high in wild type *E. coli* cells. After 100 generations, 80% of the plasmids showed deletions of multiple triplet repeats, with preferences for 20, 60, 100 or 140 repeats over all other multiples of repeats (Jaworski *et al.*, 1995). After 100 generations in mismatch repair deficient *E. coli* strains, *mutS*, *mutL* and *mutH*, only 15-40% of the plasmids analyzed contained deletions (Jaworski *et al.*, 1995). These results led to the proposal that *E. coli* mismatch repair (MMR) proteins recognized the loops formed at triplet repeats during replication, generated long single-stranded gaps where stable secondary structures may form, allowing bypass by DNA polymerase during resynthesis, resulting in deletion mutations (Jaworski *et al.*, 1995).

In plasmids containing the (CTG)₁₈₀ insert, deletions only occurred within the inserted sequence and not in the DNA sequences flanking the repeats (Jaworski *et al.*,

1995). It was proposed that the long triplet repeats may form a stable hairpin structure or some other DNA secondary structures that allow DNA polymerase to bypass the base of the deletion intermediate and continue to synthesize the daughter strand, leaving out the sequences that form the secondary structure (Jaworski *et al.*, 1995). Chastain II *et al.* (1995) found that the DNA fragments containing CTG triplet repeats derived from the human myotonic dystrophy gene migrated 20% faster than expected in non-denaturing polyacrylamide gels, further supporting the suggestion that some kind of secondary structure must be forming in the DNA helix within the region containing the triplet repeats. Pearson and Sinden (1996) found that the larger the number of CTG repeats, the more complex the secondary structure formed, and that these structures are stable at physiological pH up to a temperature of 55°C.

Several groups have conducted studies on the stability of triplet repeats, but little is known about the mechanisms by which these repeats expand or contract. The stability of CTG triplet repeats near the ColE1 origin was analyzed in *E. coli* by determining the sizes of the plasmids (extracted from deletion mutants) on 1.75% agarose gels and by analysis of the sizes of restriction fragments containing the triplet repeats on 8% polyacrylamide gels (Rosche *et al.*, 1996). In all the mutants analyzed, the decrease in plasmid size was attributable to deletions within the triplet repeats.

Using *E. coli* RM121, which contains a temperature sensitive mutation in the gene encoding single-stranded DNA-binding protein (SSB), Rosche *et al.* (1996) investigated the effect of SSB on deletion mutagenesis. When grown at 42°C, this strain loses its functional SSB. It was found that the absence of SSB resulted in an increased frequency

with which a large number of CTG triplet repeats were deleted (Rosche *et al.*, 1996). The results suggest that SSB may stabilize triplet repeats by preventing the formation of DNA secondary structures in single stranded DNA.

The evidence collected from the many studies conducted on the mechanisms of deletion indicate that deletions are a heterogeneous class of mutations that are formed by a variety of different pathways, involving both the processes of DNA replication and recombination (DasGupta *et al.*, 1987). It has also been suggested that the same type of deletion mutation may be formed by different mechanisms (Balbinder *et al.*, 1989).

1.1.2.1 Mechanisms of deletion mutation

The extensive studies of spontaneous deletions conducted at the molecular level primarily in prokaryotic systems have led to proposals of several different mechanisms.

1) In the case of inverted repeats flanked by direct repeats, misalignment between the direct repeats may be stabilized by a slipped mispaired intermediate (SMI) during DNA synthesis (Albertini *et al.*, 1982). Resolution of this structure can lead to a deletion mutation. 2) The deletion of palindromes may be a result of intermolecular or intramolecular recombination (Dianov *et al.*, 1991). 3) The formation of a cruciform by the inverted repeats and the subsequent processing of this structure by nucleases may also account for deletion of palindromes (Glickman and Ripley, 1984). 4) Errors caused during the movement of transposable elements may contribute to the generation of deletions (Balbinder, 1993). 5) The repair of double-stranded breaks may result in a deletion mutation (Schulte-Frohlinde *et al.*, 1993). These mechanisms and those through

which chemicals interact with DNA and induce deletion mutations, are still not well understood.

Since deletion endpoints have been found at directly repeated sequences, it was proposed that a slippage-misalignment model may be used to explain these observations (Balbinder *et al.*, 1989). This model predicts that during replication, DNA polymerase synthesizes the first repeat, then strand slippage occurs and the daughter strand hybridizes with the second repeat of the template strand, leaving a bulge protruding from the template strand. DNA synthesis then continues across the base of the transient deletion intermediate, leading to the deletion of one of the direct repeats and the sequences between the two repeats (Albertini *et al.*, 1982; Trinh and Sinden, 1993). DNA misalignments at direct repeats may be the deletion intermediates of either replication or recombination (Ripley, 1990). Dianov *et al.* (1991) have demonstrated that when long direct repeats are involved, deletion results mainly from unequal crossing-over between two plasmid molecules (*e.g.*, dimerization of plasmid DNA). These results support the theory that deletions arise through intermolecular recombination.

Support for the misalignment model was sought by studying the effect of the length and sequence context of repeated sequences on deletion frequency. In the *lacI* gene, it was found that the mutation of (CTGG)₄ to (CTGG)₃ occurred with a higher frequency than (CTGG)₃ to (CTGG)₂ (Ripley, 1990), suggesting that the longer repeat sequence allows more opportunity for misalignment to occur, and also increases the stability of the intermediate. Both these factors may lead to an increased deletion frequency.

It has been observed that palindromic sequences flanked by directly repeated sequences tend to increase the frequency of deletion (Albertini *et al.*, 1982). As an extension to the misalignment model, which predicts the role of direct repeats in promoting deletions, three distinct roles have been proposed to account for the generation of deletions at a palindromic sequence flanked by direct repeats. It has been suggested that the palindromic sequence: 1) exactly juxtaposes the centres of the repeats; 2) tandemly aligns the repeats; and 3) brings the repeats closer together (Glickman and Ripley, 1984). Examination of the sequencing data of spontaneous *lacI* deletion mutations revealed that 40% of the deletions could be explained by either the misalignment between direct repeats model or the palindromic juxtapositioning model (Glickman and Ripley, 1984).

Another hypothesis for the formation of deletion mutations involves double-strand breaks in plasmid DNA (Schulte-Frohlinde *et al.*, 1993). Using an *in vitro* system in which the T7 ligase gene (gene 1.3) was interrupted by a 93 bp nonsense sequence flanked by 20 bp direct repeats, Kong and Masker (1994) introduced a double-strand break into the 93 bp sequence. Repair of this strand break was found to be accompanied by deletion of the DNA in between the direct repeats, leading to the proposal that the breaking and rejoining of DNA can contribute to the formation of deletions (Kong and Masker, 1994).

1.1.3 Induced mutagenesis

Mutations in DNA that arise as a result of damage caused by physical and chemical mutagens are known as induced mutations. Damaged DNA creates a challenge for the cell's metabolic processes in that these processes must continue to function in the presence of chemically or physically induced lesions in DNA. Two physical agents that can induce a wide variety of DNA lesions are ionizing radiation and ultraviolet (UV) radiation. Radiation damage to DNA may be either direct or indirect. Direct effects of ionizing radiation result from the interaction of DNA with the radiation energy, while indirect effects result from the interaction of DNA with a reactive species produced by irradiation (Friedberg *et al.*, 1995).

Strand breaks, and damage to nitrogenous bases and sugar residues by ionizing radiation have been studied extensively. Ionizing radiation can induce single- and double-strand breaks, which may be lethal to the cell, depending on the extent of damage. When a radiation-induced single-strand break occurs in DNA, denaturation results in the vicinity of the break, increasing the chance of attack by free radicals. It has also been found that in the majority of these strand breaks, the OH group at the 3' ends are usually damaged such that simple DNA ligation is no longer possible (Friedberg *et al.*, 1995).

DNA damage by UV radiation is the most extensively studied of the two types of radiation. The exposure of cells to UV radiation has provided the best model for examination of the biological consequences of DNA damage, as well as the repair and tolerance of these lesions. Several types of lesions are produced by UV irradiation, including the thymidine dimer, which is formed when adjacent thymidines are cross

linked (Weaver and Hedrick, 1992). UV lesions and UV mutagenesis represent important paradigms upon which much of our understanding of mutagenesis, DNA repair and induced SOS functions have been formed.

Due to the industrial pollution generated in the past century, chemical mutagens have become the focus of many studies. The structures and mechanisms of action (on DNA) of a wide variety of chemical mutagens and carcinogens have been elucidated using a range of different chemical, biochemical and molecular techniques. Some chemical mutagens are able to react directly with DNA to cause mutations, whereas others require metabolic activation to react with DNA. Chemical mutagens that require metabolic activation before reacting with DNA include benzo[a]pyrene, aflatoxins, nitropyrenes and N-2-acetylaminofluorene (2-AAF). The interaction between these compounds and DNA have been studied extensively and a wealth of literature can be found on different aspects of these chemical mutagens and the variety of DNA lesions that they generated.

1.1.3.1 DNA Repair

In order to survive, cells must respond to the damage sustained from exposure to physical and chemical mutagens. Responses to the DNA lesions include: reversal of the DNA damage, excision repair and tolerance of the damage.

In principle, the simplest method for repairing DNA lesions would be one in which a single enzyme catalyzes the direct reversal of the damage (Friedberg *et al.*, 1995). The most extensively studied class of enzymes which catalyzes the

monomerization of pyrimidine dimers is known as photoreactivating enzymes or DNA photolyases. *E. coli* DNA photolyase activity was first identified *in vitro* by Rupert *et al.* (1958); however, the enzyme itself was not well characterized until very recently due to its extremely low levels in cells under normal growth conditions (Friedberg *et al.*, 1995). It has been found that *E. coli* cells contain only 10 to 20 photolyase molecules per cell (Harm *et al.*, 1968). Photolyase has been found to contain two chromophores, 1,5-dihydroflavin adenine dinucleotide (FADH₂) and 5,10-methanyltetrahydrofoly polyglutamate (MTHF) (Johnson *et al.*, 1988). The mechanism of photoreactivation by *E. coli* DNA photolyase involves a light-independent step, followed by a light-dependent step. In the first step, the enzyme binds specifically to DNA containing pyrimidine dimers. In the light-dependent step, the enzyme substrate complex absorbs a photon with MTHF, which acts as a photoantenna (Sancar and Sancar, 1987). The energy is then transferred to FADH₂, which monomerizes the pyrimidine dimer in an electron transfer reaction (Kim *et al.*, 1992).

Although direct reversal of DNA damage is the simplest process, the types of damage repaired in this manner are limited. The DNA repair process which occurs most often in nature is known as excision repair, which involves the excision of the damaged or incorrect bases from the DNA and replacing them with the correct bases (Friedberg *et al.*, 1995). There are two types of excision repair: base excision repair and nucleotide excision repair. Base excision repair involves the removal of the damaged/incorrect base as free bases; whereas nucleotide excision repair processes excise the damaged DNA as intact nucleotides.

Base excision repair is initiated by DNA glycosylases (Duncan *et al.*, 1976), which hydrolyze the N-glycosylic bond linking the base to the sugar-phosphate backbone, resulting in the formation of another type of DNA damage. The sites at which the excised bases are lost, are referred to as apurinic or apyrimidinic (AP) sites. AP endonucleases specifically recognize AP sites in duplex DNA (Lindahl, 1979), and produce nicks by hydrolysis of either of the phosphodiester bonds immediately 5' or 3' to the AP site; however, most AP endonucleases are specific for the 5' phosphodiester bond (Friedberg, 1995). This results in a 5' terminal deoxyribose-phosphate residue, which is degraded by exonucleases (also referred to as DNA-deoxyribophosphodiesterases), to generate a single nucleotide gap in the DNA duplex (Kornberg and Baker, 1992). The gap is filled in by DNA polymerase and the bonds are reconnected by DNA ligase.

Nucleotide excision repair is similar to base excision repair, but is biochemically more complex and requires many more gene products. This process has been most extensively studied in the excision repair of pyrimidine dimers in *E. coli*. Genetic studies have revealed that the gene products of *uvrA*⁺, *uvrB*⁺ and *uvrC*⁺ (the UvrABC damage-specific endonucleases) are required for the excision repair of pyrimidine dimers (Howard-Flanders *et al.*, 1966). The UvrA and UvrB proteins form a complex that binds to DNA and tracks along until the damaged site is encountered, at which point the UvrA protein dissociates, leaving a stable UvrB-DNA complex (Sancar and Sancar, 1988). UvrC binds at this site and induces a conformational change which enables the UvrB protein to nick the DNA 4 nucleotides 3' to the damaged site, followed by a nick catalyzed by UvrC protein that is 7 nucleotides 5' to the lesion (Orren and Sancar, 1990).

DNA helicase II, also referred to as UvrD protein, is responsible for releasing the oligonucleotide fragment generated by the UvrABC endonucleases, as well as displacing the UvrC protein (Orren and Sancar, 1990). The UvrB protein is released during the repair synthesis reaction, in the presence of Pol I and the substrate dNTPs (Orren *et al.*, 1992). Once the gap has been filled in with the appropriate dNTPs, DNA ligase catalyzes the sealing of the nicks to complete the process of nucleotide excision repair.

1.1.3.2 DNA damage tolerance

Although cells have a wide variety of predominantly error-free DNA repair mechanisms available for coping with DNA damage (Koffel-Schwartz and Fuchs, 1989), they have also evolved DNA damage tolerance mechanisms to deal with damage that cannot be repaired. Bulky DNA adducts usually inhibit DNA replication by blocking the replication fork, leading to non-coding or mis-coding regions. When DNA replication is arrested, there are two mechanisms which cells can use to overcome this problem. One is to reinitiate DNA synthesis downstream from the blocked site, leaving a gap that can eventually be filled in by some other mechanism. The second is to continue polymerization of the DNA past the lesion, a process known as translesion DNA synthesis (TLS). This latter process is highly mutagenic and is required for UV and most chemical mutagenesis (Friedberg *et al.*, 1995).

When DNA is damaged by UV or chemical mutagens, a complex set of physiological responses, known as the SOS responses, is induced. This response involves the expression of more than 20 genes and is regulated by the LexA and RecA proteins

(Ennis *et al.*, 1989). The SOS response has been found to promote the occurrence of translesion synthesis (Koffel-Schwartz *et al.*, 1996), a highly mutagenic process, which implies that SOS induction must play a major role in the generation of mutations. The SOS system and its role in mutagenesis will be discussed in more detail in section 1.2.

1.1.3.3 Types of induced mutations

Different types of mutation (base substitutions, frameshifts, deletions, etc.) can be induced by many different mutagens. For example, base substitutions can be caused by alkylating agents, and bulky adducts induced by chemicals such as benzo[a]pyrenes, nitroarenes, aflatoxins and N-2-acetylaminofluorene (Friedberg *et al.*, 1995). Conversely, different mutagens are often capable of inducing more than one type of mutation. N-2-acetylaminofluorene has been found to induce base substitutions, -1 and -2 frameshifts, additions and deletions (Koffel-Schwartz *et al.*, 1984). The types and mechanisms of mutation involving AAF-adducted DNA are discussed in detail in section 1.3.

1.2 SOS Response

1.2.1 SOS Regulatory network

The SOS system is a set of multigene, physiological responses in bacteria that is induced by DNA damage. This system enhances cellular capacity for DNA repair, recombination and mutagenesis (Smith and Walker, 1998). This induced response to DNA damage increases the cell's chance of survival by either directly repairing the damage or by enhancing the cell's ability to tolerate the damage. Many of the

physiological changes that result from SOS induction are due to the coordinated derepression of approximately 20 unlinked operons with widely divergent functions, collectively known as the SOS regulon (Smith and Walker, 1998). This derepression causes reinitiation of DNA replication, increased DNA excision repair, recombinational repair and SOS mutagenesis. Regulation of the SOS response is dependent upon the products of two SOS genes, *recA* and *lexA* (Ennis *et al.*, 1989).

The SOS system has been more extensively researched in *E. coli* than in any other organism. More than 20 of the SOS-regulated genes in *E. coli* have been identified on the basis of their regulatory characteristics, using gene and operon fusion technology. Some of these genes include *recA*, *lexA*, *umuDC* (Peat *et al.*, 1996), *uvrA*, *uvrB*, *uvrD*, *polB*, *dinB*, *dinD*, *dinF*, *dinG*, *dinH*, *dinI*, *dinQ*^{**}, *dinY*^{**}, *recQ*^{**}, *dnaA*^{**}, *dnaN*^{**}, *phr*^{**}, *nrdaB*^{**}, *mucAB* (Elledge and Walker, 1983). All of these genes are induced as part of the SOS response, but those indicated with a ^{**} do not appear to be regulated by LexA (Friedberg *et al.*, 1995). Although many of the SOS genes have been identified, the functions of all these genes have yet to be determined. Also, for some of the induced physiological responses observed, the genes responsible for the induction of these responses have yet to be identified.

1.2.2 Model for SOS regulation

The model for the basic regulatory mechanism of the SOS system is found in the schematic diagram in Figure 1.1. In an uninduced *E. coli* cell, the gene product of *lexA*

Figure 1.1 Schematic diagram of a model for the basic regulatory mechanism of the SOS system. When DNA is damaged, an inducing signal activates the RecA coprotease, which cleaves the LexA repressor, leading to the activation of the many SOS genes. As the damaged DNA is repaired, the cellular level of RecA coprotease decreases, resulting in the accumulation of the LexA repressor, which represses the expression of the SOS genes. This allow the cell to return to the uninduced state. (From Friedberg *et al.*, 1995).

[illegible]

The diagram illustrates the regulation of the LexA repressor. The LexA repressor (black dots) binds to the operator regions of the *lexA* gene, *recA* gene, and an inducible gene. Upon induction, the LexA repressor is cleaved (indicated by wavy lines), releasing the RecA protein (black bars). The RecA protein then binds to the DNA at the site of the bulky adduct, forming a RecA filament (thick black bar). This process is mediated by RecA* (Coprotease).

INDUCED STATE

(LexA) acts as a repressor of the more than twenty genes listed in the above section, including *recA* and *lexA* itself, by binding to similar operator sequences upstream of each gene or operon (Walker, 1984). Operator sequences to which LexA binds are commonly known as SOS boxes (Friedberg *et al.*, 1995). Some loci, such as *recA*, have only one SOS box, while others, such as *lexA* and *umuDC*, have two (Walker, 1984). Even in the uninduced state, many of the SOS genes, including *recA* and *lexA*, are expressed at significantly high levels. The RecA protein, for example, is expressed at approximately 7200 molecules per cell in uninduced cells (Sassanfar and Roberts, 1990), which is believed to be sufficient for the role(s) that RecA plays in homologous recombination (Smith and Walker, 1998).

When DNA damage occurs or DNA replication is inhibited, the cell generates a signal for SOS induction. *In vivo* and *in vitro* studies have suggested that this intracellular signal consists of regions of single-stranded DNA which are formed during replication of a damaged template or when normal replication is interrupted (Peat *et al.*, 1996). The binding of RecA to these single-stranded regions of DNA in the presence of a nucleoside triphosphate reversibly converts it into RecA*, which exhibits a coprotease activity that promotes cleavage of the LexA protein, as well as several other SOS-regulated proteins (Smith and Walker, 1998). The cleavage of LexA occurs at the alanine-glycine peptide bond near the middle of the protein, thus inactivating LexA as a repressor and promoting increased expression of various SOS genes, including *recA*. The genes whose operators are weakly bound to LexA are the first to be transcriptionally upregulated. If the induction is sufficiently strong, more molecules of RecA are activated

and more molecules of LexA are cleaved, allowing genes whose operators are tightly bound by LexA to be expressed fully.

As the cell recovers from the effects of the SOS-inducing treatment, the regions of single-stranded DNA decrease in number as a result of DNA repair processes (Smith and Walker, 1998). RecA* molecules convert back to their inactive RecA form, leading to an increase in the cellular level of LexA. The increase in the number of LexA molecules results in more binding of the operator sequences of the SOS genes, causing repression, which returns the cell to the uninduced state (Smith and Walker, 1998).

1.2.3 The role of SOS response in mutagenesis

When cells are exposed to chemical or physical mutagens, various types of DNA damage results, depending on the type of mutagen. Cells have many different mechanisms to cope with DNA damage. The most obvious is the repair of the damaged DNA by the direct reversal of the damage or by excision repair of the damage (Peat *et al.*, 1996). However, cells have evolved other methods to cope with DNA damage that may not involve any kind of DNA repair. These cellular responses are referred to as DNA damage tolerance mechanisms, which allow the cells to survive, despite the unrepaired damage to their genomes.

A significant proportion of mutations induced in prokaryotic cells following DNA damage involves a special DNA damage tolerance process referred to as translesion DNA synthesis (Napolitano *et al.*, 1997). Available evidence suggests that translesion DNA synthesis is an induced cellular response that is part of a spectrum of the inducible SOS

responses (Walker, 1984). Translesion DNA synthesis involves the insertion of one or more nucleotides directly opposite the lesion with subsequent extension of the terminus past the lesion. This process has been found to be highly mutagenic and, for most lesions, requires the functions of the *umuD* and *umuC* genes (Napolitano *et al.*, 1997).

The *recA*, *lexA* and *umuDC* gene products, as well as DNA polymerase III are all required for SOS-induced targeted mutagenesis (Koffel-Schwartz and Fuchs, 1989). Replication is generally inhibited at bulky adduct lesions because the distortion in the template strand will hinder insertion of a base by DNA polymerase III (PolIII) opposite the lesion. DNA replication must pass through the site of a lesion with the introduction of errors in order for targeted mutagenesis to occur.

E. coli strains carrying mutations in the *umuD* or *umuC* genes are virtually nondurable by a wide variety of mutagens, including UV radiation, 4-NO, methylmethanesulfonate (MMS) and neocarcinostatin (Friedberg *et al.*, 1995). The *umuDC* locus is inducible by DNA damage and is located downstream of an SOS box, suggesting that this locus is a member of the SOS regulon (Friedberg *et al.*, 1995). Unlike *recA* and *lexA* mutants, *umuD* and *umuC* mutants can still express a variety of SOS responses, leading to the conclusion that the *umuD* and *umuC* gene products are required for SOS mutagenesis by different mutagenic agents (Baptisia *et al.*, 1990).

The UmuD protein is proteolytically cleaved in a RecA*-mediated manner at the Cys-24-Gly-25 bond to yield the activated form, UmuD', which is the carboxyl terminal domain of UmuD (Peat *et al.*, 1996). UmuD' forms a homodimer, and it is postulated that this homodimer and the UmuC protein function in SOS mutagenesis by directly

modifying DNA polymerase in a way that allows the polymerase to bypass DNA replication blockages (Baptisia *et al.*, 1990).

The intact UmuD protein is not simply the inactive form of UmuD', but is a protein that performs a regulatory role in the SOS response (Baptisia *et al.*, 1990). After RecA*-mediated cleavage of UmuD, three different dimers may be formed: UmuD' homodimer, which is active in SOS mutagenesis; and UmuD homodimer and UmuD-UmuD' heterodimer, which are not active in mutagenesis (Peat *et al.*, 1995). Since UmuD protein is cleaved much less efficiently than LexA protein *in vivo* and *in vitro*, intact UmuD would accumulate before the increase in intact LexA could bring the cell back to its basal level of *umuDC* expression (Baptisia *et al.*, 1990).). It has been suggested that as the SOS response begins to shut off, an accumulation of UmuD would lead to the preferential formation of UmuD-UmuD' heterodimers, which results in the inactivation of UmuD' in mutagenesis (Baptisia *et al.*, 1990).

1.3 N-2-Acetylaminofluorene

Among the most extensively characterized of all chemical mutagens and carcinogens are N-2-acetylaminofluorene (2-AAF) and its deacetylated form, N-2-aminofluorene (2-AF) (see Figure 1.2 for structures). Although these compounds are not by-products of industrial processes that pose any environmental hazard, they still provide excellent models for studying metabolic activation, chemical mutagenesis and carcinogenicity of aromatic amines (Hoffmann and Fuchs, 1997). They also provide a

Figure 1.2 Structures of three common aromatic amines. 2-AF represents N-2-aminofluorene. 2-AAF is N-2-acetylaminofluorene. 2-AAAF is the readily reactive compound N-acetoxy-N-2-acetylaminofluorene, also known as N-Aco-AAF).

better understanding of the mechanisms through which DNA damage caused by the covalent bonding of bulky adducts to guanine (G) bases is repaired, or leads to mutations.

1.3.1 Formation of AAF and AF adducts

Some mutagens can react directly with DNA, but many, such as 2-AF and 2-AAF, require metabolic activation. In order to introduce AF adducts into DNA, 2-AF must be converted into the reactive N-hydroxy-N-2-aminofluorene (N-OH-AF) (Hoffmann and Fuchs, 1997). AAF adducts can be introduced either indirectly by the activation of 2-AAF to the derivative N-hydroxy-2-AAF, or directly through the use of the readily reactive compound, N-acetoxy-N-2-acetylaminofluorene (N-Aco-AAF) (refer to Figure 1.2 for structure) (Fuchs *et al.*, 1976). N-hydroxy or N-acetoxy species break down to nitrenium/carbenium ions that react with DNA. The majority of the adducts formed are linked between the C-8 position of guanine and the N-2 of the carcinogen (Broyde and Hingerty, 1983). The two types of adducts formed at the C-8 position are N-(deoxyguanosin-8-yl)-2-acetylaminofluorene (dG-C8-AAF), the acetylaminofluorene adduct, and its deacetylated form N-(deoxyguanosin-8-yl)-2-aminofluorene (dG-C8-AF) (Bichara and Fuchs, 1985; Schaaper *et al.*, 1990). The third type of adduct, the 3-(deoxyguanosin-N²-yl)-2-acetylaminofluorene (dG-N²-AAF), is formed by a linkage between the N-2 position of guanine and the C-3 of 2-AAF (Neft and Heflich, 1994). Approximately 85% of adducts are bound to the C8 position and 15% are bound to the N² position (Hoffmann and Fuchs, 1997).

In vitro formation of adducts can be carried out by treatment of plasmids with N-Aco-AAF. Such treatment results in the formation of the major adduct, dG-C8-AAF. Because the $\text{OC}=\text{OCH}_3$ of N-Aco-AAF is a good leaving group, the bond between the nitrogen and oxygen of this leaving group breaks and allows the nitrenium ion to react with the guanine residue of DNA (refer to Figure 1.3). Binding of the activated carcinogen to the amino group of guanine results in the formation of dG-N²-AAF.

1.3.2 DNA adduct conformation

The binding of AF and AAF adducts to DNA causes conformational changes in DNA. When AAF binds to the C-8 position of deoxyguanosine, this guanine residue rotates around the N-glycosidic bond, from the *anti* to the *syn* conformation, allowing the adduct to be inserted into the helix and displacing the guanine residue (Schaaper *et al.*, 1990; Hoffmann and Fuchs, 1997). This is referred to as the insertion-denaturation model (Fuchs *et al.*, 1976), also known as the base-displacement model (Broyde and Hingerty, 1983). The insertion results in a localized denaturation of the DNA and hinders replication (Garcia *et al.*, 1993). The extent of denaturation of the helix depends on the sequence context surrounding the adducted guanine residue and can extend between three to eight base pairs on either side of the adduct (Hoffmann and Fuchs, 1997). Nuclear magnetic resonance (NMR) studies conducted on a 9 bp AAF-modified DNA oligomer indicated the presence of a predominant conformation. This conformation shows the modified base in the major groove and the fluorene ring protruding into the minor groove,

Figure 1.3 The formation of the major adduct, N-(deoxyguanosin-8-yl)-2-acetylaminofluorene, from the reactive N-acetoxy-N-2-acetylaminofluorene. The product, dG-C8-AAF, is formed by the bond between the N-2 of the carcinogen and the C-8 of DNA.

N-acetoxy-N-2-acetylaminofluorene

N-(deoxyguanosin-8-yl)-2-acetylaminofluorene

(dG-C8-AAF)

which is consistent with the base-displacement/insertion-denaturation model (O'Handley *et al.*, 1993).

In contrast to AAF adducts, AF adducts do not cause such drastic conformational changes in the DNA. The structural change caused by an AF adduct can be described by the outside binding model, in which the fluorene ring remains outside the helix, allowing normal Watson-Crick base pairing between the G-AF and the C (Broyde and Hingerty, 1983; Cho *et al.*, 1994). In addition to the outside binding conformation, AF adducts can also be inserted into the helix without disrupting the B-DNA structure (Broyde and Hingerty, 1983).

Aside from the disruption of G:C base pairing, the conformational change caused by AAF adducts also results in destabilization of the DNA in the vicinity of the adduct. It has been found that DNA modified by N-Aco-AAF showed a decrease in melting temperature of 1.15°C per percent of modified bases (Fuchs *et al.*, 1976). Other evidence which supports the observation that AAF-adducted DNA is less stable include: decreased buoyant density and intrinsic viscosity, increased rates of formaldehyde unwinding (Fuchs and Duane, 1974), increased immunoreaction with antinucleoside antibodies (Santella *et al.*, 1981), increased sensitivity to single-stranded-specific DNA endonucleases (Fuchs, 1975), increased binding to a peptide and a protein which recognize single-stranded nucleic acids, and increased reaction with chemical probes used to detect unpaired nucleotides (Garcia *et al.*, 1993). Another sensitive method for the detection of denatured sites in DNA is the use of a radioimmunoassay which employs anti-cytidine antibodies. Santella *et al.* (1981), found that AAF-modified DNA showed

an increased reactivity with anti-cytidine antibodies, supporting the suggestion that DNA is denatured in the vicinity of the bound adduct. When this assay was conducted on AAF-modified poly(dG-dC) poly(dG-dC), results showed that this polymer adopts a conformation whose properties are similar to those of Z-DNA (Santella *et al.*, 1981). Circular dichroism analyses indicated that DNA randomly modified with 2-AAF is found in the form of B-DNA, while alternating AAF-modified purine-pyrimidine sequences (*e.g.* GCGC) take on the Z form of DNA when modified with 2-AAF (Santella *et al.*, 1981), suggesting that the conformation assumed by dG-C8-AAF is dependent upon the sequence context of the bases surrounding the adduct. In Z-DNA, the adducted dG residue is in the *syn* conformation while the complementary dC residue remains in the *anti* conformation, maintaining the G:C base pairing. Since the C-8 position of the guanine is at the surface of the Z-DNA molecule (Heflich and Neft, 1994), 2-AAF can easily access and attack the C-8 outside the helix, without causing any denaturation in this segment of DNA. This observation is supported by the fact that AAF modified DNA in the Z conformation is resistant to S_1 nuclease digestion and does not react with anti cytidine antibodies (Santella *et al.*, 1981).

1.3.3 Mutagenicity of AAF and AF

Fuchs and colleagues (Fuchs *et al.*, 1981) developed a forward mutation assay based on the detection of mutations in the tetracycline-resistance gene (*tet*) of pBR322 in *E. coli*, and examined the spectrum of mutations caused by AAF and AF adducts. A 276 base pair fragment, between the *Bam*HI and *Sal*I restriction sites in the 5' end of the *tet*

gene, was used as the target of mutagenesis. The plasmid DNA was first treated *in vitro* with the desired mutagen, then either one of two methods was used to obtain isolates in which only the target fragment contains DNA adducts. In some experiments, the fragment was cut from the treated plasmid with restriction enzymes and ligated into untreated plasmid DNA cut with the same enzyme. In other experiments, the treated plasmids were used to directly transform host cells, then an interplasmidic recombination assay was used to distinguish between mutants that carried mutations in the 276 bp fragment and those which carry mutations elsewhere on the plasmid (Koffel-Schwartz *et al.*, 1984). Plasmid DNA carrying adducts within the target 276 bp fragment was transformed into SOS-induced *E. coli* cells and selected for on the basis of ampicillin resistance. Isolates with mutations in the *tet* gene were identified by their tetracycline-sensitive phenotypes. These mutants were analyzed by sequencing.

Koffel-Schwartz *et al.* (1984) discovered that more than 90% of the mutations induced by AAF adducts in the *tet* gene were frameshift mutations, of which 41% were -1, 37% were -2, 16% were -3, and only 6% were +1. After N-Aco-AAF treatment of pBR322, a dose-dependent increase in the frequency of tetracycline-sensitive mutants was observed (Koffel-Schwartz *et al.*, 1984). The observation that most of the mutations occurred at G:C base pairs (Koffel-Schwartz *et al.*, 1984), together with the fact that AAF reacts only with guanine residues, suggests that adducted bases are more prone to mutagenesis. Using the same assay, Bichara and Fuchs (1985) found that, unlike AAF adducts, AF adducts in the *tet* gene resulted primarily in base-pair substitutions (85%

base substitutions, 15% frameshifts), most of which were G:C to T:A transversions (82%).

The work of Bichara and Fuchs (1985) also indicated that there is a correlation between the mutational specificity of AAF and AF adducts and the conformational changes which they induce. The conformational changes induced by AAF adducts (whether by the insertion-denaturation model or the induction of Z-DNA) resulted in frameshift mutations at alternating purine-pyrimidine sequences, while the less severe conformational changes induced by AF adducts tended to generate base substitution mutations.

The forward assay involving N-Aco-AAF-treated pBR322 also revealed that the mutations induced by AAF adducts were not randomly distributed among the modified guanine residues, but rather, that some sites have a higher mutation frequency than others. Two types of such mutational hotspots were identified: sequences in which a single base (primarily G) was repeated; and sequences in which a dinucleotide (*e.g.*, GC) was repeated (Koffel-Schwartz *et al.*, 1984). An example of the latter type of hotspot is the *NarI* restriction site sequence (5'-GGCGCC-3'). While the predominant type of mutation in the repetitive G's was found to be the -1 frameshift, -2 frameshifts occurred with high frequency in the *NarI* sequence (Koffel-Schwartz *et al.*, 1984). The mutation spectrum of AF adducts showed that the few frameshift mutations induced by AF adducts occurred at the same hotspots as those of AAF adducts. However, the base substitutions, which account for the majority of AF mutations, appeared to be more randomly distributed (Bichara and Fuchs, 1985).

Theoretically, mutational hotspots may occur because of selection bias, differential repair or unequal distribution of premutational lesions; however, none of these reasons were applicable to the AAF hotspots found in the *tet* gene used in the pBR322 forward assay (Hoffmann and Fuchs, 1997). Since frameshift mutations at all sites in the target sequence resulted in a tetracycline-sensitive phenotype, selection bias was not considered to be a factor in the occurrence of hotspots. If preferential repair of adducts in less mutable sites was the cause of mutational hotspots, then a *uvrA* mutant, which lacks the nucleotide excision repair mechanism (the major repair pathway for AAF adducts), should show different mutable sites from those of a repair-proficient strain (Hoffmann and Fuchs, 1997). However, this was not the case. Koffel-Schwartz *et al.* (1984), demonstrated that the *uvrA* strain did not show differing mutable sites from the wild-type strain. Lastly, in order to determine whether there exists an unequal distribution of premutational lesions, distribution of AAF adducts was measured in the genetic target. It was found that there was no direct correlation between the distribution of AAF-induced mutations and the distribution of AAF adducts in DNA (Fuchs, 1983). These data support a conclusion that the distribution of AAF-induced mutations in DNA was determined by the way the DNA damage at different sites was processed, rather than by selection bias, differential repair or unequal distribution of premutational lesions (Hoffmann and Fuchs, 1997).

Analysis of the N-Aco-AAF-induced mutations in the *E. coli lacI* gene revealed that 21% of all the mutations found were deletions (Schaaper *et al.*, 1990). A large, 276 bp deletion that was observed 17 times in the AAF-modified *lacI* gene was absent from

collections of spontaneous mutations analyzed previously (Schaaper *et al.*, 1990), suggesting that this deletion mutation was induced by N-Aco-AAF treatment. At the end points of this deletion were 10 bp direct repeat sequences, suggesting that slippage or misalignment may be responsible for generating the mutation (Schaaper *et al.*, 1990).

1.3.4 Mechanisms by which AAF adducts induce mutations

1.3.4.1 Effect of genetic control and replication hindrance on mutagenesis

Bulky adducts can hinder the replication of DNA; therefore, the survival of a cell is dependent on its ability to repair these lesions (by nucleotide excision repair) or to bypass them during replication (Hoffmann and Fuchs, 1997). Because mutations are generated during the processing of damaged DNA, the hindrance of replication by adducts and the bypass of lesions are important steps in mutagenesis. Damage avoidance and translesion synthesis are two different mechanisms by which a cell can survive chemical lesions (Figure 1.4). Damage avoidance processes, which include postreplication recombinational repair and lesion-induced strand switching, use information in the undamaged strand to circumvent the blockage induced by the adduct (Hoffmann and Fuchs, 1997). This process tends to be error-free. In translesion synthesis, the DNA polymerase inserts a nucleotide directly opposite the adducted nucleotide and continues to extend the polynucleotide chain. This latter process may be error-free; however, when it is error-prone, frameshift mutations could arise when elongation proceeds from a slipped mutagenic intermediate at the adduct site (Hoffmann and Fuchs, 1997).

Figure 1.4 Schematic diagram of two different strategies used by cells to tolerate chemical adducts in DNA. Damage avoidance can be in the form of postreplication recombinational repair or lesion-induced strand switching, both of which are error-free. The second strategy, translesion synthesis (TLS), may be error-free or error-prone. (From Hoffmann and Fuchs, 1997).

Damage Avoidance

Translesion Synthesis (TLS)

In vitro studies involving bacteriophage M13 DNA have shown that AAF adducts block replication by T7 DNA polymerase, Sequenase 2, T4 DNA polymerase and *E. coli* DNA polymerase I (Hoffmann and Fuchs, 1997). Single-turnover kinetics using an AAF adduct in a synthetic template has shown that T7 DNA polymerase incorporates dNTPs in a normal fashion up to the site of the adduct, at which point the rate of incorporation of dCTP opposite the adducted guanine becomes 4×10^6 times slower than that opposite an unmodified guanine (Lindsley and Fuchs, 1994). This lower rate of base insertion continues even after incorporation of the base opposite the adducted base (Lindsley and Fuchs, 1994). The T7 DNA polymerase usually incorporates dCTP opposite the adducted guanine; however, when misincorporation does occur, dATP is incorporated more frequently than either dGTP or dTTP (Lindsley and Fuchs, 1994), leading to the induction of a G to T transversion as shown by Bichara and Fuchs (1985).

The blockage of replication by AAF adducts has also been found to cause strand loss in plasmid DNA (Koffel-Schwartz *et al.*, 1987). The effects of AAF adducts in only one strand of a double-stranded plasmid were studied using heteroduplex pBR322 containing a single mismatch as a genetic marker in the *tet* gene. Plasmids with AAF adducts on only one strand were transformed into a *mutS uvrA* strain which is incapable of correcting the single mismatch marker or excising the adducts. In more than 80% of the transformants analyzed, the strand containing the adducts was lost, which is considerably higher than the background frequency of the loss of a strand marker in unmodified plasmids (less than 20%) (Koffel-Schwartz *et al.*, 1987). When plasmids with AAF adducts on both strands were transformed into the same strain, the

transformants contained genetic information from one strand or the other but not both, suggesting that the strand in which lesions are bypassed first are conserved (Koffel-Schwartz *et al.*, 1987).

In terms of genetic control, Koffel-Schwartz *et al.* (1984) and Bichara and Fuchs (1985) have reported that the SOS system plays a role in the mutagenicity of AAF and AF adducts in *E. coli*. As discussed in the previous section (1.2), the SOS system is a set of coordinated responses to DNA damage that is controlled by the *recA* and the *lexA* genes (Hoffmann and Fuchs, 1997). An important component of the SOS response is the activated form of the UmuD protein, UmuD', which in conjunction with UmuC and activated RecA (RecA*), enables the DNA polymerase III holoenzyme to replicate across the damaged DNA site, leading to the induction of mutations (Peat *et al.*, 1996).

The irradiation of bacteria with UV prior to transformation with AAF-modified plasmids resulted in an increase in the AAF-induced mutation frequency, suggesting that AAF-induced mutagenesis was dependent upon the SOS system (Janel-Bintz *et al.*, 1991). Using a specific reversion assay, Janel-Bintz *et al.* (1991) found that AAF-induced -2 frameshift mutations within alternating GC sequences required a LexA-controlled function that is not UmuDC and occurred in the absence of RecA protein, as long as the SOS regulon was derepressed. In contrast, AAF-induced -1 frameshift mutations in repetitive G sequences have been found to be dependent upon the presence of functional *umuDC* gene products (Napolitano *et al.*, 1994).

1.3.4.2 Induction of -1 frameshift mutations in repeated G sequences

Contiguous guanine sequences are hotspots for AAF-induced -1 frameshift mutations. To study the effect of an adduct's position on the frequency of -1 frameshifts, plasmids containing a single AAF adduct on different bases within a run of three and four guanine residues were constructed. The results showed that adducts bound to guanine residues at the 3' end of the repeated guanine sequences had a frequency of -1 frameshift mutations that was three to four orders of magnitude greater than the spontaneous rate, and by two to three orders of magnitude greater than adducts bound to the guanine at the 5' end of the repeated guanines (Lambert *et al.*, 1992; Napolitano *et al.*, 1994). These results led to the suggestion that within a run of Gs, the presence of guanine residues 5' to the adducted base is required for -1 frameshift mutagenesis (Lambert *et al.*, 1992). The use of plasmids containing the mutational sites within different sequence contexts yielded similar results, suggesting that the above observations are characteristic of AAF -induced frameshift mutations in runs of guanine residues (Lambert *et al.*, 1992; Napolitano *et al.*, 1994).

To investigate the effect of the nucleotide on the 3'-side of the G³ position on the induced mutation frequency, the sequence 5'-CG¹G²G³Y-3', where Y is A, G, C, or T, was constructed and used in the mutation assay (Napolitano *et al.*, 1994). The frequency of -1 frameshift mutations was highest when the base immediately 3' to the adducted G³ (base Y) was A, followed by G, C and T, in this order (Napolitano *et al.*, 1994).

One concept that is often used to explain the occurrence of frameshift mutations within contiguous runs of bases is the strand slippage model, proposed by Streisinger and

colleagues over thirty years ago (Streisinger *et al.*, 1966). A proposed slippage model for -1 frameshift mutagenesis induced by an AAF adduct at different positions within a monotonous run of four guanine residues is shown in Figure 1.5. When the adduct is on G², a bulge is created at that position and since the adduct hinders elongation of the new strand (Lindsley and Fuchs, 1994), the cytosine residue that was inserted opposite the adducted guanine during DNA synthesis has an increased chance for slippage (Lambert *et al.*, 1992). The slipped mispaired intermediate (SMI) is stabilized by the cytosine residue that is now paired with G¹. With the adduct on G³, two different SMIs are possible. The intermediate containing two cytosine residues paired with G¹ and G² is more stable than the one containing only one cytosine paired with G¹. When the adduct is bound to G⁴, there are three different SMIs, two of which are similar to the two formed when the adduct is on G³. The last SMI is the most stable due to the three cytosine residues that are available to form normal base pairing with G¹, G² and G³. Once the SMI is stabilized, elongation continues from this intermediate and leads to a -1 frameshift mutation in the newly synthesized strand. The adduct on G⁴ not only forms the most stable intermediate, but also allows the formation of three SMIs, while an adduct bound to G² and G³ can only form one and two SMIs, respectively. This provides an increased opportunity for slippage to occur when the AAF adduct is on G⁴, leading to increased -1 frameshift mutagenesis. The number, and stability, of slipped intermediate structures that may be derived through translesion synthesis past AAF adducts at the 3'-end of a run of Gs (i.e. G⁴) as compared to the 5'-end (i.e. G¹) may account for the marked positional effect of lesions within homopolymeric runs.

Figure 1.5 Slippage model for -1 frameshift mutagenesis induced by an AAF adduct at different positions within a monotonous run of guanine residues.
(From Napolitano *et al.*, 1994)

Heteroduplexes constructed to simulate slipped mutagenic intermediates were used to determine whether or not AAF adducts affect the stability of mispaired intermediates. It was found that the AAF-modified heteroduplex exhibited a melting temperature 6.5°C higher than that of the unmodified heteroduplex, indicating that AAF helps stabilize the mispaired intermediates (Garcia *et al.*, 1993). Using NMR to study heteroduplexes containing a bulged guanine residue (i.e. a slipped intermediate) that is either modified with AAF or unmodified, Milhé *et al.* (1994) observed an increase of 15°C in the melting temperature of the modified over the unmodified heteroduplex. Also, the unmodified heteroduplex showed extensive denaturation in the vicinity of the bulged guanine residue, while the AAF-modified heteroduplex remained paired up to 30°C (Milhé *et al.*, 1994). This stability of the AAF-modified site was suggested to be responsible for the high rate of -1 frameshift mutations at repetitive sequences (Milhé *et al.*, 1994).

In a series of experiments designed to determine the stability of adduct-induced frameshift intermediates, chemical probes such as hydroxylamine and bromoacetaldehyde (which react with single-stranded cytosines) were employed, and found to be highly reactive to cytosine residues opposite the adducted guanine in AAF-modified homoduplexes, suggesting that the AAF adduct had a denaturing effect over the three bases in the complementary strand (Garcia *et al.*, 1993). However, in heteroduplexes in which AAF is bound to a bulged guanine, the cytosines in the complementary strand were found to be non-reactive to the two probes, suggesting that the denaturation induced by AAF in the homoduplexes does not occur in the SMIs (Garcia *et al.*, 1993). These results

suggest that AAF causes a local denaturing effect within the homoduplex, which is alleviated by the formation of a bulge (as in SMIs), contributing to the increase in -1 frameshift mutations observed in repetitive sequences that have been modified with AAF (Garcia *et al.*, 1993).

1.3.4.3 Induction of -2 frameshift mutations in alternating GC sequences

The induction of -2 frameshift mutations in the *NarI* sequence by AAF could theoretically be caused by binding of an adduct to any of the three guanines in the sequence. Due to the nature of the *NarI* sequence (G¹G²CG³CC), it was postulated that three different frameshift events may take place which can generate the GGCC result: the deletion of G²C, CG³, or G³C. To determine the effect of the position of the adducted guanine on mutagenesis, Burnouf *et al.* (1989) constructed three double-stranded plasmid molecules, each carrying a single AAF adduct at one of the three guanines. When these plasmids were transformed into *E. coli* JM103 cells, it was found that only the adduct bound to the third guanine residue resulted in the deletion of a dinucleotide that leads to a -2 frameshift mutation (Burnouf *et al.*, 1989). This same specificity was observed in an experiment involving plasmid vectors, with a simian virus 40 origin of replication and one AAF adduct, that was allowed to replicate in a human cell extract *in vitro* (Hoffmann and Fuchs, 1997).

It has been suggested that a change from *anti* to *syn* in DNA conformation may increase the occurrence of -2 frameshift mutations. It was observed that AAF adducts favour a B to Z transition in synthetic polynucleotides and that alternating GpC sequences

tend to change from B-DNA to Z-DNA form (Santella *et al.*, 1981). This led to the proposal that AAF adducts cause a destabilization of the B-DNA and triggers a local transition to Z-DNA, which in turn promotes the mutational process by making the site a better substrate for proteins which process premutational lesions into mutations (Hoffmann and Fuchs, 1997).

It was suggested that AAF binding to G³ of the *NarI* sequence somehow induces a local conformational change which is responsible for the mutagenic processing that leads to -2 frameshifts. Burnouf *et al.* (1989) reasoned that the major building blocks of Z-DNA are the dinucleotides 5'-CG or 5'-TG, and since CG³ is the only such building block in the *NarI* sequence, it is suspected to be the one which induces the B to Z transition. It was suggested that aside from the presence of the 5'-CG dinucleotide, the ability of the neighbouring bases to accommodate the two B-Z junctions that must form on both sides of the CG³ unit also plays an important role in the *NarI* site being the hotspot for AAF-induced -2 frameshift mutations (Burnouf *et al.*, 1989).

Koehl *et al.* (1989) further studied this structural effect of the position of the AAF adduct within the *NarI* sequence by constructing a *NarI* containing 12 bp oligomer (ACCG¹G²CG³CCACA). The oligomers were modified with N-Aco-AAF and three guanine-AAF monoadducted isomers were isolated. Each monoadducted oligomer and the unmodified oligomer (used as a control) were annealed to the complementary strand and CD spectra of the four different duplexes were analyzed and compared. When the CD spectrum of the unmodified duplex was compared to the spectra of the duplexes containing the AAF adduct at the G¹ and G² positions, relatively small alterations of the

CD signal was found. However, the difference spectrum for the unmodified duplex and the duplex with AAF-G³ was found to be very similar to a Z-form minus B-form differential spectrum, suggesting that AAF modification of the third guanine residue in the *NarI* sequence causes a major change in the local structure of the DNA helix from the B-form to the Z-form. This led to the suggestion that the -2 frameshift mutation within the *NarI* sequence results from the processing of the unusual DNA structure induced by AAF binding to the third guanine residue in this sequence (Koehl *et al.*, 1989).

A different collection of evidence, however, suggests that -2 frameshift mutations in *NarI* and *NarI*-like sequences are associated with slippage during replication (Hoffmann and Fuchs, 1997). The observation that one single AAF adduct close to the origin of replication in the lagging strand was 20-fold more likely to generate a -2 frameshift than an adduct on the leading strand (in *E. coli*) provided support for the suggestion that -2 frameshift mutagenesis by AAF adducts is linked to replication (Veaute and Fuchs, 1993). The observation that an AAF adduct on G³ of the *NarI* sequence was very mutagenic while the adducts on G¹ and G² were not (Burnouf *et al.*, 1989) supports the theory that slippage during replication is the mechanism by which -2 frameshift mutations are generated, because only AAF adducts on G³ yielded the slipped mutagenic intermediate required to generate -2 frameshift mutations (Hoffmann and Fuchs, 1997). Therefore, it was proposed that a slippage model similar to that of AAF-induced -1 frameshift mutations can be used to describe the induction of -2 frameshifts by AAF (Hoffmann and Fuchs, 1997).

Tebbs and Romano (1994) examined the mutagenesis of a *NarI* site that was modified with an AAF adduct and found that -2 frameshifts were the predominant mutations at *NarI* sites. A -2 frameshift mutation in the *NarI* sequence must delete one of the following dinucleotides, G²C, CG³, or G³C, in order to give the resultant GGCC sequence. Of the 68 dinucleotide deletions analyzed, 62 were -CG deletions, which strongly suggests that there was a preference for the deletion of CG dinucleotides over the deletion of GC dinucleotides (Tebbs and Romano, 1994). This led to the conclusion that an AAF adduct within a *NarI* sequence causes a conformational change in the DNA which results in the deletion of a CG dinucleotide during replication (Tebbs and Romano, 1994).

A slightly different model for the formation of -1 and -2 frameshift mutations was proposed by Bertrand-Burggraf *et al.* (1994). They presented evidence that T4 Endonuclease VII, a resolvase, is able to recognize the distortion caused by an AAF adduct and create a nick across from the adduct, in the non-adducted strand, and suggested that a similar resolvase in *E. coli* also nicks double-stranded DNA containing adducts in the same manner (Bertrand-Burggraf *et al.*, 1994). Once a nick has been created, a gap is produced starting at the nick by exonucleolytic activity, and a slipped mutagenic intermediate is formed during repair synthesis of the gap (Bertrand-Burggraf *et al.*, 1994).

1.4 Objectives

Interest in mutation research has grown considerably in recent years because of the link between mutations and many diseases. One class of mutations, deletion mutations, is not very well characterized because they make up a relatively small proportion of mutations and are usually difficult to isolate.

There are presently many studies which show that AAF modification of DNA results in an increased frameshift mutation frequency. A study conducted by Schaaper *et al.* (1990) suggested that random modification of the *E. coli lacI* gene with AAF can increase the frequency of different types of mutations, including deletions. The objective of this thesis is to develop a sensitive plasmid-based reversion assay for the detection and characterization of spontaneous and AAF-induced deletion mutations. This system will be used to analyze the effects of AAF adducts, SOS induction, as well as the sequence context surrounding the deleted insert on the frequency of deletion mutation.

2. MATERIALS AND METHODS

2.1 Bacterial strains, plasmids and primers

The bacterial strains, plasmids and sequencing primers used in this study are listed in Tables 2.1, 2.2 and 2.3, respectively.

Table 2.1	<i>E. coli</i> strains used in this study	
Strain	Genotype/Phenotype	Reference
TK603	<i>uvrA6</i> derivative of AB1157 <i>thrA1 leuB6 90(gpt-proA)62 hisG4(0c)</i> <i>ilv325 lacY1 galk2 ara14 mtl1 xyl5</i> <i>uvrA6 supE44 tsx33 rpsL31=3DstrA31</i> <i>thi1 rac⁻ rfbB D1 mgl51 kdgk51 str</i> resistant	Kato and Shinoura (1977)
TG1	used both <i>uvrA-F'</i> and <i>uvrA+F'</i> Δ <i>lac</i> strain with <i>lacI</i> in a F' episome Δ (<i>lac-pro</i>) F' <i>traD36 proA⁺B⁺</i> <i>lacI^qlacZ</i> Δ M15 <i>thi supE hsd</i>	Maniatis <i>et al.</i> (1989)
DH5 α	<i>endA1 hsdR17(r_k⁻ m_k⁺) supE44 thi-1</i> <i>recA1 gyrA(Nal^r) relA1 Δ(lacIZYA-argF)</i> <i>ΔlacU169 deoR(ϕ80<i>lacZ</i> ΔM15)</i>	Hanahan, D. (1983)

Table 2.2	Plasmids used in this study		
Plasmid	Antibiotic marker	promoter/insert/ <i>Ori</i> orientation	Reference or source
pNS17p	Amp/Cm	palindromic/forward	Bichara <i>et al.</i>
pNS17np	Amp/Cm	non-palindromic/forward	Bichara <i>et al.</i>
pNSR17p	Amp/Cm	palindromic/reverse	Bichara <i>et al.</i>
pNSR17np	Amp/Cm	non-palindromic/reverse	Bichara <i>et al.</i>
pLC	Amp/Kan	<i>lac</i> /no insert/forward	Bichara <i>et al.</i>
pLC20p	Amp/Kan	<i>lac</i> /palindromic/forward	this study
pLC20np	Amp/Kan	<i>lac</i> /non-palindromic/forward	this study
pLC20T	Amp/Kan	<i>lac</i> /triplet repeat/forward	this study

pLCR	Amp/Kan	lac/no insert/reverse	Bichara <i>et al.</i>
pLCR20p	Amp/Kan	lac/palindromic/reverse	this study
pLCR20np	Amp/Kan	lac/non-palindromic/reverse	this study
pLCR20T	Amp/Kan	lac/triplet repeat/reverse	this study
pTC	Amp/Kan	ptac/no insert/forward	Bichara <i>et al.</i>
pTC20p	Amp/Kan	ptac/palindromic/forward	this study
pTC20np	Amp/Kan	ptac/non-palindromic/forward	this study
pTC20T	Amp/Kan	ptac/triplet repeat/forward	this study
pTCR	Amp/Kan	ptac/no insert/reverse	Bichara <i>et al.</i>
pTCR20p	Amp/Kan	ptac/palindromic/reverse	this study
pTCR20np	Amp/Kan	ptac/non-palindromic/reverse	this study
pTCR20T	Amp/Kan	ptac/triplet repeat/reverse	this study

Table 2.3 Sequencing primers used in this study

Primer	Sequence
pPAK1	5'-ACCTGATTGCCCCGACAT-3'
pSMYL1	5'-GGTGTAAACAAGGGTGAA-3'
ptacprime1	5'-GGAAAGCGGGCAGTGAGCG-3'

2.2 Growth of bacterial cultures

Each bacterial strain was grown overnight in Luria (LB) broth (1.0% tryptone, 0.5% NaCl, 0.5% yeast extract). Six mL of overnight cultures for each strain were pelleted, resuspended in 0.5 mL of 10% glycerol and stored at -80°C. Bacterial cultures were also maintained on LB agar (LBA) plates and slants, which were stored at 4°C.

2.3 Preparation of electrocompetent cells

An overnight culture (2 mL) was used to inoculate 250 mL of LB broth and shaken at 37°C to an OD₆₀₀ of approximately 0.5. The culture was centrifuged at 5000 x g, 2°C, for 15 minutes. The pelleted cells were resuspended in 250 mL of 10% ice cold glycerol and centrifuged as above. This washing procedure was repeated, using 125 mL and then 65 mL of 10% glycerol. The cells were resuspended in one mL of 10% glycerol, transferred to a two mL Eppendorf tube and respun. The volume was brought down to one mL, the cells were resuspended and 40 µL aliquots were distributed into 1.5 mL Eppendorf tubes. Competent cells were stored at -80°C. Dilutions (10⁻⁷ and 10⁻⁸) of these cells were plated onto LB agar plates in triplicate to determine the final cfu/mL.

2.4 Preparation of SOS-induced electrocompetent cells

An overnight culture (2 mL) was inoculated into 250 mL of LB broth and shaken at 37°C to an OD₆₀₀ of approximately 0.5. The cells were pelleted (10,000 x g, 10 min., 2°C) and resuspended in 150 mL of MgSO₄ (10 mM). A 10⁻⁶ dilution (prepared in LB) was plated (50 µL) in triplicate onto LB agar (labelled A to determine the cfu/mL before UV exposure). The resuspended cells were divided into 6 petri dishes (25 mL per plate) and irradiated with UV (70 J/m²) while rocking. A 10⁻⁵ dilution was plated (50 µL) in triplicate onto LB agar (labelled B to determine the cfu/mL after UV exposure). The cells were pelleted at 10,000 x g, 2°C for 10 minutes, resuspended in 150 mL LB and shaken at 37°C for 30 minutes to allow for expression of SOS functions. The cells were pelleted (10,000 x g, 10 min., 2°C), resuspended in 250 mL of cold 10% glycerol and centrifuged

again. The washings were repeated in 125 mL then 65 mL of 10% glycerol. The cells were resuspended in one mL of 10% glycerol, divided into 40 μ L aliquots and stored at -80°C . A 10^{-8} dilution was plated onto LB agar plates (labelled C) in triplicate to ensure a final cfu/mL of $> 10^9$. The level of survival of SOS-induced cells was calculated as follows:

$$\text{Survival} = (\text{B/A}) \times 100\%.$$

2.5 Transformation by electroporation

Plasmid DNA (0.1 μ g) was added to a 40 μ L aliquot of competent cells, mixed, transferred to a cold 0.2 cm cuvette and left on ice for one minute. The cells were pulsed at 2.5 KV and 960 μ L SOC medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl_2 , 10 mM MgSO_4 , 20 mM glucose) was immediately added to the cells, mixed, transferred to a 15 mL test tube and shaken at 37°C for 15 to 120 minutes (depending on the plasmid used). Dilutions were made in LB, then plated in triplicate on LBA containing the appropriate antibiotics for selection of the desired plasmid. The plates were incubated at 37°C overnight.

2.6 Extraction of plasmid DNA

2.6.1 Boiling miniprep method for the extraction of plasmid DNA

An overnight culture (4 mL) from a single colony carrying the desired plasmid was pelleted in a microfuge (5 min., 14,000 $\times$ g). The pellet was resuspended in 400 μ L of STET buffer (50 mM EDTA, 10 mM Tris-HCl, 8% w/v sucrose, 5% v/v Triton X).

Twenty μL of lysozyme (10 mg/mL) was added, vortexed briefly and incubated in a boiling bath for 40 seconds. The lysed cells were centrifuged (5 min., 4°C , $13,200 \times g$) in an angle rotor Eppendorf centrifuge. The gelatinous pellet was discarded, then 1/5 the volume of 10 M ammonium acetate and two volumes of ice cold ethanol were added and mixed. The solution was centrifuged (30 min., 4°C) and the pellet was washed with one mL of 70% ethanol. The pellet was dried in a vacuum oven at 50°C for 10 minutes and resuspended in 20 μL of TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

2.6.1.1 Purification of plasmid DNA by Gene Clean

Prior to double-stranded sequencing, plasmid DNA was purified using the Gene Clean method. NaI (300 μL) was added to the DNA in TE buffer, mixed, then five μL of Glass Milk was added, mixed and held on ice for five minutes. The tube was spun for five seconds to pellet the glass milk and the supernatant was discarded. The pellet was washed three times with 400 μL of New Wash solution, resuspended in 10 μL of dH_2O and incubated at 50°C for five minutes. The glass milk was again pelleted and the supernatant containing the DNA was removed and stored in a clean Eppendorf tube.

2.6.2 Extraction of plasmid DNA using Wizard™ Miniprep Kit

Cells from an overnight culture (4 mL) were pelleted in a microfuge at $10000 \times g$ for two minutes. The pellet was resuspended in 200 μL of Cell Resuspension Solution (50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 100 $\mu\text{g}/\text{mL}$ RNase A). The cells were lysed by mixing with 200 μL of Cell Lysis Solution (0.2 M NaOH, 1% SDS) and neutralized

by the addition of 200 μ L of Neutralization Solution (1.32 M potassium acetate, pH 4.8). The lysate was centrifuged at 10,000 x g for five minutes in a microfuge. The clear lysate was transferred to a fresh Eppendorf tube and one mL of resin was added. A Wizard Minicolumn was attached to the end of a syringe barrel and placed onto a vacuum manifold. The lysate-resin mixture was transferred to the syringe barrel and a vacuum was applied until the sample had completely passed into the column. The resin was washed with two mL of Column Wash Solution (80 mM potassium acetate, 8.3 mM Tris-HCl, pH 7.5, 40 μ M EDTA, 55% ethanol). The column was detached from the syringe and centrifuged at 10,000 x g for two minutes in a microfuge to remove residual wash solution. TE buffer (30 μ L) was added to the column and after five minutes, the plasmid DNA was eluted by centrifuging the column again at 10000xg for one minute. The DNA was stored at -20°C.

2.7 Large scale extraction of plasmid DNA

2.7.1 Large scale plasmid DNA extraction by CsCl-EtBr gradient

A single colony of DH5 α containing the desired plasmid was grown overnight in LB broth containing ampicillin (50 μ g/mL). The cell culture was pelleted in the Sorvall RC-5B centrifuge using the GSA rotor (5000 rpm, 5 min., 4°C) and resuspended in a total of 30 mL of TS buffer (50mM Tris-HCl, pH 8.0, 20% (w/v) sucrose). Six mL of EDTA (0.25 M, pH 8.0) was added, vortexed and two mL of lysozyme in 0.25 M EDTA, pH 8.0 (10 mg/mL) was added and gently swirled. After five minutes on ice, eight mL of Triton-X detergent mixture (2.0% (v/v) Triton X-100, 50 mM Tris-HCl, pH. 8.0, 25 mM

EDTA) was added, mixed gently and centrifuged (Sorvall GSA rotor, 15,000 rpm, 45 min., 4°C). The supernatant was transferred to a clean tube, then 15 mL of PEG/NaCl mix (40% (w/v) polyethylene glycol 6000, 2 M NaCl) was added, mixed and stored at 4°C overnight.

The DNA-PEG/NaCl solution was precipitated (Sorvall GSA rotor, 5000 rpm, 20 min., 4°C), the pellet was resuspended in 3.5 mL of TE buffer and 4.00 g of CsCl was added. Once the CsCl was completely dissolved, 320 µL of 10 mg/mL ethidium bromide (EtBr) was added and centrifuged (Sorvall SS34 rotor, 10,000 rpm, 10 min., room temp.). The supernatant was transferred to a plastic tube (Beckman Quick-Seal), the refractive index was adjusted to 1.3860, the remainder of the tube was filled with mineral oil and the top was heat sealed. The tube was centrifuged (Beckman Vti80 fixed angle rotor, 60,000 rpm, 16 hours, 15°C) under vacuum to separate desired plasmid DNA from RNA and proteins.

The lower band, containing the desired closed circular plasmid DNA, was transferred into a small plastic tube using an 18-gauge hypodermic needle (under UV light). The EtBr was removed by extraction with CsCl saturated isopropanol until the DNA phase was no longer pink. To remove CsCl and residual EtBr, the DNA solution was dialyzed against TE (6 hours, 4°C, gentle stirring). The DNA solution was precipitated in 400 µL of 10 M ammonium acetate and four mL of cold absolute ethanol at -20°C overnight. The plasmid DNA was precipitated (IEC Centra MP4R swinging bucket rotor, 8000 rpm, 50 min., 4°C), vacuum dried (15 min., 50°C) and dissolved in 100 µL of TE.

2.7.2 Wizard Midiprep System for the extraction of plasmid DNA

An overnight culture (100 mL) containing the desired plasmid was pelleted at 10,000 x g, 4°C for 10 minutes. The cell pellet was resuspended in three mL of Cell Resuspension Solution. Cell Lysis Solution (3 mL) was added, mixed, then Neutralization Solution (3 mL) was added and mixed. The cell suspension was centrifuged at 14,000 x g for 15 minutes at 4°C. The supernatant was transferred to a clean tube and 10 mL of Purification Resin was added, mixed and transferred to a Wizard Midicolumn. A vacuum was applied to pull the resin/DNA mix into the column. Two 15 mL aliquots of Column Wash Solution were used to wash the resin. The column was cut from the reservoir and centrifuged at 14,000 x g in a microfuge to remove residual wash solution. TE buffer (200 µL), preheated to 50°C, was added and allowed to be absorbed by the resin. The plasmid DNA was eluted by centrifugation at 14,000 x g for five minutes and stored at -20°C.

2.8 Double-stranded DNA sequencing

2.8.1 Annealing using double-stranded templates

For each DNA sample, 5-10µg DNA (~7 µL of DNA from boiling miniprep) was combined with 15 pmol primer (pPAK1 or pSMYL1, depending on the plasmid sequenced) and brought up to eight µL with dH₂O. The mixture was boiled for five minutes, quickly put into wet ice for transfer to 4°C microfuge, briefly spun down (~ 10 seconds), then stored on ice until needed.

2.8.1.1 Preparation and annealing of single-stranded templates

Single stranded templates for plasmid DNA (from Wizard Miniprep Kit) were made for sequencing by taking approximately two μg of DNA and making the volume up to 90 μL with dH_2O . Ten μL of NaOH (2 N) was added, mixed and incubated at 37°C for 30 minutes. Eleven μL of NaAc (3 M, pH 5.0) was added and mixed by vortex before 350 μL of cold EtOH/NaAc (0.16 M) and 10 μL of glycogen (2 mg/mL) were added and mixed by inversion. The solution was stored at -20°C overnight.

The DNA was precipitated in a 4°C microfuge for 30 minutes at 13,200 x g, washed with one mL of 70% ethanol, dried in a vacuum oven (50°C) and resuspended in seven μL of TE. The hybridization of primer to plasmid DNA was carried out by adding 1 μL of primer (1 pmol/ μL) and 2 μL of 5X Sequenase Buffer (200 mM Tris-HCl, pH 7.5, 100 mM MgCl_2 , 250 mM NaCl) to the seven μL of single stranded template. The mixture was warmed to 65°C in a water bath for two minutes, then the water bath was allowed to cool slowly to 30°C . The DNA was immediately put on ice until labelling.

2.8.2 Sequencing reactions

Reaction mixtures were made by adding two μL of Sequenase buffer (only for double stranded templates), two μL of diluted labelling mix (1.5 μM dGTP, 1.5 μM dCTP, 1.5 μM dTTP), one μL of 0.1 M dithiothreitol (DTT) and 0.5 μL α - ^{35}S -ATP (approximately 10 μCi) to the DNA/primer sample. Four units (4U) of Sequenase Version 2.0 T7 DNA Polymerase (20 mM KPO_4 , pH 7.4, 1 mM DTT, 0.1 mM EDTA, 50% glycerol) in Sequenase dilution buffer (10 mM Tris-HCl, pH 7.5, 5 mM DTT, 0.5

mg/mL BSA) was added to the side of each tube. Into separate wells of a 96 well microtitre plate, 2.5 μ L of each termination mix (80 μ M dATP, 80 μ M dCTP, 80 μ M dGTP, 80 μ M dTTP, 50 mM NaCl; plus 8 μ M of one of the dideoxynucleotides) was added and heated to 37°C for two minutes on the Hybaid Combi Thermal Reactor TR2. The reaction mixtures were centrifuged to combine all the ingredients and 3.2 μ L of each reaction mixture was added to each microtitre well containing the dideoxynucleotides. After five minutes at 37°C, four μ L of STOP solution (95% formamide, 20 mM EDTA, 0.05% Bromophenol blue, 0.05% xylene cyanol FF) was added to each well, and the samples were stored at -20°C until it was time to load onto the gel.

2.8.3 Sequencing polyacrylamide gel electrophoresis

The polyacrylamide gel was made by adding 35 μ L of N,N,N',N'-tetramethylethylenediamine (TEMED) and 280 μ L of 10% ammonium persulfate to 70 mL of a ready-made 8% prefab solution (40% acrylamide/bisacrylamide, 5X TBE, 8.5 M urea). The gel was prerun for 30 min. (60 watts) in 1X TBE (89 mM Tris-borate, 2 mM EDTA, pH 8.3). The reactions were heated at 75°C for three minutes before loading. Three μ L of each reaction was loaded into the wells and the gel was run at 60 watts (constant power) for approximately two hours. The gel was transferred onto chromatography paper and then dried under vacuum for two hours at 80°C. Once dried, the gel was exposed to Kodak Biomax MR-1 film overnight at room temperature and the film was developed the following day.

2.9 Plasmid Construction

Table 2.4 Oligomers used in this study

Oligomer	Sequence
20p	5'-AATTGGTCACCTATCAATTCGAGCTTGAGTATCGTACGATACTCAGTCACCTATC-3'
20p _C	3'-CCAGTGGATAGTTAAGCTCGAACTCATAGCATGCTATGAGTCAGTGGATAGTTAA-5'
20np	5'-AATTGGTCACCTATCAATTCGAGCTTGCAAGCGATACGATACTCAGTCACCTATC-3'
20np _C	3'-CCAGTGGATAGTTAAGCTCGAACGTTGCTATGCTATGAGTCAGTGGATAGTTAA-5'
20T	5'-AATTGGTCACCTATCAATTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCGTCACCTATC-3'
20T _C	3'-CCAGTGGATAGTTAAGCTCGACGACGACGACGACGACGACGACGACGACGAGTGGATAGTTAA-5'

2.9.1 Digestion of parent plasmids pLC, pLCR, pTC and pTCR

Each parent plasmid (2 µg) was digested in a 20 µL reaction containing 20 U of *EcoRI* enzyme (New England Biolabs), 10X *EcoRI* buffer (supplied by the manufacturer), and BSA (1 mg/mL). The reaction was incubated at 37°C for 2.5 hours. TE buffer (250 µL) and 100 µL of PCI (phenol:chloroform:isoamyl alcohol in the ratio of 25:24:1) were added, shaken for one minute and centrifuged at 14,000 x g for five minutes in a microfuge. The aqueous layer containing the DNA was transferred to a clean tube, 900 µL of EtOH/NaAc (0.16 M) and 10 µL of glycogen (2mg/mL) were added, mixed and stored at -20°C overnight. The DNA was pelleted (13,200 x g, 4°C, 30 min.), washed with 300 µL of 70% ethanol, vacuum dried and dissolved in 10 µL of TE.

2.9.2 Preparation of inserts

The oligomers 20p, 20p_C, 20np, 20np_C, 20T, 20T_C (100 pmol of each) were phosphorylated in 10 µL reactions with the addition of 30 U of T4 kinase (NEB), 10X T4

kinase buffer (supplied by the manufacturer) and ATP (1 mM). The ingredients were mixed and incubated for one hour at 37°C. Inactivation of the enzyme was carried out at 65°C for 10 minutes.

The oligomers were annealed in 20 µL reactions, using 50 pmol of each of the phosphorylated oligomers and its complement, as well as 10X T4 ligase buffer (from NEB) and made up to volume with ddH₂O. Once mixed, the reaction was incubated at 95°C for three minutes, then allowed to slowly cool to room temperature in a water bath, yielding a final concentration of 2.5 pmol/µL of each insert.

2.9.3 Ligation of digested vector and insert

The ligation reactions were carried out in 10 µL reactions. Vector DNA (0.5 µg), 0.15 pmol of insert, 100 U of T4 DNA ligase, 10X ligase buffer, BSA (1 mg/mL) and ATP (10 mM) were combined in an Eppendorf tube, mixed and incubated at 16°C overnight. The ligation mixtures were then PCI extracted, precipitated in 0.16 M EtOH/NaAc and glycogen (2mg/mL) overnight. The mixture was centrifuged (13,200 x g, 4°C, 15 min.), the pellet was washed with ethanol, dried in a vacuum oven (50°C) and dissolved in 10 µL of TE.

2.9.4 Screening for clones containing inserts in the correct orientation

The ligation mixture from each type of plasmid (3 µL) was transformed into DH5_α competent cells as described in section 3.5, and plated onto LBA plates containing ampicillin (50 µg/mL). Transformants were gridded onto LBA-Amp (50µg/mL) plates

using sterile toothpicks and incubated at 37°C overnight. These master plates were replica plated onto LBA plates containing ampicillin and kanamycin (50 µg/ mL and 10 µg/mL, respectively), which were again incubated at 37°C overnight. The colonies which failed to grow on the second set of plates carried the desired inserts. The plasmid DNA from these colonies were extracted using the Wizard Miniprep Kit (described in 2.6.2) and sequenced (as described in 2.8) to determine whether the insert had been cloned into the *EcoRI* site in the correct orientation. Large scale plasmid DNA extraction from the desired clones was carried out using the CsCl-EtBr gradient method (described in 2.7).

2.9.5 Calculation of ΔG values

The ΔG values of hairpin structures (modelled as RNA) formed by the 17 bp palindromic and 17 bp non-palindromic inserts in the pNS system were calculated using the program MacDNAsis. The ΔG values for hairpin structures (also modelled as RNA) formed by the 20 bp palindromic, 20 bp non-palindromic and 20 bp triplet repeat inserts in the plac and ptac systems were also calculated using MacDNAsis. These values were used to analyze and compare the relative stabilities of the hairpin structures formed by each type of insert.

2.10 Modification of plasmid DNA with N-acetoxy-N-2-acetylaminofluorene

Some of the pNSR (2), ptac (5) and plac (2) plasmids were modified with (N-Aco-AAF) to determine the effect of AAF-modification on deletion frequency. The

modification was carried out by the Bichara laboratory in Strasbourg, France. This modification procedure involved drying down 100 µg of the plasmid DNA and resuspending it in 200 µL of citrate buffer (2 mM, pH 7.0), yielding a final DNA concentration of 0.5 µg/µL. This DNA solution was reacted with six µL of tritium-labelled, N-Aco-AAF in ethanol (400 µg/mL) at 37°C. At one minute intervals, the equivalent of 20 µg of DNA (40 µL) was removed from the reaction tube and added to 120 µL of ethanol acetate (0.16 M) on ice to stop the reaction. Once the last aliquot of the reaction was stopped, all five aliquots were extracted four times with ice cold ethanol acetate (0.16 M) to remove any excess N-Aco-AAF. The DNA was then pelleted in a microfuge (14,000 x g, 15 min., 4°C), washed with 70% ethanol and resuspended in 500 µL of TE buffer. The number of moles of incorporated AAF was calculated as the amount of radioactivity incorporated into the plasmid DNA. The amount of plasmid DNA was determined spectrophotometrically.

2.11 Statistical analysis of deletion frequency data

The statistical analyses of the deletion frequency data were carried out using the F-test, the student's t-test and the Mann-Whitney test. Prior to performing any of these tests, the deletion frequencies of all the plasmids were transformed using the arcsine transformation ($X' = \arcsin\sqrt{X}$). Statistical theory indicates that percentages or proportions tend to form a binomial instead of a normal distribution (Zar, 1974). The data which result from the arcsine transformation will have an underlying distribution that is nearly normal. The arcsine transformation involved taking the square-root of the

deletion frequency, then calculating the arcsine of that number. Once this transformation was completed on all the frequency data, the transformed values were averaged and the variance of each set of data was determined.

The determination of whether or not a difference exists between two sets of frequency data was carried out using the student's t-test and the Mann-Whitney test. The two-sample t-test assumes that the two sets of data being compared came at random from normal populations with equal variances (Zar, 1974). The F-test was used to determine if the variances of the two sets of data to be compared were equal. This test is a simple variance ratio test employing the equation $F = S_1^2/S_2^2$, where S_1^2 and S_2^2 are the variances of the two sets of data being compared and S_1^2 is the larger of the two values. The result of this test was compared to the value given in a table containing critical values of the F distribution. If the calculated F value is greater than the critical value, the variances are not considered to be equal.

The t-test, used to compare the mean deletion frequencies of two different plasmids, was performed using the following equation:

$$t = (X_1 - X_2)/\sqrt{(S_p^2/n_1 + S_p^2/n_2)}.$$

The pooled variance, S_p^2 , was calculated using the equation:

$$S_p^2 = (v_1 S_1^2 + v_2 S_2^2)/(v_1 + v_2), \text{ where}$$

X_1 and X_2 are the mean deletion frequencies being compared,
 n_1 and n_2 are the number of samples in each data set,
 S_1^2 and S_2^2 are the variances of each data set, and
 v_1 and v_2 are the number of degrees of freedom.

For the t-test, I chose to use the 0.05 level to define significance.

When the variances of the data being compared were found to be unequal by the F-test, the t-test could not be used to compare the means. Therefore, a nonparametric procedure, the Mann-Whitney test, was used to test for differences between the dispersion, or variability, of the two sets of data. In this test, the actual deletion frequencies were not used, but rather, the numbers were ranked from highest to lowest, with the highest deletion frequency assigned a rank of 1, the second highest was assigned a rank of 2, and so on. The Mann-Whitney statistic was calculated using the equations:

$$U = n_1 n_2 + [n_1 (n_1 + 1) / 2] - R_1 \text{ and } U' = n_1 n_2 - U, \text{ where}$$

n_1 and n_2 are the number of samples in each data set, and R_1 is the sum of the ranks of one set of data.

The value U' was calculated only to determine whether U or U' was the larger value.

The larger of the two values obtained was compared to the critical values of the Mann-Whitney U distribution to determine whether or not the difference between the data sets was significant. A calculated U value that is greater than the critical value indicates that there is a significant difference between the two sets of data, at the indicated level of confidence. For the Mann-Whitney test, I chose to use the 0.05 level to define significance. However, due to the small sample sizes of some of the comparisons, it was not possible to use 0.05. In these cases, the 0.1 level was used to define significance.

The deletion frequency data was analyzed using each of the above-mentioned tests. Due to the assumptions necessary for the t-test, it was not possible to perform the t-test on all the comparisons made, therefore the Mann-Whitney test was also performed on all the comparisons. It must be noted, however, that the Mann-Whitney test is not as

strong as the t-test, in that the t-test is able to determine significant differences in data sets containing small sample numbers and the Mann-Whitney test requires larger samples in order to do so.

3. RESULTS

3.1 Plasmids Constructs

3.1.1 The pNS plasmids

The pNS parent plasmid was constructed by cloning the chloramphenicol (Cm) resistance gene, isolated within a *Bsa*AI restriction fragment, from plasmid pACYC184 into the *Ssp*I site of pUC8. This construct contains both the ampicillin and chloramphenicol resistance genes. The *Eco*RI site located in the *lacZ* polylinker region of pUC8 was destroyed, leaving a unique *Eco*RI site within the Cm gene, at position 3583, into which the 17 bp palindromic or 17 bp non-palindromic insert was cloned to yield plasmids pNS17p and pNS17np. In the case of the 17 bp palindromic insert, part of one of the flanking direct repeats can base pair with adjacent bases on the same strand to form the palindrome. The 17p and 17np insert sequences are as follows.

17p 5'-AATTGGTCATCTATCAATTCCGATAGATGACGATTCATCGTCATCTATC-3'
 3'-CCAGTAGATAGTTAAGGCTATCTACTGCTAAGTAGCAGTAGATAGTTAA-5'

17np 5'-AATTGGTCATATATCAATTCCGCAAGTCGGAAGTTCATCGTCATCTATC-3'
 3'-CCAGTAGATAGTTAAGGCGTTCAGCCTCTAAGTAGCAGTAGATAGTTAA-5'

In order to construct a plasmid with an origin of replication in the reverse orientation, a second *Afl*III site was placed at position 1599, 815 bases downstream from the first *Afl*III site (at 784). The *Afl*III fragment was excised and inserted in the opposite orientation to yield plasmids pNSR17p and pNSR17np. The characteristics of the four resultant plasmid constructs pNS17p, pNS17np, pNSR17p and pNSR17np are listed in

Table 2.2. Figure 3.1 contains plasmid maps of the two parent constructs, pNS and pNSR, with the insert sequences shown at the bottom.

3.1.1.1 Optimization of reversion assay conditions for pNS plasmids

Transformation experiments were conducted using the pNS parental plasmid and *E. coli* TK603 cells, an expression time of 15 or 30 minutes, and two types of selective media containing varying concentrations of Amp (50 and 150 $\mu\text{g/mL}$) and Cm (10, 20 and 170 $\mu\text{g/mL}$). After overnight incubation at 30°C, all the plates containing only Amp showed growth (at Amp concentrations of 50 and 100 $\mu\text{g/mL}$). The Cm plates (at a Cm concentration of 170 $\mu\text{g/mL}$) were incubated for another 48 hours, but still showed no signs of growth. However, the plates with 10 and 20 $\mu\text{g/mL}$ of Cm showed considerable growth after 72 hours of incubation at 30°C. From these initial results, it was determined that the best conditions for the reversion assay using the pNS plasmids would be 30 minutes expression time, selective media LB-Amp50 and LB-Amp50-Cm10, and incubation times of 24 and 72 hours for each type of medium, respectively. The numbers following the antibiotic names indicate the antibiotic concentration in $\mu\text{g/mL}$.

In collaboration with our laboratory, Dr. Bichara's laboratory in Strasbourg, France also conducted experiments using the pNS plasmid system. It was found that the deletion mutants were temperature sensitive. Also, several colonies growing on Cm plates were phenotypically Cm resistant, but sequencing data revealed that they were genetically Cm sensitive. These results and the fact that Cm is bacteriostatic and not bactericidal prompted the development of the plac plasmid system (below) for the

Figure 3.1 Plasmid maps of the parental pNS plasmids, pNS and pNSR. The plasmid on the left (pNS) contains an origin of replication in orientation 1, while the plasmid on the right contains an origin of replication in orientation 2. The 17p (palindromic) and 17np (non-palindromic) insert sequences are shown at the bottom of the figure. *Ori* is the origin of replication. *Cm* is the chloramphenicol resistance marker and *Amp* is the ampicillin resistance marker. Relevant restriction sites are indicated.

Inserts: Palindromic 5'-AAATTGATAGATGACGATGAATCGTCATCTATCGGAATTGATAGATGACC-3'

Non-palindromic 5'-AAATTGATAGATGACGATGAATCTCCGACTTCCGGAATTGATAGATGACC-3'

detection of deletion mutants. A bactericidal antibiotic is one capable of killing the bacteria; whereas a bacteriostatic antibiotic inhibits bacterial growth (Tortora *et al.*, 1989).

3.1.2 The plac plasmids

The parental plac plasmids, which contained both the ampicillin and kanamycin resistance genes, were constructed by the Bichara laboratory. This fusion *lacZ*-Kan plasmid system was constructed by cutting the *kan^r* gene from Tn903 at the *Xho*I and *Sal*I sites, and cloning it into the *Sal*I site in the polylinker region of the *lacZ* gene of pUC18 (which already carries an *Amp^r* gene), yielding pUC18AK. The kanamycin protein expressed from this construct lacks the first ten amino acids of the wild type protein at the NH₂-terminus and contains the first 16 amino acids from the *lacZ* gene, but still confers kanamycin resistance. A schematic diagram of the construction is found in Figure 3.2.

Several alterations were made by site-directed mutagenesis. A potential reinitiation codon (ATG) was found at position 25 of the kanamycin protein and was changed from methionine to valine. A *Sma*I site (CCCGGG), a known hotspot for induced frameshift mutations (Lambert *et al.*, 1992), was found within the first 48 bases (i.e. in the *lacZ* part of the fusion gene) and was deleted. The fusion *lacZ*-Kan gene was removed from pUC18AK using *Eco*RI and *Stu*I, then cloned into the *Hinc*II/*Eco*RI site of two pUC8 plasmids (one with the origin of replication in the normal orientation, and one with a reverse ori). The resultant parent plasmids are pLC and pLCR (reversed origin of replication). The plasmid maps of these two constructs are shown in Figure 3.3.

Figure 3.2 Schematic representation of the construction of the lacZ-kan gene and of the parent pLC and pLCR plasmids. The solid region represents the *lacZ* gene. The hatched region represents the kan^r gene. The relevant restriction sites are indicated.
a.a. stands for amino acid.

Figure 3.3 Plasmid maps of the parent plasmids pLC and pLCR. Both these plasmids are under the control of the *plac* promoter. The plasmid on the left contains an origin of replication in orientation 1, while the plasmid on the right contains an origin of replication in orientation 2. The three different types of insert sequences, 20p, 20np and 20T are shown at the bottom of the figure. 20p represents the 20 bp palindromic insert. 20np is the 20 bp non-palindromic insert. 20T is the 20 bp triplet repeat insert. *LacZkan* is the fusion *lacZ*-kan gene. Amp is the ampicillin resistance gene. *Ori* is the origin of replication. *plac* is the *plac* promoter. Relevant restriction sites are indicated.

[illegible]

3.1.2.1 Optimization of reversion assay conditions for the plac system

The parent plasmids pLC and pLCR were transformed into TG1uvrA-F' competent cells as described in section 2.5. The cells were incubated at 37°C to allow for expression of the antibiotic resistance markers, then plated (at 15 minute intervals) onto LB-Amp50 and LB-Amp50-Kan10. At early time points, it was found that the number of ampicillin resistant colonies increased much faster than the number of kanamycin resistant colonies, indicating that the expression time necessary to obtain ampicillin resistance was less than that required to obtain kanamycin resistance. It was also found that approximately 120 minutes were required for the number of kanamycin resistant colonies to equal the number of ampicillin resistant colonies. Therefore, studies involving the plac plasmids containing inserts were conducted using an expression time of 120 minutes.

The optimal concentration of kan present in LBA plates was determined by growing pLC20np transformants on kan concentrations of 20, 10, 5 and 2.5 µg/mL. Many small colonies were found on plates containing less than 10 µg/mL kan and sequencing data from the plasmids extracted from these small colonies revealed that they all contained the full insert. These small colonies were not found on plates containing 10 or 20 µg/mL kan; therefore, a kan concentration of 10 µg/mL was deemed optimal.

3.1.3 Construction of the ptac plasmids

Due to the long expression time required for the kanamycin resistance gene to be expressed in pLC and pLCR (approximately two hours), the plac promoter was replaced

Figure 3.4 Plasmid maps of the parent plasmids pTC and pTCR. Both these plasmids are under the control of the ptac promoter that was cloned in between the *SapI/EcoRI* restriction sites. The plasmid on the left contains an origin of replication in orientation 1, while the plasmid on the right contains an origin of replication in orientation 2.

The three different types of insert sequences, 20p, 20np and 20T are shown at the bottom of the figure. 20p represents the 20 bp palindromic insert. 20np is the 20 bp non-palindromic insert. 20T is the 20 bp triplet repeat insert.

LacZkan is the fusion *lacZ*-kan gene. Amp is the ampicillin resistance gene. *Ori* is the origin of replication. ptac is the ptac promoter. Relevant restriction sites are indicated.

Inserts: pallidromic-AAT TGGT CACCTATCAATT CGAGC TTGAGTATCGTACGATACICAGTCACCTATC
non-pallidromic-AATTGGT CACCTATCAATT CGAGC TTGCMGCCGATACGATACTCAGTCACCTATC
triplet-AAT TGGT CACCTATCAATT CGAGC TGC TGCTGC TGCTGC TGCTGC TGCTGC TGCTGC TGCTGC

by the ptac promoter, which was taken from plasmid pTTQ18. An approximately 300 bp fragment, containing the ptac promoter, was removed from plasmid pTTQ18 using *SapI* and *EcoRI*. Plasmids pLC and pLCR were also digested with the same two enzymes to remove the plac promoter. The ptac promoter was then ligated with the *SapI/EcoRI*-digested plasmids, pLC and pLCR, resulting in the plasmids pTC and pTCR. Plasmid maps of the two pTC and pTCR parent plasmids are found in Figure 3.4

The DNA sequence of the 342 bp fragment (upstream from the ptac promoter) cloned into the ptac plasmids was confirmed by sequencing with both a forward (pPAK1) and a reverse (ptacprime1) primer. The sequences of the two primers are listed in Table 2.3. The sequence in Figure 3.5 contains only the 342 base pairs between the *SapI* site and the *EcoRI* site.

AAG	<u>CGG</u>	<u>AAG</u>	<u>AGC</u>	GCC	CAA	TAC	GCA	AAC	CGC	CTC	TCC
	<i>SapI</i>										
CGC	GCG	TTG	GCC	GAT	TCA	TTA	ATG	CAG	AAT	TAA	TTC
TCA	TGT	TTG	ACA	GCT	TAT	CAT	CGA	CTG	CAC	GGT	GCA
CCA	ATG	CTT	CTG	GCG	TCA	GGC	AGC	CAT	CGG	AAG	CTG
TGG	TAT	GGC	TGT	GCA	GGT	CGT	AAA	TCA	CTG	CAT	AAT
TCG	TGT	CGC	TCA	AGG	CGC	ACT	CCC	GTT	CTG	GAT	AAT
GTT	TTT	TGC	GCC	GAC	ATC	ATA	ACG	GTT	CTG	GCA	AAT
ATT	CTG	AAA	TGA	GCT	<u>GTT</u>	<u>GAC</u>	<u>AAT</u>	TAA	TCA	TCG	GCT
	⁻¹⁰										
CGT	<u>ATA</u>	<u>ATG</u>	TGT	GGA	ATT	GTG	AGC	GGA	TAA	CAA	TTT
	⁻³⁵										
	RBS										
CAC	<u>ACA</u>	<u>GGA</u>	AAC	AGC	GAT	G(<i>EcoRI</i> site)					

Figure 3.5 Sequence of the 342 bp fragment containing the ptac promoter that was cloned into the pLC and pLCR parent plasmids. Recognition site of *SapI*, the putative ribosome binding site (RBS), the -10 and -35 sequences are underlined. The position of the *EcoRI* site is indicated at the end of the sequence, in brackets.

Figure 3.6 Schematic diagram showing the cloning of the 20 bp palindromic insert into a pLC plasmid. The sequences of the 20 bp palindromic insert are indicated with opposing arrows on top of the sequence. The direct repeats are shown above the sequence, with arrows in the same direction. The bases belonging to the vector are underlined.

5'-...CATGATTACG AATTCGAGCTCGGA...-3'
 3'-...GTACTAATGCTTAA GCTCGAGCCT...-5'

+

20p Insert

20p: 5'-AATTGGTCACCTATCAATTTCGAGCTTGAGTATCGTACGATACTCAGTCACCTATC-3'
 20pc: 3'-CCAGTGGATAGTTAAGCTCGAACTCATAGCATGCTATGAGTCAGTGGATAGTTAA-5'

↓

5'-...CATGATTACG AATTGGTCACCTATCAATTTCGAGCTTGAGTATCGTACGATACTCAGTCACCTATC AATTCGAGCTCGGA...-3'
 3'-...GTACTAATGCTTAACCAGTGGATAGTTAAGCTCGAACTCATAGCATGCTATGAGTCAGTGGATAGTTAAGCTCGAGCCT...-5'

pLC20p

plasmid

3.2 Spontaneous deletion frequencies

3.2.1 Origin of replication

In all three plasmid systems used, when the origin of replication is in the same direction as the direction of transcription of the genes containing the inserts, these plasmids are designated as plasmids whose origins are in orientation 1. When the fragment containing the origin of replication was excised and cloned into the plasmids in the opposite orientation, the resulting plasmids are referred to as the plasmids in which the origin has orientation 2. Refer to Figures 3.1, 3.3 and 3.4 for plasmid maps of the parental plasmids containing origins of replication in orientation 1 and orientation 2, in each of the three systems.

A change from orientation 1 to orientation 2 in the *plac* and *ptac* plasmid systems causes the distance between the origin of replication and the repeats to increase. In orientation 1, the distance from the origin to the repeats is only 367 bp. However, in orientation 2, the distance between the origin and the repeats increases to 2521 bp.

When the plasmid origin of replication is in orientation 1, transcription and replication occur in the same direction. With the origin in orientation 2, transcription and replication are in opposite directions, causing a change in the relationship between the coding strand and the leading and lagging strands of replication. When the *Ori* is in orientation 1, the transcribed strand is replicated discontinuously. However, in a plasmid containing an *Ori* in orientation 2, the transcribed strand would be replicated continuously.

3.2.2 Spontaneous deletion frequencies in pNS plasmids

Each type of pNS plasmid was transformed into TK603 cells to determine the spontaneous frequency with which the insert sequences were deleted (Table 3.1). A total of 8146 deletion mutants were scored in 29 independent experiments in SOS⁻ and SOS⁺ TK603 cells. It was found that the frequency of deletion of the 17 bp palindromic insert in plasmid pNS17p was 185-fold greater than that of the 17 bp non-palindromic insert in plasmid pNS17np when the origin of replication was in orientation 1. The deletion frequency of the palindromic insert was 22-fold higher than that of the non-palindromic insert when cloned into plasmids containing an origin of replication which was in orientation 2. A 12-fold decrease in the deletion frequency of the palindromic insert was observed in plasmids containing replication origins in orientation 2 (pNSR17p) versus origins in orientation 1 (pNS17p). In the case of the non-palindromic insert, the deletion frequency in plasmids in which the origin has orientation 1 (pNS17np) was only 1.5-fold greater than that of plasmids with the origin in orientation 2 (pNSR17np). Statistical analysis of these deletion frequencies is shown in Table 3.2.

Table 3.1 Mean spontaneous deletion frequencies of 17 bp palindromic and non-palindromic inserts in pNS plasmids in *E. coli* TK603 cells and statistical comparisons of the deletion frequencies between SOS⁻ and SOS⁺ cells

Plasmid	n	Mean deletion frequency (in SOS ⁻ cells)	n	Mean deletion frequency (in SOS ⁺ cells)	F-test	t-test	M-W test
pNS17p	6	$9.26 \times 10^{-5} \pm 3.75 \times 10^{-5}$	3	$4.53 \times 10^{-5} \pm 3.13 \times 10^{-6}$	47.5*	N/A	18*
pNSR17p	4	$7.67 \times 10^{-6} \pm 5.40 \times 10^{-7}$	3	$6.50 \times 10^{-6} \pm 1.73 \times 10^{-6}$	13.1	1.17	8
pNS17np	4	$5.01 \times 10^{-7} \pm 1.20 \times 10^{-7}$	3	$6.23 \times 10^{-7} \pm 1.34 \times 10^{-7}$	1.06	0.98	9
pNSR17np	3	$3.51 \times 10^{-7} \pm 1.24 \times 10^{-7}$	3	$3.80 \times 10^{-7} \pm 6.64 \times 10^{-8}$	4.14	0.46	5

Note: n is the number of independent experiments conducted; * indicates the difference is significant at the 5% level; N/A indicates the t-test could not be performed.

The four pNS plasmids were introduced into both SOS⁻ and SOS⁺ TK603 cells to determine whether or not SOS-induction of the cells has an effect on the deletion frequencies of the inserts. Statistical analysis of the data (Table 3.1) revealed that the pNS17p plasmid exhibited a significantly higher deletion frequency in the SOS⁻ cells than in the SOS⁺ cells. However, a comparison between the deletion frequencies of the other three plasmids in SOS-induced and uninduced cells showed no significant difference, suggesting that SOS induction does not have an effect on deletion frequency of inserts in this system.

Table 3.2 Statistical analysis of the deletion frequencies of the pNS plasmids in TK603 cells

Plasmids being compared	SOS- TK603 cells			SOS+ TK603 cells		
	F-test	t-test	M-W test	F-test	t-test	M-W test
pNS17p vs. pNSR17p	337.6*	N/A	24*	1.84	14.8*	9 ^a
pNS17p vs. pNS17np	387.2*	N/A	24*	7.69	33.32*	9 ^a
pNS17np vs. pNSR17np	1.71	1.37	9	2.57	2.29**	9 ^a
pNSR17p vs. pNSR17np	1.49	24.01*	12 ^a	36.5	8.29*	9 ^a

Note: * indicates that the difference is significant at the 5% level.

** indicates that the difference is significant at the 10% level.

N/A indicates that the t-test could not be performed.

^a indicates the calculated U value is equal to the critical value at the 10% level.

Statistical analysis of the deletion frequency data of the pNS plasmids (Table 3.2) revealed that there is a significant difference (at the 5% level) between the deletion frequencies of plasmids pNS17p and pNSR17p, suggesting that an origin of replication in orientation 2 decreases the deletion frequency. A comparison of the deletion frequencies

of plasmids pNS17p and pNS17np, as well as pNSR17p and pNSR17np, showed a significant difference (at the 5% level), suggesting that the palindromic insert is deleted more frequently than the non-palindromic insert, regardless of the orientation of the *Ori*. In the case of the comparison between pNS17np and pNSR17np, there was no significant difference between the deletion frequencies of these plasmids. Plasmid pNS17np was deleted 1.5-fold more frequently than plasmid pNSR17np.

The mean deletion frequencies of the pNS plasmids in SOS⁺ cells were analyzed using the t-test and the Mann-Whitney test (Table 3.2). When the frequencies of pNS17p and pNSR17p were compared, it was found that the insert in pNS17p was deleted with a significantly higher frequency (at the 5% level) than the insert in pNSR17p, again suggesting that an origin of replication in orientation 2 decreases the deletion frequency under the experimental conditions. However, a comparison of plasmids pNS17np and pNSR17np showed only a 2-fold decrease in deletion frequency in plasmids in which the origin has orientation 2 versus 1, which was not found to be significantly different at the 5% level by either the t-test or the Mann-Whitney test. Comparisons of the deletion frequencies of palindromic and non-palindromic inserts in plasmids containing the origin of replication in either orientation showed that, regardless of the direction of the *Ori*, the palindromic insert was deleted at a significantly higher frequency than the non-palindromic insert. This was also found to be the case in the SOS⁻ cells.

3.2.3 Spontaneous deletion frequencies in plac plasmids

The six plac plasmids were transformed into SOS⁻ and SOS⁺ *E. coli* TG1uvrA-F' competent cells as described in section 2.5. A total of 3792 deletion mutants were scored in 38 experiments and the spontaneous deletion frequencies were calculated. The mean deletion frequencies for all six plasmids were found to be very similar, as shown in Table 3.3.

Table 3.3 Mean spontaneous deletion frequencies of 20 bp palindromic, non-palindromic and triplet repeat inserts in plac plasmids transformed into *E. coli* TG1uvrA-F' cells and the statistical analyses of the deletion frequencies between SOS⁻ and SOS⁺ cells

Plasmid	n	Mean deletion frequency (in SOS ⁻ cells)	n	Mean deletion frequency (in SOS ⁺ cells)	F-test	t-test	M-W test
pLC20p	3	$1.40 \times 10^{-5} \pm 5.04 \times 10^{-6}$	3	$4.36 \times 10^{-5} \pm 2.94 \times 10^{-5}$	7.23	2.32*	9 ^a
pLC20np	4	$2.03 \times 10^{-5} \pm 3.71 \times 10^{-6}$	3	$1.00 \times 10^{-5} \pm 1.38 \times 10^{-6}$	2.90	4.42*	12 ^a
pLC20T	3	$1.52 \times 10^{-5} \pm 2.72 \times 10^{-6}$	3	$2.34 \times 10^{-5} \pm 1.60 \times 10^{-5}$	20.75	0.61	5
pLCR20p	4	$1.42 \times 10^{-5} \pm 7.01 \times 10^{-6}$	3	$4.90 \times 10^{-5} \pm 6.98 \times 10^{-6}$	2.75	4.97*	12 ^a
pLCR20np	3	$1.23 \times 10^{-5} \pm 2.91 \times 10^{-6}$	3	$1.10 \times 10^{-5} \pm 5.81 \times 10^{-6}$	4.27	0.41	6
pLCR20T	3	$1.47 \times 10^{-5} \pm 4.03 \times 10^{-7}$	3	$3.41 \times 10^{-5} \pm 1.08 \times 10^{-5}$	331.4*	N/A	9 ^a

Note: n is the number of independent experiments conducted.

* indicates that the difference is significant at the 5% level.

N/A indicates that the t-test could not be performed.

^a indicates that the calculated U value is equal to the critical value at the 10% level.

Statistical comparisons of each of the six plac plasmids in SOS⁻ and SOS⁺ cells (Table 3.3) showed that the insert in plasmid pLC20np was deleted more frequently in SOS⁻ cells than in SOS⁺ cells. However, the opposite was found for plasmids pLC20p

and pLCR20p, in which the inserts were deleted 3.5-fold more frequently in SOS⁺ than in SOS⁻ cells. Plasmids pLC20T and pLCR20np showed no significant difference in deletion frequency between SOS⁻ and SOS⁺ cells when the t-test was employed. Plasmid pLCR20T was found to be statistically significant at the 10% level by the Mann-Whitney test. These results are inconsistent and the role that SOS induction plays in the formation of deletion mutations could not be determined.

The observation that the deletion frequencies of all six plac plasmids were very similar suggested that neither the insert sequence nor the orientation of the origin of replication had an effect on the deletion frequency of inserts in plac plasmids. This observation was further supported by the results of the statistical analysis (Table 3.4). The t-test showed that pLC20np exhibited a significantly higher deletion frequency than pLCR20np. However, this apparent increase was not supported by the Mann-Whitney test. The comparison between pLC20T and pLCR20T was not performed using the t-test due to the unequal variances, and the Mann-Whitney test showed no differences in deletion frequency. The differences in the remaining four comparisons were not statistically significant by both the t-test and the Mann-Whitney test.

The mean spontaneous frequencies which resulted from the transformation of the six plac plasmids into SOS induced *E. coli* TG1uvrA-F' cells are shown in Table 3.3. The comparison between pLC20p and pLC20np could not be made using the t-test due to the unequal variances (Table 3.4). Results of this comparison using the Mann-Whitney test showed that the difference in deletion frequency may be significant at the 10% level. However, it was observed that all the deletion frequencies for pLC20p were greater than

those for pLC20np and that the mean deletion frequency of pLC20p was more than 4-fold greater than the mean deletion frequency of pLC20np. The observations and the result of the Mann-Whitney test suggest that the palindromic insert is deleted with a significantly higher frequency than the non-palindromic insert in the pLC plasmids. The remaining comparisons were not found to be statistically significant by either the t-test or the Mann-Whitney test.

Table 3.4 Statistical results from comparisons of the deletion frequencies of the palindromic, non-palindromic and triplet repeat inserts in plac plasmids in SOS⁻ and SOS⁺ TG1uvrA-F' cells

Plasmids being compared	SOS ⁻ TG1uvrA-F' cells			SOS ⁺ TG1uvrA-F' cells		
	F-test	t-test	M-W test	F-test	t-test	M-W test
pLC20p/pLC20np	3.0	1.72	9	62.6*	N/A	9 ^a
pLC20p/pLC20T	3.53	0.39	6	1.23	1.30	7
pLC20np/pLC20T	1.18	1.77	8	50.9*	N/A	8
pLC20p/pLCR20p	1.64	0.03	7	12.1	0.21	6
pLC20np/pLCR20np	1.17	2.80*	10	14.5	0.08	6
pLC20T/pLCR20T	41.5*	N/A	5	2.60	0.91	7

Note: * indicates the difference is significant at the 5% level

** indicates the difference is significant at the 10% level

N/A indicates that the t-test could not be performed

^a indicates the calculated U value is equal to the critical value at the 10% level.

3.2.4 Spontaneous deletion frequencies in ptac plasmids

The six ptac plasmids were introduced into TG1uvrA-F' cells to determine whether the three different types of inserts and the different orientations of the origin of

replication would affect the deletion frequency of each insert. The plasmids were transformed as described in section 2.5, 4398 deletion mutants were scored in 45 experiments, the deletion frequencies were determined, the mean deletion frequencies were calculated and are listed in Table 3.5.

Table 3.5 Mean spontaneous deletion frequencies of 20 bp palindromic, non-palindromic and triplet repeat inserts in ptac plasmids transformed into *E. coli* TG1uvrA-F' cells

Plasmid	n	Mean deletion frequency (in SOS ⁻ cells)	n	Mean deletion frequency (in SOS ⁺ cells)	F-test	t-test	M-W test
pTC20p	6	$1.94 \times 10^{-4} \pm 5.01 \times 10^{-5}$	3	$5.01 \times 10^{-5} \pm 2.88 \times 10^{-5}$	1.60	4.64*	18
pTC20np	5	$1.69 \times 10^{-4} \pm 4.99 \times 10^{-5}$	3	$4.95 \times 10^{-5} \pm 4.66 \times 10^{-6}$	25.9	4.70*	15
pTC20T	4	$6.73 \times 10^{-5} \pm 1.78 \times 10^{-5}$	3	$5.03 \times 10^{-5} \pm 1.70 \times 10^{-5}$	1.52	1.09	8
pTCR20p	4	$3.56 \times 10^{-5} \pm 1.34 \times 10^{-5}$	3	$5.43 \times 10^{-5} \pm 2.92 \times 10^{-5}$	3.35	0.94	12 ^a
pTCR20np	4	$5.83 \times 10^{-5} \pm 2.89 \times 10^{-5}$	3	$2.71 \times 10^{-5} \pm 2.84 \times 10^{-6}$	66.4*	N/A	9
pTCR20T	4	$2.33 \times 10^{-5} \pm 2.13 \times 10^{-6}$	3	$2.61 \times 10^{-5} \pm 2.86 \times 10^{-6}$	1.72	1.25	10

Note: n is the number of independent experiments conducted.

* indicates that the difference is significant at the 5% level.

^a indicates that the calculated U value is equal to the critical value at the 10% level.

When the deletion frequencies of the six ptac plasmids introduced into SOS⁻ and SOS⁺ cells were compared (Table 3.5), it was found that the frequency of deletion in the palindromic and non-palindromic inserts in plasmids pTC20p and pTC20np increased significantly in the SOS⁻ cells as compared to SOS⁺ cells. The remaining four plasmids showed no significant difference in deletion frequency between SOS⁻ and SOS⁺ cells.

Statistical analysis of the deletion frequency data for ptac plasmids (Table 3.6) showed that there is no significant difference between the deletion frequencies of the palindromic insert and the non-palindromic insert. However, both the palindromic and non-palindromic inserts were deleted with a significantly higher frequency than the triplet repeat insert. When the deletion frequencies of the three types of inserts in plasmids with origins in either orientation 1 or 2 were examined, it was found that the palindromic and non-palindromic inserts were deleted with a significantly higher frequency in plasmids in which the origins were in orientation 1. In the case of the triplet repeat inserts, a three-fold decrease in deletion frequency was observed in plasmids in which the origin was in orientation 2 versus orientation 1. It was not possible to use the t-test for this comparison because of the unequal variances of the two sets of data. The Mann-Whitney test showed that at a 5% significance level, the difference is just on the borderline of the critical value. The result of the Mann-Whitney test, together with the observation that all the deletion frequencies for pTC20T are at least two-fold greater than those for pTCR20T, suggest that the frequency of deletion of the 20T insert is greater in plasmids in which the origin is in orientation 1 than in plasmids in which the origin is in orientation 2. These results suggest that plasmids containing *Ori* in orientation 1 are deleted with a significantly greater frequency than plasmids containing *Ori* in orientation 2 for all three types of inserts in the ptac plasmid system.

In contrast to the results from the statistical analyses of the ptac plasmids in SOS⁻ cells, five out of six of the comparisons made for the same plasmids in SOS⁺ cells showed no significant difference in deletion frequencies (Table 3.6). Only the comparison

between plasmids pTC20np and pTCR20np showed that the insert was deleted with a significantly higher frequency in the pTC20np plasmid, suggesting that plasmids in which the origin is in orientation 2 has a decreased deletion frequency in comparison to plasmids in which the origin is in orientation 1. As shown in Table 3.6, the t-test revealed that the difference is significant at the 5% level, while the Mann-Whitney test indicated that the difference may be significant only at the 10% level.

Table 3.6 Statistical results from comparisons of the deletion frequencies of inserts in ptac plasmids in SOS⁻ and SOS⁺ TG1uvrA-F' cells

Plasmids being compared	SOS ⁻ TG1uvrA-F' cells			SOS ⁺ TG1uvrA-F' cells		
	F-test	t-test	M-W test	F-test	t-test	M-W test
pTC20p/pTC20np	1.09	0.76	23	38.0	0.17	6
pTC20p/pTC20T	2.50	5.04*	24*	2.62	0.12	5
pTC20np/pTC20T	2.73	4.01*	20*	14.5	0.05	5
pTC20p/pTCR20p	2.48	7.05*	24*	1.18	0.19	5
pTC20np/pTCR20np	1.71	3.54*	20*	1.50	6.04*	9 ^a
pTC20T/pTCR20T	24.2*	N/A	16	21.46	2.07	8

Note: * indicates that the difference is significant at the 5% level.

N/A indicates that the t-test could not be performed.

^a indicates the calculated U value is equal to the critical value at the 10% level.

3.3 Molecular nature of the deletion mutants

3.3.1 The pNS mutants

After the plasmids were transformed into TK603 cells, the bacterial cells were plated onto selective media plates, examined and deletion frequencies were calculated.

Mutant colonies growing on LB-Amp50-Cm10 plates were chosen at random and sequenced. Sequence analysis of the deletion mutations in the pNS plasmids revealed that all the mutants sequenced contained precise deletions (Table 3.7). A precise deletion is one in which the 17 bp palindromic (or non-palindromic) insert and one of the flanking direct repeats have been completely deleted, leaving only one direct repeat. When the deletion is precise, the chloramphenicol resistance gene is put back into the correct reading frame and is therefore expressed, allowing selection for deletion mutants. However, the protein expressed from this gene differs from the wild type protein by five extra amino acids. A schematic diagram of the deleted insert sequence is shown in Figure 3.7.

Table 3.7 Types of deletion mutations found in pNS plasmids introduced into both SOS⁻ and SOS⁺ TK603 cells (sequencing data)

Plasmid	SOS induction (SOS ⁻ /SOS ⁺)	Number of mutants sequenced	Type of deletion mutation
pNS17p	SOS ⁻	7	all precise deletions
pNS17p	SOS ⁺	4	all precise deletions
pNSR17p	SOS ⁻	9	all precise deletions
pNSR17p	SOS ⁺	4	all precise deletions
pNS17np	SOS ⁻	6	all precise deletions
pNS17np	SOS ⁺	4	all precise deletions
pNSR17np	SOS ⁻	6	all precise deletions
pNSR17np	SOS ⁺	3	all precise deletions

Figure 3.7 Schematic diagram showing the difference in sequence between the pNS plasmid prior to cloning of the insert, plasmid pNS17p and the deletion mutant. The *Eco*RI site is underlined in the sequence of the pNS plasmid before cloning. For the pNS17p plasmid and the deletion mutant, the vector sequences are italicized, the direct repeats are double-underlined and the bases which form the palindrome are in bold (larger font as well).

5'-...CATCCGGAATTCCGTATGGCAATG...-3'

3'-...GTAGGCCTTAAAGGCATACCGTTAC...-5'

(pNS plasmid before cloning of insert)

5'-...CATCCGGAATTGGTCATCTATCAATTCCGTATGGCAATG...-3'

3'-...GTAGGCCTAAAGGCATAGTTAAGGCATACCGTTAC...-5'

Plasmid pNS17p

⇓

5'-...CATCCGGAATTGGTCATCTATCAATTCCGTATGGCAATG...-3'

3'-...GTAGGCCTAAAGGCATAGTTAAGGCATACCGTTAC...-5'

Deletion mutant

3.3.2 The plac mutants

Kanamycin resistant mutants from all the plac deletion experiments were chosen at random for sequencing. The plasmid DNA was extracted by the Wizard Miniprep Kit as described in section 2.6.2 and sequenced as described in section 2.8. At least three mutants from each type of plasmid, in both SOS⁻ and SOS⁺ cells, were randomly chosen and sequenced with the pPAK1 primer. Several different types of deletion mutants were found. Of the 42 plac deletion mutants sequenced, 11 were precise deletions, 26 were mixed populations, 4 contained full inserts, and 1 contained two copies of the insert minus a T.

A precise deletion is one in which the 20p/20np/20T insert and one of the flanking direct repeats were deleted, putting the fusion lacZ-kan gene back into the correct reading frame so that the kanamycin resistance gene can be expressed. Figure 3.8 contains a schematic diagram showing the sequence differences between a plasmid containing the full 20p insert and the corresponding deletion mutant. Figure 3.9 shows the difference between the full insert sequences and the corresponding deleted sequences of a pLC20np plasmid.

The mixed mutants were the ones which showed both the sequence of a plasmid containing the full insert and the precise deletion sequence. Such a situation may arise through the presence of a dimer formed between a plasmid containing the full insert and a plasmid containing the precise deletion sequence. Since pUC plasmids have high copy numbers, mixed mutants may also result from colonies that carry two types of plasmids: those which contain the insert sequence and those containing precise deletions. Four of

Figure 3.8 Schematic diagram showing the difference in sequence between the pLC plasmid (prior to cloning of the insert), plasmid pLC20p and the deletion mutant. The *Eco*RI site is underlined in the sequence of the pLC plasmid before cloning. For pLC20p and the deletion mutant, the vector sequences are italicized, the direct repeats are double-underlined and the bases which form the palindrome are in bold (larger font as well).

5'-...CATGATTACGAATTCGAGCTCGGA...-3'

3'-...GTACTAATGCTTAAGCTCGAGCCT...-5'

(pLC plasmid before cloning of insert)

5'-...CATGATTACGAATTGGTCACCTATCAATTCGAGCTCGGA...-3'

3'-...GTACTAATGCTTAACCAGTGGATAGTTAAGCTCGACTCAGTGGATAGTTAAGCTCGAGCCT...-5'

Plasmid pLC20p

5'-...CATGATTACGAATTGGTCACCTATCAATTCGAGCTCGGA...-3'

3'-...GTACTAATGCTTAACCAGTGGATAGTTAAGCTCGAGCCCT...-5'

Deletion mutant

Figure 3.9 Photograph of the sequences of plasmid pLC20np (the plasmid containing the 20 bp non-palindromic sequence) and the deletion mutant. A is the sequence containing the full insert. B is the sequence of the deletion mutant. The double-headed arrow shows the insert sequence that is deleted in a precise deletion. The small single arrow pointing at B shows the position from which the insert was deleted. The loading order is TACG (5' to 3' from top to bottom).

A

B

the mutants sequenced had a wild type sequence. There was also one mutant that appeared to be a duplication rather than a deletion mutant. This pLC20p mutant had an extra copy of the insert, with one base deleted, a T at position (67). The deletion of this base did not put the kan resistance gene back into its reading frame, yet the mutants were still able to grow on medium containing 10 μ g/mL kanamycin .

Table 3.8 Types of deletion mutations found in plac plasmids transformed into SOS⁻ and SOS⁺ TG1uvrA-F' cells (from sequencing data)

Plasmid	Number of each type of deletion mutant found					
	SOS ⁻ TG1uvrA-F' cells			SOS ⁺ TG1uvrA-F' cells		
	Precise	Mix	Other	Precise	Mix	Other
pLC20p	0	2	1	0	4	0
pLC20np	1	1	1	1	1	1
pLC20T	0	3	0	2	1	0
pLCR20p	2	3	0	1	4	0
pLCR20np	1	2	0	2	0	1
pLCR20T	1	2	0	0	3	1

Note: Mix represents a mutant which shows both the insert sequences and the precise deletion sequences.

3.3.3 The ptac mutants

The mutant colonies that grew as a result of the transformation of the six ptac plasmids into SOS⁻ and SOS⁺ TG1uvrA-F' cells were randomly chosen for sequence analysis. The plasmid DNA was extracted by the Wizard Miniprep method as described in Section 2.6.2 and sequenced as described in section 2.8. In a total of 42 ptac mutants sequenced, 26 contained precise deletions, 14 contained the insert sequence as well as the

precise deletion sequences, and only two mutants were found to retain the full insert sequences (Table 3.9). It was observed that the plac mutants contained predominantly mixed populations while the ptac mutants contained mostly precise deletion mutants. It was also found that in the ptac system, the majority of the SOS⁺ mutants showed precise deletions whereas the SOS⁻ mutants showed an approximately equal number of precise deletions and mixed populations (Table 3.9). This suggests that SOS induction may play a role in the generation of deletion mutations in the ptac system. Figures 3.10 and 3.11 contain photographs of the insert sequence and the revertant sequence of plasmids pTC20p and pTC20T, respectively.

Table 3.9 Types of deletion mutations found in ptac plasmids transformed into SOS⁻ and SOS⁺ TG1uvrA-F' cells (from sequencing data)

Plasmid	Number of each type of deletion mutant found					
	SOS ⁻ TG1uvrA-F' cells			SOS ⁺ TG1uvrA-F' cells		
	Precise	Mix	Other	Precise	Mix	Other
pTC20p	1	2	0	3	0	0
pTC20np	1	4	0	3	0	0
pTC20T	0	3	0	2	1	1
pTCR20p	2	1	0	3	0	0
pTCR20np	3	0	0	5	1	0
pTCR20T	2	1	0	1	1	1
Total	9	11	0	17	3	2

Note: A mix represents a mutant which shows both the wild type sequence and the revertant sequence.

Figure 3.10 Photograph of the sequences of plasmid pTC20p (the plasmid containing the 20 bp palindromic sequence) and the deletion mutant. A is the sequence containing the full insert. B is the sequence of the deletion mutant. The double-headed arrow shows the insert sequence that is deleted in a precise deletion. The small single arrow pointing at B shows the position from which the insert was deleted. The loading order is TACG (5' to 3' from top to bottom).

A

B

Figure 3.11 Photograph of the sequences of plasmid pTC20T (the plasmid containing the 20 bp triplet repeat sequence) and the deletion mutant. A is the sequence containing the full insert. B is the sequence of the deletion mutant. The double-headed arrow shows the insert sequence that is deleted in a precise deletion. The small single arrow pointing at B shows the position from which the insert was deleted. The loading order is TACG (5' to 3' from top to bottom).

A

B

3.4 Induced deletion frequency in AAF-modified plasmids

3.4.1 Deletion frequencies in AAF-modified pNS plasmids

The two pNS plasmids, pNSR17p and pNSR17np, were modified with AAAF as described in section 2.10 to yield plasmids with different levels of modification. The level of modification is given as the average number of AAF adducts per plasmid. A comparison of the mean deletion frequencies of plasmid pNSR17p containing different levels of AAF modification, in SOS⁻ cells, indicated that AAF modification did not affect deletion frequency. When these same four plasmids were transformed into SOS⁺ cells, there appeared to be a small increase in deletion frequency with an increase in the level of AAF-modification. A total of 710 deletion mutants were scored in the 14 experiments (0-AAF and 5.17-AAF) and the mean deletion frequencies of plasmid pNSR17p at different levels of AAF-modification in both SOS⁻ and SOS⁺ cells were recorded in Table 3.10.

Plasmid pNSR17np had also been modified to different levels with AAAF, yielding plasmids with an average of 2.60, 3.44, 4.37 and 5.43 AAF adducts per plasmid. However, transformation of these plasmids into both SOS⁻ and SOS⁺ cells resulted in transformation efficiencies so low that deletions were not detectable.

Table 3.10 Mean deletion frequencies of plasmid pNSR17p containing different levels of AAF modification, transformed into SOS⁻ and SOS⁺ *E.-coli* TK603 cells

Level of modification	n	Mean deletion frequency (SOS ⁻ cells)	n	Mean deletion frequency (SOS ⁺ cells)
0-AAF	3	$2.06 \times 10^{-6} \pm 5.28 \times 10^{-7}$	4	$4.01 \times 10^{-6} \pm 6.45 \times 10^{-7}$
3.48-AAF	3	$3.70 \times 10^{-6} \pm 5.28 \times 10^{-7}$	2	$4.30 \times 10^{-6} \pm 4.44 \times 10^{-7}$
4.69-AAF	3	$2.42 \times 10^{-6} \pm 2.96 \times 10^{-7}$	2	$5.00 \times 10^{-6} \pm 6.25 \times 10^{-7}$
5.17-AAF	3	$2.58 \times 10^{-6} \pm 3.88 \times 10^{-7}$	4	$5.73 \times 10^{-6} \pm 2.03 \times 10^{-7}$

Note: n is the number of independent experiments conducted.

The highest level of AAF-modification, 5.17-AAF adducts per plasmid, did not appear to adversely affect the transformation efficiency of this plasmid. Therefore, statistical analysis was used to compare the deletion frequencies of only the unmodified and the 5.17-AAF modified plasmids. In SOS⁻ TK603 cells, both the t-test and the Mann-Whitney test showed that the deletion frequency in the AAF-modified plasmid pNSR17p was not significantly higher than that of the unmodified plasmid, suggesting either that random AAF modification does not increase deletion frequency in the pNS plasmid system or that modified plasmids have been selected against. This was also found to be true when these plasmids were transformed into SOS⁺ cells, as shown in Table 3.10. These results suggest that SOS induction does not affect the deletion frequency of the randomly AAF-modified pNSR17p plasmid.

3.4.2 Survival of N-acetoxy-N-2-acetylaminofluorene-modified plasmids

The five ptac plasmids, pTC20p, pTC20np, pTCR20p, pTCR20np and pTCR20T, were modified to different levels, ranging from 0 to 32 AAF adducts per plasmid

molecule. When the modified plasmids were transformed into *E. coli* TG1uvrA-F' cells, a decrease in transformation efficiency was found with increasing AAF modification levels. Modification levels of 11-AAF and 21-AAF adducts per plasmid of pTC20p resulted in transformation efficiencies that were 8.6% and 0.7% of the unmodified plasmid, respectively. For 10-AAF, 19-AAF and 27-AAF adducts per plasmid of pTC20np, transformation efficiencies were 14%, 1.7% and 0.07% of the unmodified plasmid, respectively. Similar results were found with the remaining three plasmids, as shown in the graph in Figure 3.12.

Figure 3.12 The effect of AAF-modification on the survival of the five ptac plasmids in-SOS+ TG1uvrA-F' cells. The transformation efficiency was used to determine the % survival, with the transformation efficiency of 0-AAF being used as 100% survival. The level of modification is the average number of AAF adducts per plasmid.

3.4.3 Deletion frequencies in AAF-modified plac plasmids

Two of the six plac plasmids, pLC20p and pLCR20p, were AAF-modified to approximately 7 and 11 adducts per plasmid, respectively, for use in this study. The unmodified and modified versions of plasmids pLC20p and pLCR20p were introduced into both SOS⁻ and SOS⁺ cells, and a total of 2043 deletion mutants were scored in the 24 experiments. The mean deletion frequencies were calculated and are shown in Table 3.11. Statistical analysis of the data (Table 3.11) showed that in SOS⁻ cells, the deletion frequency of the 7-AAF modified pLC20p was not significantly greater than that of the unmodified plasmid. However, the t-test did find that the deletion frequency of the 11-AAF modified pLCR20p was significantly higher than that of the unmodified pLCR20p.

In the SOS⁺ cells, both the t-test and the Mann-Whitney test showed that there was no significant difference in deletion frequency between the unmodified and the modified pLC20p or the pLCR20p plasmids (Table 3.11). It was observed, however, that the deletion frequency of the 11-AAF modified pLCR20p was only half that of the unmodified version. These results suggest that random modification with AAF does not affect deletion frequency in the plac plasmids containing palindromic inserts.

Table 3.11 Mean deletion frequencies of unmodified and AAF-modified *lac* plasmids transformed into SOS⁻ and SOS⁺ *E. coli* TG1uvrA-F' cells and statistical analyses of the deletion frequencies between the unmodified and AAF-modified plasmids

Plasmid	n	Mean deletion frequency (unmodified plasmid)	n	Mean deletion frequency (AAF-modified plasmid)	F-test	t-test	MW test
pLC20p	3	$1.56 \times 10^{-5} \pm 1.88 \times 10^{-6}$	3	$2.66 \times 10^{-5} \pm 1.28 \times 10^{-5}$	28.1	1.19	6
(in SOS ⁺)	3	$1.08 \times 10^{-5} \pm 3.76 \times 10^{-6}$	3	$2.23 \times 10^{-5} \pm 1.19 \times 10^{-5}$	4.25	1.34	6
pLCR20p	3	$1.03 \times 10^{-6} \pm 1.08 \times 10^{-7}$	3	$1.68 \times 10^{-6} \pm 3.03 \times 10^{-7}$	4.29	3.22*	9 ^a
(in SOS ⁺)	3	$1.16 \times 10^{-5} \pm 1.34 \times 10^{-5}$	3	$5.39 \times 10^{-6} \pm 4.31 \times 10^{-6}$	4.54	0.44	5

Note: n is the number of independent experiments conducted.

* indicates that the difference is significant at the 5% level.

^a indicates that the calculated U value is equal to the critical value at the 10% level.

3.4.4 Deletion frequencies in AAF-modified *ptac* plasmids

Five of the six *ptac* plasmids were randomly modified to different degrees with N-Aco-AAF as described in section 2.10. The highly modified plasmids did not yield any transformants when introduced into either SOS⁻ or SOS⁺ cells. Therefore, only plasmids pTC20p, pTC20np, pTCR20p, pTCR20np and pTCR20T, modified with 11, 10, 8, 8, and 9 AAF adducts/plasmid, respectively, were used to study the effect of AAF-modification on deletion frequency (Table 3.12). A total of 12080 deletion mutants were scored in 77 deletion experiments involving the unmodified and modified *ptac* plasmids.

Table 3.12 Mean deletion frequencies of unmodified and AAF-modified ptac plasmids transformed into SOS⁻ and SOS⁺ *E.-coli* TG1uvrA-F' cells and statistical analyses of the deletion frequencies between the unmodified and AAF-modified plasmids

Plasmid	n	Mean deletion frequency (unmodified plasmid)	n	Mean deletion frequency (AAF-modified plasmid)	F-test	t-test	MW test
pTC20p	4	$2.00 \times 10^{-5} \pm 1.11 \times 10^{-5}$	3	$6.20 \times 10^{-5} \pm 4.34 \times 10^{-6}$	15.51	4.51*	12 ^a
(in SOS ⁺)	3	$4.61 \times 10^{-5} \pm 1.76 \times 10^{-5}$	3	$5.79 \times 10^{-5} \pm 1.74 \times 10^{-5}$	1.52	0.69	6
pTC20np	3	$3.22 \times 10^{-5} \pm 1.09 \times 10^{-5}$	4	$2.84 \times 10^{-5} \pm 1.62 \times 10^{-5}$	2.17	0.41	8
(in SOS ⁺)	3	$2.33 \times 10^{-5} \pm 9.60 \times 10^{-6}$	3	$4.16 \times 10^{-5} \pm 5.17 \times 10^{-6}$	5.45	2.3**	8
pTCR20p	5	$3.38 \times 10^{-5} \pm 8.26 \times 10^{-6}$	6	$6.45 \times 10^{-5} \pm 3.72 \times 10^{-5}$	10.6	N/A	25 ^a
(in SOS ⁺)	6	$3.28 \times 10^{-5} \pm 1.16 \times 10^{-5}$	5	$4.17 \times 10^{-5} \pm 1.67 \times 10^{-5}$	1.17	0.91	17
pTCR20np	3	$1.54 \times 10^{-5} \pm 3.89 \times 10^{-6}$	3	$2.60 \times 10^{-5} \pm 1.99 \times 10^{-6}$	6.35	3.16*	9 ^a
(in SOS ⁺)	3	$2.78 \times 10^{-5} \pm 1.44 \times 10^{-5}$	3	$3.56 \times 10^{-5} \pm 2.11 \times 10^{-6}$	60.2*	N/A	6
pTCR20T	3	$2.10 \times 10^{-5} \pm 1.02 \times 10^{-5}$	3	$3.68 \times 10^{-5} \pm 1.52 \times 10^{-5}$	1.08	1.29	7
(in SOS ⁺)	6	$1.50 \times 10^{-5} \pm 6.77 \times 10^{-6}$	5	$3.08 \times 10^{-5} \pm 8.20 \times 10^{-6}$	1.09	13.3*	28*

Note: n is the number of independent experiments conducted.

* indicates that the difference is significant at the 5% level.

** indicates that the difference is significant at the 10% level.

^a indicates the calculated U value is equal to the critical value at the 10% level.

Observation of the mean deletion frequencies of each plasmid in SOS⁻ cells showed some increase from unmodified to modified in all cases except for plasmid pTC20np. In this case, the deletion frequency of the modified plasmid was very slightly lower than that of the unmodified plasmid. Statistical analysis of this decrease revealed that this difference was insignificant. For plasmids pTC20p and pTCR20np, the t-test found that the deletion frequencies of the AAF-modified plasmids were significantly higher than those of the unmodified plasmids.

In plasmid pTCR20p, the variances between the two sets of data were unequal and only the Mann-Whitney test was performed on them. The results indicated that the difference between the two sets of data was not significant at the 5% level. Both the t-test and the Mann-Whitney test found the difference in deletion frequency of the pTCR20T plasmids to be insignificant.

Following transformation of SOS⁺ cells with the same five modified ptac plasmids, the mean deletion frequencies were calculated and recorded in Table 3.12. Statistical analysis of the data indicated that the deletion frequency of the AAF-modified pTC20np was significantly higher than that of the unmodified pTC20np, at the 10% level. It was also found that for pTCR20T, the AAF-modified plasmid showed a greater deletion frequency of the insert than the unmodified plasmid, at a 5% significance level. These data would suggest that AAF-modification tends to increase deletion frequency. However, analyses of the remaining three types of plasmids suggest there really is not a significant difference between the deletion frequencies of the unmodified and modified plasmids.

3.5 Molecular nature of the AAF-induced deletion mutants

The mutant colonies which grew as a result of the introduction of AAF-modified pNSR, plac and ptac plasmids were chosen at random and sequenced in order to examine the molecular nature of these induced mutants. It was found that the majority (68%) of the 60 AAF-induced mutants sequenced contained precise deletions (Table 3.13). All of the pNS deletion mutants (derived from modified and from unmodified plasmids), were

found to contain precise deletions, indicating that AAF-induction does not affect the molecular nature of these mutants.

While the majority of the AAF-induced plac mutants contained a mixture of the insert sequence and the precise deletion sequences, most of the AAF-induced ptac mutants were found to contain precise deletions. This was also the case with the uninduced plac and ptac mutants. The results in Table 3.13 show that in the ptac system, the ratio of precise deletions to mixed populations is greater in the SOS⁺ mutants than in the SOS⁻ mutants. This was also observed with the ptac mutants which resulted from transformation with unmodified plasmids. Three of the ptac mutants contained the inserted sequences and were still able to grow on kanamycin plates. Also, one of the pTCR20p mutants was found to contain the full insert minus a G at position (80), which puts the kanamycin resistance gene back into frame, allowing expression of the gene and conferring resistance to the mutant colony.

Table 3.13 Types of deletion mutations found in pNS, plac and ptac plasmids transformed into SOS⁻ and SOS⁺ cells (from sequencing data)

Plasmid	Number of each type of deletion mutant found					
	SOS ⁻ cells			SOS ⁺ cells		
	Precise	Mix	Other	Precise	Mix	Other
pNSR17p	5	0	0	5	0	0
pLC20p	1	2	0	0	3	0
pLCR20p	1	2	0	1	2	0
pTC20p	2	1	0	4	1	0
pTC20np	4	0	0	5	0	0
pTCR20p	1	2	2	2	0	1
pTCR20np	2	0	1	3	0	0
pTCR20T	2	2	0	3	0	0
Total ^A	13	9	3	18	6	1

Note: A mix represents a mutant which shows both the inserted sequences and the sequences that result from a precise deletion.

^A indicates that the total number of each type of mutant does not include the ones for plasmid pNSR17p.

3.6 Dimer formation in deletion mutants

The plasmid DNA transformed into the host cells in all three plasmid systems were in monomer form. However, when the plasmid DNA extracted from the revertants was analyzed on agarose gels, it was found that considerable amounts of dimers had formed. The plasmid DNA of the revertants from the transformation of pLC20np into SOS⁻ TG1uvrA-F' cells was separated on agarose gel and the monomer and dimer concentrations were determined using a UV spectrophotometer (at 280nm). It was found that 12 out of 20 mutant colonies contained a monomer concentration which was twice

the dimer concentration, five colonies showed approximately equal monomer and dimer concentrations, two colonies contained a monomer concentration that was half the dimer concentration, and one colony had only dimers.

Further investigation of the formation of dimers in revertant colonies was carried out using various modified and unmodified plac and ptac plasmids. In these experiments, the revertant plasmid DNA was run on agarose gel to separate the monomers and dimers, which were purified from the agarose gel, quantified and used to transform SOS⁻ TG1uvrA-F' cells. Six out of seven experiments showed that a greater number of transformants were found on kan containing plates when the dimers were introduced into the host cells than when the monomers were introduced. This observation suggests that the dimers were better able to confer kan resistance to the mutant colonies than the monomers. This leads to the hypothesis that recombination may be involved in the generation of deletion mutations via dimer formation. However, these are only preliminary results and further investigation into this hypothesis is required before any conclusions may be drawn regarding recombination being a mechanism of deletion formation in these plasmid systems.

4. DISCUSSION

Three different plasmid-based reversion assays were designed in an attempt to develop a system with high sensitivity and specificity for the detection and characterization of spontaneous and AAF-induced deletion mutations. The first system consisted of the four pNS plasmids, which contained a Cm^r marker interrupted by either a 17 bp palindromic or 17 bp non-palindromic insert (refer to section 3.1.1). The revertants derived from deletion of the inserts differed from the parental plasmids in that the chloramphenicol resistance gene contained five additional amino acids (Gly, His, Leu, Ser, Ile); the mutants were found to be temperature sensitive. The temperature sensitivity of the pNS revertants and the bacteriostatic nature of chloramphenicol activity prompted the development of the plac and ptac plasmid systems.

The plac plasmids contained a fusion *lacZ*-Kan gene. The kanamycin protein expressed from this fusion gene lacks the first 10 amino acids of the kan^r gene, but is still able to confer kanamycin resistance. The kanamycin gene on the plac plasmids was interrupted by one of three different types of inserts (20 bp palindromic, 20 bp non-palindromic or 20 bp triplet repeat) flanked by direct repeats. Due to the relatively long expression time, the plac promoter was replaced with the ptac promoter. The plac and ptac plasmids differ only in the type of promoter present upstream of the fusion *lacZ*-Kan gene. It was found that the time required for the full expression of kanamycin resistance in the plac plasmids was 120 minutes, as compared to only 90 minutes in the ptac

plasmids. These results indicated that *ptac* is a more effective promoter than *plac*. The mutants from the *plac* and *ptac* systems were neither temperature sensitive, nor slow growing. Also, the *plac/ptac* plasmids contained a *Kan^r* gene instead of a *Cm^r* gene. Kanamycin is a bactericidal antibiotic. The revertants which resulted from deletion of the inserts in the *plac/ptac* systems differed from the parental plasmids by five additional amino acids (Asn, Trp, Ser, Pro, Ile), but were still able to confer kanamycin resistance to the bacteria.

4.1 Spontaneous deletions

4.1.1 The pNS plasmid system

The spontaneous deletion frequencies of the four pNS constructs were compared using statistical analyses. In this system, the palindromic insert was found to be deleted with a frequency that was 185-fold greater than deletion of the non-palindromic insert. When the *Ori* was in orientation 2, the palindromic insert was still deleted more frequently than the non-palindromic insert, but only by 22- fold. The observation that the deletion frequency of palindromic inserts was significantly greater than that of non-palindromic inserts, regardless of orientation of the *Ori*, suggests that sequences containing palindromic inserts are able to form more stable deletion intermediates, which is consistent with the finding of Balbinder *et al.* (1989). After replication of the first direct repeat, slippage may have occurred, allowing the newly synthesized direct repeat in the progeny strand to base pair with the second direct repeat in the template strand

(Figure 4.1). The presence of the palindromic sequence allows the formation of a hairpin structure, which stabilizes the misalignment of the direct repeats leading to a higher deletion frequency (Trinh and Sinden, 1993). The non-palindromic insert may form unstable secondary structures that could destabilize the misaligned intermediate, resulting in a lower deletion frequency.

In the 17p palindromic insert sequence of the pNS system, the inverted repeat is not symmetrical with respect to the direct repeats. There is a ten base pair overlap between the palindrome and one of the direct repeats (Figures 4.1a and 4.1b). This gives rise to two possible slipped mispaired intermediates, depending on the strand that is being used as the replication template. Using the top strand as the template, slippage would occur after replication of the first direct repeat, allowing the entire 17 bp direct repeat of the progeny strand to base pair with the second direct repeat of the parental strand (Figure 4.1a). This results in the deletion of the palindromic sequence and the first direct repeat (which contains the ten base overlap with the palindrome). If the other strand was used as template, slippage of the direct repeat would result in a slipped mispaired intermediate in which only eight bases of the direct repeat of the progeny strand would base pair with the parental strand, forming a secondary structure that is less stable (Figure 4.1b).

When the *Ori* is cloned into the plasmid in orientation 2, the relationship between the leading and lagging template strands of replication changes. If deletion mutation does not occur preferentially in the leading or lagging strands, the orientation of the *Ori* should not effect the deletion frequencies of the two types of plasmids (pNS17p and pNSR17p). However, statistical comparisons between the deletion frequencies of the pNS plasmids

Figure 4.1a A model for the generation of a deletion mutant in plasmid pNS17p via a slippage mechanism (using the top strand as the template). The bold, opposing arrows show the palindromic sequence. The arrows in the same direction mark the directly repeated sequences.

Insert Sequence in pNS17p:

Replication of top strand:

Slippage occurs

Slipped mispaired intermediate

Extension of slipped intermediate to form heteroduplex

Plasmid replication or heteroduplex repair

Deletion mutant:

Figure 4.1b A model for the generation of a deletion mutant in plasmid pNS17p via a slippage mechanism, using the bottom strand as template. The bold, opposing arrows show the palindromic sequence. The arrows in the same direction mark the directly repeated sequences.

Insert Sequence in pNS17p:

Replication of bottom strand:

Deletion mutant:

revealed that the inserts were deleted more frequently in the plasmids with the *Ori* in orientation 1 than in the plasmids containing an *Ori* in orientation 2, regardless of the type of insert present. These results show that the frequency of deletion mutation in this plasmid system is dependent upon the orientation of the target gene with respect to the *Ori* and suggest that deletion mutation occurs preferentially in the lagging strand.

The current DNA replication model suggests that there is a dimeric DNA polymerase III holoenzyme complex which synthesizes the leading and lagging strands simultaneously (Debyser *et al.*, 1994). During DNA replication, the leading strand may be continuously synthesized, while the lagging strand is synthesized discontinuously and temporally later than the complementary sequences in the leading strand (Rosche *et al.*, 1995). This leaves portions of the lagging strand single-stranded, which provides more opportunity for the formation of secondary structures and possibly misalignments, leading to an increase in deletion frequency in the lagging strand. Slippage of the template during lagging strand synthesis leads to the formation of the more stable slipped mispaired intermediate.

The observation that deletions occur more frequently during lagging strand synthesis is consistent with other studies. To determine whether deletion mutations occur preferentially in the leading or lagging strand, Trinh and Sinden (1991) examined the effect of reversing the direction of the *Cm^r* on the deletion frequency of various palindromic and non-palindromic inserts flanked by direct repeats in plasmids pBR325 and pBR523. It was found that reversing the *Cm^r* gene significantly decreased the deletion frequency of the inserts. This system was further investigated by Rosche *et al.*

(1995), who found that when slippage was stabilized by hairpin formation in the lagging strand, deletion of the inserts occurred up to 50-fold more frequently than when hairpin formation was in the leading strand. These results support the suggestion that during replication, slipped misalignment of DNA to form hairpin structures may occur preferentially in the lagging strand (Trinh and Sinden, 1991; Rosche *et al.*, 1995).

The stability of hairpin structures formed by the 17 bp palindromic and 17 bp non-palindromic inserts were analyzed by calculation of ΔG values. It was found that the hairpin structure formed by the palindromic insert (shown in Figure 4.1a) has a ΔG value of -12.7 kcal/mol, while the structure formed by the non-palindromic insert has a ΔG value of -1.0 kcal/mol. These ΔG values support the suggestion that the palindromic insert forms a more stable secondary structure than the non-palindromic insert, allowing more opportunity for slippage to occur, leading to a higher deletion frequency of the insert. The deletion frequency data obtained in this study is consistent with this hypothesis.

The pNS plasmid system was found to be a sensitive system for the detection of deletion mutants. Sequencing data revealed that only the revertants which contained a precise deletion grew on the selective media, allowing for detection of only the desired deletion mutants. Also, the sequence context of the two types of inserts led to the formation of secondary structures with very different ΔG values, which may explain the marked differences in deletion frequencies between the 17p and 17np inserts. However, there were some drawbacks to using this system. The mutants were slow growing, requiring 72 hours to form colonies that could be scored reliably. They were also

temperature sensitive, exhibiting no growth when incubated at 37°C. After transformation, the cells were allowed to express at 30°C instead of 37°C; and plates were also incubated at 30°C. Using this system, Dr. Bichara's laboratory in Strasbourg also found the mutant colonies to be slow growing and temperature sensitive. Their sequencing data indicated that some of the mutants were phenotypically Cm resistant, but genetically Cm sensitive. However, all of the mutants sequenced in our laboratory were found to contain precise deletions, indicating that they were all both phenotypically and genetically Cm resistant.

4.1.2 The plac and ptac plasmid systems

The six plac plasmids (pLC20p, pLC20np, pLC20T, pLCR20p, pLCR20np and pLCR20T) and six ptac plasmids (pTC20p, pTC20np, pTC20T, pTCR20p, pTCR20np, and pTCR20T) were transformed into *E. coli* TG1uvrA-F' competent cells, and the deletion frequencies of each of the three types of inserts were compared using statistical analyses. The mean deletion frequencies between plasmids pLC20np and pLCR20np (three-fold), pTC20p and pTCR20p (5.5-fold), pTC20np and pTCR20np (three-fold), and pTC20T and pTCR20T (three-fold) were found to be significantly different, with the plasmids containing the *Ori* in orientation 1 showing a significantly greater frequency of deletion of the insert than those with the *Ori* in orientation 2. These data are consistent with the pNS data, suggesting that the orientation of the *Ori* has an effect on the deletion frequency of all three types of inserts. As mentioned in section 4.1.1, reversing the orientation of the *Ori* changes the relationship between the leading and lagging strands of

DNA replication. When the *Ori* and the antibiotic resistance gene are in opposite directions, the deletion frequency has been found to decrease significantly (Trinh and Sinden, 1991). This is in agreement with the results found in the *plac* and *ptac* plasmids.

The differences in the mean deletion frequencies of the four sets of plasmids discussed above were found to be statistically significant; however, the differences were only two- to five-fold. These differences are very small in comparison to the differences observed in plasmid pNS17p and pNSR17p (which was 12-fold), suggesting that the pNS plasmid system may be more susceptible to deletion mutation when the *Ori* orientation is reversed. As shown in Figures 4.1a and 4.1b, the asymmetry of the inverted repeats with respect to the direct repeats in the 17p insert (of the pNS system) allows the formation of two different slipped mispaired intermediates, with the intermediate formed during lagging strand synthesis being more stable. This may explain the large difference in deletion frequency between pNS plasmids containing *Ori* in orientation 1 and those containing *Ori* in orientation 2. Unlike the pNS system, the inverted repeat in *plac* and *ptac* plasmids is symmetrical with respect to the direct repeats in the *plac/ptac* plasmid systems (Figure 4.2); therefore, the slipped mispaired intermediates that form during leading and lagging strand syntheses are expected to be equally stable. This may explain the relatively small differences in deletion frequency that were observed between *plac/ptac* plasmids containing *Ori* in opposite orientations, as compared to the significant differences seen in the pNS system.

Figure 4.2 A model for the generation of a deletion mutant in plasmid pLC20p (or pTC20p) via a slippage mechanism. The bold, opposing arrows show the palindromic sequence. The arrows in the same direction mark the directly repeated sequences.

Wild-type pLC20p (or pTC20p) sequence:

5'...AATTGGTCACCTATCAATTTCGAGCTTGAGTATCGTACGATACTCAGTCACCTATCAATTTCGAGCTC...-3'
 3'...TTAACCAGTGGATAGTTAAGCTCGAACTCATAGCATGCTATGAGTCAGTGGATAGTTAAGCTCGAG...-5'

Replication of top strand:

5'...AATTGGTCACCTATCAATTTCGAGCTTGAGTATCGTACGATACTCAGTCACCTATCAATTTCGAGCTC...-3'
 CAGTGGATAGTTAAGCTCGAG...-5'

Slippage occurs

Slipped mispaired intermediate:

Extension of slipped intermediate to form heteroduplex

Plasmid replication or heteroduplex repair

Deletion mutant:

5'...AATTGGTCACCTATCAATTTCGAGCTC...-3'
 3'...TTAACCAGTGGATAGTTAAGCTCGAG...-5'

Statistical comparisons between the deletion frequencies of the palindromic, non-palindromic and triplet repeat inserts showed that the palindromic and non-palindromic inserts were deleted with approximately the same frequency; however, both types of inserts were deleted significantly more frequently than the triplet repeat insert, in the ptac system. The 20p insert was deleted three-fold as frequently as the 20T insert, and the 20np insert was deleted twice as frequently as the 20T insert. This observation was not found in the plac system. These differences in deletion frequency of the different inserts are small as compared to the pNS system, in which it was found that the palindromic insert (17p) was deleted with a 185-fold or 22-fold increase in frequency over the non-palindromic insert (17np), depending on the orientation of the *Ori*. Analysis of the ΔG values of the most stable secondary structures formed by each type of insert (as predicted by MacDNAsis) showed that the 20p insert had a ΔG value of -6.0 kcal/mol, the 20np insert had a ΔG value of -0.9 kcal/mol and the 20T insert had a ΔG value of -0.3 kcal/mol. These values indicated that of the three types of inserts used, the palindromic insert has the ability to form the most stable secondary structure. However, it is less stable than the palindromic insert in the pNS plasmids ($\Delta G = -12.7$ kcal/mol). This is reflected in the insignificant differences in deletion frequency of the inserts in the plac/ptac systems. One reason may be that the difference in the ΔG values of the palindromic insert and the other two inserts is not very large, suggesting that perhaps the theoretical stability of the hairpin structure formed by the palindromic insert is not sufficient to effect an increase in deletion frequency over the other two types of inserts.

When the deletion frequencies of the inserts in the 12 plasmids were compared in SOS⁻ and SOS⁺ cells, the differences were found to vary between zero and four-fold. Some of the plasmids showed an increase in deletion frequency when transformed into SOS⁺ cells, while other plasmids showed a decrease in deletion frequency, regardless of the type of insert or the orientation of the *Ori*, in the *plac/ptac* plasmids. No pattern was apparent from the results obtained. Likewise, no pattern was observed in the deletion frequencies of the pNS plasmids transformed into SOS⁺ cells. When a cell encounters DNA-damaging agents the SOS system is induced to deal with the damage by an error-prone damage tolerance mechanism known as translesion synthesis, which leads to mutagenesis (Smith and Walker, 1998). Since the SOS system is designed to deal with damaged DNA and the plasmids in the three systems were not exposed to DNA-damaging agents, SOS induction of the cells was not expected to have an effect on the rate of deletion mutation.

4.1.3 Molecular nature of spontaneous deletion mutants

Sequencing data of revertant plasmid DNA revealed that the mutants contained either precise deletions or a mixture of two different types of sequences: the wild type plasmid containing the insert, and the precise deletion sequence. The mixed populations may have arisen in several different ways. The *plac/ptac* plasmids are pUC derivatives and have high copy numbers. Since several plasmids may be found in one cell, it is possible to have colonies which carry plasmids that contain insert sequences and plasmids that have precise deletions. If the insert-containing plasmid has gone through

one round of replication without introducing any errors, and then a deletion occurs, this will result in a cell that contains two different types of plasmids (the insert sequences and precise deletion). The mixed mutants are able to survive in the presence of kanamycin because they contain plasmids with inserted sequences and plasmids with precise deletions. The plasmids that contain precise deletions are able to express the Kan^r gene and confer resistance to the bacteria, allowing the growth of the mixed mutants. Also, a high proportion of dimers have been observed (on agarose gel) in these mixed deletion mutants, suggesting that dimers may have formed between an insert-containing plasmid and one containing a precise deletion. A precise deletion would be generated right after transformation, in the first round of replication.

4.2 Induced deletions in AAF-modified plasmids

4.2.1 Survival of N-Aco-AAF-modified ptac plasmids

The five ptac plasmids that were modified to different levels with N-Aco-AAF were found to exhibit a dose-dependent decrease in transformation efficiency when introduced into TG1 *uvrA*-F' competent cells. In these experiments, the plasmids were randomly modified with N-Aco-AAF at levels ranging from 0 to 32 AAF adducts per plasmid molecule. At approximately 10 AAF adducts per plasmid, the transformation efficiency decreased to an average of 20%. Further increase in the level of modification decreased the transformation efficiency to close to zero. These results are consistent with the findings of Koffel-Schwartz *et al.* (1984), who compared the survival of

AAF-modified plasmids in wild type and *uvrA* strains of *E. coli*. The decreased survival of AAF-modified plasmids in *uvrA* strains compared to the wild type strain suggests that the unexcised AAF lesions prevent plasmid replication, leading to lethality (Koffel-Schwartz *et al.*, 1984). Fuchs and Seeberg (1984) have found that the excision repair of AAF lesions is *uvrABC*-dependent.

The survival of the modified ptac plasmids, found to decrease dramatically with increasing modification, may be caused by the many replication blocks formed on the DNA by the AAF adducts. If there were only a small number of adducts, the cell may be able to use lesion avoidance mechanisms to bypass the lesions in a *uvrA* cell. However, if there are many bulky adducts on the plasmid, there would be a large number of stalled replication forks, not all of which could be rescued by damage avoidance mechanisms. This blockage in the replication of the plasmid does not allow expression of the antibiotic resistance genes necessary for growth on the selective media, leading to lethality. The TG1*uvrA*-F' cells transformed with plasmids containing approximately 10 AAF adducts had a 20% survival, while plasmids containing a greater number of adducts showed a survival rate that was close to zero. Bypass may still occur, but its occurrence is at such low frequencies that it would be impossible to carry out the assay.

4.2.2 AAF-induced deletion mutants

Five of the six unmodified ptac plasmids and the corresponding five plasmids containing approximately 10-AAF adducts per plasmid molecule were used in the reversion assay to determine the effects of AAF-modification and SOS induction on

deletion frequency. Statistical analyses revealed that in plasmids pTC20p and pTCR20np, the deletion frequencies in the AAF-modified plasmids were significantly higher than the unmodified versions of these two plasmids. However, the values were only marginally different, with a three-fold and a 1.5-fold increase for plasmids pTC20p and pTCR20np, respectively. For plasmids pTCR20p and pTCR20T, the increases in deletion frequency from unmodified to AAF-modified were two-fold and were found to be statistically insignificant. In plasmid pLCR20p modified with N-Aco-AAF to an average of approximately 10 AAF adducts per plasmid, a significant difference in deletion frequency was observed between the unmodified and AAF-modified plasmids; however, the actual difference was only approximately two-fold. Previous studies have found that AAF adducts induce -1 frameshift mutations at repetitive sequences and -2 frameshifts at the *NarI* restriction site (Burnouf *et al.*, 1989; Bintz and Fuchs, 1990). Bintz and Fuchs (1990) found that modification with N-Aco-AAF induced -2 frameshift mutations in alternating GC sequences with a frequency of $2-3 \times 10^{-4}$, which is 5-120 times as high as the deletion frequencies observed in our *plac/ptac* plasmid system. These results suggest that the formation of the AAF adducts on the plasmids may have an effect in inducing the formation of deletion mutations in the *ptac* system. However, the effect is extremely small.

When the unmodified and AAF-modified *plac/ptac* plasmids were transformed into SOS⁺ cells, the results were inconclusive. In plasmids pTC20np and pTCR20T, it was found that the AAF modified plasmids showed only a two-fold higher deletion frequency than the corresponding unmodified plasmids. However, statistical analysis

indicated that the differences were significant. The remaining three plasmids (pTC20p, pTCR20p and pTCR20T) showed no significant increase in deletion frequency from the unmodified to the corresponding modified plasmids. These results suggest that SOS induction does not play a role in increasing the frequency of deletion mutation in these modified plasmids.

Likewise, the mean deletion frequencies of the two AAF-modified plac plasmids, pLC20p and pLCR20p, in SOS⁺ cells were not significantly different from those of the unmodified plasmids. It has been suggested that induction of the SOS response causes an increase in the occurrence of translesion synthesis, which is believed to play a role in increasing the rate of formation of mutations (Koffel-Schwartz *et al.*, 1996). During an SOS response, the cellular concentrations of UmuD' and UmuC increase. These proteins are believed to interact with PolIII holoenzyme, producing mutations during elongation from a misincorporated nucleotide at a lesion site (Koffel-Schwartz *et al.*, 1996). It has been suggested that UmuD'C proteins allow replication to proceed through a lesion during the SOS response by preventing the polymerase from dissociating at the lesion (Koffel-Schwartz *et al.*, 1996). Alternatively, Napolitano *et al.* (1997) have provided evidence that during DNA synthesis, UmuD'C promotes the extension of 3'-OH termini situated within slipped template-primer structures. SOS induction causes a significant increase in the occurrence of frameshift mutations on plasmids carrying a single AAF adduct (Veaute and Fuchs, 1993; Napolitano *et al.*, 1997). However, the results obtained from the unmodified and modified plac/ptac plasmids suggest that if SOS induction plays any role in the formation of deletion mutations, then that contribution is relatively small.

When mutants which resulted from the transformation of AAF-modified plac plasmids were sequenced, the majority were found to contain mixed populations. As was observed in the unmodified plac plasmids, mutants derived from AAF-modified plasmids were mostly mixed (75%) and the remaining colonies (25%) contained precise deletions. When mutants from transformation experiments involving AAF-modified ptac plasmids were sequenced, it was found that the number of precise deletions was more than two-fold the number of mixed mutants in SOS⁻ cells, suggesting that AAF modification may also play a role in the induction of deletion mutations in the ptac plasmid system. These results also suggest that deletion mutations induced by AAF are generated in the first round of replication. DNA polymerase is blocked by the AAF adduct; precise deletion formation may result from translesion synthesis. If the deletion mutation was generated in subsequent rounds of replication, then the mutant would most likely be a mixed mutant. Since there were twice as many precise mutants as mixed mutants, it is believed that most deletion mutations in AAF-modified plasmids occur in the first round of replication in the ptac plasmids. It was also found that more of the SOS⁺ mutants contained precise deletions than the SOS⁻ mutants. SOS induction causes an increase in a cell's translesion synthesis activity, leading to increased mutagenesis (Koffel-Schwartz *et al.*, 1996), which may explain why so many more precise deletion mutations were found in SOS⁺ cells than in SOS⁻ cells.

The pNS plasmids, pNSR17p and pNSR17np, were modified with N-Aco-AAF to produce plasmids containing an average of 0 to 5 AAF adducts per plasmid molecule. Transformation of the modified pNSR17p plasmids into both SOS⁻ and SOS⁺ TK603

cells showed no significant difference in the deletion frequency of the insert in either SOS⁻ or SOS⁺ cells. These results indicate that at this level of modification, the number of AAF adducts were not sufficient for the induction of deletion mutations in this plasmid, or that AAF does not induce deletion mutations in this system. The transformation efficiencies of the modified and unmodified plasmids were also found to be very similar, suggesting that replication was either not blocked by the bulky AAF adducts, or that the blocked replication fork was efficiently rescued by damage avoidance mechanisms which employ the complementary strand (Koffel-Schwartz *et al.*, 1996).

In contrast to the results reported here, Schaaper *et al.* (1990) have reported a 520-fold increase in the frequency of deletion mutations in the *lacI* gene modified with N-Aco-AAF. They observed a 276 bp deletion that occurred 17 times, and suggested that the high frequency with which it was recovered and its absence from previous collections of spontaneous mutations indicate that it may have been induced by the N-Aco-AAF treatment.

In this study, no difference in deletion frequency was observed in AAF-modified plasmids with *Ori* in opposite orientations. Veaute and Fuchs (1993) investigated the effect of the orientation of a frameshift selective marker with respect to the orientation of the origin of replication. An AAF adduct was placed in the non-transcribed strand (the lagging strand) of the *lacZ* α -complementing gene of plasmid pUC8 (referred to as the lagging orientation). The *lacZ* α -complementing gene was cut with the restriction enzyme *HaeII* and cloned in the opposite orientation, placing the adduct on the leading strand of replication (referred to as the leading orientation). It was found that, regardless

of SOS induction, induced frameshift mutations occurred with a 20-fold higher frequency when the adduct was in the lagging strand than when it was in the leading strand (Veaute and Fuchs, 1993). The presence of a bulky AAF adduct may block replication, causing the cell to rely on mechanisms such as damage avoidance or translesion synthesis to bypass this blockage. A replication block formed by the AAF adduct would cause the region around the replication fork to be single stranded, increasing the opportunity for slippage to occur. If the adduct was formed near the palindromic insert, the chance of slippage may further increase due to the formation of a stable slipped mispaired intermediate. These two factors may contribute to the small increase in deletion frequency in the AAF-modified plasmid.

4.3 Dimer formation in deletion mutants

The plasmids from all three systems used for transformation experiments were analyzed by agarose gel electrophoresis prior to introduction into cells, and were found to contain only the monomeric forms. However, analysis of the revertant plasmids showed that a significant proportion of dimers were present. In some mutants, only dimers were present. In preliminary experiments, the formation of dimers in revertants was examined by separating the monomers and dimers on agarose gel and purifying these plasmids for transformation. Equal amounts of monomers and dimers were transformed into *E. coli* cells and the mutants which grew in the presence of kan were analyzed. Six out of seven monomer/dimer plasmids transformed into host cells showed that a larger number of

transformants per μg of plasmid resulted from the dimers than from the corresponding monomers. This suggests that a larger proportion of dimers, as compared to monomers, contained deletion mutations, supporting the notion that recombination may be involved in the formation of deletions. These results are in agreement with the results from earlier works by Mazin *et al.* (1991) and Dianov *et al.* (1991). Both groups found that deletions between long direct repeats (>150 bp) were always recovered as dimers or higher order oligomers. It has been suggested that mutant dimers have a greater chance of being established after transformation than monomers (Kuzminov, 1996), which may also explain the increased number of dimers over monomers in our experiments.

In a similar experiment involving the transformation of monomers and dimers extracted from AAF-modified ptac revertants, it was found that more colonies grew from the transformation of monomers than the corresponding dimers, in the presence of kan, suggesting that different mechanisms may account for deletion mutation formation in unmodified and modified plasmids. However, these are only preliminary results and further examination of the formation of dimers in revertants derived from transformation with unmodified and modified plasmids must be conducted before a conclusion about the involvement of recombination as a mechanism in the generation of deletion mutations may be drawn.

Mazin *et al.* (1991) found that the large majority of their revertant plasmids were in dimeric form, which is consistent with the results obtained in this study. It has been argued that parental plasmids inhibit phenotypic expression of revertants such that in the presence of parental plasmids, only a small proportion of revertant plasmids will generate

revertant colonies. It has been suggested that revertants in dimer form can overcome the inhibition caused by parental plasmids because dimers are able to replicate faster than monomers to eventually replace the parental monomers and colonize the cell (Mazin *et al.*, 1996). The observation that most of the revertant plasmids were in dimer form has also led to the suggestion that recombination is occurring in the transformed cells and may contribute to the generation of deletion mutations.

In recent experiments, Mazin *et al.* (1996) examined the effect of plasmid dimerization on the formation of deletion mutations between direct repeats. They found that dimerization significantly increases the efficiency of deletion of revertants, implying that a decrease in the frequency of revertant colonies would be observed if the dimers were resolved into monomers. To test this, they introduced a dimer-resolution site into their plasmids to force them into monomeric form. This resulted in a significant decrease in the frequency of revertant colonies, suggesting that maintaining a multicopy plasmid in the monomeric form interferes with the fixation of deletion mutations. This led to the suggestion that deletion formation is mechanistically associated with plasmid dimerization (Mazin *et al.*, 1996).

5. CONCLUSIONS

The sequence context surrounding the deleted insert was found to have a significant effect on the spontaneous deletion frequency of the insert in the pNS plasmid system. The 17 bp palindromic insert was deleted more frequently than the 17 bp non-palindromic insert. However, in the plac/ptac systems, no significant differences in deletion frequency were found between the three types of inserts examined.

AAF-modification at approximately ten adducts/plasmid caused a small increase in deletion frequency in several of the inserts in the plac/ptac plasmid systems. In the pNS plasmids, AAF-modification at approximately five adducts/plasmid did not have an effect on the frequency of deletion mutation. SOS induction did not have an effect on the spontaneous or AAF-induced deletion frequencies of inserts in any of the three systems developed. In the pNS system, plasmids containing *Ori* in the same orientation as the antibiotic resistance markers (orientation 1) showed a significantly greater deletion frequency than plasmids containing *Ori* in the opposite direction (orientation 2), suggesting that deletion mutations occur preferentially during lagging strand synthesis. The same trend was found in the plac/ptac plasmid system; however, the differences were very small in comparison to those observed in the pNS system.

6. REFERENCES

- Albertini, A.M., M. Hofer, M.P. Calos and J.H. Miller. 1982. On the formation of spontaneous deletions: The importance of short sequence homologies in the generation of large deletions. *Cell*. **29**:319-328.
- Balbinder, E. 1993. Multiple pathways of deletion formation in *Escherichia coli*. *Mutation Research*. **299**:193-209.
- Balbinder, E., C. Mac Vean, R.E. Williams. 1989. Overlapping direct repeats stimulate deletions in specially designed derivatives of plasmid pBR325 in *Escherichia coli*. *Mutation Research*. **214**:233-252.
- Battista, J.R., T.Ohta, T. Nohmi, W. Sun and G.C. Walker. 1990. Dominant negative *umuD* mutations decreasing RecA-mediated cleavage suggest roles for intact UmuD in modulation of SOS mutagenesis. *Proceedings of the National Academy of Sciences of the U.S.A.* **87**:7190-7194.
- Bertrand-Burggraf, E., B. Kemper and R.P.P. Fuchs. 1994. Endonuclease VII of phage T4 nicks N-2-acetylaminofluorene-induced DNA structures *in vitro*. *Mutation Research*. **314**:287-295.
- Bichara, M. and R.P.P. Fuchs. 1985. DNA binding and mutation spectra of the carcinogen N-2-aminofluorene in *Escherichia coli*: A correlation between the conformation of the premutagenic lesion and the mutation specificity. *Journal of Molecular Biology*. **183**:341-351.

- Bintz, Regine and R.P.P. Fuchs. 1990. Induction of -2 frameshift mutations within alternating GC sequences by carcinogens that bind to the C8 position of guanine residues: Development of a specific mutation assay. *Molecular and General Genetics*. **221**:331-338.
- Broyde, S. and B. Hingerty. 1983. Conformation of 2-aminofluorene-modified DNA. *Biopolymers*. **22**:2423-2441.
- Burnouf, D., P. Koehl and R.P.P. Fuchs. 1989. Single adduct mutagenesis: Strong effect of the position of a single acetylaminofluorene adduct within a mutation hot spot. *Proceedings of the National Academy of Sciences of the U.S.A.* **86**:4147-4151.
- Chastain II, P.D., E.E. Eichler, S.Kang, D.L. Nelson, S.D. Levene and R.R. Sinden. 1995. Anomalous rapid electrophoretic mobility of DNA containing triplet repeats associated with human disease genes. *Biochemistry*. **34**:16125-16131.
- Cho, B.P., F.A. Beland and M.M. Marques. 1994. NMR structural studies of a 15-mer DNA duplex from a *ras* protooncogene modified with the carcinogen 2-aminofluorene: conformational heterogeneity. *Biochemistry*. **33**:1373-1384.
- DasGupta, U., K. Weston-Hafer and D.E. Berg. 1987. Local DNA sequence control of deletion formation in *Escherichia coli* plasmid pBR322. *Genetics*. **115**:41-49.
- Debyser, Z., S. Tabor, C.C. Richardson. 1994. Coordination of leading and lagging strand DNA synthesis at the replication fork of bacteriophage T7. *Cell*. **77**:157-166.

DeMarini, D.M., H.E. Brockman, F.J. de Serres, H.H. Evans, L.F. Stankowski and A.W.

Hsie. 1989. Specific-locus mutations induced in eukaryotes (especially mammalian cells) by radiation and chemicals: A perspective. *Mutation Research*. **220**:11-29.

Dianov, G.L., A.V. Kuzminov, A.V. Mazin and R.I. Salganik. 1991. Molecular mechanisms of deletion formation in *Escherichia coli* plasmids. I. Deletion formation mediated by long direct repeats. *Molecular and General Genetics*. **228**:153-159.

Drake, J.W., B. Charlesworth, D. Charlesworth and J.F. Crow. 1998. Rates of spontaneous mutation. *Genetics*. **148**:1667-1686.

Duncan, J., L. Hamilton and E.C. Friedberg. 1976. Enzymatic degradation of uracil-containing DNA. *Journal of Virology*. **19**:338-345.

Elledge, S.J. and G.C. Walker. 1983. The *muc* genes of pKM101 are induced by DNA damage. *Journal of Bacteriology*. **155**:1306-1315.

Ennis, D.G., N. Ossanna and D.W. Mount. 1989. Genetic separation of *Escherichia coli* *recA* functions for SOS mutagenesis and repressor cleavage. *Journal of Bacteriology*. **171**:2533-2541.

Friedberg, E.C., G.C. Walker and W. Siede. 1995. Introduction to mutagenesis. In DNA repair and mutagenesis. ASM Press. Washington, D.C.

Fuchs, R.P.P. 1975. *In vitro* recognition of carcinogen-induced local denaturation sites in native DNA by S₁ endonuclease from *Aspergillus oryzae*. *Nature*. **257**:151-152.

- Fuchs, R.P.P. 1983. (Letters to the Editor) DNA binding spectrum of the carcinogen N-acetoxy-N-2-acetylaminofluorene significantly differs from the mutation spectrum. *Journal of Molecular Biology*. **177**:173-180.
- Fuchs, R.P.P. and M.P. Duane. 1974. Dynamic structure of DNA modified with the carcinogen N-acetoxy-N-2-acetylaminofluorene. *Biochemistry*. **13**:4435-4440.
- Fuchs, R.P.P., J.F. Lefevre, J. Pouyet and M.P. Daune. 1976. Comparative orientation of the fluorene residue in native DNA modified by N-acetoxy-N-2-acetylaminofluorene and two 7-halogeno derivatives. *Biochemistry*. **15**:3347-3351.
- Fuchs, R.P.P., N. Schwartz and M.P. Duane. 1981. Hot spots of frameshift mutations induced by the ultimate carcinogen N-acetoxy-N-2-acetylaminofluorene. *Nature*. **294**:657-659.
- Fuchs, R.P.P. and E. Seeberg. 1984. pBR322 plasmid DNA modified with 2-acetylaminofluorene derivatives: transforming activity and *in vitro* strand cleavage by the *Escherichia coli* *uvrABC* endonuclease. *EMBO Journal*. **3**:757-760.
- Garcia, A., I.B. Lambert, R.P.P. Fuchs. 1993. DNA adduct-induced stabilization of slipped frameshift intermediates within repetitive sequences: Implications for mutagenesis. *Proceedings of the National Academy of Sciences of the USA*. **90**:5989-5993.

- Glickman, B.W. and L.S. Ripley. 1984. Structural intermediates of deletion mutagenesis: A role for palindromic DNA. *Proceedings of the National Academy of Sciences of the U.S.A.* **81**:512-516.
- Goodman, M.F. and D.K. Fygenson. 1998. DNA polymerase fidelity: From genetics toward a biochemical understanding. *Genetics*. **148**:1475-1482.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*. **166**:557-580.
- Harm, W., H. Harm and C.S. Rupert. 1968. Analysis of photoenzymatic repair of UV lesions in DNA by single light flashes. II. In vivo studies with *Escherichia coli* cells and bacteriophage. *Mutation Research*. **6**:371-385.
- Heflich, R.H. and R.E. Neft. 1994. Genetic toxicity of 2-acetylaminofluorene, 2-aminofluorene and some of their metabolites and model metabolites. *Mutation Research*. **318**:73-174.
- Hoffmann, G.R. and R.P.P. Fuchs. 1997. Mechanisms of frameshift mutations: insight from aromatic amines. *Chemical Research in Toxicology*. **10**:347-359.
- Howard-Flanders, P., R.P. Boyce and L. Theriot. 1966. Three loci in *Escherichia coli* K-12 that control the excision of pyrimidine dimers and certain other mutagen products from DNA. *Genetics*. **53**:1119-1136.
- Janel-Bintz, R., G. Maenhaut-Michel, M. Takahashi and R.P.P. Fuchs. 1991. Roles of recA mutant allele (*recA495*) in frameshift mutagenesis. *Biochimie*. **73**:491-495.

- Jaworski, A., W.A. Rosche, R. Gellibolian, S. Kang, M. Shimizu, R.P. Bowater, R.R. Sinden and R.D. Wells. 1995. Mismatch repair in *Escherichia coli* enhances instability of (CTG)_n triplet repeats from human hereditary diseases. *Proceedings of the National Academy of Sciences of the U.S.A.* **92**:11019-11023.
- Johnson, J.L., S. Hamm-Alvarez, G. Payne, G.B. Sancar, K.V. Rajagopalan and A. Sancar. 1988. Identification of the second chromophore of *Escherichia coli* and yeast DNA photolyases as 5,10-methnyltetrahydrofolate. *Proceedings of the National Academy of Sciences of the U.S.A.* **85**:2046-2050.
- Kato, T. and Y. Shinoura. 1977. Isolation and characterization of mutants of *Escherichia coli* deficient in induction of mutants by ultraviolet light. *Molecular and General Genetics.* **156**:121-131.
- Kim, S.T., Y.F. Li and A. Sancar. 1992. The third chromophore of DNA photolyase: Trp-277 of *Escherichia coli* DNA photolyase repairs thymine dimers by direct electron transfer. *Proceedings of the National Academy of Sciences of the U.S.A.* **89**:900-904.
- Koehl, P., P. Valladier, J.F. Lefevre and R.P.P. Fuchs. 1989. Strong structural effect of the position of a single acetylaminofluorene adduct within a mutation hot spot. *Nucleic Acids Research.* **17**:9531-9541.
- Koffel-Schwartz, N. and R.P.P. Fuchs. 1989. Genetic control of AAF-induced mutagenesis at alternating GC sequences: An additional role for RecA. *Molecular and General Genetics.* **215**:306-311.

- Koffel-Schwartz, N., F. Coin, X. Veaute and R.R. Fuchs. 1996. Cellular strategies for accommodating replication-hindering adducts in DNA: Control by the SOS response in *E. coli*. *Proceedings of the National academy of Sciences of the USA*. **93**:7805-7810.
- Koffel-Schwartz, N., G. Maenhaut-Michel and R.P.P. Fuchs. 1987. Specific strand loss in N-2-acetylaminofluorene-modified DNA. *Journal of Molecular Biology*. **193**:651-659.
- Koffel-Schwartz, N., J.M. Verdier, M. Bichara, A.M. Freund, M.P. Duane and R.P.P. Fuchs. 1984. Carcinogen-induced mutation spectrum in wild-type, *uvrA* and *umuC* strains of *Escherichia coli*; strain specificity and mutation-prone sequences. *Journal of Molecular Biology*. **177**:33-51.
- Kong, D. and W. Masker. 1994. Deletion between direct repeats In T7 DNA stimulated by double-strand breaks. *Journal of Bacteriology*. **176**:5904-5911.
- Kornberg, A. and T. Baker. 1992. DNA Replication. W.H. Freeman & Company. New York.
- Kuzminov, A. 1996. Mutant fixation via plasmid dimerization and its relation to human diseases. *Trends in Genetics*. **12**:246-249.
- Lambert, I.B., R.L. Napolitano and R.P.P. Fuchs. 1992. Carcinogen-induced frameshift mutagenesis in repetitive sequences. *Proceedings of the National academy of Sciences of the U.S.A.* **89**:1310-1314.

- Lindahl, T. 1979. DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and A base excision-repair. *Prog. Nucleic Acid Research in Molecular Biology*. **22**:135-192.
- Lindsley, J.E. and R.P.P. Fuchs. 1994. Use of single-turnover kinetics to study bulky adduct bypass by T7 DNA polymerase. *Biochemistry*. **33**:764-772.
- Loeb, L.A. and K.C. Cheng. 1990. Errors in DNA synthesis: A source of spontaneous mutations. *Mutation Research*. **238**:297-304.
- Maniatis, T., J. Sambrook and E.F. Fritsch. 1989. Molecular Cloning: A Laboratory Manual. Second Edition. Cold Spring Harbor Laboratory Press. New York.
- Mazin, A.V., A.V. Kuzminov, G.L. Dianov and R.I. Salganik. 1991. Mechanisms of deletion formation in *Escherichia coli* plasmids. II. Deletions mediated by short direct repeats. *Molecular and General Genetics*. **228**:209-214.
- Mazin, A.V., T.V. Timchenko, M.K. Saparbaev and O.M. Mazina. 1996. Dimerization of plasmid DNA accelerates selection for antibiotic resistance. *Molecular Microbiology*. **20**:101-108.
- Milhé, C., C. Dhalluin, R.P.P. Fuchs and J.F. Lefèvre. 1994. NMR evidence of the stabilization by the carcinogen N-2-acetylaminofluorene of a frameshift mutagenesis intermediate. *Nucleic Acids Research*. **22**:4646-4652.
- Minnick, D.T. and T.A. Kunkel. 1996. DNA Synthesis Errors, Mutators and Cancer. *Cancer Surveys*. **28**:3-20.
- Napolitano, R.L., I.B. Lambert and R.P.P. Fuchs. 1994. DNA sequence determinants of carcinogen-induced frameshift mutagenesis. *Biochemistry*. **33**:1311-1315.

- Napolitano, R.L., I.B. Lambert and R.P.P. Fuchs. 1997. SOS factors involved in translesion synthesis. *Proceedings of the National Academy of Sciences of the U.S.A.* **94**:5733-5738.
- O'Handley, S.F., D.G. Sanford, R. Xu, C.C. Lester, B.E. Hingerty, S. Broyde and T.R. Krugh. 1993. Structural characterization of an N-acetyl-2-aminofluorene (AAF) modified DNA oligomer by NMR, energy minimization and molecular dynamics. *Biochemistry*. **32**:2481-2497.
- Orren, D.K. and A. Sancar. 1990. Formation and enzymatic properties of the UvrB DNA complex. *Journal of Biological Chemistry*. **265**:15796-15803.
- Orren, D.K., C.P. Selby, J.E. Hearst and A. Sancar. 1992. Post-incision steps of nucleotide excision repair in *Escherichia coli*. Disassembly of the UvrBC-DNA complex by helicase II and DNA polymerase I. *Journal of Biological Chemistry*. **267**:780-788.
- Owen, J.E. and G. Streisinger. 1985. Mechanisms of spontaneous and induced frameshift mutation in bacteriophage T4. *Genetics*. **109**:633-659.
- Pearson, C.E. and R.R. Sinden. 1996. Alternative structures in duplex DNA formed within the trinucleotide repeats of the myotonic dystrophy and fragile X loci. *Biochemistry*. **35**:5041-5053.
- Peat, T.S., E.G. Frank, J.P. McDonald, A.S. Levine, R. Woodgate and W.A. Hendrickson. 1996. The UmuD' protein filament and its potential role in damage induced mutagenesis. *Structure*. **4**:1401-1412.

- Pierce, J.C., D. Kong and W. Masker. 1991. The effect of the length of direct repeats and the presence of palindromes on deletion between directly repeated DNA sequences in bacteriophage T7. *Nucleic Acids Research*. 19:3901-3905.
- Ripley, L.S. 1990. Frameshift mutation: Determinants of specificity. *Annual Review of Genetics*. 24:189-213.
- Rosche, W.A., A. Jaworski, S. Kang, S.F. Kramer, J.E. Larson, D.P. Geidroc, R.D. Wells and R.R. Sinden. 1996. Single-stranded DNA-binding protein Enhances the stability of CTG triplet repeats in *Escherichia coli*. *Journal of Bacteriology*. 178:5042-5044.
- Rosche, W.A., T.Q. Trinh and R.R. Sinden. 1995. Differential DNA secondary structure-mediated deletion mutation in the leading and lagging strands. *Journal of Bacteriology*. 177:4385-4391.
- Sancar, G.B. and A. Sancar. 1987. Structure and function of DNA photolyases. *Trends in Biochemical Science*. 12:259-261.
- Sancar, A. and G.B. Sancar. 1988. DNA repair enzymes. *Annual Review of Biochemistry*. 57:29-67.
- Santella, R.M., D. Grunberger, S. Broyde and B.E. Hingerty. 1981. Z-DNA conformation of N-2-acetylaminofluorene modified poly(dG-dC) determined by reactivity with anti cytidine antibodies and minimized potential energy calculations. *Nucleic Acids Research*. 9:5459-5467.
- Sassanfar, M. and J.W. Roberts. 1990. Nature of the SOS-inducing signal in *E. coli*: the involvement of DNA replication. *Journal of Molecular Biology*. 212:79-96.

- Schapper, R.M., N. Koffel-Schwartz and R.P.P. Fuchs. 1990. N-Acetoxy-N-acetyl-2-aminofluorene-induced mutagenesis in the *lacI* gene of *Escherichia coli*. *Carcinogenesis*. **11**:1087-1095.
- Schulte-Frohlinde, D., K.H. Worm and M.Merz. 1993. Double-strand breaks in plasmid DNA and the induction of deletions. *Mutation Research*. **299**:233-250.
- Smith, K.C. 1992. Spontaneous mutagenesis: Experimental, genetic and other factors. *Mutation Research*. **277**:139-162.
- Smith, B.T. and G.C. Walker. 1998. Mutagenesis and More: *umuDC* and the *Escherichia coli* SOS response. *Genetics*. **148**:1599-1610.
- Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita, *et al.* 1966. Frameshift mutations and the genetic code. *Cold Spring Harbor Symp. Quant. Biol.* **31**:77-84.
- Tebbs, R.S. and L.J. Romano. 1994. Mutagenesis at a stie-specifically modified *NarI* sequence by acetylated and deacetylated aminofluorene adducts. *Biochemistry*. **33**:8998-9006.
- Tortora, G.J., B.R. Funke and C.L. Case. 1989. Microbiology: An Introduction. Third Edition. The Benjamin/Cummings Publishing Company, Inc. New York.
- Trinh, T.Q. and R.R. Sinden. 1991. Preferential DNA secondary structure mutagenesis in the lagging strand of replication in *E. coli*. *Nature*. **352**:544-547.
- Trinh, T.Q. and R.R. Sinden. 1993. The influence of primary and secondary DNA structure in deletion and duplication between direct repeats in *Escherichia coli*. *Genetics*. **134**:409-422.

- Veaute, X. and R.P.P. Fuchs. 1993. Greater susceptibility to mutations in lagging strand of DNA replication in *Escherichia coli* than in leading strand. *Science*. **261**:151-164.
- Walker, G.C. 1984. Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. *Microbiological Reviews*. **48**:60-93.
- Weaver, R.F. and P.W. Hedrick. 1992. Genetics. Second Edition. Wm.C. Brown Publishers. Dubuque, IA.
- Zar, J. H. 1974. Biostatistical Analysis. Prentice-Hall, Inc. New Jersey.

IMAGE EVALUATION TEST TARGET (QA-3)

APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved