NOTE TO USERS

This reproduction is the best copy available.

UMI[®]

**N-(2'-DEOXYGUANOSINE-8-YL)-N-ACETYL-2-AMINOFLUORENE INDUCED
TRANSLATION SYNTHESIS EVENTS IN *E. COLI*: ROLE OF Y-FAMILY
ERROR-PRONE POLYMERASES AND THE DNA SEQUENCE CONTEXT**

A Thesis Submitted to the Faculty of Graduate Studies
Carleton University

In Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy

Department of Biology, College of Natural Sciences
Faculty of Graduate Studies, Carleton University

Submitted by:
M. Reza Nokhbeh

© M. Reza Nokhbeh, Ottawa, Canada, 2004

Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-612-97838-9

Our file Notre référence

ISBN: 0-612-97838-9

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

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

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

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

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

*This thesis is dedicated to
my parents, my wife Shahnaz
and my son Aram*

PAGINATION ERROR.

ERREUR DE PAGINATION.

TEXT COMPLETE.

LE TEXTE EST COMPLET.

Abstract

A method was developed to construct a series of phagemid based single stranded circular vectors that carry a single N-2-acetylaminofluorene-C8-dG adduct. The method makes use of a cloned palindromic linker in the pUC118 phagemid that carries two symmetrically located *HincII* sites. Internal annealing of the cloned palindromic linker in the corresponding viral single stranded form of the vector generates a unique *HincII* site in the middle of the stem of a stem-loop structure. The larger *HincII* digestion product of the single stranded vector was used to clone short oligonucleotides that carried a single dG-C8-AAF adduct. The resultant single stranded circular vectors were used to study translesion (TLS) events past a single dG-C8-AAF adduct located at the G1 and G3 positions in a *SmaI* site, and at the G3 position in a *NarI* site. TLS was studied in SOS induced and uninduced *E.coli* JM103, a JM103 $\Delta umuC$ mutant strain, and derivatives that expressed Pol V and/or Pol RI from a plasmid. In all cellular backgrounds, TLS past all three adduct types appeared to be SOS dependent. TLS past these adducts was primarily error free. Mutagenic TLS events past the SmaG3-AAF adduct, primarily -G frameshifts followed by less frequent base substitutions, were dependent on either Pol V or Pol RI. TLS events past the SmaG1-AAF adduct were frequently error free, however, -C semi targeted frameshifts and G→T transversions were also induced in a Pol RI background. Error free TLS events past the NarG3-AAF adduct were largely dependent on SOS inducible Pol V and Pol RI; however, the major -2 frameshift TLS events were dependent on a SOS inducible polymerase other than Pol V and Pol RI. Case dependent mutagenesis models that explain the mechanistic basis of the observed TLS events past

the dG-AAF adducts involve the Y-family polymerases acting individually or in combination to manage the replication fork.

Acknowledgements

My utmost gratitude belongs to my parents who planted in my mind the love of science and the belief in the utter beauty of scientific notion. My greatest respect goes to late Dr. Rahim Nokhbeh who taught me how to stand and live for humane values; and to late Hamid Nokhbeh, my great teacher and mentor of life. Nonetheless, I could have never been able to accomplish anything in my life if I had not been supported by my wife, Shahnaz.

I would like to express my gratitude to my supervisor Dr. Iain Lambert for his support, tolerance and the opportunity of working and learning in his lab. The late Dr. Campbell Wyndham, my former advisory committee member, will always be remembered for his patient listening, great understanding and positivism. I am grateful to Dr. Linda Bonen for her invaluable guidance throughout my project and for her time to evaluate this writing. I would also like to extend my gratitude to Dr. Cupples, Dr. Golshani, Dr. Burk and Dr. Mingarelli for allocating their time to evaluate this work.

My heartfelt admiration goes to Dr. Myron Smith, with whom I shared the same laboratory, for his endless support especially at the most difficult times and for being in the lab even after midnights. Sheryl Hubbard, Cristina Micali, Raymond Tropiano, Leila Haidari, Mohamad Reza Dehghany, Nadereh Mir-Rashed, Christine Ficker, Moshi Kotierk and Amaneh Mirzaee, the former members of Dr. Smith's research group, all have contributed to my unforgettable good memories of sharing the lab environment with Myron's group.

Sebastian Kowalski and Yen Ly are especially appreciated for their kindness, great friendship, and their never ending moral support to me and my family throughout the most difficult times of our lives.

I would like to thank Dr. Lambert's former graduate, undergraduate honors students, Yen Ly, Katalin Bertenyi, Peter and Jacob Koziarz, Craig Carroll, Noelani Laycock, Shahram Boroumandi, Jacqui Whiteway, Nimani Sandhu, Parul Kumar, Sepideh Vaziri, Ron Booth, Geoff Turner, Catherine Fafrowicz, Christine Armour, Susan Richter, Sophie Douville, Laurie Duval, Ian Lawford, Nick Pokorny, Todd Hodgson, Jason Bailey, Sherri Boucher, John Veall, Brian Julian, Ami Barakat, and Dinesh Warnakulasuriyarachchi, most of whom contributed to my learning process by letting me understand and share solutions to difficulties in their research projects. Special thanks to Dr. Suzanne Paterson with whom I shared several collaborative works during my stay in Iain's lab.

Last but not the least, I would like to thank Dr. Ken Dimock, my present supervisor for my post doctoral studies, for his infinite positivism, his great support and understanding, and for opening a totally new chapter in my research career.

Table of Contents

<u>Subject</u>	<u>Page</u>
Abstract	IV
Acknowledgements	VI
Table of contents	VIII
List of abbreviations	XIV
List of figures	XVI
List of tables	XVIII

1. Introduction

1.1	General background	1
1.2	Metabolism and metabolic activation of nitrogenous aromatic hydrocarbons	4
1.3	Structural properties of AAF and AF DNA adducts as the basis for their biological effects	9
1.4	DNA conformational B to Z transition and the effect of AAF modification	16
1.5	Cellular strategies in response to DNA damage	18
1.5.1	DNA repair mechanisms	20
1.5.1.1	Direct reversal of damage	22
1.5.1.2	Repair by removal of incorrect or damaged bases	22
1.5.1.2.1	Base excision repair	22
1.5.1.2.2	Nucleotide excision repair	22
1.5.1.2.3	Nucleotide mismatch repair	24
1.5.2	Damage avoidance and recombinational bypass	26
1.5.3	SOS induction and translesion replication	29
1.5.3.1	Molecular organization of SOS regulon and regulation of SOS responses	30
1.5.3.1.1	Temporal activation of SOS functions follow a logical order	32

1.5.3.1.1.1 DNA damage dose dependent time lag for posttranslational processing of UmuD	37
1.5.3.1.1.2 Protein-protein interactions and formation of protein complexes	38
1.5.3.1.1.3 Differential autoproteolytic turnover of Umu proteins	39
1.5.3.2 Y-family and other error prone polymerases	40
1.5.3.3 TLS properties of error prone polymerases	46
1.5.3.3.1 Pol IV, Pol κ , Pol ι and Pol μ are involved in untargeted mutagenesis	50
1.5.3.3.2 TLS past bulky chemical substitutions	51
1.5.3.3.3 TLS past UV induced DNA lesions	54
1.5.3.3.4 TLS past abasic sites	54
1.6 AAF and AF induced <i>in vivo</i> mutagenesis	56
1.7 Scope and objectives of this study	59

2. Materials and Methods

2.1 Materials

2.1.1 Media, bacterial strains and maintenance	62
2.1.2 Enzymes and kits	62
2.1.3 DNA vectors and manipulations	63
2.1.4 DNA Oligonucleotides	63
2.1.5 DNA and protein sequence analysis software	65

2.2 Methods

2.2.1 Construction of single stranded phagemid based vectors that carry a single dG-C8-AAF adduct	65
2.2.2 Construction of pLHII	65
2.2.2.1 Digestion of pUC118	65
2.2.2.2 Kinasing the HII49 oligomers	67
2.2.2.3 Formation of HII49 linker	67
2.2.2.4 Cloning of the HII49 linker into pUC118	67
2.2.2.5 Preparation of the plus strand of SSCpLHII phagemid	68

2.2.2.6 Digestion of SSCpLHII and preparation of stock SSLpLHII	69
2.2.2.7 Optimization of the conditions for cloning the short DNA oligomers with and without AAF modification into SSLpLHII	70
2.2.2.8 Testing the efficiency of removal of scaffold DNA by heat denaturation and/or annealing to an anti-scaffold DNA	72
2.2.3 DNA sequencing	73
2.2.4 PAGE purification of 5'- ³² P end-labelled oligomers	75
2.2.5 Labelling of oligomers for probe hybridization	76
2.2.6 Preparation of competent cells induced for SOS functions	76
2.2.7 Preparation of competent cells without induction of SOS functions	78
2.2.8 Preparation of cellular collections transformed with single stranded constructs with and without AAF modification	78
2.2.9 Screening for mutational and non-mutational bypass events	79
2.2.9.1 Transfer of the colonies onto nitrocellulose membranes and their processing	79
2.2.9.2 Probe hybridization	80
2.2.9.3 Stripping the probes off the membranes	80
2.2.10 Construction of pDC322	80
2.2.10.1 PCR amplification of umuDC operon	81
2.2.10.2 Digestion of pBR322 with <i>EcoRI</i> and <i>PstI</i>	81
2.2.10.3 Ligation of <i>umuDC</i> PCR product and linearized pBR322 and screening for JM103 $\Delta umuC$ (pBR322/ <i>umuDC</i>) clones	82
2.2.10.4 Statistical analyses	82
2.2.10.4.1 Two-sample Student's t-test	83
2.2.10.4.2 The combined probabilities Chi-square test	84

3. Results

3.1 Experimental plan for constructing the single stranded phagemid	86
---	----

vectors that carry a site specific single dG-AAF adduct	
3.1.1 Large amount of TL1 digestion oligonucleotide was needed	86
3.1.2 TL1 and its digestion products needed to be eliminated	88
3.2 Development of the SS-DNA vectors carrying a single dG-AAF adduct	97
3.2.1 Construction of the phagemid SSCpLHII	97
3.2.2 Production and purification of the monomeric plus strand of the SSCpLHII	99
3.2.3 Digestion of the SSCpLHII	101
3.2.4 Optimization of cloning of a short oligonucleotide without an AAF adduct	101
3.2.5 Construction of the single stranded circular phagemid vectors with a single dG-AAF adduct and quantification of the SSC form	106
3.2.6 The efficiency of removal of scaffold DNA using anti-scaffold DNA	110
3.3 Mutagenesis studies using single stranded circular vectors containing an AAF modified single guanosine	112
3.3.1 Translesion synthesis events past SmaG3.AAF adduct	114
3.3.1.1 The TLS events past SmaG3.AAF adduct in <i>E.coli</i> JM103 cellular background and the effect of SOS induction	115
3.3.1.2 The TLS events past SmaG3.AAF adduct in <i>E.coli</i> JM103 Δ umuC strain and the effect of SOS induction	120
3.3.1.3 TLS events past SmaG3.AAF in <i>E.coli</i> JM103 Δ umuC(pDC322) and the effect of SOS induction	123
3.3.1.4 TLS events past SmaG3.AAF in <i>E.coli</i> JM103 Δ umuC(pGW1700) and the effect of SOS induction	127
3.3.2 Targeted G1→T and semi targeted minus C frameshift events past SmaG1.AAF in <i>E.coli</i> JM103 Δ umuC(pGW1700) and the effect of SOS induction	131
3.3.3 Translesion synthesis events past NarG3.AAF adduct	135
3.3.3.1 Translesion synthesis events past NarG3.AAF in <i>E.coli</i> JM103	136

and the effect of SOS induction	
3.3.3.2 TLS events past NarG3-AAF adduct in <i>E.coli</i> JM103 $\Delta umuC$	137
and the effect of SOS induction	
3.3.3.3 TLS events past NarG3-AAF adduct in <i>E.coli</i> JM103 $\Delta umu(pDC322)$	139
and the effect of SOS induction	
3.3.3.4 TLS events past NarG3-AAF adduct in <i>E.coli</i> JM103 $\Delta umuC$ (pGW1700)	141
and the effect of SOS induction	

4. Discussion

4.1. Advantages of using single stranded circular phagemid vectors that carry a single dG-AAF adduct	146
4.1.1. DNA repair background noise is circumvented	146
4.1.2 DNA structural topology effects are eliminated	147
4.1.3. Damage avoidance mechanisms are detoured	148
4.2. Development of single stranded phagemid vectors carrying a single dG-AAF adduct	149
4.3. Optimization of Mutation Screening	160
4.4. Mutagenesis	162
4.4.1. Error free TLS past SmaG3-AAF adduct is dependent on SOS regulated Pol V and Pol RI functions	162
4.4.2. Minus G frameshifted error prone TLS past SmaG3-AAF adduct is dependent on SOS regulated Pol V and Pol RI functions	162
4.4.3. TLS past SmaG3-AAF adduct in JM103 $\Delta umuC$ strain is dependent on an SOS regulated function other than Pol V and Pol RI	173
4.4.4. Base substitution TLS events past SmaG3-AAF adduct is dependent on SOS regulated Pol V and Pol RI functions	175
4.4.5. TLS events past SmaG1-AAF adduct are modulated by SOS regulated error prone polymerases and the position of the AAF adduct in <i>SmaI</i> sequence context	177

4.4.6. Unlike error free TLS the error prone -2 frameshifted TLS events past NarG3-AAF adduct is dependent on an SOS regulated polymerase in addition to Pol V and Pol RI	181
4.5. General model for mutagenesis and concluding remarks	186
10. List of References	193

List of Abbreviations

AAF: N-2-acetylaminofluorene
AF: N-2-aminofluorene
Amp^r: Ampicillin resistant /resistance
APS: Ammonium persulphate
BSA: Bovine Serum Albumin
CI: Chloroform Isoamyl alcohol
dG-C8-AAF: N-(2'-Deoxyguanosine-8-yl)-N-acetyl-2-aminofluorene
DMF: dimethylformamide
DTT: Dithiothreitol
EDTA: Ethylenediaminetetraacetic acid
EFB: Error free bypass
EPB: Error prone bypass
Glu-P-1: 2-amino-6-methylpyrido[1,2-a:3',2'-d]imidazole
Glu-P-2: 2-amino-dipyrido[1,2-a:3',2'-d]imidazole
IPTG: Isopropyl- β -D-thiogalactopyranoside
IQ: 2-amino-3-dimethylimidazo[4,5-f]quinoline
MeIQ: 2-amino-3,5-dimethylimidazo[4,5-f]quinoline
MeIQx: 2-amino-3,8-dimethylimidazo[4,5-f]quinoxaline
moi: Multiplicity of infection for phage
NER: Nucleotide excision repair
PCI: Phenol Chloroform Isoamyl alcohol (25:24:1 V/V)
Phe-P-1: 2-amino-5-phenylpyridine
PVP: Polyvinylpyrrolidone
Rif^r: Rifampicin resistant/resistance
SMI: Slipped mutagenic intermediate
SSC: Single stranded circular form of phagemid
ssc: Sodium chloride, sodium citrate buffer
SSL⁺: Single stranded linear form of coding strand of phagemid

SSPE: Sodium chloride, sodium phosphate, EDTA buffer

SSpLHIINarG0: Single stranded circular phagemid vector containing a cloned 15 nucleotide long oligomer, NarG0, that does not have any modification in the cloned 5'-G₁G₂CG₃CC-3' sequence context

SSpLHIINarG3.AAF: Single stranded circular phagemid vector containing a cloned 15 nucleotide long oligomer, NarG0, that carries a dG-C8-AAF adduct at the G₃ position of the 5'-G₁G₂CG₃CC-3' sequence context

SSpLHIISmaG0: Single stranded circular phagemid vector containing a cloned 14 nucleotide long oligomer, SmaG0, that does not have any modification in the cloned 5'-CCCG₁G₂G₃-3' sequence context

SSpLHIISmaG1.AAF: Single stranded circular phagemid vector containing a cloned 14 nucleotide long oligomer, SmaG1, that carries a dG-C8-AAF adduct at the G₁ position of the 5'-CCCG₁G₂G₃-3' sequence context

SSpLHIISmaG3.AAF: Single stranded circular phagemid vector containing a cloned 14 nucleotide long oligomer, SmaG3, that carries a dG-C8-AAF adduct at the G₃ position of the 5'-CCCG₁G₂G₃-3' sequence context

Str^r or rpsL: Resistance/resistant to Streptomycin due to a mutation in the 30S ribosomal subunit protein L

TE: Tris.EDTA buffer

TEMED: N, N, N', N'- tetramethyl-ethylenediamine

Tet^r: tetracycline resistant /resistance

TLS: Translesion synthesis

Trp-P-1: 3-amino-1,4-dimethyl-5-H-pyrido[4,3-b]indole

Trp-P-2: 3-amino-1-methyl-5-H-pyrido[4,3-b]indole

X.gal: 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

List of Figures

Figure no.	Description	Page
Figure 1.1	Schematic representation of Fluorene derivatives and dG-AF and dG-AAF	6
Figure 1.2	Metabolism and metabolic activation of nitrogenous aromatic	7
Figure 1.3	RecF recombinational bypass pathway	27
Figure 2.1	Schematic representation of the method used for construction of single stranded vectors carrying a dG-AAF adduct	66
Figure 3.1	Schematic representation of construction of M13mp18 based single stranded circular vectors that carry a single dG-AAF modification	87
Figure 3.2	A sample probe hybridization result	94
Figure 3.3	The verification of the sequence of the cloned HII49 linker	98
Figure 3.4	Resolution of a compressed sequence in HII49 linker	100
Figure 3.5	Phage and phagemid DNA isolation from <i>E.coli</i> JV30 (pLHII), infected with helper phage M13KO7 and purification of SSCpLHII viral strand	102
Figure 3.6	The verification of the <i>HincII</i> digestion of the stem loop structure of internally annealed SSCpLHII, through sequencing the digestion product	103
Figure 3.7	Optimization of cloning of a short oligonucleotide into SSLpLHII	105
Figure 3.8	The transformation capacities of the samples from cloning optimization experiment	107
Figure 3.9	Construction of single stranded circular vectors containing SmaG0, SmaG3.AAF and SmaG1.AAF, NarG0 and NarG3.AAF short oligonucleotides	109
Figure 3.10	The efficiency of removal of the scaffold oligonucleotide using heat denaturation and annealing to anti-scaffold oligonucleotide	113

Figure 4.1	Sequencing of pUC118, pLH118 and pLH-NarG0 constructs (pUC118 derivatives); pUC128, pLH128; M13mp18 and MLH (M13mp18 derivative)	156
Figure 4.2	Mechanism of error free and -1 frameshift mutation at <i>SmaI</i> site	172
Figure 4.3	Mechanism of error free, targeted and semi-targeted mutations induced by SmaG1-AAF adduct	180
Figure 4.4	Mechanism of error free and -2 frameshift mutation at NarI site: formation of minor and major slipped mutagenic intermediate structure induced by AAF modification	185
Figure 4.5	Schematic representation of a general model for mutagenesis hydrocarbons	190

List of Tables

Table no.	Description	Page
Table 1.1	Hierarchical induction of cellular SOS functions in response to DNA damage	34
Table 1.2	<i>In Vitro</i> translesion synthesis properties of prokaryotic and eukaryotic error prone DNA polymerases	47-49
Table 2.1	List of the oligonucleotides used in this study	64
Table 3.1	Efficiencies of incorporation of the unmodified and AAF modified oligonucleotides into single stranded circular vectors	111
Table 3.2	The TLS events past SmaG3.AAF adduct in SOS ⁻ and SOS ⁺ cells of <i>E.coli</i> JM103 wild type	117
Table 3.3	The TLS events past SmaG3.AAF adduct in SOS ⁻ and SOS ⁺ cells of <i>E.coli</i> JM103 ΔumuC	121
Table 3.4	The TLS events past SmaG3.AAF adduct in SOS ⁻ and SOS ⁺ cells of <i>E.coli</i> JM103 ΔumuC (pDC322)	125
Table 3.5	The TLS events past SmaG3.AAF adduct in SOS ⁻ and SOS ⁺ cells of <i>E.coli</i> JM103 ΔumuC (pGW1700)	129
Table 3.6	The TLS events past SmaG1.AAF adduct in SOS ⁻ and SOS ⁺ cell of <i>E.coli</i> JM103 ΔumuC (pGW1700)	134
Table 3.7	The TLS events past NarG3.AAF adduct in SOS ⁻ and SOS ⁺ cells of <i>E.coli</i> JM103 wild type	138
Table 3.8	The TLS events past NarG3.AAF adduct in SOS ⁻ and SOS ⁺ cells of <i>E.coli</i> JM103 ΔumuC	140
Table 3.9	The TLS events past NarG3.AAF adduct in SOS ⁻ and SOS ⁺ cells of <i>E.coli</i> JM103 ΔumuC (pDC322)	142
Table 3.10	The TLS events past NarG3.AAF adduct in SOS ⁻ and SOS ⁺ cells of <i>E.coli</i> JM103 ΔumuC (pGW1700)	144
Table 4.1	Reading frame of α-complementation fragment, stop codon use and phenotypes of different single stranded circular vectors	158

1. Introduction

1.1 General background: The highly ubiquitous polycyclic aromatic amines, their amide derivatives and nitro-substituted homo- and hetero-polycyclic aromatic hydrocarbons (nitro-PAHs) are pollutants that represent significant health and environmental hazards (Rosenkranz and Mermelstein 1983; Tokiwa and Ohnishi 1986). Aromatic amines comprise a group of mutagenic and carcinogenic products, closely related to nitro-PAHs, to which humans are exposed in both environmental and occupational settings. Cigarette smoke contains a plethora of alkylated aromatic arylamines that may account for an increased incidence of bladder cancer observed among smokers (Patrianakos and Hoffmann, 1979; Luceri *et al.*, 1993). In addition to environmental sources, various mutagenic and carcinogenic arylamines are formed as amino acid or sugar pyrolysates during routine daily cooking processes. Ingestion of compounds such as 2-amino-3-dimethylimidazo[4,5-f]quinoline (IQ), 2-amino-3,8-dimethylimidazo[4,5-f]quinoxaline (MeIQ), 2-amino-6-methylpyrido[1,2-a:3',2'-d]imidazole (Glu-P-1), 2-amino-dipyrido[1,2-a:3',2'-d]imidazole (Glu-P-2), 3-amino-1,4-dimethyl-5-H-pyrido[4,3-b]indole (Trp-1) and 3-amino-1-methyl-5-H-pyrido[4,3-b]indole (Trp-2) is considered an important source of chronic human exposure to these carcinogenic compounds; daily consumption of these compounds through foods reaches as high as 0.04-7.0 mg/day/person (Sugimura, 1985 and 1986; Johansson and Jägestad, 1994). Aromatic amines, which can be formed through either reduction of the nitro moiety of nitroarenes or pyrolysis of amino acids, cause unscheduled DNA synthesis in human cultured hepatic cells, are bladder carcinogens in humans, and produce tumors in

the liver, intestine and mammary glands in experimental animals (Beland *et al.*, 1983; Beland and Kadlubar, 1985; Beamand *et al.*, 1998).

As model arylamine and arylamide compounds for chemical mutagenesis and carcinogenesis, the extensively studied 2-aminofluorene (AF) and N-2-acetylaminofluorene (AAF) are the products of the metabolic nitroreduction of 2-nitrofluorene and its subsequent N-acetylation, respectively. These compounds were developed and patented by the US Department of Agriculture as potential fungicides and insecticides in 1940, but their use was banned when the potential carcinogenicity of AAF was demonstrated (Wilson *et al.*, 1941; Miller *et al.*, 1955). AF and AAF are genotoxic and mutagenic in both bacteria and eukaryotic cells (Casarett *et al.*, 1986; Heflich and Neft, 1994; Hoffmann and Fuchs, 1997), are able to induce unscheduled DNA synthesis, transform bladder epithelial cells and hepatocytes in different animal models, and their subcutaneous administration leads to the development of sarcomas (Bartsch *et al.*, 1977; Ide *et al.*, 1981; Beland *et al.*, 1983; Casarett *et al.*, 1986; Wang *et al.*, 1987; Gerlyng *et al.*, 1994; Ross and Leavitt, 1998). Besides genotoxicity, cytotoxic effects of AAF treatment that have been recently documented include interference with and modulation of the cell cycle (Trautwein *et al.*, 1999; Lindeman *et al.*, 2000), modulation of the apoptotic response and cellular proliferation in rat hepatic tissue, morphological and physiological alterations in the mitochondrial membrane, interference with the electron transfer chain in treated cells (Bitsch *et al.*, 1999; 2000), and elicitation of immunosuppressive effects (Koh *et al.*, 1995; Lee and Yang., 2000).

Most of the genotoxic and cytotoxic effects of arylamines, arylamides and their nitro-derivatives are achieved through their metabolic activation in both eukaryotic and

prokaryotic cells. The relatively short-lived intermediates that arise during the detoxification of AAF and AF through oxidative metabolic pathways, can react with biomacromolecules like DNA and proteins resulting in formation of adducts with these molecules. Alternatively, the bioactivated species may interfere with organellar protein components, e.g. mitochondria, leading to impairment of electron transfer chain and uncoupling of oxidative phosphorylation. More elaboration on metabolic activation of AAF and AF is presented in section 1.2.2.

Nitro-PAHs, which are structurally and metabolically related to arylamines, are produced by nitration of polycyclic aromatic hydrocarbons from fossil fuels and other organic matter during incomplete combustion processes, by atmospheric phototransformation (Okinaka *et al.*, 1984), and during cooking processes in high heat or by direct exposure to flame. Among more than 200 types of environmentally significant nitro-PAH's, 1-nitropyrene, 2-nitrofluorene and 2-nitrofluoranthene are produced in massive amounts annually through exhaust gases, and are released to the urban air as ambient particles (Schuetzle, 1983; Arey *et al.*, 1988; Beije and Möller , 1988; Möller *et al.*, 1993). Other fluorene derivatives, including 2,7- and 2,5-dinitrofluorene and 2,4,7-trinitrofluorene and their 9-oxo- derivatives, are also found in ambient air where the 2,7-dinitrofluorene concentration in some highly industrialized cities reaches about 30-fold that of the ambient 2-nitrofluorene concentration (Matsushida and Iida, 1986). The nitrated fluorene derivatives are frameshift mutagens and carcinogens (McCoy *et al.*, 1981). The potent carcinogenicity of 2,7-dinitrofluorene, 2-nitrofluorene, 2,5-dinitrofluorene, 9-oxo-2-nitrofluorene, 9-oxo-2,4,7-trinitrofluorene and 9-hydroxy-2-

nitrofluorene in animal models has been well documented (Malejka-Giganti *et al.*, 1999 and references therein).

The mutagenic activities present in carbon blacks that were used for production of xerographic toners prior to 1980 have been attributed primarily to nitropyrenes (Löfroth *et al.*, 1980; Rosenkrantz *et al.* 1980). Besides 1-nitropyrene (1-NP) the well-known representatives of nitropyrenes, 1,3-dinitropyrene (1,3-DNP), 1,6-dinitropyrene (1,6-DNP) and 1,8-dinitropyrene (1,8-DNP), are also found in grilled meats and cigarette smoke condensates. All of these compounds are potent mutagens in bacteria and eukaryotic cells and cause sister chromatid exchange in Chinese hamster ovary cells, chromosomal aberrations in rat epithelial cells, unscheduled DNA repair and neoplastic transformation in rodents and human cells (Sinha and Dao, 1974; Rosenkrantz and Mermelstein 1983; Tokiwa and Ohnishi 1986; Cavalieri *et al.*, 1988; El-Bayoumy *et al.*, 1988 and 1993; Lambert *et al.*, 1991; Lambert *et al.*, 1998, Imadia *et al.*, 1991a and 1991b; Tagliaferri *et al.*, 2000, Lambert *et al.*, 2001).

1.2 Metabolism and metabolic activation of nitrogenous aromatic hydrocarbons:

As with many xenobiotic compounds, no perfect metabolic pathway has evolved among all organisms that can degrade nitrogenous aromatic compounds to contribute to the energy requirements of the organism, and at the same time avoid hazardous effects due to the bioactive intermediary metabolites formed. In some prokaryotes and fungal species, metabolism of aromatic compounds is frequently centered on the aromatic nucleus, which is subject to ring oxidation through the action of mixed function monooxygenases and dioxygenases, followed by ring scission and mineralization of the products (Van der Meer *et al.*, 1992). Alternatively, phenolic, nitro- and carboxyl substituted homo- and

heterocyclic aromatic compounds can be degraded anaerobically through ring reduction followed by ring cleavage (Berry et al., 1987). The central intermediates resulting from ring cleavage are usually mono- or dicarboxylic fatty acids that can further be degraded into Acetyl COA through β -oxidation. A variety of nitroaromatic compounds can be used as growth substrates by some species of *Pseudomonas*, *Nocardia* and *Arthrobacter*. The nitro moiety in these compounds is released in an oxidative reaction replacing it with a hydroxyl group; further oxidation and ring fission results in mineralization of these compounds (Higson F.K., 1992a and b). Alternatively, nitrogenous functional groups such as nitro and nitroso groups can be rapidly reduced to hydroxylamine and/or amine species by the function of bacterial nitroreductases (Messier et al., 1981; Karpinsky et al., 1982; Rosenkranz et al. 1982; McCoy et al., 1981; Bryant et al., 1984; Howard et al., 1987; Zenno et al., 1996 a and b; Zenno et al., 1998). The DNA reactive intermediates, such as hydroxylamines, formed through reductive pathway (Fig. 1.1) are believed to contribute to the mutagenic/carcinogenic potential of these compounds and the therapeutic efficacy of the nitroheterocyclic antibiotics (Kedderis et al., 1988; Tocher, 1997). The hydroxylamine intermediate is transformed to acetyl- or sulfo-conjugated metastable compounds, which can spontaneously produce electrophilic arylnitrenium species that can attack nucleic acids (Fig. 1.2). However, the reverse reaction, i.e. oxidation of amines to the hydroxylamine, does not appear to be efficiently carried out in prokaryotes.

In eukaryotes, oxidation of the aromatic nucleus or the amine functional group, followed by conjugation and cellular excretion, is the major strategy to reduce the level of

Fig. 1.1: Schematic representation of fluorene derivatives and their guanine nucleotide adducts

Fig 1.2: Metabolic activation of 2-nitrofluorene and formation of dG-C8-AF adduct

the cellular damage caused by nitrogen substituted polycyclic aromatic compounds. The functional enzymatic accessory includes a range of cytochrome P-450 or flavin containing monooxygenases and peroxidases, as well as conjugative enzymes (reviewed in: Weisburger and Weisburger, 1958; Cramer et al., 1960; Miller et al., 1961 and 1964; Miller and Miller, 1969; Asokan et al., 1985; Kato, 1986; Butler et al., 1989; Kleman et al., 1990; Fulco, 1991; Sakai et al., 1992; Nelson et al., 1993 and 1996; Zhang et al., 1996; Jez et al., 1997; Gupta and Abou-Donia 1998; Goksoyr and Husoy, 1998; Danielson et al., 1999; Morel et al., 1999; Landi et al., 1999; Paquette et al., 2000; Arinc et al., 2000; Longo et al., 2000; Meunier et al., 2000; Gelardi et al., 2001; Menzel et al., 2001; Gonzalez et al., 2001). The P450 dependent ring or N-hydroxylation products of the arylamines and arylamides are further subject to esterification modifications by conjugative enzymes such as sulfotransferases, acetyltransferases and glucuronyltransferases, which accept a wide spectrum of substrates (reviewed in: Kreik, 1974; Jakoby et al., 1980; Mulder, 1981; Bartsch, 1981; Kato, 1986; Chung et al., 1993; Kadlubar, 1994; Babu et al., 1995; Mackenzie et al., 1997; Hsieh et al., 1998; Green and Tephly, 1998; Warnecke et al., 1999; Li et al., 2001; Sim et al., 2000). The resulting conjugates show slightly increased water solubility, and can be excreted through either bile or urinary routes. In this respect these reactions are considered as the general cellular detoxification strategy. This is particularly the case with glucuronidation reactions, which significantly increase the solubility of aromatic amines in bodily fluids resulting in their efficient renal or hepatic clearance. In contrast, both sulfonation and acetylation reactions have been implicated in the metabolic activation of a wide range of xenobiotics,

including aromatic amines (Watabe et al., 1982; Glatt., 1997). The intermediate species formed during these reactions can be easily transformed, non-enzymatically, into more electrophilic species, such as arylnitrenium ions, that can attack macromolecules, including nucleic acids and proteins to form their respective adducts (King and Philips, 1968; Weiseburger and Weiseburger, 1973; ; King et al, 1979; Kadlubar et al., 1977; Bartsch, 1981; Allaben et al., 1982; Miller and Miller, 1981; Miller et al., 1985; Babu et al., 1992; Hein, 1988a and 1988b; Levy et al., 1994; Kato and Yamazoe, 1994; Zenser et al., 1998; Lewis et al., 2000). Formation of nucleotide adducts is primarily responsible for the mutagenic and carcinogenic effects of these chemicals. However, the mutagenic/carcinogenic fate of the xenobiotic chemicals is determined by the efficiency of the different enzymes involved in activation and detoxification reactions. The presence of competing reactions, organ and tissue specific polymorphic distribution of these enzymes with varying activities, and the nature and persistence of the intermediary metabolites are factors that contribute to the level of the cellular damage, their mutagenicity and carcinogenicity (reviewed in: Kriek, 1974; Miller, 1978, Bartsch, 1981; Miller and Miller, 1983; Beland and Kadlubar, 1985; Heflich and Neft, 1994; Parker et al., 1996; Rothman et al., 1997).

1.3 Structural properties of AAF and AF DNA adducts as the basis for their

biological effects: The relatively short lived electrophilic intermediates derived from metabolism of AF and AAF, can attack nucleophilic centers in biomacromolecules such as proteins and nucleic acids (Heflich and Neft, 1994). In DNA the covalently bound AF and AAF adduct formed at the C8 position of guanine residues represent the principal

lesion (Westra and Visser, 1979; Tang and Lieberman, 1983) (Fig. 1.1). Despite the similarity of their chemical structures and the similar target for adduction, the dG-C8-AF (i.e. *N*-(2'-deoxyguanosine-8-yl)-*N*-2-aminofluorene) and dG-C8-AAF (i.e. *N*-(2'-deoxyguanosine-8-yl)-*N*-acetyl-2-aminofluorene) adducts assume completely different conformations in DNA. The distinct biological processing of the dG-C8-AF and dG-C8-AAF adducts are dependent on the conformation of the lesions which produce different mutational outcomes (Bichara and Fuchs, 1985), and have differential persistence in the cell (Kreik, 1974; Beland and Kadlubar, 1985). Hingerty and Broyde (1986) adopted a computerized analysis to calculate the minimized semi-empirical potential energies for the conformers generated by the AF and AAF adducts in the context of alternating dimeric, tetrameric and octameric GC sequences in single stranded and duplex DNA. In the predicted energy minimized structures containing the dG-C8-AAF adduct, the AAF moiety preferentially protrudes to the exterior of the undeformed left handed Z-DNA helix (Santella *et al.*, 1981a; Hingerty and Broyde, 1986; See also section 1.4). In contrast, to a lesser extent, the AAF moiety acquires base stacking by displacement of the guanine residue in the *syn* conformation - a rotation that places the AAF moiety in the major groove of the B-DNA helix- resulting in local denaturation and helix bending (Hingerty and Broyde, 1986). However, the lowest energy conformer of the C8-AF modified guanine residue in the context of alternating GC nucleotides, is in the *anti* conformation with the AF situated in the major groove stacked with neighboring bases in the local B-DNA context. In contrast to the AAF lesion, the AF lesion does not produce a significant helix distortion, rendering the AF adduct more refractive to repair functions. These theoretical analyses are consistent with the solution structures determined by

proton NMR studies of random DNA sequences carrying a dG-C8-AF adduct that produced a major AF conformer (60%) in a relatively unperturbed B-helix in which the arylamine moiety is projected into the major groove. The minor conformer contained an AF-modified guanine in the *syn* conformation with the arylamine group positioned at the minor groove (Cho *et al.*, 1994).

The presence of an extra N-acetyl group in AAF compared to AF is responsible for the structural differences noted above. Acetylation of arylamines, as in AAF, appears to have a similar effect to alkylation of the aromatic moiety, which can produce changes in the structural and biological properties of the compounds. For example, Beland *et al.* (1997) have shown that in a number of dG-C8 adducts of alkyl-substituted arylamine derivatives, i.e. aminofluorene, aniline and aminobiphenyl, the ortho alkyl-substituted adducts preferentially assume the *syn* conformation about the glycosidic bond resulting in projection of the alkylarylamine residue into the helix of the double stranded DNA. The ortho-substituted arylamine analogs that form *syn* dG adducts show a significantly higher mutagenicity in the Ames test and higher tumorigenicity in animal models compared to the analogous nonalkylated arylamine (El-Bayoumy *et al.*, 1981; Nussbaum *et al.*, 1983).

Most of the multiple kinds of mutations induced by a specific type of DNA adduct, including those derived from the carcinogenic arylamines, occur at the replication fork when DNA polymerases attempt to replicate past the lesion. Cho and Zhou (1999) conducted a NMR study to address the structural properties of the dG-AAF adduct in the context of single stranded and double stranded DNA, reminiscent of the single-double stranded junction of the replication fork where the DNA polymerase encounters the

adducted guanine residue. They used a 7-fluoro substituted dG-C8-AAF (dG-C8-FAAF) adduct in three different structural contexts- as a monomeric single nucleotide, and incorporated into single- and double stranded DNA. Their results showed the presence of four conformers for monomeric adduct and two similar conformers for DNA incorporated forms in both of single and double stranded contexts (Cho and Zhou, 1999). The nonacetylated counterpart of dG-C8-FAAF exhibited a single resonance signal under the same experimental conditions. The presence of the bulky acetyl group in the monomeric dG-C8-FAAF adduct produced four torsional conformers resulting from a zero and a 180° rotation around the guanyl- nitrogen bond to form two conformers, each of which produced a *cis* and *trans* torsional conformer due to zero and 180° rotation of the acetyl group around the amide bond. However, the guanine residue in the dG-C8-FAAF adduct in the context of both the single and double stranded DNA showed a strong preference for the *syn* conformation around the glycosidic bond. In addition, the stable stacking property of the fluorene moiety with the aromatic rings of the adjacent bases in the same strand imposes an increased torsional restriction on the rotation around the guanyl- nitrogen and fluorene-nitrogen bonds (Evans and Levine, 1986 and 1988). This leaves the acetyl group to assume either the *cis* or *trans* conformation with torsional angles of zero and 180° around the amide bond. The energy barrier between acetyl *cis* and *trans* conformations is greater in the double stranded DNA context (14.3 kcal/mol) (Cho and Zhou, 1999) compared to that in single stranded DNA (2.1 kcal/mol) (Shapiro *et al.*, 1994); however, the *trans* conformer is preferred over the *cis* conformer due to its lower energy state in both contexts. The importance of the report by Cho and Zhou (1999) is

two fold. First, they provided additional molecular structural evidence for the propensity of the dG-AAF adduct to assume a DNA intercalated base displaced conformation; however, the study emphasized the similarity of the conformation of the adduct regardless of its physical presence in single or double stranded DNA. Second, the propensity and persistence of the modified guanine residue in the displaced *syn* conformation, in the single stranded DNA context, provided structural support for the existence of a similar *in vivo* condition at the double-single stranded junction of the replication fork when the DNA polymerase pauses at the site of the dG-AAF lesion.

Molecular structural studies that address the premutagenic intermediate structures formed during mutagenesis can shed some light on the role of the structural determinants of these mutational pathways. A number of studies have addressed the structural and physicochemical properties of slipped mutagenic intermediate (SMI) structures within reiterated G's and/or alternating GC dinucleotides that contain a single dG-AAF modification, formation of which is required for frameshift mutagenesis (Streisinger and Owen, 1966; Streisinger et al., 1985; Lambert et al., 1992a and 1992b; Napolitano *et al.*, 1994 and 1997; see also section 1.4). The formation of a SMI is thermodynamically favored, a fact that has been confirmed by NMR studies, physicochemical stability measurements, and chemical probing of short heteroduplex DNA fragments bearing putative SMI structures (Garcia *et al.*, 1993; Milhe *et al.*, 1994). An AAF molecule covalently bound to a bulged guanine residue, in the context of three G residues opposite two C residues, in short (14-mer) double stranded heteroduplex DNA molecules, results in an increase of the melting temperature of the modified heteroduplex compared to either unmodified heteroduplex or modified homoduplex structures, suggesting that the

presence of dG-AAF enhances stability of the structure that resembles the SMI of the -1 frameshift mutagenesis pathway (Garcia *et al.*, 1993). The nuclear magnetic resonance spectroscopy of DNA heteroduplexes containing a bulged guanine residue, either unmodified or modified at dG-C8 with a covalently bound AAF, showed that the extra guanine residue without modification stacked into the helix while the major conformation of the AAF-modified guanine is the one with the AAF residue situated external to the helix with a higher stability compared to unmodified structure (Milhe *et al.*, 1994). The chemical probing of the cytosine residue across from dG-AAF in the homoduplex structure showed an increased reactivity of this residue, indicating a local adduct induced denaturation. However, no reactivity was detected in AAF-modified SMI in heteroduplex structures, indicating that the base pairing of the C residues with unmodified G residues located 5' and/or 3' to the AAF adduct were intact (Garcia *et al.*, 1993).

Heteroduplexes containing a *NarI* site (5'-G₁G₂CG₃CC-3') with an unmodified or AAF-modified G at the site of an extrahelical GC bulge resemble the slipped intermediates that might be formed during replication to produce spontaneous and induced -2 frameshifts, respectively. A NMR study of heteroduplexes containing a *NarI* site with a bulged GC dinucleotide in one strand, which was either unmodified or modified with AAF at the G₃ position, showed that all bases in the unmodified intermediate stacked within the helix without disruption of the B-like conformation (Milhe *et al.*, 1996). However, in the AAF modified heteroduplex the C residue in 5'-G₃_{AAF}C-3' bulged segment projected out and the AAF moiety was inserted within the helix (Milhe *et al.*, 1996). DNA melting studies showed that the modified slipped intermediate structure had a melting temperature equal to that of the unmodified

homoduplex of the same size (Koehl *et al.*, 1989a). Both the AAF modified heteroduplex containing a bulge, and the unmodified homoduplex, exhibited considerably higher stability than the unmodified heteroduplex, which had a 12°C lower melting temperature.

Both error-free replication through the dG-AAF adduct and replication resulting in frameshift mutations, require the insertion of a cytosine residue across from the modified guanine residue. Since, the *syn* dG-AAF adduct cannot base pair with cytosine and thus blocks replication, correct base pairing requires the dG-AAF residue to assume an *anti* conformation, at least transiently. The presence of multiple conformations has been shown through one and two dimensional NMR studies using a B-form 9-mer heteroduplex containing a dG-C8-AAF residue positioned opposite a C residue on the complementary strand. Among the multiple conformations, including the *anti* conformation, the presence of the bulky acetylaminofluorene moiety induced one as the predominant conformation (>70%) (O'Handley *et al.*, 1993 and 1992). In this major conformation, the modified G residue was displaced from the helix by assuming *syn* conformation and was located extrahelically in the major groove, reminiscent of base-displacement (Grunberger *et al.*, 1970) and insertion-denaturation models (Fuchs and Daune, 1972). One of the guanine amino protons was hydrogen bonded with a phosphate oxygen in the sugar-phosphate backbone, and the C residue across from dG-AAF was stacked within the duplex. The AAF moiety was inserted into the helix and stacked with the adjacent bases while the distal ring of the fluorene moiety protruded into minor groove, causing a flexible hinge-like local bending in the helix.

A variety of studies have shown that G→T and G→C transversions are frequently induced by dG-AAF and dG-AF in both eukaryotes and prokaryotes (Bichara and Fuchs,

1985; Carothers *et al.*, 1989 and 1993; Melchior *et al.*, 1994; Roy and Fuchs, 1994; Shibutani *et al.*, 1998). These observations emphasize the importance of the formation of mismatches between modified guanosine residue with dAMP and dGMP at the replication fork. The solution NMR structure of dG-C8-AF mismatched to a G residue (Abuaf *et al.*, 1995) (a structure that would lead to a G→C transversion), revealed a base displaced structure similar to that in which an A residue was situated across from the dG-C8-AF adduct (a structure causing G→T transversion) (Norman *et al.*, 1989). The modified G assumed a *syn* glycosidic alignment resulting in rotation of the AF ring into the B-DNA minor groove and hydrogen bond disruption at the modified base site while the rest of the base pairing remained intact (Abuaf *et al.*, 1995). The insertion of a T nucleotide across from the modified guanine residue results in a G→A base substitution (Bichara and Fuchs, 1985; Melchior *et al.*, 1994). Unfortunately, there is no reported study addressing the NMR structural properties of the T:dG-AAF mismatch. The intrinsic nature of DNA polymerases, the modulatory effect of the DNA sequence context, and the *in vivo* structural conformational polymorphism of the bulky DNA adduct at the replication fork are assumed to play important roles in determining the nature of translesion replication (Loechler, 1996). However, the nature of the specific molecular conformational determinants that influence the nature of the inserted mismatching base across from the modified guanine residue is not clear.

1.4 The effect of AAF modification on DNA conformation: The conformation of DNA in alternating purine-pyrimidine sequences may undergo an alteration that results in a left handed form of DNA called Z-DNA (Wang *et al.*, 1979; Arnott *et al.*, 1991). In this

conformation the glycosidic bond torsional angle of the guanine residues adopt the *syn* conformation, the torsional angles in the phosphodiester bonds change to alter the sugar-phosphate backbone, and the deoxyribose puckers alternate from C(3')-endo in purine residues to C(2')-endo in pyrimidine residues. B-DNA to Z-DNA transitions are favoured under high salt conditions, or in ethanol. In contrast to monotonic poly (dG:C) sequences, the early circular dichroism spectroscopic studies on alternating poly (dGC:CG) containing DNA sequences showed a marked inverted CD spectrum (Pohl and Jovin, 1972) , in high salt solutions or in ethanol. This change in the spectrum was attributed to the capability of the alternating copolymer to adopt the Z-conformation (Sage and Leng, 1981; Santella *et al.*, 1981a). However, an inverted CD spectral pattern was not sufficient to suggest a B to Z transition without disruption of base pairing, since the denatured poly (dGC:CG) was also able to show a clear inverted CD spectral pattern similar to double stranded copolymer (Narasimhan and Bryan, 1976). Nevertheless, the resistance of the poly (dGC:CG) double stranded copolymer to S1 nuclease digestion, in contrast to the marked susceptibility of the denatured form (Sage and Leng, 1981; Santella *et al.*, 1981b), was sufficient to conclude that the B to Z conformational transition does not lead to loss of local base pairing. However, such a conformational change allows for formation of distorted base pairing that result in a slight deviation from the Watson-Crick base pairing characteristic of B-form.

The chemical modification of guanine residues with AAF in a run of poly (dGC:CG) resulted in an inverted CD spectral pattern which, in a similar way to unmodified copolymer, showed complete resistance to S1 nuclease digestion (Santella *et al.*, 1981a). In addition, anti-cytidine antibody, which detects any unpaired cytidine

residues in single stranded regions of a DNA polymer, failed to react with the AAF modified poly (dGC:CG) in a radioimmunoassay (Santella *et al.*, 1981b). Together, the results of these studies, which were carried out in the absence of ethanol and/or high salt solutions, are consistent with the induction of Z-DNA conformation, suggesting that the presence of a bulky AAF moiety was sufficient to induce the Z-DNA conformation, and that it does not cause any denaturation of the DNA in Poly (dGC:CG).

In contrast to poly (dGC:CG), AAF modification of the guanine residues in either poly (dG:C) or in random DNA sequences does not result in any CD spectral inversion, but shows high sensitivity to S1 nuclease digestion and increased reactivity with anti-cytidine antibody. These observations suggest that the AAF modification does not induce any B to Z transition in these sequence contexts; however, AAF appears to cause a local denaturation of DNA (Grunberger *et al.*, 1970; Fuchs and Daune, 1972). The relevant biological significance of these structural transitions has been the focus of several lines of study since the mid 1980s, that will be addressed in the following sections.

1.5. Cellular strategies in response to DNA damage: DNA, as a highly dynamic molecule, is subjected to continuous alteration by its exposure to a multitude of factors from endogenous and/or exogenous sources. The nitrogenous bases that contain either unsaturated chemical bonds or a variety of reactive nucleophilic centres are more likely to be modified by their exposure to these effectors as compared to deoxyribose-phosphate backbone. Among the contributors to DNA damage from endogenous sources, oxidative damage, depurination and depyrimidination (reviewed in Marnett, 2000, Williams and Jeffrey, 2000 and Jackson and Loeb, 2001), cytosine and 5methyl-cytosine deamination

are considered as major sources of spontaneous mutations (Coulondre et al., 1978; Yebra and Bhagwat, 1995; Friedberg et al., 1995; Macintyre et al., 2001). However, spontaneous sequence alterations can also be achieved in the absence of any DNA modification by the instantaneous base tautomeric transitions that may arise during DNA synthesis (Douhal et al., 1995). In addition, DNA is under a continuous assault by a multitude of highly reactive chemicals from environmental sources that can either act as ultimate mutagens/carcinogens or be metabolically activated to reactive species capable of covalently modifying bases and forming chemical DNA adducts (see sections 1.2 and 1.3). The immediate effect of DNA lesions is experienced at the level of the loss of the coding capacity of the modified or lost bases to serve as template for DNA replication. As a result, the replicating polymerase may stall in front of non-instructional bases causing the replication fork to collapse (Moore and Strauss, 1979; Moore et al., 1981; Bessho et al., 1992; Woodside and Guengerich, 2002). In cells exposed to massive chemical DNA damage the effects of replicational arrest may be lethal.

From an evolutionary point of view, acquisition of several cellular mechanisms that assist removal of DNA damage, avoidance of damage, and synthesis through the blocking lesion, which might contribute to the rescue of the halted replication fork and increase survival at the possible expense of increased mutation rate, are of positive fitness importance. At least three cellular mechanisms are shared among prokaryotic and eukaryotic cells that can fulfill this need: DNA repair, damage avoidance mechanisms, and translesion synthesis. The following sections provide a brief introduction to bacterial cellular strategies which form a response to DNA damage.

1.5.1 DNA repair mechanisms: The most effective solution to deal with DNA damage is removal of the DNA lesion or misincorporated bases across from the undamaged bases as replication errors. Enzymes involved in different DNA repair pathways utilize double stranded DNA as their substrates and restore the normal sequence and structure of the damaged DNA by direct reversal of damage, complete removal of the damaged nucleotide, or removal of several nucleotides including the DNA adduct followed by repair synthesis through the gap introduced by such removal (Friedberg et al., 1995). The DNA repair mechanisms contribute to genome stability, and hence, in mammals, defective enzymatic components of many of these major repair pathways have been revealed in recent years to form the basis for several cancer predisposition syndromes (reviewed in Hoeijmakers, 2001).

1.5.1.1. Direct reversal of damage: Enzymatic photoreactivation (PR) of UV damaged DNA and removal of the alkyl group from bases adducted by alkylating agents are the classical examples of direct reversal mechanisms (reviewed in: Pegg and Byers, 1992; Pegg, 2000; Sancar, 2000; Margison and Santibanez-Koref, 2002). Exposure of DNA to ultraviolet radiation results in formation of cyclobutane dipyrimidine photodimers (CDP) between adjacent pyrimidines. The most prevalent isomers include *cis-syn* and *trans-syn* dimers that are the substrates for the photoreactivating enzyme, DNA photolyase. However, the minor photoproduct, the pyrimidine- pyrimidone (6-4) lesion, that is formed more frequently when the exposure intensity exceeds 500 J/m^2 , is not a substrate for the same enzyme. The prokaryotic *E.coli* DNA photolyase is a monomeric enzyme of molecular mass of about 50 kDa that uses FADH and 5,10-methenyltetrahydrofolate (5,10-MTHF) as co-factors. The enzyme is photoreactivated by absorption of a photon at

385 nm by the MTHF as the primary light harvesting chromophore and subsequent transfer of the acquired energy onto FADH, followed by capture of an electron and proton from a tryptophan residue in the apoenzyme. Transferring the electron to the cyclobutane in the pyrimidine dimer resolves the dimer into monomers by direct reversal of the photoadduct (Sancar G.B., 1990; Heelis et al. 1993; Sancar A, 1994).

The reversal of the O⁶-alkylguanine and O⁴-alkylthymine adducts is achieved through an inducible adaptive response by direct enzymatic removal of the alkyl group from these lesions and transfer of the alkyl moiety onto a cysteinyl residue on the enzyme (reviewed in Walker, 1985; Volkert, 1988; Linahl et al., 1988; Shevelle et al., 1990; Pegg and Byers, 1992, Pegg, 2000). Two different prototypic prokaryotic methyltransferases, O⁶-methylguanine-DNA methyltransferase I (O⁶-MGT I, encoded by the *ada* gene) and O⁶-methylthymine- DNA methyltransferase II (O⁶-MGT II, encoded by the *ogt* gene), are responsible for this function. While O⁶-MGT I shows preference for O⁶-metG, the O⁶-MGT II has a distinct preference for O⁴-metT lesions (Friedberg et al., 1995; Sassanfar et al., 1991). Both enzymes function by suicidal transfer of the alkyl group covalently onto a cysteine residue located on the C-terminal region of the enzyme.

Recently it has been shown that AlkB protein, which is also induced as part of the adaptive response is involved in demethylation of 1-methyladenine and 3-methylcytosine by a novel mechanism. AlkB belongs to the ferrous and α -ketoglutarate dependent dioxygenase superfamily that releases the methyl group from 1-metA and 3-metC by coupling the oxidative decarboxylation of the α -ketoglutarate to hydroxylation of the methylated bases (Trewick et al., 2002; Falnes et al., 2002).

1.5.1.2. Repair by removal of incorrect or damaged bases: DNA repair mechanisms that remove DNA lesions or mismatches include three repair systems shared by prokaryotic and eukaryotic cells that accomplish the removal of a damaged single nucleotide or stretch of nucleotides containing a damaged/mismatched base: base excision repair (BER), nucleotide excision repair (NER) and mismatch repair (reviewed in Walker, 1985; Van Houten, 1990; Sancar and Tang, 1993; Sancar, 1995 and 1996, Wood, 1996; Lindahl et al, 1997; Bouziane et al., 1998).

1.5.1.2.1. Base excision repair: The base excision repair (BER) functions recognize a wide range of substrates including misincorporated uracil residues, 8-hydroxyguanine, 3-methyladenine, 3- and 7-methylguanine, 5-methylcytosine and hypoxanthine residues, and G.T, G.A and 8-oxo-G.A mispairs. The preliminary starting step is recognition of these modified bases or mismatched bases by substrate specific glycosylases that catalyze the hydrolysis of the N-glycosidic bond between the deoxyribose moiety and the damaged base to produce an AP site. The phosphodeoxyribose residue at the AP site is excised from the sugar-phosphate backbone by the sequential actions of an AP-endonuclease and a deoxyribophosphodiesterase (dRpase) that introduce phosphodiester cleavage at the 5' and 3' sides of the AP site, respectively. The resultant one nucleotide gap is repaired by fill-in repair synthesis followed by ligation.

1.5.1.2.2. Nucleotide excision repair: In contrast to BER, nucleotide excision repair (NER) recognizes and repairs DNA lesions that originate primarily from exogenous sources. A multitude of covalent chemical bulky adducts including nitro-substituted aromatic compounds and their arylamine and arylamide derivatives, CDP and 6-4 pyrimidine-pyrimidone photoproducts, bulky alkylation base lesions, non-bulky AP sites

and thymine glycols, and DNA inter- and intra-strand cross linking lesions such as mitomycin bifunctional adducts, nitrogen mustard (HN₂) and *cis*- and *trans*-platinum adducts, are substrates for NER in prokaryotic and eukaryotic cells (Van Houten, 1990; Sancar, 1995 and 1996; Prakash and Prakash, 2000). These DNA lesions usually result in topological alterations in DNA such as helix kinks and bending that are recognized by the NER enzymes. However, there are DNA lesions, such as AP sites, thymine glycol and O⁶-MetG adducts, that do not distort DNA but are substrates for NER. In addition, intercalating agents, single nucleotide mismatches, single nucleotide bulges and larger loops, introduce DNA structural distortions but are not substrates for NER. In prokaryotes the enzymatic accessory includes UvrABC, UvrD (helicase II), DNA polymerase I and DNA ligase (reviewed in Sancar, 1996). The excision repair carry out genome wide surveillance and hence this repair pathway is called global nucleotide excision repair to contrast it against transcription coupled excision repair (TCR), which involves the recruitment of the UvrABC components to a stalled RNA polymerase at lesion site by a match making protein called Mfd, the mutation frequency decline factor (reviewed in Selby and Sancar, 1993, Freidberg, 1995 and 1996, and Selby et al., 1991). UvrA dimerizes *in vivo* and complexes with UvrB to form a heterotrimeric UvrA₂B complex that has a limited DNA dependent ATPase activity. This latter complex scans the DNA by a tracking mechanism (Grossman and Thiagalingam, 1993) and recognizes the damaged bases (Batty and Wood, 2000). Upon binding, the protein/DNA complex undergoes a conformational change that results in the UvrA₂ component being released from the complex while the more stable UvrB-DNA complex introduces unwinding and bending in the local DNA. UvrC is then recruited to the damage site by interaction with

UvrB. Upon binding of UvrC to UvrB-DNA complex the cryptic endonuclease activity of UvrB is unmasked and dual 3' and 5' incisions are introduced to the strand that carries the damaged base. The 3' end incision is introduced by UvrB at 4-5 base pairs downstream from damaged base and the 5' incision is done by the endonuclease activity of UvrC at the 8th base upstream from the damage site, resulting in release of a 12-13 base oligomer. By the action of UvrD (helicase II) the UvrC-oligomer complex is released and the 12-13 nucleotide gap is filled in by Pol I followed by ligation that seals the nick (reviewed in Sancar, 1996).

1.5.1.2.3. Nucleotide mismatch repair: Mismatch repair (MMR) is considered as a post replicational genome surveillance mechanism (reviewed in: Modrich, 1989 and 1991; Modrich and Lahue, 1996; Marti et al., 2002). Prokaryotic and eukaryotic mismatch repair functions can recognize single base substitutions and extrahelical bulges of one to three nucleotides, and correct them on a newly synthesized strand. In general, the affinity of mismatch repair enzymes toward mismatched bases follows the order G:T and A:C > G:G and A:A > T:T and T:C > C:C (Su et al. 1988). Bacterial MMR is devoid of C:C mismatch repair and in eukaryotes this mismatch is repaired very inefficiently (Su et al., 1988; Holmes et al., 1990). Mismatches formed across from chemically altered bases are not recognized efficiently by these functions; hence, repair of such mismatched bases is generally poor. Recognition and removal of the mismatched bases from the nascent strand is achieved by a methylation dependent mechanism. Strand discrimination in methyl directed mismatch repair is based on recognition of the state of methylation of the A residues in 5'-GATC-3' sites. The A residues on the template strand are methylated at C-7 by the action of *Dam* (DNA- adenosine methylase) while the newly synthesized

strand is transiently devoid of methylation. MMR functions recognize the mismatched base and carry out incisions on the newly synthesized strand at a phosphodiester bond 5' to 5'-GATC-3' on either side of the mismatched base. In *E.coli* the protein responsible for the initial recognition of the MMR substrates is MutS, a protein with intrinsic ATPase activity (Junop et al., 2001). The presence of the mismatched bases is sensed by a conserved motif composed of GXFY in domain I of MutS (Obmolova et al., 2000; Lamers et al., 2000). Binding of an ATP to MutS causes a conformational transition in MutS that results in its dissociation from the mismatched bases and its conversion to a DNA sliding clamp that initiates the downstream repair process (Gradia et al., 1999; Junop et al., 2001). MutS translocates on DNA in an ATP dependent manner (Blackwell et al., 1998; Junop et al., 2001), pulling the DNA in two directions inward to form an α -loop until it reaches the methylation site (GATC) from both directions. The 70 kDa MutL protein binds to the MutS/hemimethylated heteroduplex DNA in the presence of ATP. The resulting ternary complex serves as a molecular dock for recruiting and activating MutH, a protein with latent endonuclease activity (Yang, 2000; Schofield et al., 2001). The latent endonuclease active site located at the interface of two subdomains of the MutH (Ban and Yang, 1998) is unmasked via physical interaction with MutL (Ban and Yang, 1998; Hall and Matson, 1999; Spampinato and Modrich, 2000). MutH introduces a nick at the 5' phosphodiester bond penultimate to GATC and leaves the phosphate group on the G residue. The resultant 5'-P-GATC and dNP-OH at the nick site on daughter strand are substrates for the *mutU* (*uvrD*) encoded helicase II that unwinds the DNA in an ATP dependent manner and exposes the single stranded DNA to exonucleases (Ali and Lohman, 1997; Mechanic et al., 2000). The exonuclease VII

(product of *xse*⁺) and RecJ (product of *recJ*) carry out the 5'→3' exonucleolytic degradation of the nascent strand and the 3'→5' exonuclease activity is carried out by ExoI (product of the *sbc*⁺ gene), and possibly by ExoVII and a recently isolated ExoX (Viswanathan and Lovet, 1999; Viswanathan et al., 2001). As a result of the combined actions of MutU helicase and exonucleases a gap of 100 to over 1000 bases is formed. The gap is filled in by DNA polymerase III in the presence of SSB and the covalent continuity of the repaired daughter strand is achieved by the action of ligase.

1.5.2. Damage avoidance and recombinational bypass: The fact that DNA damage stimulates homologous recombination in bacteria indicates that recombination may have evolved for repair purposes that provide the organism with short and long term survival capability (Mudgett et al., 1991). The genetic requirement for recombinational repair includes not only functions used for recombination but also those involved in replication (Marians, 2000). Stalling the replisome in front of a noncoding DNA damage (e.g a bulky chemical lesion) (Moore and Strauss, 1979; Moore et al., 1981) most often results in reinitiation of replication downstream from the stalling point, leaving behind a single stranded gap on the template strand, i.e. the daughter strand gap (Rupp and Howard-Flanders, 1968). Formation of the non-extendable terminus serves as a checkpoint where the communication between the replicational and recombinational functions occur (Kogoma, 1997; Marians, 2000). In *E.coli* recombinational bypass of DNA damage and repair of daughter strand gaps are achieved by the RecF pathway (Fig. 1.3).

RecF pathway requires an undamaged homologous double strand copy of the DNA and functions through three phases: I) a presynaptic phase during which the DNA is prepared for homology search and synapsis when homologous pairing and strand

Fig 1.3: RecF recombinational bypass pathway. A) formation of daughter strand gap and reinitiation of replication, B) Presynapsis and synapsis, C) Extension, D) Holliday junction resolution. See text for explanations.

exchange occurs between the damaged duplex and its homologous sister duplex, resulting in formation of a Holliday junction; II) DNA replication restart; III) a postsynaptic phase during which the Holliday junctions are resolved (Kuzminov, 1999).

The RecF pathway, also called the post replication DNA repair, requires RecF, RecO, RecR, RecA, SSB, RuvABC or RecG, Topoisomerase I and II, DNA polymerase I and DNA ligase. It is noteworthy that the RecF pathway leaves the chemical damage untouched, hence this pathway can be regarded as a damage avoidance/tolerance pathway rather than a repair pathway.

The working model for the RecF pathway assumes that during the presynaptic phase the RecFOR and RecA mediated molecular switch converts the collapsed replication fork into a recombination initiation intermediate (Clark and Sandler, 1994). The complex formed between RecO and RecR (RecOR) directs and stabilizes RecA onto the SSB nucleoprotein and aids RecA in displacing the SSB to form the RecA/ssDNA nucleoprotein filament in the 5'→3' direction (Shan et al., 1997; Cox, 2000). The RecFR complex prevents the RecA filament from extending into the double stranded regions (Webb et al., 1997). In the synaptic phase the RecA nucleoprotein filament searches for those sequences that are homologous to the gap region on the sister duplex and, by pairing the gap sequence with the homologous double stranded region, forms a transitional triplex. Further melting the triplex results in displacement of the complementary strand to which the gap sequence is annealed. This leaves the intact strand homologous to the gap sequence available for annealing by the 5' and 3' free ends of the daughter strands at the boundaries of the gap. The synapsis is facilitated by the action of two topoisomerases: DNA gyrase (Topo II) relieves the positive supercoils in

the sister duplex and Topo I relieves the negative supercoils in the newly generated duplexes after strand invasion. The gap is filled in by DNA polymerase I, which catalyzes strand extension from the 3' terminus of the invading strand, and the nick is sealed by DNA ligase. In the postsynaptic phase the Holiday junction is resolved by the action of the RuvABC or RecG helicase while the damaged base is retained on the daughter strand (West, 1996; Kowalczykowski, 1991 and 2000).

1.5.3. SOS induction and Translesion Replication: It is nearly thirty years since Miroslav Radman suggested that bacteria have a back up strategy to cope with the environmental stress that results in DNA damage (e.g. UV irradiation) (Radman M., 1975). He hypothesized that the SOS response that is induced following DNA damage is also a RecA and LexA dependent mutagenic process. Despite the initial conservatism in the scientific community, this working model has increasingly gained support (reviewed in Radman M., 1975, d'Ari R, 1985, Walker et al., 1985, Sedgwick SG, 1986, Witkin EM, 1989, Battista et al., 1990, Murli and Walker, 1993, Walker GC, 1995, Friedberg et al., 1995, Woodgate and Levine, 1996, Shinagawa H., 1996, Smith and Walker, 1998, Sutton et al., 2000, and Pham et al., 2001; Goodman, 2002; Gonzalez and Woodgate, 2002). The SOS response is now known to result from a concerted and finely tuned temporal expression of over 30 pleiotropic genes (Fernandez De Henestrosa et al., 2000; Courcelle et al., 2001) scattered throughout the bacterial genome, that together comprise a regulon. The SOS regulated genes are activated in response to inhibition of DNA synthesis (Sassanfar and Roberts, 1990). The early SOS dependent responses are centred around mechanisms that are virtually error-free. The organism attempts to either repair a DNA lesion that hinders replication (section 1.5.1), or tolerate it by switching

from replication to recombinational bypass at the site of a stalled replisome, followed by resumption of replication (section 1.5.2). However, the mutagenic SOS response is not elicited immediately; rather it occurs with a delay of about 50 minutes post damage when the error-free mechanisms prove inadequate for repair of massive DNA damage (Sommer et al., 1998; Rangarajan et al., 1999; Rangarajan et al., 2002). This prolonged lag period before ensuing activation of mutagenic pathways, i.e. translesion synthesis bypass (TLS), is controlled directly by inhibition of the TLS (Yasuda et al., 1998 and 2001) and by a primitive form of cell cycle control mechanism that by itself is SOS regulated, and functions by inhibition of the cell division and DNA synthesis (Holland and Jones, 1985; Autret et al., 1997; Opperman et al., 1999; Murli et al., 2000; Sutton and Walker, 2001). Ultimately, if the severity of the DNA damage exceeds the capacity of TLS functions and the replication of the chromosome cannot be rescued then some cryptic cellular elements are activated that result in cell lysis and death (Ebina et al., 1983; Lu and Chak, 1996). This might be regarded as “bacterial apoptosis” (Brunskill and Bayles, 1996; Wang et al., 2000; Lewis K, 2000; Grundling et al., 2001; Bayles KW, 2003).

1.5.3.1 Molecular organization of SOS regulon and regulation of SOS responses:

The immediate effect of physical and chemical damage to DNA is manifested in hindrance of the replication machinery (Khidhir et al., 1985). In a wild type *E.coli* strain growing in rich nutritional conditions a rapid replication halt lasts only for about ten minutes, and is followed by slow origin-independent replication restart (Sommer et al., 1998; Rangarajan et al., 1999; Rangarajan et al., 2002). The origin independent replication restart leaves numerous single stranded gaps, each of 1000-2000 nucleotides

in length all over the chromosome between the replisome stalling site and the new replication restart point. These transiently exposed chromosomal single stranded regions are believed to be the inducing signal for SOS responses (Sassanfar and Roberts, 1990; Higashitani et al., 1995). It has long been known that the LexA protein, the universal transcriptional repressor of the SOS regulon, acts as the stress signal transducer to the downstream genes in the regulon (Mount et al., 1972; Brent and Ptashne, 1981; Little, 1984). Under normal physiological conditions, the homodimeric LexA protein (22 kD/monomer) stays bound to the *lex* boxes of genes in SOS regulon and prevents the transcription of its own gene and more than 30 other pleiotropic genes scattered on the chromosome (Fernandez De Henestrosa, 2000; Courcelle et al., 2001). The *lexA* palindromic consensus 5'TACTG(TA)₅CAGTA3' sequence can be found in a nucleotide range of +1 to -330 from the start codon among the known SOS genes (Fernandez De Henestrosa, 2000). The degree of relative divergence from the consensus sequence, designated by heterology index (HI), and the number of the *lex* boxes for genes of the SOS regulon appear to determine the extent to which the gene is repressed by the LexA repressor (Fernandez De Henestrosa, 2000; Courcelle et al., 2001). The lower the HI value for a given gene, the less divergent it is from the consensus, and hence the higher would be the LexA affinity (Schnarr et al., 1991; Lewis et al, 1994; Robinson et al., 1998; Fernandez De Henestrosa et al., 2000). In addition to transcriptional regulatory mechanisms, the functional presence of the basal and induced intracellular pools of SOS proteins are determined by their turn over rates, the rate of posttranslational processing, and intimate protein-protein interactions among some of these proteins. The fine tuning of SOS responses results in a concerted temporal activation of cellular functions that

result in the rescue of DNA replication, and commencement of normal cell division.

1.5.3.1.1 Temporal activation of SOS functions follow a logical order: The single stranded patches of DNA formed following DNA damage are immediately covered by single stranded DNA binding protein (SSB) monomers. The SSB covered regions are also the sites for nucleation of RecA protein. Under normal conditions the cytoplasmic RecA protein pool is about 10,000 copies per cell. This amount is sufficient for its oligomerization on single stranded DNA in the 5' to 3' direction while replacing the SSB. Nucleation of the RecA on ssDNA starts by formation of contacts between each RecA monomer and three backbone sugar-phosphate units resulting in a right handed helical nucleoprotein filament from which bases are protruded into the deep inner groove of the helix. Interaction between the RecA and ssDNA causes a conformational transition of the RecA protein to its co-proteolytically active form commonly known as RecA* (Little, 1991). The inner helical groove of the RecA nucleoprotein forms a scaffold where protein-protein interaction between the RecA and LexA protein takes place (Freitag and McEntee, 1989; Yu and Egelman, 1993). Contacts between the C-terminal residues of LexA and RecA* favours a conformational transition of the LexA protein and activation of its otherwise cryptic serine autoproteolytic function (Luo et al., 2001). Mechanistically, the serine protease activity of LexA is unmasked by lowering the pKa value of the Lys156 residue from 10 to 6 (Roland et al. 1992), its interaction with Ser119 and the spatial proximity of both Lys156 and Ser119 residues with the peptide bond between Ala84 and Gly85 in LexA. The Lys156 captures a proton from Ser119 leading to formation of an oxyanion nucleophilic centre on Ser119 that is eventually used to attack the carbonyl group of the peptide linkage between A84 and G85 (Little, 1991 and 1993). As a result,

an 84 amino acid N-terminal fragment is cleaved off the LexA protein. This autoproteolytic digestion prevents the dimerization of the LexA protein and reduces the active intracellular LexA pool. Decreased cytoplasmic LexA pool results in shifting the equilibrium between the free and DNA bound LexA toward unbound form by releasing the repressor from the *lexA* boxes of genes that have particularly high HI values. Interestingly, genes with high HI values are those that code for proteins that are believed to be activated immediately: PolB, RecN, RuvAB and DinB (Pol IV) (Table 1.1). Derepression of *polB* and *ruvAB* genes result in 7 and 8-fold increase in their protein levels, respectively (Qiu and Goodman, 1997), and the derepression of *recN* and *dinB* results in 10-fold increase in β -galactosidase activity in gene fusion experiments (Table 1.1). While PolB is required for origin independent assembly and restart of the otherwise disintegrated replication fork after DNA damage, the DinB (Pol IV) might be helpful in reading through chemically modified or unmodified repetitive sequences resulting in targeted and untargeted mutations (Kim et al., 1997 and 2001). The *recN* and *ruvAB* gene products are involved in recombinational damage avoidance pathways. The activation of replication restart and recombinational damage avoidance pathways help to restore the continuity of the disrupted DNA strands without necessarily removing the DNA lesions.

The second group of genes that are derepressed almost immediately after replication restart and activation of recombinational pathways are those of the nucleotide excision repair pathway, and regulatory proteins that prevent mutagenic pathways. The target genes are *uvrA*, *uvrB*, *uvrD* and *dinI* with HI values in the range of 6.1-8.8 (Table 1.1). It is noteworthy that at this stage the *lexA* gene itself would be completely derepressed (Table 1.1). As a result of such derepression the cytoplasmic levels of UvrA, UvrB,

<u>Gene</u>	<u>Function</u>	<u>Heterology index*</u>	<u>Protein copy per cell **</u>		<u>Increased fusion**</u>
			<u>Basal level</u>	<u>Induced level</u>	<u>B.gal activity</u>
<i>Regulation</i>					
lexA	repression/derepression	6.34, 8.32	1300	1	6
recA	derepression	4.31	1000-10000	100000	12
<i>Recombination and Replication restart</i>					
polB	replication restart	12.09	30-50	210-350	7
ruvAB	Holliday junction resolvase	9.19	700/200	5600/1600	3
recN	Recombination	5.16, 9.38, 11.47	?	?	10
<i>Repair</i>					
uvrA	Nucleotide Excision repair	6.98	20	250	5
uvrB	Nucleotide Excision repair	6.11	250	1000	4
uvrD	Nucleotide Excision repair	8.8	5000-8000	25000-65000	6
<i>Cell cycle and mutagenesis</i>					
dinB	TLS	12.84	250	~2500	~10
sfIA	Cell division inhibition	4.65	?	?	125
dinI	RecA TLS function inhibitor	6.24	<500	2300	?
umuD/umuC	TLS, RecA recomb. inhibition	2.77	180/10	2400/200	22/?

Table 1.1: Hierarchical induction of cellular SOS functions in response to DNA damage. Recombination, replication restart and repair functions are likely to be activated first. Cell cycle inhibition and TLS functions are the last cellular resorts to be activated.

* See "Fernandez De Henestrosa, 2000"

** See "Kuzminov A., 1999"

UvrD and DinI proteins are elevated in the range of 4 to 12-fold (Table 1.1). The UvrABD together with UvrC are components of nucleotide excision repair pathway that is involved in error-free removal of a broad spectrum of DNA lesions (see also section 1.5.1.2.2). DinI is a small protein of 9 kD that serves as antirecombination and antimutagenic agent. It is involved in inhibition of the coprotease and recombination activities of RecA protein by binding to the inner groove of the RecA nucleoprotein filament (Yasuda et al, 1998). Such binding does not disturb the structure of the filament but competitively prevents the access of the LexA and UmuD to the filament (see below for proteolytic activation of UmuD) (Yasuda et al., 2001). However, it is noteworthy that competitive inhibition of the coprotease activity of the RecA* is more directed toward UmuD rather LexA, since it was shown that the degree of LexA processing in the *dinI* null mutant, and wild type strains of *E.coli* are similar. These observations were explained by the differential affinities of LexA, DinI and UmuD toward the activated RecA filament in the order of LexA>DinI>UmuD (Yasuda et al., 1998 and 2001). Intermediate affinity of the DinI to RecA* is consistent with its role as a transient inhibitor of UmuD processing; this would result in inhibition of the mutagenic pathways while the decline in the cellular LexA pool continues. The antirecombinogenic function of the DinI protein was evident by sharp decline of the frequency of recombination mediated by P1 phage transduction that resulted in about 100-fold reduction of recombination frequency (Yasuda et al., 1998). The biological importance of the DinI mediated negative regulatory circuit is in the preferential allocation of the cellular resources into the repair of DNA lesions that have already been bypassed by recombinational mechanisms, at the same time avoiding the costly consequences of

premature full activation of mutagenic mechanisms by minimizing the intracellular pools of the mutagenic proteins i.e. UmuD', the proteolytic product of UmuD. If the remaining DNA lesions can be removed effectively at this stage and both strands of the chromosome made ready for successive rounds of replication, then the cell cycle can be restored.

If the extent of the DNA damage is so great that fully derepressed recombinational and repair mechanisms are unable to rescue the replication, then the SOS signal persists and propagates down to the third group of genes. The *recA* and *sfiA* genes are likely to be derepressed fully at this stage with HI values of 4.31 and 4.65, respectively (Table 1.1). Derepression of the *sfiA* gene results in accumulation of the SfiA protein, which interacts with the FtsZ protein, an essential cell division protein that forms cytokinetic tubulin like rings involved in septum formation, and inhibits its polymerization by interaction with GTPase domain of FtsZ (Begg et al., 1995; Huisman and D'Ari, 1981; Mukherjee et al., 1998). Cell division inhibition provides the cell with additional time to continue repairing the DNA lesions. Derepression of the *recA* gene results in an over 10-fold increase of intracellular RecA protein that may be used for binding to the lingering ssDNA, depletion of the remaining LexA pool, and titrating out the DinI pool.

Complete depletion of the LexA pool results in derepression of the most tightly regulated genes in the SOS regulon e.g. the *umuDC* operon (HI of 2.77). Expression of the *umuDC* operon causes the intracellular level of UmuD and UmuC to rise by 10- to 12-fold (Woodgate and Ennis, 1991). Titrating out the DinI by the presence of excess RecA relieves the inhibitory effect of DinI on the access of UmuD to the RecA nucleofilament. As a result of derepression of the *umuDC* operon the intracellular level of the UmuD increases and then predominates at about 20 minutes post SOS induction.

UmuD is able to dimerize to UmuD₂ and form the UmuD₂C heterotrimeric complex by binding to UmuC. The UmuD is posttranslationally processed to its mutagenically active form, UmuD', by RecA facilitated autoproteolytic cleavage of the N-terminal 24 amino acids, which occurs with similar kinetics to those observed for proteolytic cleavage of LexA (Burckhardt et al., 1988; Shinagawa et al., 1988; Nohmi et al., 1988; MacDonald et al., 1998). Occupation of the inner groove of the RecA nucleofilament by both DinI, UmuD and its derivatives continues to deter the RecA from its recombinogenic function. UmuD' dimerizes and binds to UmuC to form a heterotrimeric complex of UmuD'₂C (Woodgate et al., 1989). Formation of this complex provides the basis for the onset of mutagenic TLS at around 50-60 minutes post SOS induction. It is important to note that the induction of the *umuDC* operon does not immediately upregulate the mutagenic replication; rather, transition from more accurate mechanisms to full error prone pathway progresses gradually. A negatively operating circuitry, orchestrated by the *umuDC* operon, controls such a gradual transition by a primitive form of damage check point and cell cycle control (Operman et al., 1999; Ferentz et al., 2001). The timing of the post translational processing of the UmuD to UmuD' appears to play the central role as a molecular switch in this process. At least three regulatory phenomena are believed to contribute to the fine tuning of timing of cell cycle arrest followed by transition to full TLS. These are discussed more fully in the following paragraphs.

1.5.3.1.1.1 DNA damage dose dependent time lag for posttranslational processing

of UmuD: At about 15-20 minutes post stress the cytoplasmic UmuD pool starts rising. However, the level of the active UmuD' pool, the RecA* mediated autoproteolytic product of UmuD, only peaks at 20-30 minutes after UmuD reaches its maximum level

under moderate level of DNA damage e.g. 20 J/m² UV dose . The delay between these two time points has been reported to be dependent on the severity of the DNA damage and becomes prolonged, reaching up to 60 minutes post stress at 50 J/m² UV dose (Opperman et al., 1999). The time lag between derepression of the *umuDC* operon and commencement of full mutagenic TLS commitment provides the cell with additional time to use its repair capacity to remove as much DNA damage as possible.

1.5.3.1.1.2 Protein-protein interactions and formation of higher order protein

complexes: UmuD₂, UmuDD' and UmuD'₂, the homo and heterodimeric derivatives of UmuD and UmuD', are present in equilibrium as the damage signal persists (Battista et al., 1990b; Jonczyk and Nowicka.,1996). The equilibrium is gradually shifted in favour of the UmuD'₂, the only mutagenically active species, at the later stages of the operon activation, while the UmuD₂ predominates only at the early stages. The temporal isolation of these species has at least two major biological consequences. First, the distinct physiological role of each species is exerted at different time windows when they are most needed. Second, the fine tuning of the effective intracellular concentrations of each species becomes possible as a result of their differential proteolytic turnover (see below).

It has been shown that both UmuD₂ and UmuD'₂ are able to make protein-protein contacts with the α , β and ϵ subunits of Pol III (Tadmore et al., 1992; Sutton et al., 1999; Sutton et al., 2001a,b,c). Despite the fact that the relative affinities of UmuD₂ and UmuD'₂ toward the Pol III subunits are different, only when such protein complexes are formed with UmuD₂ they are able to block the progression of the Pol III and bring the DNA synthesis to halt. This might be viewed as a primitive form of DNA damage check

point that results in temporary cell cycle cessation, providing the cell with an extended damage repair time interval before the mutagenic functions kick in.

Besides the role of these protein complexes in cell cycle control, it is also known that the mutagenic role of UmuD'₂C can only be facilitated by formation of protein complexes with RecA-ssDNA, SSB, and Pol III subunits including, the γ clamp loader and the β sliding clamp (Tang et al., 1998 and 1999; Sutton et al., 2001a,b,c). Interaction of UmuD'₂C with the 3' tip of RecA-ssDNA not only functions as a directional chaperone to target the UmuD'₂C to the DNA damage site (Frank et al., 1993; Frank et al., 2000), but also serves as a switch for TLS by inhibition of the recombinogenic function of RecA (Sommer et al., 1993; Boudsocq et al., 1997). Positioning the UmuD'₂C at the lesion site provides it with proximity to the subunits of the stalled Pol III. UmuD and UmuD' compete against the α subunit of Pol III to bind to the same face of the β clamp (Sutton et al., 2001) that can result in prevention of reassembly of the Pol III at the DNA lesion site and installation of the UmuD'₂C instead. In addition, the β clamp provides the UmuD'₂C with mechanistic tethering to the DNA and enables it to start synthesis of DNA past the lesion.

1.5.3.1.1.3 Differential autoproteolytic turnover of Umu proteins: It also appears that the differential proteolytic turnover of the UmuD and UmuD' plays a role in regulating the transition timing from error free strategies to full mutagenic TLS (Gonzalez and Woodgate, 2002). It has been shown that UmuD₂, UmuC and their heterotrimeric complex, the UmuD₂C, are substrates for Lon protease in *E.coli* and *Salmonella*, while UmuD' and its derivatives are not (Frank et al., 1996a; Gonzalez et al., 1998). However, UmuD' can be degraded by another protease called ClpXP, but only when it is present in

its heterodimeric form with UmuD i.e. UmuDD'. The biological significance of the differential proteolytic turnover of the Umu proteins can be viewed as follows. As time passes the relative concentration of the UmuD' increases and the UmuD is gradually pulled out of UmuD- β/ϵ complexes, which block DNA synthesis, to form UmuDD' species that have neither mutagenic nor cell cycle control properties, but are subject to proteolytic degradation by ClpXP protease (Frank et al., 1996a; Gonzalez et al., 2000). This trend provides the cell with yet another opportunity for smooth transition from cell cycle cessation mode to slow mutagenic DNA synthesis recovery mode. It appears that the collective action of the cellular proteases keep the Umu protein concentrations low, which avoids excessive mutagenic background (Gonzalez et al., 2000).

Activation of the TLS pathway is considered as final cellular strategy for handling DNA damage and rescuing the otherwise collapsed replication forks. However, by being forced into the TLS pathway, cells may increase their survival at the cost of an increased frequency of mutation.

The collective error-free and error-prone pathways allow DNA damage to be repaired or bypassed, with or without mutation. This restores the continuity of the DNA, and allows resumption of normal replication. Eventually, single stranded regions that originally served as the SOS inducing signal disappear and the RecA*-ssDNA nucleofilaments become less abundant as DNA loses the bound RecA protein. This results in accumulation of the LexA protein dimers that can restore the repressed state of the genes in the SOS regulon.

1.5.3.2 Y-family and other error prone polymerases: Over the past three decades, genetic analyses have indicated that wild type *E.coli* strains exposed to various mutagens

exhibit significant level of enhanced mutagenesis while *umuDC* mutants do not (Witkin, 1976; Kato and Shinoura, 1977; Steinborn, 1978; Walker and Dobson, 1979). Together the mutagenesis data has suggested that the *umuDC* gene products play a crucial role in DNA translesion synthesis in bacteria. Experiments with UV or ethyl methanesulfonate treated temperature sensitive mutants of DnaE i.e. the α catalytic subunit of the major *E.coli* replicating polymerase, Pol III, showed no significant increase in induced mutagenesis when cells were grown at restrictive temperature (43°C). The situation seemed reversible when cells were returned to non restrictive temperature, suggesting a direct role for Pol III in mutagenic TLS (Bridges et al., 1976; Hagensee et al, 1978). Delayed photoreversal experiments in *umuC36* mutant *E.coli* cells, which resulted in increased histidine revertants, were interpreted as additional evidence for involvement of Pol III in TLS past UV photodimers (Bridges and Woodgate, 1984). It was also reported that several DNA polymerases, including Pol III, could not synthesize past UV and acetylaminofluorene treated DNA at the lesions *in vitro* (Moore et al., 1981). These observations, together with the notion that UmuDC had a role in enhancement of the UV induced mutagenesis (Bridges and Woodgate, 1985b) and increased Weigle reactivation of the λ phage (Walker and Dobson, 1979), prompted Bridges and Woodgate (1985a) to propose the “two-step” model of mutagenesis. According to this model Pol III was involved in the “misincorporation” step of a base across from a TT dimer, and a subsequent “bypass” step is “facilitated” by the UmuDC proteins (Bridges and Woodgate, 1985a). This mutagenesis model was revisited following subsequent observations that *dnaQ* mutants, with a defective ϵ proof reading subunit, showed enhanced mutability for some TLS types in *umuDC* background (Woodgate et al, 1987; Vandewiele et al., 1998;

Fuchs and Napolitano, 1998). These observations led to a mechanistic refinement of the model by attributing an auxiliary role of UmuDC to down regulation of the proof reading activity of the ϵ subunit during SOS induced TLS. Although, Pol III remained as the major mutagenic polymerase in this model, the ill-defined role of the UmuDC proteins and their interactions with ϵ subunit needed more intensive biochemical support. This required expression and purification of UmuDC proteins and subsequent testing using an *in vitro* TLS assay.

Numerous attempts were made to express and purify the UmuDC proteins with the aim of reconstituting translesion bypass *in vitro*. However, difficulties in expression and purification of UmuC in particular, limited the amount of functional protein that was obtained using an inefficient procedure that required multiple denaturation/renaturation cycles (Woodgate et al., 1989). However, even with limited quantities, the UmuD'₂C complex was used in an *in vitro* TLS assay past an AP site in a reaction that included SSB, RecA, β , γ and core Pol III* (Rajagopalan et al., 1992). Bypass was observed in the presence, but not in the absence of UmuD'₂C; however, no control reactions were carried out to exclude Pol III* from the reaction mixture. This was a historical mistake that reflected the apriori reliance of researchers on the “two step” model described previously. These results, at best, could only verify the previous assignment of an auxiliary role for UmuDC. Improved UmuDC purification was achieved by Hsp70/60 treatment of cell lysates containing UmuC (Petit et al., 1994), and subsequent use of classical multi-step column chromatographical procedures to purify the functional UmuD'C complex (Bruck et al., 1996). Eventually, in 1998, Tang *et al* used this latter method to isolate relatively high purity preparations of the whole UmuD'C complex in sufficient quantities to be used

in an *in vitro* TLS assay (Tang et al., 1998). It was shown that the UmuD'C was able to copy DNA on normal and AP containing templates in the presence of SSB, RecA, γ and β subunits, in the presence and absence of Pol III. In an independent study Livneh and coworkers expressed and purified the maltose binding protein (MBP) linked UmuC and UmuD' fusion proteins --interestingly, with much less trouble-- and used them in a translesion synthesis assay (Reuven et al., 1998). The TLS assay included the expressed recombinant UmuC and UmuD' proteins, RecA, SSB and Pol III holoenzyme. Since their experiments were based on the "two step" model mentioned above, Livneh *et al.* did not carry out a control experiment in the absence of Pol III; however, they were able to show increased bypass past AP site in the presence of UmuD'C functions as compared to limited bypass by Pol III alone. In a follow up experiment, subsequent to the publication of the results of Tang *et al.* (1998), they confirmed the *bona fide* polymerase functions associated with the MBP-fused UmuD'C proteins (Reuven et al., 1999). The UmuD'₂C complex is now referred to as polymerase V (Pol V). Subsequently the polymerase function and TLS of Umu-like proteins that are expressed on naturally occurring plasmids has been successfully tested for the purified MucA' and MucB proteins, the known enterobacterial UmuDC homologs on pKM101 plasmid. The MucA'₂B complex has been called the Pol RI (Goldsmith et al., 2000). It is noteworthy that in both experiments a limited capacity of Pol III to bypass synthetic AP sites was observed; however, it was enhanced only when UmuD'₂C and other components of the reaction were added. TLS had an absolute requirement for activated RecA* and was further stimulated in the presence of β / γ and SSB proteins. The real challenge to the "two-step" model came from the observation that in the absence of any cellular polymerases, both MucA'₂B and

UmuD'₂C were able to carry out DNA synthesis in a distributive manner and with a clearly reduced fidelity, an observation that undermined the major mutagenic role of Pol III that had dominated the field for almost 25 years (see discussion for current views on the model). It is noteworthy that the idea of characterizing these proteins for their polymerase activities did not develop through a systematic scientific approach; rather, it was based on serendipity and inferences made from other works not related to any of the Umu or Muc proteins (see below).

In addition to UmuC purification problems, research into the role of UmuD'₂C and its plasmid encoded homologs through the study of eukaryotic orthologs had been hampered, at least in part, by the limited homology that these proteins have with their eukaryotic orthologs and specifically with the known polymerases of the A-D polymerase families. As a result, in database searches, there was a strong possibility that any putative eukaryotic ortholog would have been considered insignificant. The only documented observation was made by Kulaeva *et al.* who noticed a few insignificant sequence alignments in the human EST database against umuC/dinB, which did not result in immediate serious consideration (Kulaeva *et al.*, 1996). However, the discovery of the template directed deoxycytidyl transferase activity of yeast Rev1 (Nelson *et al.*, 1996a) and polymerase activity of the cloned yeast Rad30 protein (McDonald *et al.*, 1997), and demonstration of their TLS capabilities and their partial homologies with prokaryotic UmuC and DinB proteins, resulted in a revolutionary and fundamental change in the way that the research community as a whole regarded these proteins. The idea that these proteins and their plasmid-encoded homologs might themselves be *bona fide* polymerases directed the research in a totally different and a new avenue. As a result the time period

between 1998 to 2002 witnessed the exciting discovery and characterization of new gram negative prokaryotic and eukaryotic polymerases. This trend has continued between 2002 and now, albeit with a slower pace, to expand the list to include newer members of error prone polymerase in gram positive bacteria (see reviews: Woodgate, 1999; Goodman and Tippin, 2000a, Goodman and Tippin, 2000b; Kuntz et al., 2000; Sutton et al., 2000; Sutton and Walker, 2001; Friedberg, 2001a; Friedberg et al., 2001b; McDonald et al., 2001; Pham et al., 2001; Wang, 2001; Boudsocq et al., 2002; Goodman, 2002; Gonzalez and Woodgate, 2002; Bruk et al., 2003; Christmann et al., 2003; Rattray et al., 2003; Tippin et al., 2004; Le Chatelier et al., 2004). The list includes two prokaryotic polymerases DinB (Pol IV) and Pol V (UmuD'C), and several eukaryotic polymerases: Yeast Rev1; the yeast Pol η (Rad30), a DinB ortholog; human and mouse homologs of Rad30 (Rad30A); Polk (DINB1), another human and mouse DinB ortholog; and a DinB ortholog called Pol ι (Rad30B). Together, these proteins comprise a unique superfamily of error prone polymerases (Nelson et al., 1996a; McDonald et al., 1997; Johnson et al., 1999a; McDonald et al., 1999; Gerlach et al., 1999; Ogi et al., 1999; Tissier et al., 2000b). A recent phylogenetic analysis has revealed that the members of this superfamily include over 52 proteins that are present in all domains of life from bacteria to human; therefore, by consensus they were categorized under a new name, the Y-family of error-prone polymerases and subcategorized under UmuC/DinB/Rev1/Rad30, each typifying a subfamily (Ohmori et al., 2001). It is important to note that the error prone polymerases are not limited to Y-family and the list has been expanded to include the prokaryotic B-family member Pol II (*polB*) (Becherel and Fuchs, 2001; Shen et al., 2002), gram positive C-family members DnaE polymerases of *S.pyogenes* and *B.subtilis* (Bruk et al., 2003; Le

Chtelier et al., 2004), and DnaE2 of *M.tuberculosis* (Boshoff et al., 2003), eukaryotic B-family member Pol ξ (Yuan et al., 2000; Johnson et al., 2000a; Rechkoblit et al., 2002), X-family members Pol β (Servant et al., 2002; Bassett et al., 2002), Pol μ (Aoufouchi et al., 2000; Duvauchelle et al., 2002; Zhang et al., 2001a; Zhang et al., 2002a; Havener et al., 2003), and Pol λ (Aoufouchi et al., 2000).

1.5.3.3 TLS properties of error prone polymerases: The hallmark properties of the Y-family of polymerases are not only their lack of significant sequence homology with other known classical polymerases from the A, B, C and D families, but also their biochemical properties including low fidelity in nucleotide selection, lack of intrinsic 3'→5' exonuclease proof reading activity and low processivity (Goodman and Tippin, 2000b). These characteristics provide them with the capacity to copy the undamaged and damaged bases in an error prone manner resulting in untargeted and targeted mutations. The *in vitro* DNA replication properties of the prototypic members of Y-family polymerases, together with the B- and X-family members of error prone polymerases, have been compiled in Table 1.2. Interestingly, the TLS inferences from these *in vitro* studies are in agreement with the available *in vivo* data. These polymerases show a range of bypass properties when dealing with a specific type of DNA lesion including error free bypass, erroneous base misinsertion and in some cases base frame shifting. The relative efficiency of extension from the misincorporated nucleotides (expressed by calculated f_{ext} values) compared to correct base insertion is in the range of 30-fold to 10^{-4} (data not shown, see references in Table 1.2). Error prone polymerases accept a wide range of structurally diverse substrates from undamaged DNA to damaged bases, including AP sites, chemically cross linked bases, UV photodimers and chemically modified bulky base

Table 1.2: *In Vitro* translesion synthesis properties of prokaryotic and eukaryotic error prone DNA polymerases

	Gene	Family	DNA Lesion	Base insertion	Mutation	References
Prokaryotic						
DinB (Pol IV)	E.coli dinB	Y	Undamaged	C/slippage	-1	Wagner et al., 1999
			AP	A>G	AP dependent	Tang et al., 2000
			dG,AF and dG,AAF	C/slippage	1-2 nt deletions	Suzuki et al., 2001
			(+)-trans-B[a]P-N2-dG	C>>T not G,A	Error free>>G→A	Shen et al., 2002
			(-)-trans-B[a]P-N2-dG	C>>T>G	Error free>>G→A>G→C	"
UmuD'C (Pol V)	E.coli umuDC	Y	(+)-trans-B[a]P-N6-dA	T>C>A>G	Error free>>A→G>A→T>A→C	"
			(-)-trans-B[a]P-N6-dA	T>C>A>G	Error free>>A→G>A→T>A→C	"
			<i>cis-syn</i> TT CPD	A across 3T>G	Error free>T→C	Reuven et al.,1999; Tang et al., 2000
			<i>cis-syn</i> TT CPD	A across 5T>G	Error free>T→C	Tang et al., 2000
			(6-4)TT	A across 5T>G	Error free>T→C	"
Pol η	Yeast Rad30 and Human Rad30A (XPV)	Y	(6-4)TT	G across 3T>A	T→C>Error free	Reuven et al., 1999; Tang et al., 2000
			AP	G>A>T>C	AP dependent	Tang et al.,1998 and Reuven et al.,1998
			(+)-trans-B[a]P-N2-dG	A>G>C not T	G→T,C>G→A>Error free	Shen et al.,2002
			(-)-trans-B[a]P-N2-dG	A>G>C>T	G→T>G→C>Error free>G→A	"
			(+)-trans-B[a]P-N6-dA	T>A>G>C	Error free>A→T>A>C,G	"
Pol η	Yeast Rad30 and Human Rad30A (XPV)	Y	(-)-trans-B[a]P-N6-dA	T>A>G not C	Error free>A→T>A→C	"
			<i>cis-syn</i> TT CPD	AA	Error free	Johnson et al.,2000b
			<i>cis-syn</i> TT CPD	3T: A>>T>C=G	Error free>>T→A>T→G,C	"
			<i>cis-syn</i> TT CPD	5T: A>>T,C,G	Error free>>T→A,G,C	"
			(6-4)TT	G across 3T	T→C	Zhang et al., 2000d
Pol η	Yeast Rad30 and Human Rad30A (XPV)	Y	G,AAF	C	Error free	Vaisman et al., 2000
			GpG[cisplatin]	C across 3'G	Error free	Masutani et al., 2000
			GpG[cisplatin]	A across 5'G	G→T	Zhang et al., 2000d; Yuan et al., 2000
			8-Oxoguanine	C>A>G	Error free>G→T,C	Yuan et al., 2000; Masutani et al., 2000
			AP	G>A	AP dependent	Zhang et al., 2000d
Pol η	Yeast Rad30 and Human Rad30A (XPV)	Y	5T-AP	A	-1	Zhang et al., 2000d
			B[a]P-N2-dG	A>T>G>C	G→T>G→A>G→C	Zhang et al., 2000d; Chiapperino et al., 2002
			(+)-trans-B[a]P-N2-dG	A>G>T,C	G→T,C>A>Error free	Rechtkoblit et al.,2002
			(-)-trans-B[a]P-N2-dG	A>G,T>C	G→T>C>A>Error free	"
			(+)-trans-B[a]P-N6-dA	T>A>G,C	A→T	"
Pol η	Yeast Rad30 and Human Rad30A (XPV)	Y	(-)-trans-B[a]P-N6-dA	A>T,C,G	A→T,C	"

Table 1.2: *In Vitro* translesion synthesis properties of prokaryotic and eukaryotic error prone DNA polymerases (continued 2)

Gene		Family	DNA Lesion	Base insertion	Mutation	References
Eukaryotic						
Pol ϵ	Human Rad30B	Y	8-Oxoguanine	C	Error free	Zhang et al., 2001b
			G.AAF	C	Error free	Zhang et al., 2001b
Pol κ	Human DinB1	Y	AP	G>T>A>C	AP dependent	Zhang et al., 2001b; Johnson et al., 2000a
			(6-4)TT	A across 3'T	Error free	Zhang et al., 2001b; Johnson et al., 2000a; Tissier et al., 2001a
			<i>cis-syn</i> TT CPD	T across 3'T	T→A	Zhang et al., 2001b; Tissier et al., 2001a
			(+)-trans-B[a]P-N2-dG	G	G→C	Rechtkoblitt et al., 2002
			(-)-trans-B[a]P-N2-dG	A, G	G→C, T	"
			B[a]P-N6-dA	None	T→C	"
			Undamaged	G across T	T→C	Zhang et al., 2000b
			Untargeted	All 12 mipairing	All kinds, more T→G	Tissier et al., 2000b
			Untargeted		1-3 base deletions	Ohashi et al., 2000a; Zhang et al., 2000c
			(6-4)TT	None	-	Ohashi et al., 2000a
Rev1	Yeast Rev1	Y	<i>cis-syn</i> TT CPD	None	-	Zhang et al., 2000a; Ohashi et al., 2000b
			GpG[displatin]	None	-	"
			8-Oxoguanine	None	-	"
			AP	A	G→T	Zhang et al., 2000a
			5'T-AP	A	AP dependent	Ohashi et al., 2000b; Zhang et al., 2000a
			G.AAF	A	-1	"
			(+/-)-trans-B[a]P-N2-dG	T>C>G	G→A, C	"
			(+/-)-trans-B[a]P-N6-dA	C	Error free	Zhang et al., 2000a; Rechtkoblitt et al., 2002
			dG.AF	None	Error free=G→A=G→T	Rechtkoblitt et al., 2002
			dG.AAF	A=T=C	G→A>>G→T and G→C	Suzuki et al., 2001
Pol η	Human Pol η	Y	8-Oxoguanine	C>>T>G>A	Error free>>G→A>G→C>G→T	Zhang et al., 2002b
			(+/-)-trans-B[a]P-N2-dG	C>>G>T>A	Error free>>G→C>G→A>G→T	"
			1, N6-ethenoadenine	C>>T>G>A	Error free>>G→A>G→C>G→T	"
			G.AAF	C	Error free	"
			AP	C>>T>G>A	AP dependent	Zhang et al., 2002b; Nelson et al., 1996
			(6-4)TT	None		Zhang et al., 2002b
			<i>cis-syn</i> TT CPD	None		"

Table 1.2: *In Vitro* translesion synthesis properties of prokaryotic and eukaryotic error prone DNA polymerases (continued 3)

Other Pols					
Gene	Family	DNA Lesion	Base insertion	Mutation	References
Pol ξ Needs Pol I or Pol η for extension	Yeast Rev3/Rev7	B	AP	G>A	Yuan et al., 2000; Johnson et al., 2000a
			(6-4)TT dG.AAF	C C	
Pol II	E.coli polIB	B	B[a]P-N2-dG B[a]P-N6-dA	None None	Rechtkoblitt et al., 2002 Rechtkoblitt et al., 2002
			dG.AAF (NarG3 adduct) (+)-trans-B[a]P-N2-dG (-)-trans-B[a]P-N2-dG (+)-trans-B[a]P-N6-dA (-)-trans-B[a]P-N6-dA	C G>>C not A,T G>A>T>C T>>A>G not C T>>G>A not C	
Pol β	Human Pol beta	X	(6-4)TT <i>cis-syn</i> TT CPD GpG[cisplatin]	A across 3'T>C>G A across 3'T>C>G	Servant et al., 2002 Bassett et al., 2002
Pol μ	Human Pol mu	X	Semi targeted past CCCdG.AAF Untargeted at AA,GG,TT,CC repeats	G>T,C	Duvauchelle et al., 2002 Zhang et al., 2001a
			AP 1,N6-ethenoadenine (+/-)-trans-B[a]P-N2-dG dG.AAF (6-4)TT	Context dependent Context dependent Context dependent Context dependent None AA	Zhang et al., 2002a " " " " " "
			<i>cis-syn</i> TT CPD GpG[cisplatin]	Error free 2-4 nt deletions	
					Havener et al., 2003

adducts. In contrast to their considerably higher error rates as compared to other high fidelity polymerases (10- to 100-fold), some (e.g. Pol η and Dpo4) are able to bypass some DNA lesions (e.g. TT cyclobutane dimers) that normally block classical polymerases with stunning accuracy and efficiency (Johnson et al, 1999a; Ling et al., 2001).

1.5.3.3.1 Pol IV, Pol κ , Pol ι and Pol μ are involved in untargeted mutagenesis: The prokaryotic Pol IV (*dinB*) gene that was originally identified as part of the SOS regulon (Kenyon and Walker, 1980), has been shown to be involved in untargeted mutagenesis in the λ_{c1857} gene (Brotcorne-Lannoye and Maenhaut-Michel, 1986). An increase of about 800-fold in frameshift mutation in homopolymeric runs of guanine residues in the *lacZ* gene was observed when the DinB protein was overexpressed from a plasmid (Kim et al., 1997). The *dinB* gene, initially cloned by Ohmori et al. (1995), was re-cloned by Wagner et al. (1999), expressed and purified, and used in an *in vitro* mutagenesis assay. Pol IV showed template dependent, distributive primer extension past a slipped misaligned primer/template substrate that confirmed the involvement of Pol IV in untargeted frameshift mutagenesis.

Human Pol IV ortholog, the Pol κ , and Pol ι and Pol μ are the only known eukaryotic error prone polymerases that exhibit high mutation rates when copying undamaged DNA with distinct mutagenic outcomes. Lethal deletions of 1-3 nucleotides were observed when reiterated undamaged DNA sequences were copied by Pol κ , and frequent -1 deletions were detected when homodimeric AA, TT, GG and CC template sequences were copied by Pol μ . In addition, Pol κ mediates all 12 types of base misinsertions, more prevalently C insertion across from T; while Pol ι misincorporates G across from T residues resulting in T $\rightarrow$ C transitions (Table 1.2). Although, error prone

polymerase mediated occurrence of base substitutions and the distinctive functional differences among these polymerases might be considered as contributing factors to genetic diversification, the biological significance of frameshifts and the regulatory mechanisms that control the untargeted mutagenesis are not well understood.

1.5.3.3.2 TLS past bulky chemical substitutions: Pol IV dependent induction of *in vivo* frameshift mutations by 4-nitroquinoline-N-oxide (NQO) indicates the ability of Pol IV in error prone synthesis past the NQO DNA lesions (Kim et al., 2001). The (+)-trans-B[a]P-N²-dG enantiomer is the major biologically relevant adduct formed by the trans ring opening of benzo[a]pyrene-7,8-diol 9,10-epoxide. Involvement of either Pol V or Pol IV in the induction of (+)-trans-B[a]P-N²-dG induced -1 frameshift *in vivo* in the *Sma*I site (5'CCCGGG_{BaP}3') by insertion of a cytosine nucleotide across from the B[a]P modified guanine residue and its subsequent extension from a slipped misaligned primer/template mutagenic intermediate has been suggested by the Fuchs lab (Napolitano et al., 2000). However, it was not clear, in the *polB* background used for these experiments, whether Pol IV or Pol V alone, or in combination, are responsible for the C nucleotide insertion or extension, or both. Nonetheless, the *in vitro* studies by Shen et al. (2002) showed that Pol IV, but not Pol V, could exclusively insert C residues across (+)-trans-B[a]P-N²-dG adduct in a template that did not contain reiterated guanine residues 5' to the lesion site, resulting in error free bypass of the B[a]P adducts. The requirement for correct C insertion as the prerequisite step in the slippage model, and the congruent *in vitro* data for preference of C insertion exclusively by Pol IV is consistent with a major role for Pol IV, but not Pol V, in the B[a]P induced -1 frameshift mutagenesis in reiterated sequences observed *in vivo* (see Table 1.2) (Napolitano et al., 2000). This

represents an example of how the *in vitro* data can resolve inconclusive interpretation of the *in vivo* data. The fact that Pol IV is able to insert a correct base across from B[a]P guanine adduct is clear evidence that this adduct can be bypassed by at least one of the prokaryotic Y-family polymerases in an error free manner (i.e. Pol IV); however, frameshifted mutagenic bypass is conditional on the existence of a reiterated sequence in the vicinity of the adduct. In a similar manner among the eukaryotic polymerases only Pol κ is able to bypass this adduct without error (Table 1.2); hence, Pol κ has been postulated to protect mammalian cells against the lethal and mutagenic effects of B[a]P (Ogi et al., 2002). However, in contrast, both enzymes are highly inaccurate in synthesizing undamaged sequences (see above and Table 1.2).

During *in vitro* TLS of bulky dG-AF and dG-AAF adducts, Pol IV and its human ortholog (Pol κ) are blocked one base before these adducts (Suzuki et al., 2001). Correct C insertion across from the modified bases was observed only when higher amounts of the enzymes were used; mutagenic bypass yielded primarily -1 and -2 frameshift mutations in the case of Pol IV. Pol κ was able to incorporate C, A and T nucleotides across from dG-AF adduct, causing error free and mutagenic G→A and G→T bypasses with almost equal frequencies (Table 1.2). Synthesis past dG-AAF by Pol κ was achieved primarily by T insertion across from the modified base, resulting in G→A transitions (Suzuki et al., 2001). Involvement of Pol V in *in vivo* mutagenic and nonmutagenic TLS past the dG-AAF has been documented; Pol V appears to catalyse the extension from a correct C insertion step against the AAF adduct. There is no *in vitro* report for Pol V at the moment. While Pol V and Pol IV are involved in mutagenic -1 frameshift mutations in reiterated guanine sequence contexts, Pol II is considered as the main mutagenic

polymerase that yields -2 frameshifts at AAF modified heterodimeric CG repeats and in *NarI* site (5'-GGCGCC-3') (Napolitano et al., 2000). This property of Pol II has also been observed *in vitro* (Becherel and Fuchs, 2001). It appears that the frameshift mutations induced by AF and AAF occur primarily in repeated sequence contexts and that, regardless of involvement of either Pol V and Pol II or any other polymerase, insertion of a correct C across from the AAF bulky DNA lesions is a necessary step in TLS. A similar trend can be seen in most of the eukaryotic polymerases i.e. Pol η , Pol ι , Pol ξ /Rev1; nonetheless, the final outcome might be modulated by the sequence context in which the adduct is located (Table 1.2). However, handling of these bulky adducts by Pol μ , and to a certain extent by Pol κ is most probably mutagenic. Together, these observations are consistent with the notion that both prokaryotic and eukaryotic cells have evolved multiple polymerases to deal with bulky DNA lesions like B[a]P, AF and AAF, which are also potent mutagens and carcinogens. These polymerases share the following properties: i) they are tolerant of the geometrical constraints imposed by these adducts; ii) they are capable of preferential accurate base insertion and synthesis past these adducts, however, they exhibit a higher rate of mutagenic bypass when compared to high fidelity enzymes; iii) error free TLS is modulated by the sequence context of the target DNA; and iv) they show DNA lesion dependent selective fidelity. The sequence and DNA lesion dependent specificity of incorporation by error prone polymerases might be interpreted as an indication of some sort of specialization among these enzymes that enables the cells to bypass specific types of DNA adducts, avoid their lethal consequences, and contribute to survival.

1.5.3.3.3 TLS past UV induced DNA lesions: Bacterial Pol V and eukaryotic Pol η , Pol ι , Pol β and Pol μ all are able to copy the *cis-syn* cyclobutane TT (*cis-syn* TT CPD) and (6-4)TT dimers (Table 1.2). In contrast, Pol κ , Rev1 and Pol II stall in front of these adducts and cannot synthesize past them. *In vitro* TLS past a CPD lesion by Pol η , Pol β and Pol μ , as the most inaccurate enzyme, proceeds primarily by the accurate insertion of two correct adenine nucleotides; bases other than adenine are inserted at much lower frequency ($\sim 10^{-3}$) (Table 1.2). Pol V and Pol ι also bypass CPD lesions by insertion of A nucleotides but also synthesize past these adducts by insertion of incorrect G or T residues at frequencies comparable to that observed *in vivo* (Lawrence et al., 1990; Smith et al., 1996; Zhang et al., 2001b; Tissier et al., 2001a; Reuven et al., 1999; Tang et al., 2000). However, TLS past (6-4)TT adducts is considerably more error prone compared to the CPD TT dimer, a fact that is also consistent with *in vivo* data (LeClerc et al., 1991; Smith et al., 1996). The biological significance of the *in vitro* data become immediately apparent if one considers that human Pol η , which is involved in accurate synthesis past TT dimers, is the defective enzyme among xeroderma pigmentosum variant (XP-V) patients (Masutani et al., 1999; Johnson et al., 1999b). These patients are proficient for NER but show increased sun light sensitivity, UV hypermutability, and are predisposed to skin cancer (Cleaver et al., 1973). The requirement of Pol η for accurate and efficient bypass of the TT dimers *in vivo* has been verified in XPV cells by restoration of UV resistance following trans expression of Pol η in these cells (Yamada et al., 2000; Stary et al., 2003).

1.5.3.3.4 TLS past abasic sites: Abasic sites (AP) arise due to spontaneous loss of bases, by glycolyase dependent excision of damaged bases (see section 1.5.1.2.1), or by

excision of hydrolytically deaminated cytosine bases that serve as an impediment to replication (Friedberg et al., 1995). Pol IV and its eukaryotic ortholog, Pol κ , read through abasic sites with the preferential insertion of an A residue that can also result in a -1 frameshift when the base 5' to the AP site is a T (Table 1.2). Pol V, Pol η , Pol ι and Pol ξ all carry out distributive synthesis past AP sites, and show a clear preference for G insertion over A across from the AP sites. Unlike other polymerases Rev1 inserts a C residue against the AP sites, creating a terminus that is believed to be extended by the Pol ξ (Zhang et al., 2002b). The distinctive preference of G insertion over A by Pol V and at least three out of five error prone eukaryotic polymerases might be an evolutionarily conserved characteristic of these polymerases if one considers that the majority of the AP sites originate from spontaneous hydrolytic deamination of cytosine residues. Preferential insertion of a G residue across from AP sites increases the probability that the correct information will be restored at sites of damage.

The unusual properties of error prone enzymes have been best explained by their crystallographic structural properties (Silvian et al., 2001; Ling et al., 2001; Zhou et al., 2001; Trnko et al., 2002; Washington et al., 2003; Ling et al., 2004). The available crystal structures for a number of these polymerases reveal that the previously identified five commonly conserved sequence motifs (Kulaeva et al., 1996) are arranged in palm, finger and thumb domains similar to classical polymerases with an additional domain identified as "little finger" involved in making contacts with the backbone of the template strand (Ling et al., 2003). However, the Y-polymerases lack the so called "O" and "O₁" helix loops that, in high fidelity polymerases, protrude from the fingers domain and make contact with the flat surface of the perfectly paired nascent and template bases. Such a

contact is required as a fidelity check, and to promote the “induced fit” closed conformation of the polymerase. It might be concluded that since the Y-family polymerases make minimal contact with the paired or mispaired bases, they exhibit less geometric selectivity for incoming nucleotides. Therefore, the relatively relaxed conformation of the active site binding pocket not only allows accommodation of mispaired bases but also, in extreme cases like UV photoadducts, accommodates two template TT-dimer bases in the active site and inserts two AA bases opposite them, resulting in their accurate bypass by polymerases such as Pol η and Dpo4 (Ling et al., 2001; Friedberg et al., 2001a; Boudsocq et al., 2002; Kokoska et al., 2002).

1.6 AAF and AF induced *in vivo* mutagenesis: Earlier work using the *Salmonella* reversion assay to assess the mutagenic potential of a group of N-hydroxyarylamines including N-hydroxy-2-aminofluorene revealed that these compounds are frameshift mutagens. It was also shown that the aromatic amines, aminofluorene (AF) and 2-acetylaminofluorene (AAF), bind covalently to the C(8) position of guanine to form the major DNA adduct (Beland *et al.*, 1983; Beranek *et al.*, 1982). Analysis of the AF- and AAF-induced *in vivo* forward mutational spectra shows that the majority of AAF induced mutations are frameshifts. In contrast, base substitutions comprise about 50% of the AF-induced mutations of which 85% are G→T transversions (Koffel-Schwartz *et al.*, 1984; Bichara and Fuchs, 1985; Lambert *et al.*, 1992b). Moreover, the fact that the damage distribution profile of AAF does not correlate with its mutational spectrum is consistent with the targeted nature of the induced

mutations to reiterated guanine sequences and alternating GC sequences as the main hot spots for frameshifts (Fuchs, 1984). This notion has been supported by forward mutation assay using tetracycline gene of plamid pBR322 where more than 90% of mutations induced by AAF were frameshifts that occurred as -1 frameshifts in runs of G or -2 mutations at *NarI* site (Fuchs *et al.*, 1981; Koffel-Schwartz *et al.*, 1984). A study of forward mutations induced by AAF in the *lacI* gene of *E.coli* (Schaaper *et al.*, 1990) showed a similar spectrum to that induced by AAF in the tetracycline resistance gene (Koffel-Schwartz *et al.*, 1984). In addition, a positional effect of the AAF modified guanine residue has been shown with single adduct mutagenesis studies using double stranded vectors carrying an AAF modified guanine residue at either of G1, G2 and or G3 positions of the *NarI* site. This study revealed that -2 frameshifts are exclusively induced by G3.AAF adduct but not G1.AAF and G2.AAF (Burnouf *et al.*, 1989). A similar positional effect has been shown at reiterated 5'-G₁G₂G₃-3' sequence where -G frameshift frequency is maximal when AAF is carried on G3 residue (Lambert *et al.*, 1992).

In addition to frameshifts, the forward mutational spectra induced by dG-AAF and dG-AF in prokaryotes and eukaryotes using double and single stranded vectors also include frequent G→T and G→C transversions, and to a lesser extent G→A transitions. Such base substitution mutations that arise by insertion of a mismatching base opposite the dG-AF or dG-AAF residue, are induced to a greater extent by AF as compared to AAF. The mutational spectrum induced by AF in the tetracycline resistance gene of pBR322 in *E.coli* included the G→T and G→C transversions, which represented ~70%

and <6% of total base substitutions, respectively. In the same target gene and cells, AAF treated pBR322 plasmid DNA produced G→T and G→C transversions which comprised the ~30% and ~20% of total base substitutions, respectively (Bichara and Fuchs, 1985; Melchior *et al.*, 1994). Grunberger and coworkers (Carothers *et al.*, 1989 and 1993) showed that G→T transversions (75% of base substitutions) and G→A transitions (10% of base substitutions) are induced by AAF and N-hydroxy-aminofluorene in the HPRT gene of Chinese hamster ovary cells. Moreover, induction of G→T transversion as the major base substitution mutation, and minor G→A and G→C mutations, were shown in simian kidney cells transformed with a single stranded shuttle vector carrying a single dG-C8-AAF adduct (Shibutani *et al.*, 1998). In *Saccharomyces cerevisiae*, the AAF treated plasmid DNA carrying the URA3 gene produced a forward mutational spectrum that included the G→C and G→T transversions, which represented 33% and 26% of total base substitutions, respectively (Roy and Fuchs, 1994).

The bulky nature of the AF, and especially the AAF adducts cause an impediment for progression of replication. This has been further verified by *in vitro* studies, using AAF and AF adducts and different polymerases, including Pol III, in which it has been shown that replisome stalls 3' to adduct (Belguise *et al.*, 1994; Lindsley and Fuchs, 1994). Induced *in vivo* mutational spectra emphasize the importance of insertion of a correct or a mismatched base against the AF and AAF modified guanosine residues in a polymerase dependent manner as the required preliminary step for synthesis past these adducts. This obviously would require cellular functions other than Pol III to alleviate the geometrical constraints

that both the modified base and the incoming mismatching base impose on the replicational machinery.

The genetic data regarding the cellular functions required for AAF induced mutations suggest that the nature of the proteins involved depends on target sequence and the mutational outcome. Induction of -2 frameshift mutations in AAF treated plasmids containing GC repeats was studied in *E.coli* strains, under SOS induced or uninduced conditions, which carried different alleles of *recA* and *lexA* or were defective for *umuC* function (Maenhaut-Michel *et al.*, 1992). In a *umuC* background, an almost 50-fold elevation in -GC mutation was observed by SOS induction, implying the UmuDC independent nature of AAF induced -GC mutations in GC repeats, which was also consistent with previous observations (Koffle-Schwartz *et al.*, 1984; Burnouf and Fuchs, 1985). An elevated level of -GC mutation frequency, comparable to the SOS induced wild type strain, was observed in a *lexA*(Def) background indicating that the expression of some of the *LexA* repressed functions of the SOS regulon were required for the -2 frameshift pathway. However, it was also shown that the -2 mutation was maximally induced in a *recA730* background, in which the coprotease function of RecA is constitutively expressed, but not in *recA99*, a null mutant allele, or in *lexA⁻ recA430* (defective for coprotease activity) backgrounds. These latter observations imply that the coprotease activity of RecA, i.e. RecA*, is required for induction of SOS functions other than UmuDC, that are required for the -2 pathway. All together these observations are consistent with the following properties for the -2 mutational pathway: i) it is umuDC independent, ii) it requires an SOS inducible function, iii) RecA is required only for induction of the SOS regulon. However, AAF induced -1 frameshifts that were studied

using double stranded vectors carrying monomodified G in homopolymeric runs of guanine residues were not only SOS dependent but also required RecA and UmuDC functions (Koffel-Schwartz *et al.*, 1984; Lambert *et al.*, 1992).

1.7 Scope and objectives of this study: At the time that the proposal for this study was submitted no information regarding the superfamily of error prone polymerases was available and involvement of the UmuDC and similar proteins in the mutagenic processes was explained by their auxiliary roles in the framework of the “two step model” described in section 1.5.3.2. Information regarding the mutagenicity of AAF in *E.coli* had been obtained from studies that had used either random modification of chromosomal or plasmid DNA, or single adduct strategies; in either case the DNA target was in the double stranded form. Our working model regarding the cellular responses to replication stalled at the site of a dG-AAF adduct was based on the assumption that *in vivo* bypass of such a lesion would require a replication dependent translesional synthesis mechanism. The work described in this thesis examined this model using a single stranded phagemid vector bearing a single site specific dG-AAF residue. This experimental set up was designed to completely alleviate any interference by cellular repair or recombinational mechanisms, since the outcome should be solely dependent on the capacity of the cells to replicate past the adduct. Thus, the initial objective was to optimize the design, and methodology, for construction of single stranded vectors that carry a site specifically located single AAF adduct, and test whether these vectors were suitable for the mechanistic studies of induced mutations *in vivo*. Second, we hypothesized that TLS past dG-AAF adducts, regardless of the position of modified guanosine residue and the sequence context surrounding adduct would be dependent on the SOS functions. Thus,

we used the single stranded vectors to examine mutagenesis in different cellular backgrounds. Third, the principal AAF induced mutations in single stranded vectors were expected to be frameshifts in reiterated sequences, as described earlier with double stranded vectors. However, we also wished to examine the relative contribution of error-free bypass, and misincorporation events which would yield base substitutions, to total TLS. These latter considerations were not part of the design of earlier experiments. Fourth, we hypothesized that the UmuDC and MucAB functions can contribute to TLS events and would serve to elevate survival through their differential effect on mutagenic and non mutagenic bypass events. Fifth, we wished to examine, in single stranded vectors, whether the type and the relative abundance of mutations can be influenced by the position of adduct and the sequence context within which the adduct is carried. Sixth, on the basis of earlier data, we expected that -2 mutation in *NarI* site would be dependent on an SOS inducible function other than UmuDC and MucAB. The role of these functions in the -2 mutation pathway can be reliably investigated using the experimental setup proposed for this study.

2. Materials and methods:

2.1 Materials:

2.1.1 Media, bacterial strains and maintenance: SOS induced and/or uninduced

E.coli strains JM103 (F' *traD36 lacI^q ΔlacZM15 proA⁺B⁺/ endA1 supE sbcBC thi-1 rpsL Δ(lac pro) (P1) (r_k⁺m_k⁺ r_{P1}⁺m_{P1}⁺) [Yanish-Perron *et al.*, 1985] , JM103Δ*umuC* , JM103Δ*umuC*(*pGW1700*), JM103Δ*umuC*(*pDC322*) and JM103 F⁻ strains were used for mutagenesis studies. *E.coli* DH5α (F⁻/ *endA1 hsd17 (r_k⁺m_k⁺) supE44 thi-1 recA1 gyrA (Nal^r) relA1 Δ(lacIZYA-argF) U169 deoR φ80dlacΔ(lacZ)M15*) (Woodcock *et al.*, 1989) and *E.coli* JV30, Worthington Biochemical Corporation, were used for maintenance, propagation of all plasmids or monomeric double stranded phagemids, and preparation of single stranded phagemids, respectively. Bacterial cultures of the JM103 series were grown in Luria Bertani broth or solid media supplemented with the appropriate reagent or antibiotic at the following concentrations: IPTG (60 μg/ml), X-gal (30 μg/ml), ampicillin (150 μg/ml), streptomycin (50 μg/ml), rifampicin (50 μg/ml) and tetracycline (6.25 μg/ml), as required by the given protocol. Stock cultures of the strains were maintained on LB plates supplemented with the proper antibiotic for daily use, and were subcultured on the same type of plates on a monthly basis. The long term stock cultures were maintained in 10% glycerol at -80°C, which were renewed every 2 years.*

2.1.2 Enzymes and kits: T4 DNA ligase, T4 polynucleotide kinase, all restriction enzymes, ddNTP's and dNTP's were purchased from New England Biolabs. Pwo DNA polymerase PCR kit was supplied by Boehringer Mannheim. Sequenase version 2.0 was purchased from Amersham Life Sciences. IPTG and X-gal were supplied by Diagnostic

Chemicals Ltd. and Research Organics Inc., respectively. All antibiotics were purchased from Sigma Co. $\gamma^{32}\text{P}$ -ATP (6000 Ci/mmol) and $\alpha^{35}\text{S}$ -dATP (3000 Ci/mmol) were obtained from Amersham Life Sciences. All other chemicals were purchased from BDH.

2.1.3 DNA vectors and manipulations: Phagemid pUC118, purchased from Worthington Biochemical Corporation, was used for construction of single stranded vectors containing a cloned oligomer, with and without a single acetylaminofluorene (AAF) modified guanosine residue. Plasmid pBR322 was used to clone the *umuDC* operon (pDC322). Plasmid pGW1700, a gift from Dr. Graham Walker [Perry and Walker, 1982], is a pBR322 derivative carrying a cloned chromosomal fragment that contains the *mucAB* operon. Supercoiled DNA was isolated using the alkaline lysis method (Birnboim, 1983) and purified using Wizard DNA purification resins supplied by Promega Co. PCR products were purified using Wizard PCR purification kits from Promega. All DNA stocks were stored in TE buffer at -20°C .

2.1.4 DNA Oligonucleotides: A list of the different DNA oligonucleotides designed and used in this study as cloning scaffold or for the purpose of probe hybridization, DNA sequencing, and PCR amplification is presented in Table 2.1. In all cases a high level of purity was required for preparation of the oligomers. The oligomers were either HPLC or PAGE purified before use. The dG-AAF modification containing oligonucleotides SmaG3.AAF and NarG3.AAF, were gifts from Dr. Robert P.P. Fuchs from Institute de Biologie Molculaire du CNRS Strasbourg, France, and were prepared according to the procedures described in Lambert *et al.* (1992), Koehl *et al.* (1989) and Belguisse-Valladier and Fuchs (1991).

Table 2.1: List of oligonucleotides used in this work and their applications.

Oligonucleotide Designation	Sequence	Application	Source
SmaG3.AAF (14 mer)	5' ATACCCG ₁ G ₂ G ₃ (AAAF)ACATC 3'	Construction of pLHISmaG3.AAF	Lambert et al., 1992
SmaC1.AAF (14 mer)	5' ATACCCG ₁ (AAAF)G ₂ G ₃ ACATC 3'	Construction of pLHISmaG1.AAF	Lambert et al., 1992
Sma G0 (14 mer)	5' ATACCCGGGACATC 3'	Construction of pLHISmaG0 control, probing for non mutational bypass events past SmaG3.AAF lesion	This work
Sma T (14-mer)	5' ATACCCGGTACATC 3'	Construction of pLHISmaT control clones and probing for G3->T mutational bypass events past SmaG3.AAF lesion	"
Sma C (14 mer)	5' ATACCCGGGACATC 3'	Construction of pLHISmaC control clones and probing for G3->C mutational bypass events past SmaG3.AAF lesion	"
SmaG1T (14 mer)	5' ATACCCCTGGACATC 3'	Construction of pLHISmaG1T control clones and probing for G1->T mutational bypass events past SmaG1.AAF lesion	"
SmaC1 (13 mer)	5' ATACCCGGGACATC 3'	Construction of pLHISmaC1 control clones and probing for -C semi-targeted mutational bypass events past SmaG1.AAF lesion	"
Sma-1 (13 mer)	5' ATACCCGGGACATC 3'	Construction of pLHISma-1 control clones and probing for -G frameshift mutational bypass events past SmaG3.AAF lesion	"
HII+ (53 mer)	5' AATCTATCATGTCGACACCTGTGACTG TACAGTCACAGGTGTCGACTGATAGA 3'	HincII linker's plus strand, used for construction of pLHII	"
HII- (53 mer)	5' AGCTTCTATCATGTCGACACCTGTGACT GTACAGTCACAGGTGTCGACTGATAG 3'	HincII linker's minus strand, used for construction of pLHII	"
HIIISmaG0 Scaffold (60 mer)	5' CCAGTGCCCAAGCTTCTATCATGTCGATGTCGCCGGTATG ACTGATAGAATTCTGTAATCATG 3'	Scaffold oligo for construction of SSCpLHIIISma series	"
Anti-HIIISmaG0 (60 mer)	5' CATGATTACGAATCTCATGTCATACCCGGGACATCG ACTGATAGAAGCTTGGCACTGG3'	Antiscaffold for removal of the scaffold from SSCpLHIIISma constructs	"
NarG3.AAF(15 mer)	5' ATACCCG ₁ G ₂ G ₃ (AAAF)CCACA 3'	Construction of pLHIIINarG3	Dr. Fuchs
NarG0 (15 mer)	5' ATACCCGGGCGCCACA 3'	Construction of pLHIIINarG0 control clones and probing for non mutational bypass events past NarG3.AAF lesion	This work
Nar-2 (13 mer)	5' ATACCCGGGCGCCACA 3'	Construction of pLHIIINar-2 control clones and probing for -CG frameshift mutational bypass events past NarG3.AAF lesion	"
HIIINarG0 Scaffold (62 mer)	5' CCAGTGCCCAAGCTTCTATCATGTCGTGCGCGCGGTGAT GACTGATAGAATTCTGTAATCATAT3'	Scaffold for construction of SSCpLHIIINar series	"
Anti-HIIINarI (62 mer)	5' ATATGATTACGAATCTATCATGTCATCACCGGCGGCCAC AGACTGATAGAAGCTTGGCACTGG3'	Antiscaffold for removal of scaffold from SSCpLHIIINar constructs	"
PRUmu (29 mer)	5' ACAGAAATTCAGCAGCCTATGCAGCGACA 3'	Reverse PCR primer for amplification of the unuDC operon	"
PFP _{unu} (29 mer)	5' CTGCTGCAGTATTATTGACCCCTCAGTA 3'	Forward primer for amplification of unuDC operon	"
-35lacZ (17 mer)	5' CCCAGTCACGACGTTGT 3'	Forward universal sequencing primer for lacZ	"
Rev17 (17 mer)	5' TAAACAATTCACACAGG 3'	Reverse sequencing primer for lacZ	"

2.1.5 DNA and protein sequence analysis software: MacDNASIS Pro V3.0 from Hitachi Software Engineering Co., LTD., and DNA Strider Version 1.1 were used for routine computerized sequence manipulations. Oligo version 4.0.6 for Macintosh from National BioScience, Inc. was used for designing and analyzing the primers used for PCR, sequencing and probing. RNAstructure software version 3.71 for Windows (Mathews et al., 1999) was used for secondary structure predictions and free energy calculations, MacDraft Version 2.0 from Innovative Data Design, Inc., Microsoft PowerPoint 2002, and MacPlasmap V1.8 were used for constructing plasmid maps, and figures.

2.2 Methods:

2.2.1 Construction of single stranded phagemid based vectors that carry a single

dG-C8-AAF adduct: Single dG-C8-AAF adduct-containing oligomers SmaG3-AAF, SmaG1-AAF and NarG3-AAF (Table 2.1), were cloned into the plus strand of a phagemid pUC118 derivative pLHII using the approach briefly outlined in Figure 2.1.

2.2.2 Construction of pLHII: The phagemid pLHII was constructed by fragment swapping of the HincII synthetic palindromic linker with the *EcoRI-HindIII* polylinker of the phagemid pUC118 as explained in detail below.

2.2.2.1 Digestion of pUC118: A total of 2.5 µg of phagemid pUC118 was double digested with 20 units each of restriction enzymes *EcoRI* and *HindIII* in *EcoRI* buffer in the presence of 0.1 µg/µL BSA in a total reaction volume of 20 µL. The reaction mix was incubated in a 37°C water bath for 2-3 hours. The completeness of the digestion was confirmed by running a 0.1 µg sample from the mix on a 0.7% agarose gel in parallel with undigested sample. The linearized phagemid was extracted with 100 µL PCI

Fig. 2.1: Schematic representation of the method used for construction of the single stranded circular vectors carrying a single dG-AAF adduct

(Phenol:Chloroform:Isoamyl alcohol, 25:24:1 v/v) after bringing the total volume of the mix to 300 μ L with TE. The DNA was precipitated from the aqueous phase at -20°C overnight following the addition of 1 mL of EtOH/ 0.16 M sodium acetate and 10 μ L of glycogen (2 μ g/ μ L) as carrier. The DNA pellet was recovered by spinning at 18000xg on an Eppendorf benchtop centrifuge model 5403 at 4°C for 10 minutes. The pellet was washed in 70% cold ethanol and dried in a Centrivap Concentrator, then dissolved in 10 μ L of TE to a final concentration of 0.25 μ g/ μ L.

2.2.2.2 Kinasing the HII49 oligomers: A total of 100 pmoles of each of the HPLC purified HII49 complementary strands were phosphorylated at their 5' end in 10 μ L final reaction volume using 10-fold molar excess of ATP in 1xT4 polynucleotide kinase (PNK) buffer in the presence of 10 units of the enzyme PNK. The reaction mixture was incubated in a water bath adjusted to 37°C for 1 hour. The enzyme was inactivated at 65°C for 15 minutes at the end of the incubation. The condensation from the walls of tubes was spun down and the DNA was stored in -20°C freezer until use (10 pmoles/ μ L).

2.2.2.3 Formation of HII49 linker: The HII49 linker was formed by annealing 5 μ L (50 pm) of each of the 5'-end phosphorylated complementary strands in a total volume made up to 50 μ L by water. The mixture was heated at 90°C for 2-3 minutes in a 250 mL beaker containing 100 mL hot water. Annealing was initiated by cooling the water to room temperature over an hour.

2.2.2.4 Cloning of the HII49 linker into pUC118: The *EcoRI/HindIII* double digest and the HII49 linker were mixed in molar ratio of 1:2.5 (vector:insert) using 0.47 pmoles of linear vector and ligated at 16°C overnight in 1x ligase buffer in final volume of 10 μ L using 1 μ L of T4 DNA ligase (400 Units). One microliter of the ligated mix was used to

transform the DH5 α competent cells through electroporation. The cells were plated on the LB agar plates supplemented with 150 μ g/mL Ampicillin and 30 μ g/mL X-gal . Several white colonies were probed with one of the linker strands as probe. Positively probed clones were sequenced. One of the clones with a correct sequence was designated as pLHII142-2 and used for preparation of single stranded circular form (SSC) for further analysis. A -80°C stock of DH5 α (pLHII142-2) was prepared and stored properly. For simplicity this clone will be denoted pLHII.

2.2.2.5 Preparation of the plus strand of SSCpLHII phagemid: Single stranded pLHII was prepared using the method described by Viera and Messing (1987), with some modifications. *E.coli* JV30 was transformed with phagemid pLHII. One pure colony was grown to OD₆₀₀=0.2 in 6 mL LB broth medium containing 150 μ g/mL of ampicillin. A 4 mL sample of this culture was infected with 200 μ L of helper phage M13KO7 at moi=10. After 5 minutes the whole infected culture was transferred aseptically into 200 mL of LB.Amp broth. This culture was grown to A₆₀₀= 0.6 at 37°C on a rotary shaker water bath before kanamycin was added to the final concentration of 70 μ g/mL i.e. 280 μ L from 50 mg/mL stock. The culture was then grown for 14-18 hours in a rotary shaker 37°C water bath vigorously shaking at 250 rpm. The cells were harvested by spinning the culture in 250 mL screw capped containers in a GSA rotor in Sorvall Centrifuge model RC-5B at 10,000 rpm for 15 minutes. The supernatant (100 mL) was transferred into each of 2x250 mL containers. A total of 35 mL of PEG (15% in 2.5M NaCl) was added to each portion, mixed and kept on ice for 2 hours.

The phage particles were precipitated by centrifugation at 12000xg for 15 minutes at 4°C using a GSA rotor. The supernatant was discarded carefully without dislodging the phage pellet. The pot was respun for a short period of time in order to gather the rest of the medium from the walls of the pot, and the medium was discarded. The phage pellet was resuspended in 2 mL of TE. The phage suspension was divided into 2 x 1 mL in 2 mL Eppendorf tubes and was extracted by 800 µL PCI twice and with CI once with intermittent centrifugations at 12000xg in a benchtop centrifuge for 5 minutes. The aqueous phase was pooled in a 15 mL Corex tube and the DNA was precipitated with 3 volumes of EtOH/0.16M NaCl at -20°C overnight followed by centrifugation at 12000 x g in SS-34 rotor. The pellet was washed with cold 70% EtOH, dried in a Centrivap concentrator (LabConco Inc.), and then dissolved in 100 µL TE. The whole isolate was fractionated on 0.6% agarose gel and the band corresponding to the SSCpLHII was excised from the gel (proper measures were taken in order to keep the UV exposure of the DNA to the minimal possible level) and purified on home made filters by freezing and spin-squeezing the sliced gel pieces. The DNA was precipitated from the filtrate by adding 3 volumes of EtOH/0.16 M NaAc, washed in 70% ethanol and resuspended in 200 µL of TE. The concentration of DNA was determined using UV absorbance at 260 nm, and samples were aliquoted in 50 µL portions and stored in the -20°C freezer until use.

2.2.2.6 Digestion of SSCpLHII and preparation of stock SSLpLHII: The stock linear single stranded plus strand (SSL⁺) of the pLHII was prepared by digesting the palindromic structure (Fig. 2.1) formed after annealing the intrastrand complementary sequences with *HincII* twice as follows: intrastrand annealing of 20 µg of the SSCpLHII was carried out in 1x NEB buffer #3 in the presence of BSA (0.1 µg/µL) in a total

volume of 45 μ L by incubating the mixture in a 95°C water bath (100 mL) for 2 minutes and then cooling it down to room temperature over 30-45 minutes. At the end of annealing the tubes were spun quickly to remove the condensation from the walls, and then 5 μ L (50 Units) *HincII* was added and mixed gently. The reaction mixture was incubated at 37°C overnight. The following day the mixture was extracted with 100 μ L PCI after bringing the total volume to 300 μ L with TE. DNA was precipitated from the aqueous phase by ethanol precipitation as mentioned earlier. The pellet was dissolved in 50 μ L of TE and redigested as before, and then concentrated in a final volume of 50 μ L in TE and quantified by UV absorbance at 260 nm. Equivalent samples were taken from SSC, and first and second digests were used for verification on the agarose gel and also to transform the JM103 competent cells in order to determine the digestion efficiency. Also, the accuracy of the digestion was confirmed by sequencing the linear and circular forms of the single stranded vectors from both ends.

2.2.2.7 Optimization of the conditions for cloning the short DNA oligomers, with and without AAF modification, into SSLpLHII: The 14 nucleotide long *SmaGO*

oligomer, 5'-end labelled with 32 P and PAGE purified (see section 2.2.2.8), and *HIISmaGO* scaffold DNA (HPLC grade) were used to optimize the cloning conditions.

The optimal molar ratio of the vector:scaffold:insert was determined by annealing the varying ratio of these components i.e. 1:1:2, 1:2:4, 1:4:8 and 1:8:16 in 1x NEB T4 DNA ligase buffer in a final volume of 20 μ L. For all conditions 0.4 pmoles of the vector was used. The mixtures were annealed by heating the tubes in 100 mL 95°C water bath for 2 minutes and then cooling down to room temperature gradually. After spinning the condensation from the walls of the tubes 1.0 μ L of T4 DNA ligase (400 units) was added

to the mixtures. The tubes were incubated at 16°C overnight. The following day the volume of each tube was adjusted to 300 µL with TE before extracting with 50 µL of PCI followed by CI. DNA was precipitated from the aqueous phase with 3 volumes of EtOH/0.16 M sodium acetate and 10 µL of glycogen (2µg/µL). The pellet was washed with 70% ethanol and, after drying in a Centrivap, was dissolved in 9 µL of TE. The scaffold DNA was removed by annealing the scaffold DNA to anti-scaffold, Anti-HISmaG0 (molar ratio of 1:10), that has complementary sequences to the scaffold. A negative control was also included that had all components at a 1:2:4 ratio but did not contain DNA ligase. From these final mixtures 2 µL samples were taken for electrotransformation of the JM103 competent cells. The transformation capacity gained through ligation was taken as the indication of the ligation efficiency. The transformants were selected on LB agar plates supplemented with ampicillin. The cloning of the SmaG0 oligomer was confirmed by sequencing several clones. In addition 4 µL samples from the ligation mixtures were run on the 0.7% agarose gel using a mini cell electrophoresis apparatus in parallel with SSC and minus enzyme controls. The gels were run for approximately 1-2 hours until a clear resolution between SSL and SSC forms of the vector was achieved. The gel was dried by sandwiching it between several layers of 3MM Whatmann filter paper at the bottom and one layer of Saranwrap on the top. The whole preparation was transferred onto the gel dryer, which was set to 60°C, where it was dried by application of gentle suction. The dried gel was carefully detached from the sandwich, wrapped with Saranwrap, and exposed to Kodak BioMax X-ray film in a metal autoradiography cassette at -80°C for 12 hours before developing the film. The ligation conditions for the 1:2:4 set were adopted throughout this study to clone oligomers with

and without AAF modification, and to construct control plasmids. The concentration of the SSC form that carries the cloned oligomer for a stock construct was calculated by reading the Cerenkov counts for the corresponding excised bands from the triplicate agarose gels, and entering the averaged value into a function obtained by the regression analysis of the Cerenkov counts read for a dilution series of the same $\gamma^{32}\text{P}$ -ATP source that was used for end labeling the cloned oligomer. Correction for the reduced cpm values due to the decay of the $\gamma^{32}\text{P}$ source between the date of kinasing the oligonucleotides and the time of reading, was carried out using the formula given below:

$$N^{\circ} = B \cdot \text{cpm} \cdot e^{(0.0485T)} \cdot \text{df} \cdot 1/V$$

Where:

N° = the final concentration of the oligomer in pmole/ μL

B = the slope of the regression line of the standard curve that relates the Cerenkov cpm readings to the corresponding pmole/ μL values of the half dilution series prepared for the $\gamma^{32}\text{P}$ -ATP source

cpm = the Cerenkov count reading for the eluted sample

e = the base of the natural logarithm = 2.718281828

T = number of the days elapsed since labeling

df = the dilution factor for the $\gamma^{32}\text{P}$ -ATP with non-radioactive ATP used in kinasing reaction, that can be obtained by calculating the molar ratio of ATP/ $\gamma^{32}\text{P}$ -ATP

V = volume of the sample taken for reading in μL

2.2.2.8 Testing for the efficiency of removal of scaffold DNA by heat denaturation and/or annealing to an anti-scaffold DNA: A ligation mixture for cloning the 5'-end

phosphorylated SmaG3.AAF oligonucleotide into SSLpLHII was prepared in 1x T4 DNA ligase buffer (New England Biolabs), using the optimal molar ratios of the DNA components (i.e. 1:2:4 for vector:scaffold:insert, respectively) as described in section 2.2.2.7, with the exception that the ^{32}P -labelled HIISmaG0 scaffold oligonucleotide was used in this experiment. At the end of the incubation period, 10 μL of ligation reaction mixture was distributed to each of two 0.5 mL Eppendorf tubes. One of the tubes was heated to 95°C for 2 minutes followed by rapid cooling on an ice bath, while the second tube received 1 μL of 8 μM anti-HIISmaG0 in TE (i.e. molar ratio of 1:10 for scaffold:anti-scaffold). The annealing of scaffold DNA to anti-scaffold DNA was carried out by heating this tube at 95°C for 2 minutes in a water bath, followed by gradual cooling to room temperature over 30 minutes. The efficiency of removal of the scaffold DNA using either heat denaturation followed by rapid cooling, or annealing of the scaffold DNA to a molar excess of the anti-scaffold DNA, was evaluated by agarose gel electrophoresis. Two μL of 6x loading buffer containing bromophenol blue was distributed into each of the tubes, and the whole volume of the samples was loaded into the wells of a 0.7% agarose gel. Samples from SSCpLHII and ^{32}P -HIISmaG0 scaffold DNA were used as controls. Electrophoresis was carried out for 15 minutes at 8 volts/cm, at which time a clear resolution occurred between the control bands. The gel was photographed, dried and autoradiographed as described previously in section 2.2.2.7.

2.2.3 DNA sequencing: The sequence of the *EcoRI-HindIII* linker cloned into pLHII, its linearized form, all of the control constructs, and sample positive clones from mutant collections were verified with dideoxy chain termination protocol using the T7

Sequenase DNA polymerase version 2.0, purchased from Amersham Life Sciences. All of the required solutions for sequencing, except for Sequenase reaction and dilution buffers, were prepared in the lab. The -35lacZ or Rev17 primers, and the forward and reverse sequencing primers for the *lacZ* gene, were annealed in 1x Sequenase buffer using 2-3 µg of the denatured plasmid and one pmole of primer in a total volume of 10 µL. The annealing was done by incubating the tubes in a 100 mL water bath, warmed to 75-80°C, which was then cooled down to room temperature over 30-45 minutes. The labelling reactions were carried out at room temperature following addition of 1 µL of 0.1M DTT (dithiothreitol), 2 µL of 1:5 diluted labelling mix (5x concentrated mix contains 7.5 µM each of dGTP, dCTP and dTTP), 0.5 µL of $\alpha^{35}\text{S}$ -dATP (6000 Ci/mmol) and 2 µL (3.25 U) of T7 Sequenase DNA polymerase diluted in dilution buffer by a factor of 1:8. The chain termination reactions were carried out by transferring 3.5 µL portions from the labelling reaction mixtures to the 4 wells of a microtiter plate containing 2.5 µL of termination mixes for each of the A, T, G and C mixes (dGTP, dCTP and dTTP 80 µM each, 50mM NaCl, and one of the ddNTP's). The termination reactions were carried out on a thermocycler adjusted to 37°C for 5-10 minutes and stopped by adding 4 µL of the Sequenase stop solution (95% Formamide, 20mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol FF). The polymerization products were denatured, and 2-3 µL of the samples were applied to the wells formed by 0.1 mm thick sharktooth combs. The products were separated on a denaturing 8% polyacrylamide slab gel containing 7M urea in 1xTBE through electrophoresis in 1x TBE running buffer at 70 watts (1700-2200 volts and 30-35 mA) for 2-3 hours. The gel was lifted by transferring it onto a 3MM Whatman filter paper sheet, covering with Saranwrap and drying under vacuum at 80°C

over 1-2 hours. The dried gel was autoradiographed by exposing it to BioMax Kodak X-ray film for 1-2 days followed by processing.

2.2.4 PAGE purification of 5'-³²P end-labelled oligomers: All of the oligomers used for the purpose of cloning were phosphorylated at their 5' -end with a mixture of $\gamma^{32}\text{P}$ -ATP and ATP (1:300 molar ratios) using the reaction conditions described in section 2.2.2.2 . At the end of the labelling the reaction mixture was purified on 20% denaturing polyacrylamide slab gels. The 20% gel was prepared by mixing 37.5 mL of 40% stock acrylamide/bis-acrylamide (19:1 w/w) solution with 15 mL of 5x TBE and 31.5 grams of urea to a final volume of 75 mL. After completely dissolving the urea, 450 μL of 10% APS (ammonium persulfate) and 50 μL of TEMED were added, the mixture was swirled, and poured into a mold prepared by the assembled sequencing plates. A large tooth preparatory comb (0.1 mm thick and 1 cm wide) was used to make the wells. The labelled samples were loaded into the wells after adding a volume of sequencing stop solution equivalent to 40% of the total volume, and heating the samples at 95°C for 2 minutes. The gel was run at 70 watts (1700-2200 volts and 30-35 mA) for 2-3 hours. At the end of the run the mold was opened and one of the glass plates was removed while the other glass together with gel was covered with Saranwrap. Proper position markers of dotted ³²P filter papers were mounted to the corners of the plate in an asymmetrical manner. The preparation was exposed to BioMax Kodak X-ray film for a maximum of 30 minutes and developed. The position of the band of interest was determined by overlaying the film on the gel so that the position markers and related signals could be coincided. The band was cut out of the gel and eluted from the gel after incubating it in 400 μL of TE at 37°C over night. The eluted DNA was precipitated in EtOH/0.16 M

sodium acetate as described previously (section 2.2.2.1). The pellet was dissolved in an appropriate volume and the concentration of the DNA was calculated by comparing the Cerenkov counts read for the sample against a standard curve obtained by regression analysis of the Cerenkov counts read for a dilution series of the same $\gamma^{32}\text{P}$ -ATP source that was used for the end labelling. The eluted DNA was stored at -20°C until use.

2.2.5 Labelling of oligomers for probe hybridization: Large scale 5'-end labelling of different oligomers was carried out in a total volume of 100 μL by use of 200 pmoles of the oligonucleotide in 1x T4 polynucleotide kinase buffer containing 83.5 pmoles of $\gamma^{32}\text{P}$ -ATP (6000 Ci/mmol, Amersham Life Sciences), 1 nmoles of ATP and 50 units of T4 polynucleotide kinase. The reaction mixture was incubated at 37°C for one hour. At the end of the incubation the enzyme was inactivated at 75°C for 10 -15 minutes. The condensation on the wall of the tube was spun down and the whole volume was passed through a Sephadex G-50 spun column packed into a 1 mL syringe barrel. The filtrate was recovered into a decapped 1.5 mL Eppendorf tube inserted into a 15 mL Falcon tube that was spun at ~ 2000 rpm for 10 minutes in a Janetzki T30 clinical benchtop centrifuge purchased from Dena Instruments LTD. The volume of the recovered end labeled oligonucleotide was adjusted to 400 μL with TE (final concentration of 0.5 pmoles / μL) and stored at -20°C until use. A total of 5 μL of the probe was used per membrane for probe hybridization.

2.2.6 Preparation of competent cells induced for SOS functions: The electrocompetent cells of wild type *E.coli* JM103 , JM103 $\Delta umu\text{C}$, JM103 $\Delta umu\text{C}$ (pDC322) and JM103 $\Delta umu\text{C}$ (pGW1700), induced for the SOS functions were prepared as follows. A flask containing 250 mL sterile LB broth was inoculated with 1 mL of an

overnight culture, and grown to a final A_{600} of 0.5 in a 37°C shaker water bath set to 200 rpm. The cells were recovered by spinning the whole culture in 250 mL pots in a GSA rotor (Beckmann) at 8000 rpm, 4°C for 10 minutes. The cellular pellet was resuspended and washed in 150 mL of 10 mM sterile and cold $MgSO_4$ solution. The cells were repelleted as before and resuspended completely in 150 mL of the same solution and kept on ice. Multiple platings were done from this suspension by spreading 100 μ L portions of 10^{-5} - 10^{-6} dilutions onto LB agar plates. The plates were incubated in a 37°C incubator overnight. The bacterial suspension (25 mL) was distributed into 6 petri plates and irradiated on a rocking nutator under a UV lamp (254 nm) to an average exposure of 70 Joules/ m^2 for wild type strain and strains with plasmid encoded UmuDC or MucAB functions, and at 8 Joules/ m^2 for the $\Delta umuC$ strain. The irradiated cells were transferred into 250 mL pots and multiple 100 μ L samples were taken from 10^{-4} - 10^{-5} for plating on LB agar plates. The cells were pelleted by centrifugation as described previously, resuspended in 150 mL LB broth, and allowed to recover and express the SOS functions by incubation in a rotary water bath for 30 minutes at 37°C and 200 rpm. The cells were concentrated by pelleting and resuspending them in 250 mL of 10% sterile and ice-cold glycerol, followed by a second centrifugation and resuspension in 125 mL of 10% glycerol. Finally the cells were pelleted one more time and were taken up in 2 mL of cold 10% glycerol in a 2 mL sterile Eppendorf tube. This tube was spun in an Eppendorf benchtop centrifuge for 3 minutes at the highest speed. The volume of the contents of the tube was adjusted to 1 mL and the cells were resuspended completely and (40 μ L/tube) distributed into ice-cold labeled sterile 1.5 mL Eppendorf tubes. Multiple 100 μ L

samples of 10^{-7} - 10^{-8} dilutions were taken for plating on the LB agar plates. The percentage of the cellular density corresponding to the culture before irradiation as compared to that immediately after irradiation was considered as the survival rate, usually 20-30%, while the colony counts from the last plating (after recovery) indicated the cellular density of the stock competent cells. The tubes were stored in -80°C until use.

2.2.7 Preparation of competent cells without induction of SOS functions: The procedure described in the previous section was employed to prepare the SOS⁻ electrocompetent cells for the same strains. The only modification was the omission of the resuspension of the cells in MgSO_4 solution and UV-irradiation. The washing steps were carried out immediately after harvesting the cells from the LB broth culture.

2.2.8 Preparation of cellular collections transformed with single stranded constructs with and without AAF modification : SOS-induced and uninduced competent cells prepared above were transformed in triplicates by electroporation using 1 μL (13-50 ng in SSC form) of the single stranded constructs with and without AAF modification. Cells were subjected to 2500 volts at 25 μF capacitor setting in a 0.2 mm electrode gap cuvette from BioRad that generated an electrical field of 12,500 V/Cm that decayed in a square wave shape over the 5.0-5.2 mS time interval. A total of 960 μL SOC medium was transferred into the cuvette in less than 5 seconds after shocking and the entire 1 mL volume was then transferred to a sterile tube and incubated in a 37°C water bath shaking at 250 rpm for 15 minutes. The transformants were selected on fresh LB agar plates supplemented with 150 $\mu\text{g/mL}$ of ampicillin by spreading 100 μL samples

from 10^{-1} - 10^0 dilutions. The transformation frequencies were calculated for each experimental set through the use of the averaged colony numbers for the multiple plates.

2.2.9 Screening for mutational and non-mutational bypass events: A probe hybridization strategy was adopted in order to screen for the possible outcomes of the *in vivo* replicational processing of the AAF adducts introduced into the cells on single stranded constructs. The general guidelines of the supplier of the membranes and the protocols of Maniatis (1989) were followed with some modifications when necessary.

2.2.9.1 Transfer of the colonies onto nitrocellulose membranes and their processing:

The whole plate colony lift was performed in order to transfer the colonies from plates onto Hybond-N nylon membrane discs (Amersham Life Sciences) of 82 mm in diameter. The positive and negative control strains that contained the constructs carrying the cloned oligomers with sequences related to the mutagenic or nonmutagenic bypass events, were transferred onto the membranes by putting 1 μ L of the concentrated cultures of these strains in 10% glycerol. The colonies were lysed by laying the membrane down (colonies on the upper surface) onto 3MM Whatman filter papers soaked in denaturing solution (1.5M NaCl, 0.5M NaOH) for 5-7 minutes. The gradually increasing glossy appearance of the colonies was an indication of completion of the lysis. The alkalinity of the lysis solution was neutralized by transferring the membranes onto filter papers soaked in probing neutralization solution (0.5M Tris.HCl, 1.5M NaCl, 0.001M EDTA) for 3-5 minutes followed by the buffering on the filter paper soaked in 2 x ssc pH 7.4, for 3 minutes. The denatured DNA released from the cells was fixed onto the membranes covalently by exposing the membranes to UV light at 254 nm for 15 minutes on a model FBTI 88 transilluminator, Fisher Scientific Inc.

2.2.9.2 Probe hybridization: The membrane discs prepared in section 2.2.10.1 were subjected to probe hybridization using the ^{32}P 5'-end labeled probes prepared as described in section 2.2.4. The probe hybridization of each individual membrane was carried out in a 50 mL screw capped Falcon tube in a hybridization solution comprised of 5 x SSPE, 0.5% SDS and 5xDenhardt's solution to which 0.5 mg fragmented Herring sperm DNA was added. A total of 5 μL (2.5 pm, 6×10^6 cpm/pm) of the labelled probe was added to the tube, the screw caps were tightened carefully, and the tubes were transferred into hybridization screw capped glass tubes in pairs. The tubes were incubated overnight at 40°C in a rotating hybridization oven Model 2710 (VWR Scientific), adjusted to 4 rpm.

The membranes were washed twice at room temperature for 15 minutes in the same Falcon tubes in 50 mL of low stringency washing solution, 1x SSPE and 0.1% SDS. The stringency washes for probes with low melting temperatures i.e., Sma-1, Nar-2 (13-mer) and SmaG0, SmaT, SmaC and SmaG1T (14-mers), were carried out in 50 mL of 0.1 x SSPE, 0.1% SDS at room temperature for 15 minutes, followed by another higher stringency washing in the same solution but warmed to 40-45°C for 10 minutes. For the NarG0 and all other probes with higher melting temperatures the last stringency washing was carried out at 50-55°C.

2.2.9.3 Stripping the probes off the membranes: The membranes were stripped by incubating them in 0.4 M NaOH for 15 minutes at 65°C. The membranes were then soaked in stripping buffer (0.1xssc, 0.1% SDS, 0.2M Tris.HCl pH 7.5), for 15 minutes.

2.2.10 Construction of pDC322: The *umuDC* operon was cloned into the ampicillin resistance gene of pBR322 by fragment swapping of a 1.8 Kbp PCR product containing

umuDC, flanked with built-in *EcoRI*/*PstI* sites, with the 752 bp *EcoRI* and *PstI* fragment from pBR322 that also deletes 477 bases from the 5' end of the ampicillin resistance gene.

2.2.10.1 PCR amplification of *umuDC* operon: A 1.8 Kbp fragment from plasmid pSE117 (Marsh and Walker , 1985), containing an *E.coli* genomic fragment that carries the *umuDC* operon, was amplified using PCR primers (PRE_{umu} and PFP_{umu}, Table 2.1) with built-in *EcoRI* and *PstI* sites and PWO polymerase. Mix I (25 µL) was prepared by adding the following components into an sterile 200 µL PCR tube: 1.5 µL of 10 pm/µL of each of primers, 1 µL pSE117 (0.1µg/µL), 1 µL dNTP's (10 mM each) and 20 µL ddH₂O. Mix II was prepared by adding the following into an Eppendorf tube: 5 µL of 10x PCR buffer containing 20 mM MgSO₄ , 19.5 µL ddH₂O and 0.5 µL of Pwo polymerase (5U/µL). Mix II was added to mix I, and the reaction was started on a thermocycler using the following program: predwell at 94°C for 5 minutes; 25 cycles of denaturation at 94°C for 1 min., annealing at 55°C for 30 sec and polymerization at 72°C for 1 minute; postdwell incubation at 72°C for 5 minutes; and chilling at 4°C. The negative control without template was also run. The quality of the products was verified by running a 5 µL sample from each reaction in parallel with a 1 Kbp linear ladder. The reaction products were purified by use of Wizard PCR purification resin, and eluted in TE to a final concentration of 0.16 µg/µL.

2.2.10.2 Digestion of pBR322 with *EcoRI* and *PstI* : A total of 2 µg of plasmid pBR322 was digested in a total volume of 20 µL in 1x *EcoRI* buffer using 20 units each of enzymes *EcoRI* and *PstI* . The reaction was carried out at 37°C for 2-3 hours. A 2 µL sample was checked on an agarose gel for the completeness of the digestion. The

linearized plasmid was purified using PCI extraction in a total volume of 300 μL , and was alcohol precipitated as described earlier, and concentrated to 0.2 $\mu\text{g}/\mu\text{L}$ in TE. The digestion of the *umuDC* PCR product was carried out in a similar manner except that a total of 1.6 μg (10 μL) was used and was finally concentrated to 0.16 $\mu\text{g}/\mu\text{L}$ in TE.

2.2.10.3 Ligation of *umuDC* PCR product and linearized pBR322 and screening for

JM103 Δ *umuC*(pBR322/*umuDC*) clones: Ligation was carried out in a total volume of 10 μL by mixing 0.6 μg of the linear vector with 0.64 μg of the *umuDC* PCR product (molar ratio of 1:2.5) in T4 DNA ligase 1x buffer to which 1 μL of ligase was added and incubated at 16°C overnight. One microliter of the ligation mix was used to transform the *E.coli* JM103 Δ *umuC* strain, and the transformation mixture was plated on LB agar plates supplemented with 6.25 $\mu\text{g}/\text{mL}$ of tetracycline. The plates were probed with 5'-end labelled PRE_{umu} primer, and 17 positively probed clones were checked for their sensitivity to ampicillin. The UV induced rifampicin resistance test was carried out on these clones by exposing the prewashed LB broth cultures of the clones in 10mM MgSO_4 and irradiating them at 70 J/m^2 followed by overnight enrichment of the mutants and plating on LB agar containing 50 $\mu\text{g}/\text{mL}$ of rifampicin. Parallel platings were done on LB.agar.tet plates in order to calculate the mutation frequencies. The mutation frequencies were compared to those of JM103 Δ *umuC* (pSE117) as positive control and JM103 Δ *umuC* as negative control. A positive clone designated as JM103 Δ *umuC*(pDC322) was chosen, sub-cultured and stored at -80°C for preparing the competent cells.

2.2.10.4 Statistical analyses: The significance of the effect of the varying cellular physiological conditions i.e. induction of SOS functions and/or episomal expression of

UmuDC or MucAB proteins, on the translesion synthesis outcomes were tested by performing the two sample significance t-test. In addition, the combined probabilities Chi-square test was adopted to examine the degree of the overall effect of a variable factor i.e. induction of the cellular SOS functions, on the translesion processing of the AAF adduct, by combining a series of probability values furnished by the significance Student's t-test applied to each of the different classes of bypass events.

2.2.10.4.1 Two-sample Student's t-test: The heteroscedastic Student's t-test was used to examine the null hypothesis that the means of the two sets of data are equal i.e. homogeneity of two samples, assuming that the variances of both samples are unequal. The effect of induction of the SOS functions on a single translesion synthesis event (i.e. error-free, -G frameshifts, etc.), was tested by comparing the TLS frequencies for that specific event obtained under the SOS induced conditions to the TLS frequencies under uninduced conditions. In a similar manner, the t-test was used to examine the effect of the expression of the UmuDC and MucAB functions on TLS frequencies. The statistical Toolpak of Microsoft Excel 2000-2002 from Microsoft Corporation, was used to assist the computerized analysis. The test is based on calculation of a sample “ t_s ” value, which is a function of the means and variances of two sample data as shown below:

$$t_s = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}}$$

where, $\bar{Y}_1$ and $\bar{Y}_2$ are the means and S_1^2 and S_2^2 are the variances of two data sets. The symbols “ m ” and “ n ” resemble the size of each data set.

The degrees of freedom were computed using the formula below:

$$df = \frac{\left(\frac{s_1^2}{m} + \frac{s_2^2}{n} \right)^2}{\frac{\left(\frac{s_1^2}{m} \right)^2}{m-1} + \frac{\left(\frac{s_2^2}{n} \right)^2}{n-1}}$$

The calculated “ t_s ” values were compared to the critical t-value, “ $t_{0.05[v]}$ ”, corresponding to $\alpha=5\%$ confidence limit and the calculated degrees of freedom. When $t_s > t_{critical}$ the null hypothesis of $H^0: \bar{Y}_1 - \bar{Y}_2 = 0$, was rejected and the probability for the homogeneity was obtained as the α -value that corresponds to the t_s from the t-distribution table.

2.2.10.4.2 The combined probabilities Chi-square test: The combined probabilities

Chi-square test is a complementary and conclusive statistical test that is applied to the probabilities of a series of significance tests that have examined the same hypothesis.

The test gains its special value when one or more of the preliminary significance tests fail to furnish probability values low enough to have a suggestive power. Theoretically, the test takes advantage of the fact that the $-2 \ln P$ is distributed as χ^2 values and so does the

$-2 \sum_1^K \ln P_i$, where K is the number of the preliminary tests, and P_i is the probability of

each of the preliminary significance tests. The degrees of freedom for this test is taken as

$df=2K$. The test was applied to examine the overall effect of SOS induction using the

individual P_i values obtained for the t-tests, which were carried out to test the significance

of the effect that induction of SOS functions has had on each of the individual translesion

bypass events. If the calculated $-2 \sum_1^K \ln P_i$ value appears to be greater than the $\chi^2_{0.05[2K]}$

then the null hypothesis that there has been no difference between the SOS induced and

uninduced cells, would be rejected. In this case the probability value for the test was

determined as $P < \alpha$ where α is the next highest probability value corresponding to the

$-2 \sum_{i=1}^K \ln P_i$ value with $2K$ degrees of freedom.

3. Results

3.1. Experimental plan for constructing the single stranded phagemid vectors that carry a site specific single dG-AAF adduct: Our initial experimental approach to construct the single stranded circular phagemid vectors carrying a single dG-AAF adduct is depicted in Figure 3.1. According to this plan the circular plus strand of a modified M13mp18 phage would be digested using *HincII*. A short HPLC purified oligonucleotide carrying a single dG-AAF adduct would be cloned into the modified polylinker region of the linearized single stranded vector, yielding an AAF-modified single stranded M13 vector containing a frameshift in the downstream coding sequence of the α -complementation fragment of β -galactosidase gene. Transfection of these constructs into α -complementable *E.coli* strains would allow us to monitor for revertant mutants that could restore the reading frame of the α -fragment of β -galactosidase in a blue-white phenotypic assay. For almost two and a half years I tried this approach and encountered numerous complications at almost every stage, which convinced me and the advisory committee to reassess the plan, redesign and adopt the experimental approach that is presented in Materials and Methods (Fig. 2.1). The difficulties experienced in our initial design are described in the following sections.

3.1.1. Large amount of TL1 digestion oligonucleotide was needed: The approach was based on the use of MLH, a modified single stranded circular viral strand of M13mp18 filamentous phage that was modified by the incorporation of a synthetic DNA duplex that contained a *HincII* restriction site. MLH contains a -1 frameshift in the *lacZ* α -complementation fragment coding sequence that brings two successive stop codons into

Fig. 3.1: A) Sequence context at partially duplex modified polylinker region that is formed by annealing TL1 to SSC-MLH. B) Schematic representation of construction of M13mp18 based single stranded circular vectors that carry a single dG-AAF modification.

frame. A 38-mer oligonucleotide (TL1) complementary to the plus strand modified polylinker region of MLH that carries a single *HincII* site, had been designed by Dr. Lambert before I joined his lab. Annealing this oligonucleotide to pure SSC-DNA generated a partially double stranded region that was used for linearization of the vector by *HincII* digestion. *HincII* was chosen since it does not have star activity and hence can be used for prolonged incubation periods to maximize the digestion efficiency. The extent of the digestion was examined by annealing the gel purified 5'-³²P -end labeled 38-mer TL1 oligonucleotide in varying molar ratios relative to the vector. The banding pattern was examined by running the samples on an 8% denaturing polyacrylamide gel. A 19-mer band co-migrating with the bromophenol blue band was detected which was absent in the control set. This band represents the *HincII* digestion product. According to these results the maximum digestion was achieved in a 100-fold excess molar ratio of TL1 compared to the vector.

3.1.2. TL1 and its digestion products needed to be eliminated: The product of *HincII* digestion of the MLH/TL1 duplex included unused TL1, 19-mer digestion products of the TL1 hanging on the two ends of the linearized vector and undigested SSC vector, in addition to the desired SSL⁺ MLH. The presence of the undigested TL1 and 19-mer digestion products of the TL1 could potentially interfere with the next stage of the plan that required annealing of a scaffold oligonucleotide in order to form a gapped partial duplex into which the AAF modified oligomers would be ligated. The remaining undigested vector can potentially contribute to elevated vector background transformation, and therefore, the linearized vector had to be purified from the competing

oligomers and undigested circular vector. In principle, any additional step in an experimental setup has the drawback of reducing the final yield and making the procedure more labor intensive. Significant problems were encountered in resolving the SSC and SSL^+ bands and during purification of SSL^+ cleavage product. Extensive effort was put into differential electrophoretic separation of the SSC and SSL^+ by use of both agarose and polyacrylamide gels. On polyacrylamide gels the highest resolution was achieved on a 4% gel (49:1 of acrylamide:Bis acrylamide). Nonetheless, further electroelution of the SSL^+ band from the polyacrylamide gel was extremely inefficient (~2-5% recovery).

A very narrow range of agarose gel concentration (1.5%-1.8%) was shown to be effective in separating the SSL^+ and SSC forms. Different procedures were attempted for purification of the SSL^+ band from agarose gels, including freeze-spin-squeeze, phenolizing, electroelution and finally the Gene Clean kit. Among all of these approaches the best recovery (approximately 40-50%) was obtained by use of the Gene Clean Kit. The Gene Clean purified SSL^+ was used for cloning the radioactively labeled SmaG0 and NarG0 unmodified oligonucleotides at an optimal molar ratio of 1:2:4 of vector:scaffold:oligo, respectively. The formation of the recircularized band was confirmed by autoradiography. However, no plaques were obtained when the constructs were used to transfect *E.coli* JM103 electro-competent cells. In contrast to this observation the same competent cells that were transfected with the RF and SSC forms of the MLH, were able to produce plaques at relatively high frequencies (10^7 - 10^8 pfu/ μ g for RF and 10^6 pfu/ μ g for SSC). The procedure was repeated several times; however, in each case the transfection was unsuccessful. On the other hand, the electroporation conditions

were the same compared to that of the transforming isolates ($\tau=4.9-5.3$ mSec, $E=12500$ volts/cm, $C=25\mu F$). Increasing the concentration of the DNA in the transformation mixture didn't result in any substantial increase of the transformation efficiency. We concluded that the use of the glass milk (positively charged powdered glass particles) included in the Gene Clean Kit might have been an interfering factor. The supporting observation for this conclusion was the fact that the SSL⁺ source that was gel purified using freeze-spin-squeeze method, could yield recircularized bands with transforming capability at $\sim 10^4$ pfu/ μg DNA from ligation reaction.

As a substitution for the Gene Clean Kit, we decided to use the Costar microporous filters (Spin-x UF¹⁰⁰, cut off size of 300 bases) to eliminate the TL1 and its cleaved product from the linearization reaction mixture of the SSC DNA. However, the use of these filters could not eliminate the undigested vector. As a result there was a need to maximize the linearization efficiency of the *HincII* digestion reaction. Linearization of the SSC by extended and double *HincII* digestions reduced the remaining transforming ability of the digestion mixture to about 1%. The remaining SSC together with the nonmutational bypass products of the AAF modified construct were expected to form a white plaque background.

Despite the extensive trouble shooting described above, several disadvantages remained in this cumbersome experimental design. First, by adding steps to the protocol the work became more costly and labor intensive. Second, despite the use of the newly purchased fresh filters and using the manufacturer's recommended conditions, the filtration of TL1 was not complete and examining the retentate revealed about 5%-20%

remaining TL1 contamination that could potentially contribute to regeneration of the wild type background during the following stages that required a ligation reaction. Third, the transformation efficiency of the SSC DNA was relatively low (i.e. $\sim 10^4$ pfu/ μ g). The relatively low transformation efficiency of the SSC product was a particular problem. Improving the transformation efficiency, considering that the modified constructs would not be available in large quantities, was an obligation.

The SSL⁺ stock was used to clone the radioactively labeled NarG0 (15-mer) oligonucleotides and the formation of the circular SSC band was verified on agarose gel and by autoradiography before transfection into JM103 cells. Screening the replicative forms (RF) isolated from these transformants by *HincII* and *NarI* restriction enzymes, and sequencing, revealed that 70% contained the wild type vector background, 20% were the expected positive clones with the correct sequence of NarG0 oligonucleotide inserted into the *HincII* site, and 10% were unexpected *HincII*-resistant but *NarI* sensitive clones. The dominant wild type background was considered to be partially due to residual undigested SSC and religation product of the SSL⁺ using the contaminating residual TL1 that would regenerate the SSC. It was also noticed that there was only a 16-fold increase in transforming efficiency of the ligation mix (1.6×10^4 pfu/ μ g total DNA) over ligase negative control (1.0×10^3 pfu/ μ g total DNA).

The relatively low transformation efficiency of the M13 SSC construct and the high background levels among the transformants suggested that a higher transformation efficiency would be necessary to monitor the expected frameshift mutants and also rare mutational bypass events such as base substitution. Since the maximum transformation

efficiency obtained with M13mp18 phage was insufficient, following Dr. Bonen's suggestion, we decided to use a smaller size vector, i.e. a pUC118 phagemid (3.2 kbp) derivative. In addition, this would mitigate the protocol by the fact that the bacterial colonies would be dealt with instead of phage plaques. The pLH118 phagemid was constructed by replacing the entire polylinker region of the pUC118 phagemid with a linker similar to that used in the M13 based phage vector. The inserted linker provided a unique *HincII* site and the sense strand sequence was complementary to the central sequence of the TL1 digestion oligonucleotide. The SSL⁺-pLH118 was prepared as before by a filtration step using Spin-x UF¹⁰⁰ filters. The NarG0 unmodified oligonucleotide was cloned into the SSL⁺-pLH118 and the formation of the circularization band was verified before transforming the *E.coli* JM103 cells. A substantial improvement in transformation efficiency approaching 100-fold ($\sim 10^6$ cfu/ μ g) as compared to that of phage vectors was achieved. Screening the transformants by probe hybridization and sequencing showed that the relative abundance of the positive clones carrying the inserted NarG0 sequence into *HincII* site was elevated to 50%-60%. A number of clones were sequenced, and the correct sequence of the linker region, in both positive and negative clones was confirmed. Although the relative frequencies of the pLH118 background was lowered from 70% to 40%-50% as compared to that of M13 phage vectors, the background was still sufficiently high to confirm our previous concern regarding the interference of the residual undigested SSC and the role of the remaining TL1 contamination in promoting religation of the SSL⁺ to regenerate the SSC form. These values were obtained by use of an unmodified oligonucleotide; however, based on

previous experiments it was known that the dG-AAF modification would reduce the transformation efficiency of the construct and subsequently increase the relative share of the wild type vector background and hence substantially decrease the probability of detecting the mutational bypass species. Thus, further improvement in the protocol was necessary to 1) improve the mutation screening strategy; and 2) reduce the wild type vector background.

A screening method was developed that was based on colony probe hybridization. The technique was developed by designing and using ^{32}P 5'-end labeled short DNA oligonucleotides that carried a sequence that was complementary to the sequences of the templates corresponding to nonmutagenic bypass and each of the possible mutagenic bypass events that may result when the vectors were introduced to the cells and processed *in vivo*. Preliminary optimization of hybridization conditions, and adjustment of the washing stringency conditions, was achieved by comparing the signal intensities from transformants to those of the negative and positive controls included on the same membranes. In practice the stringency of washes was controlled by varying the temperature and the ionic strength of the washing solutions such that the radioactive signal associated with the negative controls would diminish significantly, while the positive controls would retain a detectable signal with an intensity sufficient for autoradiographic detection over 1-2 days exposure to a sensitive X-ray film (Fig. 3.2). In addition, random samples of the detected negative and positive clones were sequenced to verify the accuracy of the probe hybridization results. In all cases the probe hybridization results were consistent with the sequencing results. The technique proved to be so

A

B

Fig 3.2. A sample probe hybridization result. A) JM103 SOS- cells transformed with sspLHIINarG3-AAF construct were probed with NarG-2 oligonucleotide. B) JM103 SOS+ cells transformed with sspLHIINarG0 construct were probed with NarG0 oligonucleotide. 103, negative control mock JM103 cells; -2, control JM103 cells that carry the pLHIINar-2 construct for detection of -2 frameshifts; G0, control JM103 cells that carry the pLHIINarG0 construct for detection of error free TLS; H, control JM103 cells that carry pLHII construct for detection of undigested parental construct.

powerful that it was able to detect a single base pair alteration in a short sequence context in the range of 13-20 nucleotides using a moderate washing stringency conditions (2X SSPE to 0.1X SSPE) at ambient temperature. The technique proved also to be highly effective in handling a large number of clones and could accommodate 1000-2000 colonies per membrane lifted directly from the plates containing transformants following overnight growth. The signal was unambiguously detectable from any positive single bacterial colony with a size as small as 0.5 mm in diameter. The other advantage of the probe hybridization method lies in the fact that the same membranes could be reused to examine the hybridization of other probes specific for detection of other possible classes of bypass events. This ensured us that the mutagenic and nonmutagenic events could be studied within the same population, so that deviations due to sampling errors could be avoided. Considering the sensitivity, accuracy and capacity of the method to handle large sample sizes, this method was selected over blue/white phenotypic screening that inherently suffers from several limitations, as explained below.

- i) Phenotypic screening is able to directly detect only one type of bypass event per construct. Detection of other mutational and non mutational bypass classes requires the design of additional constructs specific for detection of each class. This would not only necessitate designing more than one type of construct but would also result in the unavoidable introduction of a neighboring sequence context effect. In addition, sampling errors would be introduced due to the obligation of using multiple populations.

ii) The interpretation of the extent of the phenotypic expression of β -galactosidase is inherently error prone and suffers from subjective judgment of the appearance of blue color with varying intensities. For example, we have shared the observation with Dr. Fuchs's group in France, who in collaboration with Dr. Lambert have employed phenotypic screening in a similar parallel project in the same time period, where the constructs containing the SmaG0 (-2 frameshifted) and Sma-1 (in frame) sequence contexts, were able to produce blue colonies with varying color intensities. The blue phenotype developed to higher intensity by incubation in 4°C fridge for more than 24 hours making the distinction between them progressively difficult. The interpretation of these observations will follow in section 4.2.

iii) Phenotypic screening cannot directly discriminate the false negative and the false positive clones from real negative and positive clones. False negatives and positives can arise due to cloning artifacts and may also appear due to a secondary mutation outside the adducted base. In contrast, the use of probe hybridization directly detects any targeted mutation and eliminates the need for additional molecular analysis.

The reduction of vector background required that we devise a new cloning strategy that would eliminate the negative impact of residual amounts of components from previous stages in the vector construction process. In addition, the ideal strategy should be simple and cost effective. With these strategic expectations, I designed and developed a new cloning methodology that facilitated the construction of the single stranded circular vectors containing a single dG-AAF adduct, and a set of relevant, unmodified, isogenic control constructs.

3.2 Development of the SS-DNA vectors carrying a single dG-AAF adduct: The schematic representation of the strategy used to construct the ss-DNA vectors carrying a single acetylaminofluorene adduct (dG-C8-AAF), that were used to obtain the results presented in this thesis, is presented in Figure 2.1. A similar approach was employed for construction of a set of SS-DNA vectors with the same sequence context but without any AAF adduct. The latter served as a positive control to detect spontaneous mutations and also to screen for error-free bypass events at the site of the DNA lesion. Positive control constructs for screening for possible mutagenic bypass events i.e. frameshifts and base substitutions, were also constructed following the same protocol.

Development of an efficient protocol for construction of these vectors required troubleshooting and optimization of each step involved in the process of construction. In the following sections evidence for construction of these vectors, and a brief explanation of the observations made during the development process, will be presented.

3.2.1 Construction of the phagemid SSCpLHII: The polylinker region of the phagemid pUC118 (Viera and Messing, 1987) was replaced with a 49 base pair palindromic *EcoRI-HindIII* linker (Materials and Methods section 2.2.2) that carries two *HincII* sites, one in-frame Opal stop codon and a second Opal stop codon in the +1 reading frame.

The ligation mixture was used to transform *E.coli* DH5 α competent cells; transformants were screened for white colonies on the indicator LB agar plates supplemented with ampicillin and X-gal. Plasmids were isolated from several colonies, and the sequence of the HII linker was verified by sequencing the isolated plasmids. The correct sequence of the cloned fragment is shown in Figure 3.3. The sequence context of

Fig. 3.3: Verification of the sequence of the cloned HII49 in the pLHII construct. The double stranded (ds) and single stranded (ss) forms of pLHII were sequenced using the -35lacZ forward sequencing primer. A compression pattern was observed in the 5'-end *HincII* site, shown in the highlighted blue rectangle.

5'-CAGTCGAC-3', which includes the *HincII* site, showed a compression pattern. In order to verify the accuracy of the sequence, the double stranded and single stranded forms of the selected clone, pLHII, were sequenced using both dITP (inositol triphosphate) and dGTP termination mixtures provided by the Sequenase version 2.0 sequencing kit. The observed compression pattern was clearly resolved and the correct sequence of the cloned HII linker was verified (Fig. 3.4).

3.2.2 Production and purification of the monomeric plus strand of the SSCpLHII:

The production of the single stranded circular form of the plus (viral) strand of the phagemid construct pLHII was achieved by infecting *E.coli* JV30 in early growth phase with helper phage M13KO7 (Viera and Messing, 1987).

The helper phage M13KO7 is an engineered M13 derivative that carries a Kanamycin resistance gene. The phage viral strand origin of replication (Rasched and Oberer, 1986) is mutated so that it can maintain its own replication at a low rate, resulting in a shift in favor of the replication of the phagemid pLHII that carries the intact intergenic region (IG500) from the M13 phage. The helper phage encoded packaging proteins are used to produce phage particles that carry the single stranded circular form of the plus strand of the phagemid. Besides the monomeric form of the packaged phagemid particles, the supernatant of the infected cells usually contains phage particles from the helper phage origin and dimeric or multimeric forms of both phagemid and helper phage. The agarose gel run for the phage DNA, isolated from the supernatant of the infected *E.coli* JV30 cells carrying the pLHII phagemid, shows a major band for the phagemid monomeric plus strand, and some minor bands that represent the helper phage DNA and the multimeric forms of both of the helper phage and phagemid (Fig. 3.5, lane 1).

Fig 3.4: Resolution of the compression pattern of the 5'-*HincII* site of HII cloned into pLHII. The HII region of the SSCpLHII was sequenced using termination solutions containing dGTP (left) and dITP replacing dGTP (right). The compression pattern was partially resolved through a long period electrophoresis (2 hours) in dGTP reaction (highlighted region on the left), while it was completely resolved in the dITP reaction, (highlighted region on the right).

In order to eliminate sources of possible interference in subsequent steps, it was necessary to gel purify the major band corresponding to monomeric SSCpLHII from the agarose gel. Purification of the SSCpLHII band was carried out using the freeze-spin-squeeze method (Material and Methods section 2.2.2.5). The purification method was sufficiently efficient to produce 60-70% yield without any damage to the purified DNA (Fig. 3.5, lane 2).

3.2.3 Digestion of the SSCpLHII: The single stranded linear form of SSCpLHII i.e. SSLpLHII, was prepared by *HincII* digestion of the stable stem and loop structure ($\Delta G = -28.9$ kcal/mole predicted by RNAstructure software, Mathews et al, 1999) that formed through intra-strand annealing of the cloned palindromic sequence. Digestion was carried out twice by incubating the digestion mix at 37°C overnight. The first digestion yielded about 97% single stranded linear form, and this was increased to 99% by the second digestion. The yield was determined by measuring the remaining transformation capacity of the equivalent samples drawn from the digested stocks compared to undigested control. The resulting linear SSLpLHII has slower mobility compared to the circular form and within an hour of electrophoresis at $E = 8$ volts/cm can be resolved on a 0.7% agarose gel as a sharp band lagging behind the SSC form (Fig. 3.5, lane 3). The accuracy of the digestion was verified by sequencing samples from SSC and SSL stocks using the Sequenase kit (Fig. 3.6).

3.2.4 Optimization of cloning of a short oligonucleotide without an AAF adduct:

The efficiency of cloning of a short oligonucleotide into a single stranded vector DNA is dependent on the optimized molar ratios of the single stranded DNA components in the ligation reaction. The optimum relative concentrations of the SSL-vector, scaffold DNA,

Fig. 3.5: Agarose gel electrophoresis, using 0.7% gel, of crude phage and phagemid DNA isolation from *E.coli*JV30(pLHII), infected with helper phage M13KO7(Lane 1); Lane 2, Gel purified SSC form of pLHII viral strand; Lane 3, HincII digested SSL form of pLHII viral strand. Impurities in the crude extract include the phage DNA from the helper phage and the probable multimeric forms of the pLHII (Lane 1).

Fig 3.6 : The HincII digestion of the stem-loop structure, formed by internal annealing of the palindromic HII49 sequence in SSCpLHII, was verified by sequencing SSLpLHII and SSCpLHII. The sequence in SSLpLHII does not extend beyond the HincII digestion site, where a major non instructional A nucleotide, shown in green, is incorporated at the site of digestion.

and the oligonucleotide to be cloned, were determined by preparing ligation mixtures that contained increasing two-fold molar ratios of the scaffold DNA and the short oligonucleotide per unit single stranded vector. A total of four sets of ligation mixtures containing SSLpLHII, SmaG0 scaffold and ^{32}P -labelled SmaG0 oligonucleotide (Vector: Scaffold: oligo), were prepared with molar ratios of 1:1:2, 1:2:4, 1:4:8 and 1:8:16, respectively. A minus enzyme control set equivalent to the first reaction was also prepared. After ligation and removal of the scaffold DNA with anti-scaffold, samples were taken from each set and examined by agarose gel electrophoresis (Fig. 3.7a). When a noticeable resolution between the SSL and SSC bands was achieved, the gel was photographed and then dried under vacuum. The efficiency of incorporation of the ^{32}P -SmaG0 oligonucleotide into the single stranded DNA vector was detected by exposing the dried gel to X-ray film. The results shown in Figure 3.7a, demonstrate no incorporation of the oligonucleotide in the control set. However, incorporation of the ^{32}P -SmaG0 into the SSPLHII results in the appearance of two bands in all four experimental sets. The upper band, which co-migrates with the SSL band, represents the one-tailed, either 3' or 5' end, single ligation event resulting in a final linear ligation product. However, the lower band, which co-migrates with the SSC band, is the product of two-tailed ligation event that resulted in a single stranded circular structure. The presence and persistence of the radioactive signal in these products after removal of the scaffold DNA is direct evidence for the successful ligation reaction (Fig. 3.7b). It is noteworthy that the intensity of the radioactive signal increases from reaction set 4 through 7 with a clear shift toward the SSL band, while less obvious differences are detectable for the intensities of the same bands on agarose gel image. In considering the optimal conditions of cloning,

Fig 3.7: Optimization of cloning of a short ^{32}P labeled oligonucleotide, ^{32}P -SmaG0, into SSLpLHII. Different molar ratios of linear vector, scaffold DNA and ^{32}P -SmaG0, 14-mer oligonucleotide, were mixed and ligated. Panel A, illustrates the electrophoretic separation of the ligation products sampled from each of the reactions along with controls. Lane 1, SSCpLHII; Lane 2, SSLpLHII; Lane 3, reaction mixture without ligase; Lanes 4-7, different molar ratios of vector:scaffold:oligomer i.e. Lane 4, 1:1:2; Lane 5, 1:2:4; Lane 6, 1:4:8 and Lane 7, 1:8:16. Panel B, is the autoradiograph of the gel in panel A that shows the relative incorporation of the ^{32}P -SmaG0 into linear and circular ligation products. Panel C, Schematic representation of the 3' or 5'-end single ligation and the double ligation events, that resulted in formation of SSLpLHIIISmaG0 and SSCpLHIIISmaG0 ligation products, respectively.

two factors were taken into account. First, the productivity of ligation that yields a reasonably high proportion of SSC form, which is the only species that has the capacity to transform the competent cells, hence resulting in reasonably high transformation efficiency. Second, cost effectiveness of the ligation condition, which can be achieved by use of the reasonably low amounts of the ligation components, especially the AAF modified oligonucleotides which are the most expensive components.

In order to compare the transformation efficiencies of each of the ligation sets, samples were taken from each of the ligation sets 4 to 7 and the control set, and used for transformation of the *E.coli* JM103 competent cells. The transformant cells were selected on LB agar plates supplemented with ampicillin. The transformation efficiency values were plotted for each set (Fig. 3.8). As can be seen in the graph the minus enzyme control set produced the least number of the transformants (1.2×10^3 cfu/ μ g), which is due to the presence of the 1% undigested SSCpLHII. The transformation efficiency values increased by 10-fold, 50-fold and 51-fold for the reaction sets 4,5 and 6, respectively, as compared to the control set. The transformation efficiency for the reaction set 7 is decreased to 3.7×10^4 cfu/ μ g which accounts for about 30% reduction in transformation capacity compared to sets 4 and 5. Considering the aforementioned principles for optimal conditions for cloning, the reaction conditions for set 5 (1:2:4) were considered to be optimal for preparation of the stock constructs that carry a cloned oligonucleotide with the AAF modification.

3.2.5 Construction of the single stranded circular phagemid vectors with a single dG-AAF adduct and quantification of the SSC form: Using the determined optimum conditions for cloning, the single stranded circular phagemid vectors, SSCpLHIIISmaG1

Fig 3.8: The restoration of the transformation capacity of ligation mixtures used for optimization of the ligation conditions for cloning of the short oligonucleotide, SmaG0, into SSLpLHII. -E, minus ligase control set; ligation mixtures 1-4 contain varying molar ratios of SSL vector:Scaffold DNA:SmaG0 Oligomer as follows; ligation mix 1, 1:1:2, ligation mix 2, 1:2:4, ligation mix 3, 1:4:8 and ligation mix 4, 1:8:16. The transformation efficiency for each set was measured as the number of *E.coli* JM103 transformants per microgram of total vector DNA from each ligation reaction used in the transformation. Each transformation reaction contained an equivalent amount of the SSL vector used in ligation.

and SSCpLHIISmaG3, were constructed using short 5'-³²P labeled oligonucleotides of 14 nucleotides in length, SmaG1.AAF and SmaG3.AAF, that carry a single dG-C8-AAF adduct at position G1 and G3 of the cloned *SmaI* site, 5'-CCCG₁G₂G₃-3', respectively. The linearized SSCpLHII vector was annealed with the 60-mer HIISmaG0 scaffold DNA and either of the SmaG0, SmaG1.AAF or SmaG3.AAF oligonucleotides (Fig. 3.9, Table 2.1). The annealed assemblies were ligated overnight followed by removal of the scaffold DNA by annealing it with an excess amount of 60-mer anti-scaffold, anti-HIISmaG0 oligonucleotide. The SSCpLHIINarG0 and SSCpLHIINarG3.AAF vectors were prepared in a similar way except that the 15-mer 5'-³²P labeled NarG0 and NarG3.AAF, and 62-mer HIINarG0 scaffold and anti-HIINarG0, anti-scaffold oligonucleotides were used (Fig. 3.9, Table 2.1).

Samples from the three independently constructed stocks of the single stranded vectors carrying a cloned ³²P-labeled oligomer, with and without an AAF modified guanosine residue, were run on an agarose gel until a clear resolution between the SSC and SSL forms was achieved. The bands corresponding to the SSL and SSC forms were excised from the gel, the respective Cerenkov counts were determined, and the concentration of the SSC forms was calculated using the method described in chapter 2.

The relative incorporation of either the unmodified or AAF modified oligonucleotides into the SSL and SSC forms for each of the three stocks was calculated using the corresponding Cerenkov readings. The relative percentage incorporation values for the SSLpLHIISmaG0 and SSCpLHIISmaG3.AAF were calculated to be 43.1±5.6 and 37.0±4.1, respectively. The incorporation of ³²P-SmaG1.AAF into

Fig 3.9: Construction of single stranded circular vectors by cloning a short ^{32}P -labeled oligonucleotide with and without a single dG-AAF adduct. Ia, Ib and Ic are the agarose gel electrophoresis of the constructs carrying SmaG0 and SmaG3.AAF (Ia lanes 3 and 4, respectively); SmaG1.AAF (Ib lane 3); and NarG0 and NarG3.AAF (Ic lanes 3 and 4, respectively). In gels Ia-Ic the SSCpLHII and SSLpLHII controls were included in lanes 1 and 2, respectively. IIa, IIb and IIc are the autoradiographs of the Ia, Ib and Ic gels, respectively, that illustrate the extent of incorporation of the ^{32}P -labelled oligonucleotides into SSL and SSC forms of the ligation products.

SSCpLHIISmaG1.AAF was shown to be $41.9\% \pm 1.0$. In a similar way, the incorporation of ^{32}P -NarG0 and ^{32}P -NarG3.AAF oligonucleotides into SSCpLHIINarG0 and SSCpLHIINarG3.AAF were calculated as $48.8\% \pm 0.7$ and $47.0\% \pm 1.8$, respectively. An analysis of variance for the homogeneity of the incorporation of unmodified and modified oligonucleotides did not produce any significant difference for any of three types of constructs. The statistical analysis results for the SmaG0/G3.AAF pair were computed as $t=1.55$, $t_{\text{critical}}=2.78$ and $P=0.19$; for the SmaG0/G1.AAF pair as $t=0.368$, $t_{\text{critical}}=4.303$, $P=0.748$ and for the NarG0/G3.AAF pair as $t=1.561$, $t_{\text{critical}}=3.182$ and $P=0.216$, respectively (Table 3.1).

3.2.6 Efficiency of removal of scaffold DNA using anti-scaffold DNA: The presence of a short double stranded region formed by annealing the scaffold DNA and the single stranded vector during the ligation reaction is not favorable for two reasons. First, such a structure provides an extendible 3'-end for the polymerases. Second, it can be a substrate for repair enzymes such as nucleotide excision repair. The first problem can be circumvented by including a few mismatching nucleotides at the 3'-end of the scaffold DNA. However, both problems can be solved by complete removal of the scaffold DNA from the ligation product. The removal of the scaffold DNA after ligation can be achieved by either boiling and rapid cooling (Banerjee et al., 1988) or by annealing the scaffold to an oligonucleotide, the anti-scaffold, that contains complementary sequence to the scaffold DNA. The efficiency of both methods was examined by annealing the SSLpLHII to ^{32}P -labelled HIISmaG0 scaffold DNA and 5'-SmaG3.AAF, followed by ligation. One half of the ligation reaction mixture was heated to 95°C for 2 minutes

Oligonucleotide	Incorporation% $\pm$ SD	t_s (t-test)	$t_{0.05[v]}$	Probability
SmaG0	43.1 $\pm$ 5.6	1.559	2.776	0.194
SmaG3.AAF	37.0 $\pm$ 4.1			
SmaG0	43.1 $\pm$ 5.6	0.368	4.303	0.748
SmaG1.AAF	41.9 $\pm$ 1.0			
NarG0	48.8 $\pm$ 0.7	1.561	3.182	0.216
NarG3.AAF	47.0 $\pm$ 1.8			

Table 3.1: Comparison of the efficiency of incorporation of oligonucleotides, with and without the AAF adduct, into single stranded circular constructs. The average percentage of incorporation of each pair of oligonucleotides into the single stranded circular structure, from three independent cloning experiments, were compared using a t-test for comparison of means and variances. t_s , calculated student t-value for each pair of oligonucleotides; $t_{0.05[v]}$, student t-statistic for $P=0.05$ and degrees of freedom of v .

followed by quick cooling on an ice bath, while the second half was annealed to anti-HIISmaG0 oligonucleotide at a molar ratio of 1:10. Samples were examined on an agarose gel in parallel with SSCpLHII and ^{32}P -labeled HIISmaG0. The electrophoresis was carried out for 15 minutes until resolution occurred between the SSC and the ^{32}P -labelled scaffold DNA (Fig. 3.10a). The gel was dried and exposed to X-ray film for autoradiography (Fig. 3.10b). In Figure 3.10a a clear resolution is visible between the SSC band and the scaffold DNA (lanes 1 and 4, respectively). A clear radioactive signal was detected for the SSC band corresponding to the ligation reaction that was subjected to heat denaturation followed by rapid cooling, which implies that the removal of the ^{32}P -labelled scaffold was not complete using this method. However, trapping and removal of the scaffold by annealing it to the anti-scaffold, which can be detected as a single radioactive band at the bottom of lane 3 with no corresponding signal detectable for the SSC band, was pronounced (Fig. 3.10b). One noteworthy observation is that the band corresponding to the scaffold/anti-scaffold radioactive band in lane 3 shows slightly increased relative mobility compared to the scaffold band in lanes 2 and 4, (Fig. 3.10b). This observation can perhaps be explained by the formation of a double helical structure between scaffold and anti-scaffold DNA. A double stranded structure, despite its doubled molecular weight, might be more compact as compared to the single stranded scaffold, resulting in slightly increased relative mobility of the double stranded form.

3.3 Mutagenesis studies using single stranded circular vectors containing an AAF

modified single guanosine: Single stranded circular vectors, containing a single dG-AAF adduct, were transformed into SOS-induced or uninduced *E.coli* JM103 competent

Fig 3.10: The efficiency of removal of the ^{32}P -labeled scaffold DNA by heat denaturation followed by rapid cooling is compared to the use of anti-scaffold DNA. (a) illustrates the agarose gel electrophoresis of the control SSCpLHII (lane 1); sample from SSpLHIISmaG3 ligation mixture was heat treated followed by rapid cooling, or annealed to anti-scaffold (lanes 2 and 3, respectively); ^{32}P -labeled SmaG0 scaffold (lane 4); (b) shows the autoradiograph of the gel in (a).

cells and its *umuC* deletion mutant derivative (i.e. $\Delta umuC$). In addition, the *E.coli* JM103 $\Delta umuC$ derivatives that carried plasmid encoded UmuDC or MucAB functions, were also transformed with the same constructs. In parallel, control transformant collections were also prepared using the single stranded circular constructs that did not contain any AAF modification in the same DNA sequence contexts. The *in vivo* processing of the transformed constructs was studied through analysis of those translesion synthesis events that resulted in either error-free bypasses (EFB), or fixation of mutations as a result of error-prone bypass events (EPB). Each class of bypass events was screened by probe hybridization of a large number of the transformants using sequence specific oligonucleotide probes. The probe hybridization conditions and the level of the stringency of washing of membranes were optimized for each individual probe. Under optimized conditions the radioactive signal from the positive colonies matched that of the positive controls, while no signal was retained by any of the negative controls. Translesion synthesis frequencies were determined by calculating the number of positive signals for each class of translesion event (i.e. wild type, frameshift or base substitution) per nanogram of the modified construct, normalized against the number of positive signals for error free bypass events per nanogram of the construct without modification. The survival or total TLS is defined as the sum of the translesion synthesis frequencies of all classes of bypass event.

3.3.1 Translesion synthesis events past SmaG3.AAF adduct: The single stranded vector pLHIISmaG3.AAF and the control vector SSCpLHIISmaG0 were used to transform the SOS-induced and uninduced *E.coli* JM103, *E.coli* JM103 $\Delta umuC$, *E.coli*

JM103 Δ muc(pGW1700) and *E.coli* JM103 Δ muC(pDC322) strains. The SOS induced and uninduced collections were studied to determine the effect of the cellular background and SOS induction on *in vivo* translesion synthesis past the AAF adduct.

3.3.1.1 The TLS events past SmaG3.AAF adduct in *E.coli* JM103 cellular

background and the effect of SOS induction: The *in vivo* translesion events past the SmaG3.AAF adduct were studied by introducing the SSCpLHIISmaG3.AAF construct into SOS-induced or uninduced wild type *E.coli* JM103 competent cells. In parallel, the SSCpLHIISmaG0 construct, that lacks the AAF modification in the same sequence context, was also transformed into the same strain. Transformations were carried out in triplicate, and 5-8 platings from each transformation experiment were carried out on selective LB agar media supplemented with ampicillin. The survival of the transformed cells on the selective medium is dependent on complete replication of the ssDNA plasmid, which can only be achieved by successful replication past the AAF adduct in the cells transformed with the construct containing an AAF modification. In the control set, replication through the same sequence context does not encounter any impedance. The relative frequencies of the synthesis past the AAF adduct, per nanogram of the construct, compared to the control construct were recorded as translesion synthesis frequencies i.e. TLS. The averaged *in vivo* TLS frequencies for each class of bypass event, in SOS induced and uninduced *E.coli* JM103 cells, are shown in Table 3.2. A collection was made by transferring a total of 14,835 colonies of the SOS uninduced *E.coli* JM103 cells transformed with SSCpLHIISmaG3.AAF construct onto hybridization membranes. A similar collection was also prepared by transferring a total of 9,555 colonies of the SOS induced *E.coli* JM103 cells (UV-irradiated at 70 J/m²) transformed with the same

construct. A total of 4,521 independent colonies from triplicate SOS⁻ collections and 2,270 colonies from the triplicate SOS⁺ collections were screened for the error-free bypass events through probe hybridization using the SmaG0 probe (Table 2.1). The error-free synthesis past the AAF adduct was the most prevalent event that accounted for 13.7% and 29.1% of synthesis past the control set in SOS uninduced and SOS induced cells, respectively. Induction of SOS functions increased the propensity of the cellular functions to bypass the adduct, without any error, 2.1-fold more efficiently as compared to uninduced cells. Analysis of variance for comparison of the means and variances of the respective triplicate results for each data set produced a student “t” statistic value of 10.73, that was significantly different from the critical value of 3.16 at 0.05 confidence limit. The probability of the homogeneity of the data from SOS induced and SOS uninduced populations was calculated as $P=0.0009$, which is clearly much less than the confidence limit of 0.05 considered in the test. As a result, it is plausible to assume that the induction of the SOS functions generated an inherent capacity to bypass the replicational block formed by the AAF, without making any error at a level that was significantly different from uninduced cells.

Error-prone bypass events were detected as targeted - G₃ frameshifts or base substitution mutations. A total of 4,865 and 2,415 colonies from the SOS induced and uninduced collections transformed with SSCpLHIISmaG3.AAF, respectively, were screened for -G₃ frameshift bypass events by probe hybridization using the Sma-1 probe (Table 2.1). Among the targeted mutations, the minus G₃ frameshifts comprise the majority of the error prone TLS events, with TLS frequencies of $1.78\% \pm 0.39$ and $5.2\% \pm 0.52$ of the unmodified control construct in SOS uninduced and SOS induced cells,

TLS Event	TLS frequency(%)		Pi(t-test)	-2LnPi
	SOS ⁻	SOS ⁺		
EFB	13.7±1.02 (4521) *	29.1±2.26 (2270)	0.0008655	14.104
-G3	1.78±0.39 (4865)	5.2±0.52 (2415)	0.0003961	15.668
G3→T	0.06±0.04 (5139)	0.11±0.05 (2288)	0.1285337	4.104
G3→A	0.01± 0.01 (4865)	0.06±0.01 (2415)	0.0151322	8.382
G3→C	0.026±0.001 (9386)	0.16±0.02 (4794)	0.0060428	10.218
Survival	15.6±0.97	34.6±2.7	0.0007115	14.496
				-2ΣLnPi= 66.972 χ²_{0.05[12]= 21.026} P< 0.001

Table 3.2: Translesion synthesis frequencies past the SmaG3.AAF adduct in SOS-induced and uninduced *E.coli* JM103 strain. Pi, probabilities from t-test for the effect of SOS induction per class of TLS; -2ΣLnPi, statistic for combined probabilities Chi-square test.

* Numbers in parentheses are the total number of transformants, from triplicate experiments, screened by probe hybridization, using sequence specific probes. In some cases membranes were stripped off the probes and used for hybridization with a second probe.

respectively (Table 3.2). The induction of the SOS functions appeared to have a significant impact on promoting the minus G₃ frameshifts ($P=0.0004$) as determined by a t-test comparing the means of the data sets for SOS induced and uninduced populations.

Base substitutions occurred about 30-100 fold less frequently than targeted frameshifts. The TLS frequencies for three classes of base substitution mutations in SOS uninduced and SOS induced cells were determined by probe hybridization using probes specific to each base substitution event. A total of 5,139 colonies from the triplicate SOS uninduced collections and 2,288 colonies from the triplicate SOS induced collections (Table 3.2), were screened for G₃→T base substitution events using the SmaT probe, (Table 2.1). The TLS frequency values for G₃→T bypasses for SOS⁻ and SOS⁺ cells were $0.06\% \pm 0.05$ and $0.11\% \pm 0.05$ of total synthesis through the unmodified control construct. The t-test for comparison of the means did not produce a significant difference between the SOS induced and uninduced conditions ($P=0.13$). However, screening of 9,386 colonies from the SOS uninduced collection and 4,794 colonies from the SOS induced collection for G₃→C base substitution events, by probe hybridization using the SmaC probe, resulted in TLS values of $0.026\% \pm 0.006$ and $0.16\% \pm 0.02$ of total synthesis through unmodified control construct. The statistical test for analysis of variance showed a significant difference between the SOS⁻ and SOS⁺ collections ($P=0.0060$). In addition, screening the 4,865 colonies from the SOS uninduced collection and 2,415 colonies from the SOS⁺ collection for G₃→A base substitutions showed 0.01 ± 0.01 and 0.06 ± 0.01 percent bypass events compared to unmodified construct. The G₃→A mutations are the most infrequent bypass events and the TLS frequency values may be prone to error due to

insufficient sampling. However, the comparison of the means for the SOS⁻ and SOS⁺ collections showed a significant difference between these collections (P=0.015), (Table 3.2).

The overall TLS or survival rate represents the sum of the TLS frequencies corresponding to the error-free and error-prone bypass events past the AAF adduct. The total survival for the SOS⁻ and SOS⁺ collections were calculated as 15.6%±0.97 and 34.62%±2.7 of the synthesis through unmodified construct, respectively. The t-test for comparison of the means of the survival rates of these two collections revealed that the induction of the SOS functions promotes a significant increased survival of the modified vector in SOS⁺ bacteria compared to that of the SOS uninduced bacteria (P=0.0007), (Table 3.2).

A combined probabilities Chi-square test was used to examine the combined probabilities of all possible classes of translesion synthesis events among SOS induced and uninduced collections. The combined probabilities Chi-square test is considered as a complementary and conclusive statistical test for cases in which a series of different significance tests on different sets of data have tested the same hypothesis i.e. effect of SOS induction, (Sokal and Rohlf, 1981). If some of these tests produce probabilities that are not low enough to be suggestive of statistical significance, then combining the individual probabilities of all tests, as $-2\sum \ln P_i$, generates a new statistic that is distributed as a χ^2 statistic with degrees of freedom equal to twice the number of the preliminary tests (2K). Applying the combined probabilities test to the EFB, -G3, each of base substitutions and survival rates of SOS⁻ and SOS⁺ collections resulted in -

$2\sum \ln P_i = 66.972$ that corresponds to $P < 0.001$, which is significantly small, compared to a critical value of $\chi^2_{0.05[12]} = 21.026$, to reject the null hypothesis of homogeneity of the data in SOS^- and SOS^+ collections (Table 3.2).

3.3.1.2 The TLS events past SmaG3.AAF adduct in *E.coli* JM103 $\Delta umuC$ strain and

the effect of SOS induction: The chromosomal *umuC* gene in the host strain, *E.coli* JM103 $\Delta umuC$, is inactivated by P1 transduction of a deletion mutant of *umuC* gene that replaced the chromosomal gene. The SOS induced (UV irradiated at 8 joules /m²) and uninduced competent cells of this strain were transformed with SSCpLHIISmaG3.AAF and the control construct, SSCpLHIISmaG0. The transformed cells were selected on ampicillin plates. The translesion synthesis events were studied by probe hybridization in a manner similar to that described in section 3.2.1.1. A total of 17,378 colonies of SOS uninduced and 8,682 colonies of SOS induced *E.coli* JM103 $\Delta umuC$ transformants, obtained in triplicate independent experiments with AAF modified single stranded vector, were transferred onto hybridization membranes.

Error-free bypass past the AAF adduct comprises the majority of the translesion synthesis events among the SOS induced and uninduced collections in *E.coli* JM103 $\Delta umuC$. A total of 5,624 colonies from SOS^- and 2,665 colonies from SOS^+ collections, were analyzed with the SmaG0 probe. The error-free TLS frequencies of $0.95\% \pm 0.06$ and $1.39\% \pm 0.27$, were calculated for SOS uninduced and SOS induced cells, respectively. This constitutes a 14 and 20 fold decline in the capacity of the SOS uninduced and induced $\Delta umuC$ strain for error-free bypass of the AAF adduct, compared to the wild type SOS uninduced and induced strain, respectively. In JM103 $\Delta umuC$

TLS Event	TLS frequency (%)		Pi(t-test)	-2LnPi
	SOS ⁻	SOS ⁺		
EFB	0.95±0.06 (5624)*	1.39±0.27 (2665)	0.055899	5.768
-G3	0.32±0.05 (5571)	0.5±0.09 (2796)	0.029264	7.062
G3→T	0.002±0.0004 (11807)	0.013±0.005 (3222)	0.0393057	6.472
G3→A	<0.001 (5571)	<0.001 (2796)	NA	NA
G3→C	0.001±0.001 (11195)	0.004±0.003 (5845)	0.1289901	4.096
Survival	1.28±0.11	1.90±0.36	0.0515973	5.928
				-2ΣLnPi= 29.326 χ²_{0.05[10]= 21.026} P< 0.005

Table 3.3: : Translesion synthesis frequencies past the SmaG3.AAF adduct in SOS-induced and uninduced *E.coli* JM103 $\Delta umuC$. Pi, probabilities from the t-test for the effect of SOS induction per class of TLS; -2ΣLnPi, statistic for combined probabilities Chi-square test.

* Numbers in parentheses are the total number of transformants, from triplicate experiments, screened by probe hybridization, using sequence specific probes. In some cases membranes were stripped off the probes and used for hybridization with a second probe.

induction of cellular SOS functions appeared to have a weak impact on facilitating the cellular mechanisms to bypass the AAF adduct ($P=0.056$), (Table 3.3).

A total of 5,571 SOS uninduced and 2,796 colonies from the SOS induced JM103 *ΔumuC* collections, sampled in triplicate experiments, were screened for - G3 frameshift mutations through probe hybridization using the Sma-1 probe. The minus G3 frameshifts comprise the major class of mutational bypass events, with the TLS frequency values of $0.32\% \pm 0.05$ and $0.5\% \pm 0.09$ in SOS⁻ and SOS⁺ collections, respectively. Statistical analysis of the means and variances for minus G3 frameshift TLS frequency values in SOS⁻ and SOS⁺ collections showed a moderate and marginally significant difference between these two collections ($P=0.03$), (Table 3.3). The capacity of the *ΔumuC* strain to bypass the AAF adduct through -G3 frameshift events is reduced by a factor of 6-10 fold compared to the SOS uninduced and induced wild type strains, respectively.

Base substitution mutations comprise the second, and less frequent group of the bypass events that occurred during bypass of the AAF adduct. A total of 11,807 clones from the SOS uninduced collection and 3,222 clones of the SOS induced collection were screened for G→T transversion mutations. Translesion synthesis past the AAF adduct in SOS uninduced cells resulted in a TLS frequency of $0.002\% \pm 0.0004$. However, the TLS frequency in SOS induced cells was elevated to $0.013\% \pm 0.006$. These values are reduced by a factor of 8-30 fold compared to the wild type background. The statistical significance test for means and variances of the SOS⁻ and SOS⁺ collections showed a marginal difference among these two collections ($P= 0.04$), (Table 3.3).

Screening for G→A base substitutions among 5,571 and 2,796 clones of uninduced and induced collections, respectively, did not produce any detectable signal. However, screening of 11,195 clones from the SOS⁻ collection and 5,845 clones from the SOS⁺ collection for G₃→C transversions resulted in calculated TLS frequencies of 0.001%±0.001 and 0.004%±0.003, respectively, (Table 3.3). The apparent increase in TLS for G₃→C of 4-fold in SOS-induced versus SOS uninduced cells was not significant (P=0.129), (Table 3.3).

The survival of the SOS uninduced collection was calculated as 1.28%±0.11, that was increased slightly in SOS induced cells to 1.90%±0.36, (Table 3.3). This represents a 12 and 18-fold reduction, respectively, in survival in *E.coli* JM103 Δ *umuC* as compared to the SOS⁻ and SOS⁺ wild type collections. The t-test for comparison of the means of survival rates for SOS⁻ and SOS⁺ collections resulted in a probability of P=0.051 that was of borderline significance. Statistical analysis of the combined probabilities generated a $-2\sum \ln P_i = 29.326$ value, that corresponds to P<0.005, which was taken as an indication of the heterogeneity of the SOS⁻ and SOS⁺ collections, (Table 3.3).

3.3.1.3 TLS events past SmaG3.AAF in *E.coli* JM103 Δ *umuC*(pDC322) and the

effect of SOS induction: The *E.coli* JM103 Δ *umuC* strain was transformed with a low copy number pDC322 construct that carries a PCR fragment, amplified from plasmid pSE117 (Marsh and Walker , 1985), containing the *E.coli umuDC* operon cloned into the ampicillin resistance gene of pBR322, (Materials and Methods 2.2.10). The resulting strain is able to encode for UmuD and UmuC proteins extrachromosomally using its natural LexA repressible promoter. The SOS uninduced and induced (UV irradiated at 70

Joules/m²) competent cells of *E.coli* JM103 $\Delta umuC$ (pDC322) were transformed, in triplicate, with SSCpLHIISmaG3.AAF and the control construct SSCpLHIISmaG0 . Transformants were selected on ampicillin plates and sampled for screening. A total of 8,209 colonies from the SOS⁻ collection, sampled from triplicate experiments, and 672 colonies from the SOS⁺ collection, were transferred onto hybridization membranes. The membranes were screened for error-free and minus G3 frameshift mutation through probe hybridization using SmaG0 and Sma-1 probes, respectively (Table 3.4).

A total of 4,047 clones from the SOS⁻ cells transformed with modified vector and 672 colonies from triplicate SOS⁺ collection were screened for error-free bypass events. The TLS frequencies of error-free bypasses past the AAF modification were calculated as 16.5%±0.6 and 42.3%±1.3 for SOS⁻ and SOS⁺ collections, respectively, (Table 3.4). The comparison of the means and variances between SOS⁻ and SOS⁺ collections showed a statistically significant difference (P=0.00004), (Table 3.4).

Minus G3 frameshifts were screened using 4,162 colonies from SOS⁻, and 672 colonies from SOS⁺ collections. The clones were obtained in triplicate transformation experiments. The -1 frameshift TLS frequencies of 2.4%±0.26 and 4.9%±0.6, were calculated for SOS⁻ and SOS⁺ collections, respectively. The statistical significance test for the minus G3 frameshifts among SOS⁻ and SOS⁺ collections suggested significant heterogeneity (P=0.002). The total translesion synthesis frequencies for SOS⁻ and SOS⁺ collections were calculated as 18.9%±0.86 and 47.2%±1.7, respectively. Total bypass, both error-free and error-prone, past the AAF lesion in the SOS induced cells was significantly enhanced as compared to SOS uninduced cells (P=0.00006). In addition, the

TLS Event	TLS frequency (%)		P _i (t-test)	-2LnP _i
	SOS ⁻	SOS ⁺		
EFB	16.5±0.6 (4047)*	42.3±1.3 (672)	0.00003796	20.358
-G3	2.4±0.45 (4162)	4.9±0.6 (672)	0.0027817	11.768
Survival	18.9±0.86	47.2±1.6	0.00006484	19.288
				-2ΣLnP _i = 51.414 χ ² _{0.05[6]} = 12.592 P< 0.001

Table 3.4: Translesion synthesis frequencies past the SmaG3.AAF adduct in SOS-induced and uninduced *E.coli* JM103 $\Delta umuC$ (pDC322). P_i, probabilities from t-test for the effect of SOS induction per class of TLS; -2ΣLnP_i, statistic for combined probabilities Chi-square test.

* Numbers in parentheses are the total number of transformants, from triplicate experiments, screened by probe hybridization, using sequence specific probes. In some cases membranes were stripped off the probes and used for hybridization with a second probe.

combined probability Chi-square test for EFB , minus G3 and survival rate, produced a $-2\sum \ln P_i = 51.414$ that is much higher than the critical value of $\chi^2_{0.05[6]} = 12.592$, which indicates that the overall difference among SOS⁻ and SOS⁺ collections was significant with a probability of $P < 0.001$, (Table 3.4).

The *E.coli* JM103 $\Delta umuC$ (pDC322) strain contains multiple copies of the pDC322 carrying the cloned *umuDC* operon. The presence of the multiple copies of the operon in this strain increases the intracellular copy numbers of the plasmid encoded UmuD and UmuC. In order to investigate the effect of expression of UmuD/C functions on the translesion synthesis profile, the TLS frequencies of EFB, -G3 frameshift events and survival rates in the $\Delta umuC$ strain were compared to that of the $\Delta umuC$ /pDC322 strain. The t-test for comparison of the means of the TLS frequencies corresponding to error-free bypasses, minus G3 frameshifts and survival rates, in SOS⁻ collections of $\Delta umuC$ and $\Delta umuC$ /pDC322 strains, revealed that the TLS events in $\Delta umuC$ /pDC322 strain were increased significantly as compared to the $\Delta umuC$ strain with probability values of 0.0002, 0.003 and 0.0004, respectively. A similar test performed among the SOS⁺ collections of $\Delta umuC$ and $\Delta umuC$ /pDC322 strains suggested that the TLS frequencies for EFB, minus G3 frameshifts and survival rates in the $\Delta umuDC$ /pDC322 strain were significantly different from those of the $\Delta umuC$ strain with probabilities of 0.0002, 0.003 and 0.0002, respectively.

The combined probabilities Chi-square test for error-free bypasses, minus G3 frameshifts and survival rates in SOS⁻ collections of $\Delta umuC$ and $\Delta umuC$ /pDC322 strains, resulted in $-2\sum \ln P_i = 69.935$, a value that was significantly higher than the critical Chi-

square value of $\chi^2_{0.05[6]}=12.592$. This fact suggests that the presence of the plasmid pDC322 in $\Delta umuC$ /pDC322 strain caused increased TLS past the AAF adduct, that was significantly different from wild type strain ($P<0.001$). Performing the combined probabilities Chi-square test for TLS events, including the -G3 frameshifts, EFB and survival rates in SOS⁺ collections of the $\Delta umuC$ and $\Delta umuC$ /pDC322 strains, showed a significant difference between these strains ($-2\sum \ln P_i=71.7$ compared to $\chi^2_{0.05[6]}=12.592$) with a probability of $P<0.001$. The overall result of all statistical tests point to the fact that the episomal production of the *umuD/C* functions restore the TLS capacity to the $\Delta umuC$ strain to bypass the AAF adduct through error-free and mutagenic frameshift pathway to the levels comparable to the wild type strain.

3.3.1.4 TLS events past SmaG3.AAF in *E.coli* JM103 $\Delta umuC$ (pGW1700) and the

effect of SOS induction: The *E.coli* strain JM103 $\Delta umuC$ was transformed with plasmid pGW1700, a pBR322 derivative, that carries a fragment containing the *mucAB* operon from pKM101, cloned into the Amp^r gene of plasmid pBR322 (Perry and Walker, 1982). The SOS induced and uninduced competent cells of the *E.coli* JM103 $\Delta umuC$ (pGW1700) were transformed with SSCpLHIISmaG3.AAF and the control construct SSCpLHIISmaG0. Transformants were selected on ampicillin plates. Multiple platings were done for each experiment, which were carried out in triplicate. A total of 12,920 and 9,193 colonies from the SOS⁻ and SOS⁺ *E.coli* JM103 $\Delta umuC$ (pGW1700) transformed with the modified construct, were transferred onto the hybridization membranes. The TLS events past the SmaG3.AAF adduct were examined by probe hybridization using sequence specific probes.

A total of 2,912 and 1,590 colonies were screened for error-free bypass events among the SOS⁻ and SOS⁺ collections, respectively, by probe hybridization using the SmaG0 probe. The error-free TLS frequencies for SOS⁻ and SOS⁺ collections were calculated as 21.5%±2.4 and 34.5%±3.2, respectively. The increased EFB events among the SOS⁺ collection compared to the SOS⁻ collection appeared to be statistically significant using the t-test for comparison of the respective mean values (P=0.002), (Table 3.5).

In *E.coli* JM103 $\Delta umuC$ (pGW1700) the -G3 frameshifts comprise the main class of mutational bypass events past the AAF adduct. The screening of a total of 3,110 colonies from SOS⁻ and 1,853 colonies from SOS⁺ collections for -G3 frameshifts, resulted in TLS frequencies of 5.09%±0.4 and 7.4%±1.1, respectively. The t-test for comparison of the means of the frameshifts, showed a modest but significant difference between SOS⁺ and SOS⁻ collections (P=0.02), (Table 3.5). Base substitutions among a total of 12,605 transformants from the triplicate SOS⁻ and 8,598 transformants from the triplicate SOS⁺ collections were examined by probe hybridization. The G3→T transversions were screened using 3,320 and 3,107 colonies from the triplicate SOS⁻ and the SOS⁺ collections, respectively. The TLS frequencies of 0.44%±0.16 and 0.89%±0.08 were calculated for SOS⁻ and SOS⁺ collections, respectively. The statistical analysis revealed a significant increase in the SOS⁺ collection compared to the SOS⁻ collection (P=0.01). Screening for G3→A transitions using 3,110 colonies from SOS⁻ collection and 1,853 colonies from the SOS⁺ collections, did not produce any signal. However, screening for G3→C transversions among 6,175 colonies from the SOS⁻ collection and

TLS Event	<u>TLS frequency (%)</u>		Pi(t-test)	-2LnPi
	SOS ⁻	SOS ⁺		
EFB	21.5±2.42 (2912)*	34.5±3.2 (1590)	0.0024337	12.036
-G3	5.09±0.4 (3110)	7.4±1.1 (1853)	0.0211288	7.712
G3→T	0.44±0.16 (3320)	0.89±0.08 (3107)	0.0105610	9.102
G3→A	<0.001 (3110)	<0.001 (1853)	NA	NA
G3→C	0.24±0.03 (6175)	0.25±0.1 (3638)	0.4260779	1.706
Survival	27.2±2.71	43.1±4.4	0.0068875	11.694
				-2ΣLnPi= 42.25 $\chi^2_{0.05[10]}= 18.307$ P< 0.001

Table 3.5: Translesion synthesis frequencies past the SmaG3.AAF adduct in SOS-induced and uninduced *E.coli* JM103 $\Delta umuC$ (pGW1700) strain. Pi, probabilities from t-test for the effect of SOS induction per class of TLS; -2ΣLnPi, statistic for the combined probabilities Chi-square test.

* Numbers in parentheses are the total number of transformants, from triplicate experiments, screened by probe hybridization, using sequence specific probes.

3,638 colonies from the SOS⁺ collections, resulted in TLS frequencies of 0.24%±0.03 and 0.25%±0.1, respectively. The t-test for comparison of the means of the G3→C between SOS⁺ and SOS⁻ collections did not show any significant difference (P=0.43).

Comparison of the total TLS events among SOS⁻ and SOS⁺ collections, 27.2%±2.7 and 43.1%±4.4, respectively, showed a significant difference (P=0.007).

Despite the insignificant difference found for G3→C transversions, however, performing the combined probabilities Chi-square test resulted in $-2\sum \ln P_i = 42.25$ that was significantly higher than the critical value of $\chi^2_{0.05[10]} = 18.307$, with a probability of $P < 0.001$ (Table 3.5).

Expression of MucA and MucB functions from plasmid has restored the TLS capacity to the $\Delta umuC$ strain. The significance of the effect of expression of the MucA/B functions on the translesion synthesis past dG₃-AAF adduct was tested by comparing the TLS frequencies of EFB, mutagenic -G3 frameshift and base substitution events, and survival rates to those of the $\Delta umuC$ strain. The t-test for comparison of the means of the TLS frequencies corresponding to error-free bypasses, minus G3 frameshifts, G→T, G→C and survival rates, in SOS⁻ collections of $\Delta umuC$ and $\Delta umuC/pGW1700$ strains, revealed that the TLS events in $\Delta umuC/pGW1700$ strain were increased significantly as compared to the $\Delta umuC$ strain with probability values of 0.002, 0.001, 0.007, 0.045 and 0.002, respectively. Similarly, among the SOS⁺ collections of $\Delta umuC/pGW1700$ the TLS frequencies for EFB, minus G3 frameshifts, G→T, G→C and survival rates were significantly different from those of the $\Delta umuC$ strain with probabilities of 0.001, 0.004, 0.045, 0.018 and 0.002, respectively.

The combined probabilities Chi-square test for error-free bypasses, minus G3 frameshifts, G→T, G→C and survival rates in SOS⁻ collections of *ΔumuC* and *ΔumuC/pGW1700* strains, resulted in $-2\sum \ln P_i = 97.524$, a value that was significantly higher than the critical Chi-square value of $\chi^2_{0.05[10]} = 18.307$. This might suggest that the background expression of MucA/B was the cause for a significant increased TLS past the AAF adduct as compared to that of the *ΔumuC* strain ($P < 0.001$). Performing the combined probabilities Chi-square test for TLS events in SOS⁺ collections of the *ΔumuC* and *ΔumuC/pGW1700* strains, showed a significant difference between these strains ($-2\sum \ln P_i = 97.963$ compared to $\chi^2_{0.05[10]} = 18.307$) with a probability of $P < 0.001$. In a similar manner to the effect of the UmuD/C expression from pDC322, the episomal expression of the MucAB functions was able to restore the overall TLS capacity to the *ΔumuC* strain to bypass the AAF adduct through error-free and mutagenic pathways to the levels even exceeding the wild type strain.

3.3.2 Targeted G1→T and semi-targeted minus C frameshift events past

SmaG1.AAF in *E.coli* JM103 *ΔumuC*(pGW1700) and the effect of SOS induction: In the previous section it was shown that G₃→T mutation was the main class of base substitution induced in the presence of the MucAB functions. In order to address the possibility that the position of the AAF modification within a contiguous run of guanine residues may affect the frequency of induced G→T transversions, the single stranded circular vector SSCpLHIISmaG1.AAF was constructed, which contained the AAF adduct at the G1 position of the 5'-CCCG₁G₂G₃-3' sequence. The SSCpLHIISmaG1.AAF construct, and the control construct SSCpLHIISmaG0, were used to transform the SOS

induced and uninduced *E.coli* JM103 $\Delta umuC$ (pGW1700) competent cells in triplicate experiments. The transformants were selected on ampicillin plates. Multiple samplings were made from each of triplicate transformation experiments by transferring a total of 10,670 and 5,363 transformants from the SOS⁻ and the SOS⁺ collections, respectively, onto the hybridization membranes. The TLS events past the G1.AAF adduct were examined by hybridization, using event specific probes.

The error free bypass events were screened, using the SmaG0 probe (Table 2.1), among a total of 3,379 and 1,019 transformants from the SOS⁻ and SOS⁺ collections, respectively. The error-free translesion synthesis frequencies of $7.68\% \pm 1.1$ and $38.1\% \pm 3.3$ past the G1.AAF adduct were obtained for the SOS uninduced and induced cells, respectively (Table 3.6). The t-test for the effect of the SOS induction on enhancing the error-free bypass events, using means and variances of the TLS values for SOS⁻ and SOS⁺ collections, showed a significant difference between these collections ($P=0.002$), (Table 3.6).

The G1→T mutational bypass past the SmaG1.AAF adduct was examined among 7,291 transformants from the SOS⁻ and 4,658 transformants from the SOS⁺ collections by probe hybridization using the SmaG1T probe (Table 2.1). The TLS frequencies for G1→T transversions among the SOS⁻ and SOS⁺ collections were calculated as $0.046\% \pm 0.006$ and $0.23\% \pm 0.04$, respectively. The t-test for comparison of the means of the G1→T TLS frequencies among the SOS⁻ and SOS⁺ collections showed a significantly increased G1→T bypass among the SOS⁺ collection ($P=0.007$) (Table 3.6). However, the TLS frequencies for G→T induced by G1.AAF adduct were reduced by 9.6 and 3.9-

fold in the SOS⁻ and SOS⁺ cells, respectively, as compared to the G3.AAF adduct. The increased TLS frequencies of G3→T events compared to that of the G1→T events was significant in both the SOS⁻ and SOS⁺ cells, with probability values of 0.02 and 0.0005, respectively.

The *umuDC* dependent induction of semi-targeted -C frameshift events in the run of contiguous cytosine residues situated 5' to a G^{AAF} adduct in the 5'-CCCG₁G₂G₃-3' sequence, has been documented using double stranded vectors carrying an AAF modification at the G1 position of the same sequence context (Lambert *et al.*, 1992). In order to investigate whether MucAB dependent semi-targeted -C frameshift events are induced by G1.AAF adducts in ssDNA vectors, the SOS induced and uninduced collections of SSCpLHIISmaG1.AAF-transformed *E.coli* JM103 $\Delta umuC$ (pGW1700) cells were examined by probe hybridization using the SmaC1 probe (Table 2.1). A total of 10,670 and 5,363 transformants from the SOS⁻ and SOS⁺ collections, respectively, were screened for semi-targeted -C frameshift events. The TLS frequencies for semi-targeted -C frameshifts were calculated as 0.013%±0.001 and 0.12%±0.02 for the SOS⁻ and the SOS⁺ collections, respectively. The statistical analysis for comparison of the means of the TLS frequencies for semi-targeted -C frameshifts corresponding to the SOS⁻ and SOS⁺ collections, showed a significant enhancement among the SOS⁺ collection compared to the SOS⁻ collection (P=0.006), (Table 3.6).

The survival of the single stranded vector carrying a single SmaG1.AAF adduct among the SOS⁻ and SOS⁺ collections was determined as 7.74%±1.1 and 38.5%±3.4, respectively. The increased survival of the modified vector in the SOS⁺ collection

TLS Event	TLS frequency (%)		P_i (t-test)	$-2\text{Ln}P_i$
	SOS ⁻	SOS ⁺		
EFB	7.68±1.1 (3379) *	38.1 ±3.3 (1019)	0.002222	12.218
G1→T	0.046±0.006 (7291)	0.23±0.04 (4658)	0.006871	9.960
-C1	0.013±0.001 (10670)	0.12±0.02 (5363)	0.006566	10.516
Survival	7.7±1.1	38.5±3.4	0.0022016	12.236
				$-2\sum \text{Ln}P_i = 44.93$ $\chi^2_{0.05[8]} = 15.507$ $P < 0.001$

Table 3.6: Translesion synthesis frequencies past the SmaG1.AAF adduct in SOS induced and uninduced *E.coli* strain JM103 $\Delta umuC$ (pGW1700). P_i , probabilities from t-test for the effect of SOS induction per class of TLS; $-2\sum \text{Ln}P_i$, statistic for combined probabilities Chi-square test.

* Numbers in parentheses are the total number of transformants, from triplicate experiments, examined by probe hybridization, using sequence specific probes.

compared to that of SOS⁻ collection was statistically significant (P=0.002), (Table 3.6).

A statistical test for the combined probabilities Chi-square test using the probability values for the effect of the SOS induction on EFB, minus C frameshift and G1→T produced a value of $-2\sum \ln P_i = 44.93$ that was significantly higher than the critical value of $\chi^2_{0.05[8]} = 15.507$. This indicates an overall significant effect of the MucAB functions in promoting bypass events past the G1.AAF adduct (P<0.001), (Table 3.6).

3.3.3 Translesion synthesis events past NarG3.AAF adduct: The single stranded vector pLHIINarG3.AAF, containing a single acetylaminoflourene adduct at position G3 of the NarI sequence located in NarG3 synthetic oligonucleotide, 5' ATCACCG₁G₂CG₃(AAF)CCACA 3', was constructed by cloning the modified oligonucleotide into the single stranded linear pLHII vector. In parallel, a control vector containing a cloned NarG0 synthetic oligonucleotide with an identical sequence but lacking any modification, was also constructed. The translesion synthesis events past the NarG3.AAF adduct, and the effect of the SOS induction on the translesion bypass in different cellular backgrounds, were studied by transforming the SOS induced and uninduced competent cells of the *E.coli* strains JM103, JM103 $\Delta umuC$, JM103 $\Delta umuc$ (pGW1700) and JM103 $\Delta umuC$ (pDC322). The SOS⁻ and SOS⁺ collections were prepared for each strain transformed with single stranded circular vectors, and the collections were screened for error-free and mutagenic translesion synthesis events by probe hybridization using sequence specific oligonucleotide probes.

3.3.3.1 Translesion synthesis events past NarG3.AAF in *E.coli* JM103 and the effect of SOS induction: The SOS induced and uninduced competent cells of *E.coli* JM103 wild type strain were transformed, in triplicate, with the SSCpLHIINarG3.AAF and the control construct SSCpLHIINarG0 vectors. The transformants were selected on ampicillin plates. The SOS⁻ and SOS⁺ collections were prepared by transferring a total of 2,930 and 2,556 transformants, respectively, onto the hybridization membranes.

The error-free bypass events were studied in the SOS⁻ and SOS⁺ collections by hybridization screening of 1,578 and 1,317 colonies, using the NarG0 probe (Table 2.1). The error-free translesion frequencies of $3.0\% \pm 0.2$ and $13.8\% \pm 1.4$ were obtained for the SOS⁻ and SOS⁺ collections, respectively (Table 3.7). The induction of the SOS functions appeared to have a statistically significant effect on the enhancement of error-free translesion synthesis ($P=0.003$) (Table 3.7). Screening of a total of 1,359 colonies from the SOS⁻ and 1,239 colonies from the SOS⁺ collections, using the Nar-2 probe (Table 2.1), resulted in -CG3 frameshift TLS frequencies of $1.64\% \pm 0.29$ and $5.88\% \pm 0.24$, respectively. The t-test for comparing the means of the TLS frequencies of -CG3 deletions among the SOS⁻ and SOS⁺ collections showed a significantly increased capacity of the SOS⁺ bacteria to bypass the AAF adduct ($P=0.00002$). The survival rate in the SOS⁻ collection was determined as $4.65\% \pm 0.51$; however, the survival among the SOS⁺ collection was elevated to $19.7\% \pm 1.4$, which was significantly different from SOS⁻ collection ($P=0.0002$). The overall effect of the induction of SOS functions in increasing bypass of the NarG3.AAF adduct, as compared to SOS⁻ collection, was tested using a

combined probabilities Chi-square test. The $-2\sum \ln P_i$ value of 50.28 was significantly higher than the critical value of $\chi^2_{0.05[6]}=12.592$, with a probability of $P<0.001$.

3.3.3.2 TLS events past NarG3.AAF adduct in *E.coli* JM103 $\Delta umuC$ and the effect

of SOS induction: The TLS events past the NarG3.AAF adduct in the absence of the chromosomally encoded *umuC* function were studied by transformation of the SOS induced and uninduced competent cells of *E.coli* JM103 $\Delta umuC$, with SSCpLHIINarG3.AAF and SSCpLHIINarG0 constructs. All experiments were done in triplicate. The transformed cells were selected on ampicillin plates. The SOS⁻ and SOS⁺ collections of the cells transformed with SSCpLHIINarG3.AAF were prepared by multiple samplings from each of the triplicate transformations; a total of 2,894 and 3,230 colonies from the SOS⁻ and SOS⁺ collections, respectively, were transferred onto hybridization membranes.

The error-free TLS events past the NarG3.AAF adduct among SOS⁻ and SOS⁺ collections, were measured by probing a total of 1,441 and 1,607 transformants, respectively, using the NarG0 probe (Table 2.1). The error-free TLS frequencies of $3.1\% \pm 0.33$ and 6.2 ± 1.12 were obtained for SOS⁻ and SOS⁺ collections, respectively. The increased error-free bypass in the SOS⁺ collection was determined, by t-test analysis, to be different from SOS⁻ collection ($P=0.02$) (Table 3.8).

In the $\Delta umuC$ strain the *in vivo* -CG₃ translesion bypass of the AAF adduct, was higher than error-free bypass in both SOS⁻ and SOS⁺ collections. A total of 1,453 and 1,623 colonies from SOS⁻ and SOS⁺ collections, respectively, were screened for -CG₃ frameshifts by probe hybridization using the Nar-2 probe (Table 2.1). The TLS

TLS Event	TLS frequency (%)		$P_i(\text{t-test})$	$-2\text{Ln}P_i$
	SOS ⁻	SOS ⁺		
EFB	3.0±0.22 (1571) *	13.8±1.38 (1317)	0.0027616	11.784
-CG3	1.6±0.29 (1359)	5.9±0.24 (1239)	0.0000207	21.572
Survival	4.6±0.51	19.7±1.42	0.0002113	16.924
				$-2\sum \text{Ln}P_i = 50.280$ $\chi^2_{0.05[6]} = 12.592$ $P < 0.001$

Table 3.7: Translesion synthesis frequencies past the NarG3.AAF adduct in SOS induced and uninduced *E.coli* JM103 strain. P_i , probabilities from t-test for the effect of SOS induction per class of TLS; $-2\sum \text{Ln}P_i$, statistic for the combined probabilities Chi-square test.

* Numbers in parentheses are the total number of transformants, from triplicate experiments, screened by probe hybridization, using sequence specific probes.

frequencies for the minus CG3 frameshifts in SOS⁻ and SOS⁺ collections, were 10.2%±0.85 and 66.0%±1.3, respectively. The t-test for comparison of the means and variances between SOS⁻ and SOS⁺ collections, revealed that the six-fold increase in -2 frameshifts due to induction of the SOS functions was statistically significant (P=0.000004).

The survival of the vector containing an AAF adduct in SOS⁻ and SOS⁺ collections of the *ΔumuC* strain, was 13.3%±1.2 and 72.2%±1.4, respectively. The t-test for comparison of the mean and variance values for survival rates of modified vector in SOS⁻ and SOS⁺ collections, showed that the 5.4-fold increased survival rate among the SOS⁺ collection was statistically significant (P=0.0000003). In addition, the combined probabilities Chi-square test for the error-free, minus CG3 and survival rates in the SOS⁻ versus SOS⁺ collections produced a value of $-2\sum \ln P_i = 63.374$, that was significantly higher than the critical value of $\chi^2_{0.05[6]} = 12.592$; the probability of homogeneity in the two collections was P<0.001 (Table 3.8).

3.3.3.3 TLS events past NarG3.AAF adduct in *E.coli* JM103 *Δumu*(pDC322) and the effect of SOS induction: The effect of the episomally encoded UmuD/C functions on *in vivo* translesion bypass of a NarG3.AAF adduct was studied by transforming the SOS induced and uninduced competent cells of the *E.coli* JM103 *ΔumuC*(pDC322) strain with SSCpLHIINarG3.AAF and the control construct SSCpLHIINarG0. SOS⁻ and SOS⁺ collections of the cells transformed with modified vector were prepared by transferring a total of 12,244 and 3,386 colonies, respectively, onto hybridization membranes.

TLS Event	TLS frequency (%)		P_i (t-test)	$-2\text{Ln}P_i$
	SOS ⁻	SOS ⁺		
EFB	3.1±0.33 (1441)*	6.2±1.12 (1607)	0.0224734	7.590
-CG3	10.2± 0.85 (1453)	66.0±1.27 (1623)	0.0000044	24.678
Survival	13.3±1.15	72.2±1.43	0.00000032	31.106
				$-2\sum \text{Ln}P_i = 63.374$ $\chi^2_{0.05[6]} = 12.592$ $P < 0.001$

Table 3.8: Translesion synthesis frequencies past the NarG3.AAF adduct in SOS induced and uninduced *E.coli* JM103 $\Delta umuC$ strain. P_i , probabilities from t-test for the effect of SOS induction per class of TLS; $-2\sum \text{Ln}P_i$, statistic for the combined probabilities Chi-square test.

* Numbers in parentheses are the total number of transformants, from triplicate experiments, screened by probe hybridization, using sequence specific probes.

The frequency of error-free bypass events among 5,817 colonies from the SOS⁻ and 3,386 colonies from the SOS⁺ collections, were 6.7%±0.26 and 24.7%±2.3 for SOS⁻ and SOS⁺ cells, respectively (Table 3.9). The increased error-free bypass in the SOS⁻ induced collection, was statistically significant as compared to the SOS⁻ collection (P=0.003). The frequency of mutational bypass via -CG₃ frameshifts, 3.0%±0.1 and 6.5%±0.23 in SOS⁻ and SOS⁺ collections, respectively, was less than that of error-free bypass events. The higher -2 frameshift bypass frequency in the SOS⁺ collection, as compared to the SOS⁻ collection, was statistically significant (P=0.00008). The survival rates of the AAF modified vector in the SOS⁻ and SOS⁺ collections were calculated as 9.7%±0.21 and 31.2%±2.2, respectively. The t-test for comparison of the survival between SOS⁻ and SOS⁺ collections showed that the increased survival among SOS⁺ collection was significant (P=0.002), (Table 3.9). In addition, the combined probabilities Chi-test produced a value of $-2\sum \ln P_i = 43.138$, corresponding to $P < 0.001$, which was significantly higher than the critical value of $\chi^2_{0.05[6]} = 12.592$, (Table 3.9).

3.3.3.4 TLS events past NarG3.AAF adduct in *E.coli* JM103 Δ umuC (pGW1700)

and the effect of SOS induction: SOS induced and uninduced competent cells of *E.coli* JM103 Δ umuC(pGW1700), containing extrachromosomally encoded MucAB functions, were transformed in triplicate with SSCpLHIINarG3.AAF and the control construct SSCpLHIINarG0. The transformed cells were selected on multiple ampicillin plates. A total of 7,029 colonies from the SOS⁻ and 3,946 colonies from the SOS⁺ cells transformed with the modified vector, were transferred onto hybridization membranes.

TLS Event	<u>TLS frequency (%)</u>		$P_i(\text{t-test})$	$-2\text{Ln}P_i$
	SOS ⁻	SOS ⁺		
EFB	6.7±0.26 (5817)*	24.7±2.3 (3386)	0.0028257	11.738
-CG3	3.0±0.1 (6427)	6.5±0.23 (3386)	0.0000846	18.756
Survival	9.7±0.21	31.2±2.2	0.0001796	12.644
				$-2\sum \text{Ln}P_i = 43.138$ $\chi^2_{0.05[6]} = 12.592$ $P < 0.001$

Table 3.9: Translesion synthesis frequencies past the NarG3.AAF adduct in SOS induced and uninduced *E.coli* JM103 $\Delta umuC$ (pDC322) strain. P_i , probabilities from t-test for the effect of SOS induction per class of TLS; $-2\sum \text{Ln}P_i$, statistic for the combined probabilities Chi-square test.

* Numbers in parentheses are the total number of transformants, from triplicate experiments, screened by probe hybridization, using sequence specific probes.

Error-free bypass of the NarG3.AAF adduct in 3,471 colonies from SOS⁻ collection and 1,857 colonies from SOS⁺ collection, were examined by probe hybridization using the NarG0 probe (Table 2.1). The error-free translesion synthesis frequencies of 21.3%±2.3 and 33.1%±1.4 were obtained for the SOS⁻ and SOS⁺ collections, respectively. Statistical analysis of the increased error-free bypasses among the SOS⁺ collection, compared to SOS⁻ collection, showed a significant difference between their means, with a probability for homogeneity equal to 0.002, (Table 3.10).

The occurrence of the -CG3 frameshifts among the SOS⁻ and SOS⁺ collections, which was considerably lower than that of the *E.coli* strain JM103 Δ umuC, was 4.4%±0.32 and 6.09%±0.39, in the SOS⁻ and SOS⁺ collections, respectively, (Table 3.10). The t-test analysis revealed that the increased TLS frequencies for the -2 frameshifts among SOS⁺ collection in comparison to SOS⁻ collection was statistically significant at P=0.002, (Table 3.10).

An interesting observation was that the delivery of MucAB functions to a Δ umuC background in *E.coli* JM103 Δ umuC strain resulted in reversal of the translesion synthesis profile with a preference for the error-free bypasses compared to mutational -CG frameshifts under both SOS⁻ and SOS⁺ conditions. This was also observed when UmuDC functions were encoded on a plasmid in the Δ umuC strain. As a result, the total survival of the modified vector in the MucAB background reached 25.71%±2.4 and 39.17%±1.69 in SOS⁻ and SOS⁺ collections, respectively, with a relative preference for error-free

TLS Event	TLS frequency (%)		$P_i(\text{t-test})$	$-2\text{Ln}P_i$
	SOS ⁻	SOS ⁺		
EFB	21.3±2.26 (3471)*	33.1±1.35 (1857)	0.0022393	12.204
-CG3	4.4±0.32 (3558)	6.1±0.39 (2089)	0.002220	12.224
Survival	25.7±2.4	39.2±1.69	0.0006642	14.634
				$-2\sum \text{Ln}P_i = 39.062$ $\chi^2_{0.05[6]} = 12.592$ $P < 0.001$

Table 3.10: Translesion synthesis frequencies past the NarG3.AAF adduct in SOS induced and uninduced *E.coli* JM103 $\Delta umuC$ (pGW1700) strain. P_i , probabilities from t-test for the effect of SOS induction per class of TLS; $-2\sum \text{Ln}P_i$, statistic for the combined probabilities Chi-square test.

* Numbers in parentheses are the total number of transformants, from triplicate experiments, screened by probe hybridization, using sequence specific probes.

bypass, (Table 3.10). The overall increased capacity of the SOS⁺ cells to bypass the AAF adduct in comparison to SOS⁻ cells was statistically significant (P=0.0007).

In addition, performing the combined probabilities Chi-Square test for EFB, EPB and survival rates resulted in $-2\text{LnPi}=39.062$, a value that was considerably higher than the critical value of $\chi^2_{0.05[6]}=12.592$, suggesting that probability of homogeneity of the data for SOS⁻ and SOS⁺ collections is less than 0.001, (Table 3.10).

4. Discussion

4.1. Advantages of using single stranded circular phagemid vectors that carry a single dG.AAF adduct: Translesion replication past a noninstructional damaged base in *E.coli* and other eubacteria is an SOS dependent process and requires an SOS inducing signal, RecA, SSB, at least one of the Y-family polymerases e.g Pol V or their extrachromosomally encoded homologs, β -clamp and γ clamp loader functions (See section 1.5.3 for details). In the absence of effective repair and recombinational repair mechanisms, translesion synthesis past the noninstructional bulky dG-AAF DNA adduct is the only and final cellular resort to replicating the single stranded circular DNA vector carrying this lesion. Essentially, the *E.coli* cells are used as *in vivo* miniature test tubes that are transformed with single adduct containing ss-vectors; the survival of the vector can be measured by use of a selectable antibiotic marker on the vector, and the rates of mutagenic and non mutagenic translesion synthesis can be monitored. Furthermore, the effect of different cellular genetic backgrounds can also be assessed for their contribution to the TLS outcome. Use of the single stranded circular vectors in single adduct mutagenesis studies greatly simplifies the analysis of translesion synthesis *in vivo*. The reasons for this are described in the sections below.

4.1.1. DNA repair background noise is circumvented: DNA repair mechanisms, as described in section 1.5, are intriguingly sophisticated networks of highly conserved functions that deal with most chemical and physical DNA damage by either damage reversal or removal of the altered backbone and nitrogenous bases. Efficient DNA repair contributes strongly to a background of non mutagenic outcomes *in vivo*, and greatly

complicates the study of translesion bypass. Since all known DNA repair enzymes require a double stranded substrate, the use of single stranded circular vectors in single adduct mutagenesis studies virtually circumvents the effect of all repair functions and simplifies the analysis of data.

4.1.2 DNA structural topology effects are eliminated: B→Z topological transitions arise within sequences of alternating purine-pyrimidine residues in double stranded DNA. A high frequency of -2 frameshift mutations in the context of the *NarI* sequence (5'-GGCGCC-3') that carries a single dG-AAF adduct has led to a highly speculative hypothetical assumption, that has existed for over a decade, that these observations may be related to the ability of the adduct to induce a B→Z DNA topological transition. An additional reason to use the single stranded DNA vectors stems from the fact that in the absence of a double stranded DNA backbone the effect of the presumed DNA topological alterations on the mutagenesis process would be eliminated; thus mutational events would become solely dependent upon the nature of the adduct, the sequence context in the vicinity of dG-AAF DNA adduct, and the cellular functions that facilitate adduct bypass. In a similar manner other kinds of structural alterations such as kinks, bends and extra helical structures that might serve as sites for recruiting DNA repair proteins, would be eliminated in single stranded vectors, and the adducted residues will be directly accessible to TLS functions. On the other hand, in a single stranded DNA sequence context formation of secondary structures such as stem-loops, cruciforms, and template or nascent strand slipped intermediates are more likely to occur, reminiscent of the single-double stranded junction at the replication fork. This provides the single stranded vectors with

additional capacity as ideal vector systems to study the effect of direct and inverted DNA sequence repeats in the vicinity of the DNA lesion on mutagenic events.

4.1.3. Damage avoidance mechanisms are detoured: Replicating enzymes, such as DNA polymerase III, generally stall, and DNA synthesis is inhibited (Doudney, 1990), when a bulky DNA lesion such as dG-AAF is encountered on the template strand; as a result the replication fork collapses and the replisome disintegrates (Hanawalt, 1966; Rajagopalan et al, 1992; Belguise-Valladier et al., 1994). Frequently, disintegration of the replisome results in activation of the recombinational repair mechanisms, i.e. daughter strand gap repair (see also the RecF pathway, section 1.5.2). Replication is believed to be reinitiated at 800-1000 bp upstream from the lesion on the damaged template, in an origin independent fashion, i.e. through the PriA dependent pathway (Marians, 2000; Sandler and Marians, 2000). This leaves a single stranded gap that contains the damaged nucleotide on the template (Rupp and Howard-Flanders, 1968; Howard-Flanders and Rupp, 1972). RecA dependent homologous pairing of the complementary strand with the gap sequence, and pairing of the stalled nascent strand with the complementary daughter strand produces an extendible 3'-terminus for DNA synthesis within the gap region. The RecG, or RuvABC resolvosome dependent resolution of the Holliday junction(s) restores the replication fork and allows the DNA lesion to persist for at least one round of replication, during which it can be repaired by nucleotide excision repair. The other recombinational repair pathways of *E.coli*, i.e. the RecBCD and RecE pathways, which deal primarily with double strand breaks, are not functional on bulky DNA base modifications and hence their interference is not relevant. Recombinational repair, as a damage avoidance mechanism, requires both of the complementary DNA strands in order

to be effective in restoring replication capability at the site of the DNA lesion. The use of the single stranded circular vectors that carry a site specific single adduct in mutagenesis studies circumvents recombinational lesion avoidance pathways, and forces the lesion through translesion bypass pathway.

4.2. Development of single stranded vectors carrying single dG-AAF adduct: The successful approach used to construct the single stranded vectors was based on the use of a cloned palindromic sequence that, in the context of a single stranded circular vector, would generate a hairpin structure within which a single restriction site is formed (Fig. 2.1). This would allow linearization of single stranded DNA with a restriction enzyme (Banerjee et al., 1988). The linearized vector can subsequently be used to clone the short AAF modified oligonucleotides. Detailed information on the stepwise procedures can be found in Materials and Methods (see section 2.2).

A palindromic 49 bp HII synthetic linker, containing *EcoRI* and *HindIII* cohesive ends and two internal *HincII* sites, was replaced with a 47 bp *EcoRI-HindIII* polylinker fragment of phagemid pUC118 (Fig 2.1). The resulting pLHII construct would contain a -1 frameshift in *lacZα*, and also contains two opal stop codons and one amber codon in tandem with the second opal. The first opal stop codon is in frame with respect to the start codon and served as a measure to facilitate the phenotypic screening of the pLHII clones among parental pUC118 transformants. The second opal stop codon, together with a tandem amber codon ensures premature translational arrest in pLHII-SmaG0 construct. Preparation of the SSCpLHII construct and internal annealing of its 52 bases within the palindromic region results in the formation of a unique *HincII* site within the hairpin

structure. The hairpin structure is stabilized by an 11 base pairs A:T rich region at the bottom portion and a 15 bp G:C rich region at the top portion of the hairpin structure (Fig. 2.1). The presence of a more G:C rich (53%) region at the top of the hairpin facilitates the rapid formation and stabilization of the hairpin structure from the tip under annealing conditions, while the less stable bottom portion (36% G:C) provides for lesser internal annealing tendency within the *HincII* digestion product of SSCpLHII (i.e. SSL⁺-pLHII). The latter feature leads to an enhanced external annealing of the competing scaffold to the SSL⁺-pLHII. The released top 15 bp portion of the *HincII* digestion product, lacks homology to any part of the vector; therefore there is no interference of this portion with the next steps of the protocol, which also eliminates the necessity for any further purification. The stem portion of the hairpin also does not have any homology with the rest of the vector which circumvents the major problem that Banerjee et al. (1988) had encountered in their M13mp7 phage based design by formation of two alternative annealed isoforms, one of which was resistant to digestion.

The use of *HincII* was preferred for two reasons. First, *HincII* does not have star activity and hence prolonged incubations could be carried out to maximize digestion. Second, *HincII* does not produce cohesive ends, which minimizes self-ligation, and ligation between the top and bottom portions of the digested hairpin during the next stages, resulting in greatly reduced vector background. An unexpected problem was the presence of a high level of deletion mutant clones containing the *HincII* linker that carried deleted regions of 2-25 bp bordered with *EcoRI* and *HindIII* ends. This problem was traced to the presence of impurities in the synthetic oligomers. The problem was solved

by the use of PAGE purified linker oligonucleotides. A correct clone of pLHII was obtained and the SSC form of one of the clones was prepared for further application and molecular analyses. The sequencing of the double stranded and single stranded forms of pLHII showed a compression pattern at 5'GTCGAC3' sequence of the 5'-end *HincII* site (Fig. 3.3). This compression pattern was resolved by use of dITP in sequencing reaction and the correct sequence of the inserted linker was verified (Fig. 3.4).

Gel purified and internally annealed SSCpLHII was subjected to double *HincII* digestion that resulted in the appearance of a linearized band that lags behind the circular form on 0.7% agarose gel (Fig. 3.5, lane 3). The efficiency of digestion was assessed by measuring the remaining transforming efficiency of the digestion product normalized against that of the undigested stock. The residual transforming capacity was determined to be 1% of undigested stock (minus enzyme control : 1.2×10^6 , 1.3×10^6 transformants/ μg DNA and digestion set: 3.6×10^4 , 1.6×10^4 tr/ μg DNA). The accuracy and completeness of the digestion was verified by sequencing the linearized vector and comparing it to the sequence of the single stranded circular vector (Fig. 3.6). The polymerase stopped at the end of the linear vector (notice the GAC sequence as the 3' half of the GATGAC *HincII* site) and added a non instructional nucleotide to the end of the nascent strand with a preference given to the T nucleotide.

The SSL⁺pLHII linear DNA was used to optimize the cloning condition for the short oligonucleotides. The ³²P end-labeled SmaG0 14-mer oligonucleotide was cloned into the linear single stranded pLHII vector by use of varying molar ratios of vector:scaffold:oligo. The scaffold DNA was removed by annealing to anti scaffold

oligonucleotide and samples were run on an agarose gel (Fig. 3.7 panel A). The 5'- and 3'-end double ligation product appeared as a band running ahead of the linear band that represents the single stranded circular pLHISmaG0 construct. No SSC band was observed in the negative control set that did not include ligase in the reaction mixture (Fig. 3.7A, lane 3), and consequently, no corresponding radioactive band was also detected for control set in autoradiogram of the gel (Fig. 3.7B, lane 3). In addition, removal of the scaffold DNA did not abolish the circular integrity of the SSC band. This suggested that the required two-end ligation had successfully converted the linear band to covalently closed circular form that also co-migrated with the control SSC band in lane 1 (Fig. 3.7, panel A). The intensity of the radioactive SSC bands increased as the ratio of the vector:scaffold:oligo increased and leveled off in lane 6 and 7 (Fig. 3.7, panel B). In contrast, single ligation products of the radiolabelled SmaG0 oligomer at either 3' or 5'-ends of the linear vector, as schematically shown in panel C, resulted in a radioactive linear product that co-migrated with the SSL⁺ band. However, unlike SSC product, the intensities of the linear products did not level off as higher ratios of the scaffold and oligomer were used. This can be explained by the simultaneous incorporation of the radiolabelled SmaG0 oligomer to both 3' and 5'-ends of a single vector molecule as increasingly excessive amounts of scaffold and oligo were used (Fig. 3.7B, lanes 4-7). Equal samples from each experimental set were used for transforming the *E.coli* JM103 cells and the average transformation efficiencies were plotted (Fig. 3.8). Based on these results the optimal molar ratios of vector:scaffold:oligo for cloning was determined to be 1:2:4. This decision was made on the basis of conservative usage of AAF modified oligomers and the transformation efficiency comparable to reaction mixture in set #3

(Fig. 3.8). These results are in agreement with the previous experiments using pLH118 and MLH constructs (see also section 3.2.4).

Along with the SSCpLHII-SmaG0, the SSCpLHII-NarG0 and dG-AAF modified constructs were also prepared using the optimal conditions as described above. The ^{32}P end labeled NarG0, NarG3.AAF, SmaG1.AAF and SmaG3.AAF oligonucleotides were cloned into single stranded linear vector and the recircularized band was verified on an agarose gel (Fig. 3.9). The relative incorporation of the oligonucleotides and the concentration of the SSC forms were determined (see section 3.2.5). The statistical analyses showed insignificant differences between the efficiency of incorporation of the unmodified and modified oligonucleoties (Table 3.1, probabilities of 0.194-0.748). The insignificant statistical difference between the ligation reactions for cloning of either of modified or unmodified oligonucleotides implies that the presence of the bulky AAF modification carried on the SmaG1.AAF, SmaG3.AAF and NarG3.AAF oligonucleotides did not impose any substantial spatial hindrance on the function of ligase (Table 3.1).

Removal of the scaffold DNA from the circularized constructs is required to eliminate the possibility that extension from 3'-end of the scaffold would result in a mixed population of mutant and nonmutant phagemids, and also to eliminate the possibility that the partially double stranded region that contains the bulky modification might be a substrate for repair enzymes. This can be achieved by either boiling and rapid cooling, or by annealing the scaffold DNA to its complementary oligonucleotide, the antiscaffold. The efficiency of both methods were investigated by use of a ^{32}P labeled scaffold DNA (Section 3.2.6). The results shown in Figure 3.10a and 3.10b indicate that the boiling method that was used by Banerjee et al. (1988) was not able to remove the

scaffold completely since a clear signal from the labeled scaffold, that remained on the construct was still detectable in lane 2, while it was completely removed by antiscaffold (Fig. 3.10b lane 3). The high level of the wild type background (93%) that Banerjee et al. (1988) reported would be partially due to the combined effects of the complementary strand synthesis by extension from the scaffold, and recombinational bypass of the T-T dimers. The use of antiscaffold to remove the scaffold DNA assured us that the recombinational repair of the dG.AAF adducts would not be an interfering factor and any nonmutational bypass of these adducts would be dependent on translesion bypass.

The *E.coli* JM103 cells transformed with pLHII and pLHII-NarG0 constructs showed a persistent white phenotype on X.gal plates. While the pLHIINar-2 construct was phenotypically blue, the -2 frameshifted pLHII-SmaG0 control construct and its -1 frameshifted derivative, the pLHIISma-1, both appeared blue with varying intensities on X.gal plates. The expression of β -galactosidase activity from the -2 frameshifted construct, the pLHIISmaG0, was an unexpected observation. This observation is probably not a result of a translation reinitiation from the additional ATG in the third codon position. Reinitiation from this codon, which is in frame with the first ATG codon, will be in the -2 reading frame with respect to the downstream α -complementation fragment coding sequence, and would encounter the same two stop codons as the translation from the first start codon. Moreover, scanning the downstream sequences did not detect any in frame ATG or GTG codons.

A second explanation for the unexpected β -galactosidase activity derived from the same constructs is ribosomal frameshifting, which provides a mechanism for alternative

expression of two polypeptides encoded by a unique mRNA, or one protein product from two overlapping ORFs (Gesteland and Atkins, 1996). Ribosomal frameshifting occurs in a minority of genes among almost all organisms (see the database at <http://recode.genetics.utah.edu> ; Baranov et al., 2001 and 2002) and requires upstream *cis* acting elements, for example slippery reiterated bases (most often A nucleotides), and a near canonical RBS, and a downstream ribosome hindering topological factor such as RNA pseudoknots (Larsen et al., 2000). Codon redefinition or recoding is another mechanism by which the meaning of the stop codons is altered and the ribosome bypasses them by insertion of an amino acid e.g. bypass of opal stop codon by insertion of a tryptophan. The unexpected β -galactosidase expression in our -2 frameshifted constructs cannot be explained by ribosomal frameshifting mechanism as mentioned above, since the required *cis* acting elements for translational frameshifting are not present and the region in question is just at the beginning of the coding sequence. Moreover, rarely, redefining the two successive stop codons would still yield a -2 frameshifted polypeptide that would not result in α -complementation.

Examination of the DNA sequence of the M13mp18 phage and commercially available pUC118 (Worthington Inc.) and pUC128 (Keen et al., 1988) phagemids that were used to construct the single stranded circular vectors provides an alternative explanation for the unexpected phenotypic expressions seen in some of our constructs. Examining the sequence at the 5' end of the pUC series showed that the second codon, ACC that codes for a threonine residue, had a deletion in one of the cytosine residues (See Fig. 3.3 and Fig. 4.1). However, examining the M13mp18 derivatives did not show the same deletion. As a result, in pUC118 and pUC128 vectors the initiation from the

Fig. 4.1: Sequencing of pUC118, pLH118 and pLH-NarG0 constructs (pUC118 derivatives); pUC128, pLH128; M13mp18 and MLH (M13mp18 derivative). The constructs were sequenced using -35 *lacZ* universal primer extended past the *EcoRI* site. The ACC codon that codes for the threonine as the second amino is shown in red rectangle boxes. A cytosine residue is deleted in all pUC phagemid series but not in M13mp18 and its derivative, the MLH.

first ATG would produce a -1 frameshifted protein; however, translation from the second ATG at the third codon position will result in a functional enzyme missing only two amino acids from the N-terminus. The phenotypic behavior of the single stranded circular constructs following initiation from either start codon is given in Table 4.1. pLHII and pLHII-NarG0 constructs generate -1 and -2 frameshifted products if initiation starts from the second and/or first start codons, respectively. Translation from the first start codon in both constructs is prematurely terminated at a TGA stop codon in 3rd codon position. If this codon were recoded, e.g. by tryptophan insertion, the frameshifted product will still be terminated at the second downstream opal stop codon. In the improbable event that the second stop codon were recoded, functional α -fragment would still not be produced. Initiation of translation from the second start codon will generate -1 frameshifted polypeptide and will be terminated at a built-in opal codon in 11th codon position in pLHII. In pLHII-NarG0 a premature stop codon will not be encountered close to the initiation site but the polypeptide would be frameshifted and nonfunctional. The persistent white phenotypes of the pLHII and pLHII-NarG0 constructs are consistent with their genetic make up (Table 4.1).

The pLHII-Sma-1 and pLHII-Nar-2 constructs and pUC118 generate an in-frame α -complementation product when translation is started from the second start codon. However, starting from the first start codon brings an opal stop codon in the 3rd position into frame. As a result these constructs conferred a consistent blue phenotype to the *E.coli* host on X.gal plates (Table 4.1).

Construct	Frame when started from 2 nd ATG	Stop Codon used when initiating from 2 nd ATG	Frame when started from 1 st ATG (with -C mutation in second codon)	Stop Codon used when initiating from 1 st ATG	Overall Phenotype
pUC118	0	-	-1	+ ^a	B
pLHII	-1	+ ^b	-2	+ ^{a,c}	W
pLHII-SmaG0	-2	+ ^c	0	+ ^a	PB
pLHII-Sma-1	0	-	-1	+ ^a	B
pLHII-NarG0	-1	- ^f	-2	+ ^{a,c}	W
pLHII-Nar-2	0	-	-1	+ ^a	B

+a: Stops at TGA in 3rd codon position

+b: Stops at a built-in TGA codon in 11th codon position

+c: Stops at TGA and TAG codons built-in 3'-end of HII linker

- : No stop, reads in correct frame

-f: No stop, reads in -1 frame

Table 4.1: Reading frame of -complementation fragment, stop codon use and phenotypes of different single stranded circular vectors. The initiations are considered from the second and first ATG start codons. For initiation from the first start codon the detected -C mutation in second codon has been considered.

The pLHII-SmaG0 construct is -2 frameshifted for initiation from the second start codon and translation is terminated at a stop codon downstream from the SmaG0 insertion site. Reading from the first start codon is in-frame; however, it encounters the opal stop codon at 3rd codon position. The pLHII-SmaG0 construct would only be able to express functional α -complementation from the first start codon if it bypasses the TGA stop codon by recoding. Considering the fact that the opal stop codon is recoded by eukaryotes widely and to a lesser extent by bacteria (Firoozan et al., 1991; Karow et al., 1998; Nilsson and Ryden-Aulin, 2003), then appearance of a pale blue color among *E.coli* JM103 cells carrying pLHII-SmaG0 construct is plausible. The functional α -fragment accumulates over time and causes progressive digestion of X.gal and accumulation of chromophore, i.e. 5-bromo-4-chloro-3-indol (Table 4.1). Nonetheless, the MLH did not behave in a similar way to pUC derivatives, which is probably due the fact that the coding sequence for the α -complementation fragment is intact.

The pLHII series of single stranded circular vectors carrying a dG-AAF adduct might seem to be suitable for phenotypic mutagenesis studies that involve detection of induced -1 and -2 frameshift mutations. However, these constructs suffer from the following drawbacks and limitations; a) The nonmutational bypass clones in pLHIISmaG3 construct that appear as pale blue colonies cannot be scored reliably when the incubation period is extended past 48 hours at which time the distinction between these clones and the blue -1 mutants become difficult. However, the pLHII vector background that accounts for less than 10% of transformants stays persistently

distinguishable from other colonies by virtue of its white appearance; b) The pLHIINarG0 nonmutational bypass progeny of the NarG3.AAF adduct appear as persistent white colonies indistinguishable from the vector pLHII background transformants. The determination of the relative frequencies of these two types requires additional molecular characterization of the clones; c) The pLHII series are not suitable for mutagenesis studies involving phenotypic screening for base substitution mutations. Such an attempt would require designing additional constructs specific for each type of mutation.

As a result of the difficulties associated with phenotypic analysis of mutagenesis, it became clear that applying the probe hybridization screening method would be the best choice for the rest of the study. Such an approach would facilitate detection of the whole population of outcomes resulting from the survival of the single stranded vector that carries the AAF adduct by means of error free and error prone translesion synthesis. Such an approach, therefore, eliminates any requirement for identification of the pLHII vector background and application of correction factors.

4.3. Optimization of Mutation Screening: The SOS induced and uninduced competent cells of wild type JM103, JM103 Δ umuC and JM103 Δ umuC/MucAB were transfected by SSCpLHIISmaG0 and SSCpLHIISmaG3 constructs and plated on LB.Amp plates. Random colonies were probed by SmaG0, Sma-1 and SmaT probes, which correspond to the wild type, -1 frameshifted and G $\rightarrow$ T transversion mutations, respectively.

Adopting the probe hybridization screening method initially produced a false positive signal among G0 clones among all strains. The presence of false positive and

negative clones was attributed to the presence of impurities in SmaG0 and scaffold and antiscaffold oligonucleotide stocks. The use of PAGE purified stocks not only completely eliminated these artifacts but also increased the cloning efficiency for the control sets. No impurities were found in the SmaG3 stock since originally this stock was prepared through HPLC purification. A great deal of care was taken to eliminate possible sources of experimental variation associated with bacterial transformation. a) Competent cells were prepared and were used for transformation within a short period of time with both control and test sets. b) Replicas for each experimental set were run at the same time intervals, and the cellular and DNA components among the replicas were equalized by making master mixes. c) The pre-pulse incubation and post-pulse recovery expression periods were set equal for all experiments and were controlled carefully. d) In each case freshly prepared plates from the same batch were used for plating. These precautions ensured high transformation efficiency (10^6 - 10^7 cfu/ μ g SSC) and a high level of reproducibility.

In summary, a major objective of this thesis was to develop a mutagenesis system that used single stranded vectors that contain single AAF lesion at defined positions within mutational hotspots, and to use this system to study translesion synthesis *in vivo*. The developmental phase of this project was both complex and time consuming. The experimental objectives were achieved by optimization of each of three essential elements of the mutagenesis experiments. First, conditions for construction of the SSC vectors were developed with extreme care taken to eliminate contaminating byproducts that contribute to unacceptable background or artifacts. Second, a reproducible and high efficiency transformation protocol was applied. Third, a nonphenotypic mutation

screening method was developed that allowed accurate determination of error free bypass and all of the different mutational events associated with error prone bypass.

4.4. Mutagenesis: Induced translesion synthesis past an AAF adduct using the constructed single stranded circular vectors containing a single dG-C8-AAF modification located in the -1 and -2 frameshift hotspot sequence contexts, the *SmaI* and *NarI* sites, respectively, were studied in *E.coli* JM103 derivatives. The use of the wild type *E.coli* JM103 strain allowed us to assess the error free and mutagenic TLS events in the presence of functional SOS regulated factors. The JM103 Δ *umuC* mutant strain enabled the reductive assessment of TLS events in the absence of functional Pol V, meanwhile, it provided the opportunity to carry out complementation assays for extrachromosomally expressed Pol V and Pol RI functions. The inducible nature of these SOS controlled functions provided additional power to contrast the TLS outcome between background and fully activated expression levels of these functions; and hence to investigate the cause-effect relationship between them.

4.4.1. Error free TLS past SmaG3-AAF adduct is dependent on SOS regulated Pol V and Pol RI functions: The majority of the SmaG3-AAF adducts are bypassed in an error free manner in JM103 wild type strain and its derivatives. In all cellular backgrounds the error free bypasses comprise 73%-90% of total TLS events, regardless of induction of SOS responses (Tables 3.2, 3.3., 3.4 and 3.5). This would imply that regardless of the involvement of any specific polymerase, the general tendency of the cellular functions is aimed at correct recognition of the adducted guanosine in *SmaI* target and insertion of a correct complementary base across from the damaged residue. These

values are in agreement with findings of Napolitano et al. (1997), who also examined the error free bypass past the dG3-AAF adduct in a similar sequence context, according to which the EFB were 60-95% of the total TLS in wild type JM103 strain and its *ΔumuC* mutant strain. Considering the conservative behaviour of the cells, even at highly mutagenic hotspots such as SmaG3-AAF, one appreciates the fact that such an approach in a broader perspective, e.g. at genomic level, should ultimately contribute to preservation of the genetic integrity of the organism and would result in its increased survival. As a result, TLS provides additional opportunity for the repair mechanisms to remove the damaged base when it becomes available to the repair enzymes in duplex DNA following replication past the DNA lesion. Nonetheless, the relative efficiency of any specific function in contributing to the cellular survival can only be measured by comparing the TLS frequencies associated with the presence or absence of such functions under specific physiological conditions. The error free TLS frequencies past the SmaG3-AAF in JM103 wild type strain, regardless of the state of the cells for SOS induction, were 14 to 20-fold higher than EFB values in JM103 *ΔumuC* mutant. Moreover, EFB frequencies among SOS⁺ JM103 *ΔumuC* mutants that were expressing either the Pol V or Pol RI (i.e. MucAB) from pDC322 and pGW1700 plasmids, respectively, showed an overall increase in the range of 17- to 30-fold ($P < 0.002$) (Tables 3.2, 3.3, 3.4 and 3.5). This indicates that the presence of either Pol V or Pol RI, even at background levels, was able to restore the cellular capacity to read through the dG3-AAF adduct at levels comparable to the wild type strain. The residual level of error free TLS observed in JM103 *ΔumuC* strain that comprises about 74% of total bypasses might be due to Pol III

or an SOS function other than Pol V. We conclude that the error free TLS past the dG3-AAF adduct is largely dependent on Pol V and/or its plasmid encoded homolog, Pol RI.

Induction of SOS functions in JM103 wild type strain, and JM103 *ΔumuC* strains expressing Pol V or Pol RI resulted in 1.5 to 2.6-fold increase in error free TLS frequencies compared to their respective SOS⁻ collections, which was statistically significant ($P < 0.002$). These observations can be considered as an additional line of evidence for involvement of Pol V and Pol RI in the process of error free readthrough past the dG-AAF adduct. The relative increase in error free TLS past AAF adducts among SOS induced JM103 wild type strain compared to SOS⁻ cells did not grow beyond 2.6-fold. However, Napolitano et al. (1997) reported an almost 10-fold increase in error free TLS for JM103 wild type SOS⁺ cells compared to SOS⁻. While our TLS frequencies past SmaG3-AAF in SOS⁺ wild type JM103 and JM103 *ΔumuC* mutant are in agreement with Napolitano et al. (29.1 ± 1.0 versus 22.8 ± 3.7 in wild type strain and 1.4 versus 2.07 in *ΔumuC* strain, respectively), our SOS⁻ JM103 cells showed about 6-fold higher error free TLS frequency compared to Napolitano et al. (1997) (Table 3.2 and 3.3). The relatively high TLS frequency in the SOS⁻ JM103 wild type strain might be explained by the activation of the SOS response due to the persistence of the single stranded DNA that can function as the SOS inducing signal. The SOS inducing capacity of the persisting single stranded undamaged and damaged DNA vectors has been previously shown as explained briefly below.

Induction of SOS responses by filamentous ssDNA coliphage f1 mutant R377 that carries a defective minus strand origin of replication has been monitored by showing the

efficient cleavage of the LexA protein, the repressor of SOS regulon, within 5-10 minutes post infection (Higashitani et al, 1995). In addition, induction of expression of *dinD::lacZ* encoded fusion protein following infection of *E.coli* cells by coliphage R377 and M13ΔE101, an M13 derivative defective for minus strand origin of replication, resulted in about 5-fold increased β -galactosidase activity (i.e. 165 and 160 U/A₆₆₀ for R377 and M13ΔE101, respectively) as compared to the β -galactosidase activity in uninfected cells or cells that were infected with wild type phage (47 U/A₆₆₀). In a parallel experiment, *E.coli* cells treated with mitomycin C, which would cause genome wide DNA damage, i.e. a much larger target than phage genome, produced β -galactosidase activity of 281 U/A₆₆₀ (Higashitani et al., 1992). This suggests that ssDNA of the phage was able to induce the SOS responses at about 58% of mitomycin treated cells. It was also shown that in a similar manner to mitomycin C treated cells, the cells infected with defective phage were able to accumulate activated RecA protein at a significantly higher level than uninfected cells (Higashitani et al., 1992). Following all these observations together with detection of lingering viral single stranded DNA in the cytoplasm of the infected cells at 10 minutes post infection, Higashitani et al. concluded that the persistent cytoplasmic presence of the viral ssDNA was the source of the signal for induction of the SOS functions. The association between the cytoplasmic persistence of the single stranded wild type M13 phage DNA and the DNA damage that resulted in SOS induction has also been documented. D'Ari and Huisman (1982) showed that UV treatment of the single stranded M13 phage DNA prior to transfection resulted in induction of SOS response in the host *E.coli* cells. It was concluded that the presence of UV damaged bases in M13 ssDNA had been an impediment for normal replication of the M13 ssDNA

that prevented its conversion to double stranded replicative form (RF), and hence the DNA damage dependent delayed existence of ssDNA in the host *E.coli* cells had resulted in SOS induction. It is noteworthy that UV irradiation of *E.coli* chromosomal DNA at doses comparable to what was used in this study to induce SOS response, i.e. 40-50 J/m², results in introduction of one cyclobutane TT dimer per 3 kb of each strand (Crowley et al., 1998), which strongly inhibits replication (Banerjee et al., 1988; Courcelle et al., 1997). In comparison to CPD the bulkier AAF adduct, in pLHII derived vectors of ~3200 bases, would impose a stronger impediment to replicative Pol III. Therefore, it is plausible to consider that the AAF adduct in our single stranded vectors might have efficiently caused the Pol III to pause at the dG-AAF nucleotide resulting in prolonged exposure of the SSC vector to RecA, and hence initiation of the SOS signaling. The phage ssDNA induced expression of the *dinD* (HI=3.35), as noted in *dinD::lacZ* fusion assay above, and its close heterology index value to *umuDC* (HI=2.77), might suggest the transcriptional activation of the *umuDC* operon among SOS⁺ collections in response to AAF adducted phagemid vectors in the absence of an external inducing agent. Consequently, we suggest that the TLS events past the dG-AAF adduct among UV uninduced collections are due to the residual activation of Pol V and/or Pol RI.

Insertion of a correct C residue across from the dG-AAF adduct and extension from the dG^{AAF}:dC terminus are the two steps required for facilitating error free TLS. Whether Pol III, Pol V or Pol RI play a role in insertion and/or extension steps, has been a matter of concern. Originally, on the basis of “two step model” Pol III was considered as the main polymerase in promoting TLS past TT dimers. The delayed photoreversal experiments (DPR) were designed by Bryn Bridges to test the “two step model”. In DPR

experiments it was hypothesized that if the irradiated *umuC36* cells, defective for UmuC function, were given some time before exposure to visible light the cells would be able to bypass the TT dimers and subsequently would result in a rise in mutation frequency (Bridges and Woodgate, 1984). Based on these results, and since these experiments were carried out in the absence of a functional UmuC, Bridges and Woodgate concluded that the major polymerase responsible for both misinsertion across from a 3'T and extension thereafter was Pol III. However, subsequent observations were not in a universal agreement with the model. *In vitro* primer extension studies using Pol III holoenzyme and Pol I Klenow fragment have shown that both enzymes stall one base prior to the AAF adducted guanosine residue. The use of 3'→5' *exo*⁻ mutant Pol I Klenow fragment that lacks the proof reading function resulted in insertion of a C residue across from the dG-AAF without any further extension (Belguise-Valladier et al., 1994). These experiments suggested that even abolishing the proof reading function of these replicative polymerases was not sufficient to promote translesion replication past AAF adducts as has been the case for other polymerases e.g. T7 DNA polymerase (Napolitano and Fuchs, 1997). Moreover, the "two step model" of mutagenesis that was originally proposed by Bryn Bridges and Roger Woodgate (1985a), has recently been scrutinized in the light of findings from Roger Woodgate's lab in collaboration with Myron Goodman's lab (Tang et al., 1998; Tang et al., 1999; Tang et al., 2000; Pham et al., 2001; Woodgate R., 2001). *In vitro* primer extension studies using single TT (6-4), TT *cis-syn* cyclobutane dimer and an AP adducts, all capable of strongly impeding DNA replication, Tang et al., (2000) showed that in the presense of RecA*, SSB and β/γ complex, neither Pol III HE nor Pol IV (Din B) was able to synthesize past these adducts efficiently, while Pol V (UmuD₂C)

copied the same adducts within 30 seconds with 30%, 20% and 80% efficiencies, respectively. The Pol III HE that lacks the proof reading function was only able to synthesize past these adducts after 16-fold longer incubation time with considerable lower efficiencies, i.e. 1.2%, 1% and 0.4%, respectively. This observation, which might explain the residual low frequency TLS in UV irradiated *umuC36* mutants, reported by Bridges in delayed photoreversal (DPR) experiments, undermined the major mutagenesis role that was originally assigned to Pol III. In addition, the fidelity gel analyses showed that incorporation of a G residue across from the 3'T in a TT (6-4) adduct was preferred over an A insertion by 6:1 when Pol V was used, while A was inserted by 27 and 3.3-fold compared to G insertion by Pol III and Pol IV, respectively. The differential preference for incorporation of G instead of A by Pol V is also consistent with the *in vivo* mutation rates corresponding to T→C transition as the main mutation targeted to 3'T position that results from G insertion (Tang et al., 2000). This observation clearly questions the role for Pol III in “two step model” and reassigns the same role to Pol V. The 5'T in TT (6-4) is considerably less mutagenic and extension by Pol V from a G or A residue situated across from the 3'T is preferentially carried out by A insertion by 13- and 7.7-fold over G insertion. However, Pol III HE and Pol IV show little capability to extend from either terminus (Tang et al., 2000). It is also important to consider that under *in vivo* conditions the relative copy number of Pol V considerably outweighs that of the Pol III (Woodgate and Ennis, 1991), and that unlike the *in vitro* experiments, under *in vivo* conditions the proof reading function of Pol III is fully active; these facts are likely to reduce the involvement of Pol III in TLS. Taken together, the steady state kinetic measurements reported by Tang et al., (2000) and the differences between Pol V and Pol III HE in

specificity of base selection in copying past these adducts, are consistent with the model that Pol V but not Pol III is the major polymerase involved in TLS past the otherwise replication impeding adducts at both insertion and extension steps.

Additional support for this understanding comes from the structural studies of DNA adducts and error prone polymerases homologous to Pol V. The presence of interconvertible *syn* and *anti* conformers of dG-C8-AAF adduct at single- and double stranded DNA junctions - resembling the replication fork- and their dynamic equilibrium has been well documented (see section 1.2). In contrast to the major *syn* glycosidic alignment that causes a base displacement stacked stable conformer that also impedes the DNA polymerase, the transient *trans* alignment provides the polymerase with an opportunity for template base recognition and insertion of a correct complementary nucleotide. However, the ability of the polymerase to insert a correct complementary base is influenced by the very nature of the polymerase. In contrast to most of the polymerases in groups A to D that also include Pol III, the structural data available for prototypic error prone polymerases belonging to the Y-family of polymerases show a large and relaxed binding pocket with no capacity for base selection fidelity checking mechanism and no ability to form closed conformation (see section 1.5.3.3.4). This property of Pol V and other Y-family polymerases enables them to accommodate the template strand containing a bulky adduct in the binding pocket of these enzymes before and after insertion of a base across the modified base. In addition, the lack of 3'→5' exonuclease proof reading function provides the error prone polymerases with a unique additional opportunity to extend from the complementary strand terminus.

4.4.2. Minus G error prone TLS past SmaG3-AAF adduct is dependent on SOS regulated Pol V and Pol RI functions: The overall -G mutagenic TLS events past the SmaG3-AAF adduct in single stranded vectors account for 10%-27% of total bypasses regardless of SOS induction in JM103 cells and its derivatives. This indicates that the 5'-GGG_{AAF}-3' target sequence is a mutagenic hotspot for -G frameshifts; less frequently base substitutions may also arise within the 5'-GGG_{AAF}-3' sequence (Tables 3.2, 3.3, 3.4 and 3.5). In wild type JM103 cellular background induction of the SOS regulon resulted in about a 3-fold increase in -G TLS frequency (Table 3.2), i.e. from 1.78% to 5.2% in SOS⁻ versus SOS⁺ backgrounds (P=0.0004), respectively. In a JM103 Δ umuC strain that is defective for UmuC function, however, the -G TLS frequencies were reduced by the factor of 6- and 10-fold in SOS⁻ and SOS⁺ backgrounds (Table 3.3), respectively. Expression of UmuDC *in trans* from pDC322 construct in the JM103 Δ umuC strain resulted in restoration of the -G frameshift TLS by 8- and 10-fold, in SOS⁻ and SOS⁺ backgrounds, compared to the same strain that lacks pDC322 plasmid. This confirmed that the SmaG3-AAF induced -G frameshift mutations required Pol V, which is in agreement with the conclusions reached by Napolitano et al (2000). In a similar manner expression of the MucAB functions from pGW1700 plasmid restored the -G frameshift TLS to a level that was 15-fold higher in both SOS⁻ and SOS⁺ backgrounds as compared to the parental JM103 Δ umuC strain (Table 3.5). This observation confirms the involvement of Pol RI function in promoting synthesis past the AAF adduct, and further substantiates the involvement of the Umu-like error prone polymerases in the mutagenesis process. A relatively higher TLS frequency of -G events in JM103 strain

under SOS⁻ condition compared to what was reported by Napolitano et al. (1997) is considered to be the result of residual activation of the SOS functions in JM103 strain due to persistent ss-vectors that carry the AAF adduct. In addition, a relatively higher level of the -G frameshift TLS frequency in SOS⁻ background in JM103 *ΔumuC* that expresses Pol RI function as compared to expression of Pol V is consistent with the notion that the relatively higher efficiency of posttranslational processing of MucA versus UmuD correlates with higher mutagenic potential of Pol RI (Hauser et al., 1992), and further confirms the presence of a residual active Pol RI in SOS⁻ cells.

Induction of -G frameshift mutations in 5'-GGG_{AAF}-3' target requires the insertion of a C residue across from the AAF modified G by either Pol V or Pol RI in a similar manner to that explained above for error free TLS pathway (see section 4.4.1). However, the nature of the translesion synthesis outcome is determined by the structure of the extendible nascent strand primer terminus in relation to the template strand. The nature of the primer/template structure that underlies the mechanism of -G frameshifts can be best explained by the slippage model (Fig. 4.2) (Streisinger and Owen, 1966; Streisinger et al., 1985, Lambert et al., 1992a, Lambert et al., 1992b). The relatively relaxed conformation of the binding pocket of the error prone polymerases such as Pol V and Pol RI allows the C residue at C:G3-AAF on the primer strand to slip one nucleotide in 3'→5' direction on the template strand and base pair with the next 5'G on template. Unlike the insertion and extension steps, the formation of the slipped mutagenic intermediate (SMI) is not directly dependent on Pol V or Pol RI; rather this process is due to the tendency of the C:G-AAF base pair to denature and assume a slipped conformation that is shown to be favored thermodynamically (Garcia *et al.*, 1993; also see section 1.3).

Fig 4.2: Non-mutagenic and -G mutagenic TLS events past an AAF modified guanine residue (G₃) in the *SmaI* site. A cytosine residue is inserted across from the AAF modified G; extension from the inserted cytosine results in error free TLS. The -G frameshift mutagenic pathway is initiated by slippage of the primer C residue and base pairing with the 5'G next to the modified G that results in formation of the slipped mutagenic intermediate (SMI). Extension from the SMI results in a complementary strand that is short by one nucleotide compared to the template strand.

Formation of the SMI results in leaving the G3-AAF moiety in the extrahelical space while the newly formed C:G terminus is used as a primer on the nascent strand for further extension by the either Pol V or Pol RI (Fig. 4.2). The resulting complementary strand will be one base shorter compared to the template strand. The use of the deletion containing strand as the template during the subsequent rounds of replication produces plasmid progeny bearing the -G frameshift, while the complementary strand that contains the AAF modification will most likely be cured and eliminated from the population.

4.4.3. TLS past SmaG3-AAF adduct in JM103 Δ umuC strain is dependent on an

SOS regulated function other than Pol V and Pol RI: The overall survival of the JM103 Δ umuC strain that carries the ss-vector containing the SmaG3-AAF adduct was reduced to 1.28% and 1.90% in SOS⁻ and SOS⁺ backgrounds, respectively (Table 3.3), which corresponds to 12- to 18-fold lower survival as compared to the wild type JM103 strain (see Table 3.2). Induction of the SOS regulon resulted in a slight increase in error free and mutagenic bypass events past the AAF adduct, however, such an effect seems to be only marginally important in the case of -G frameshift TLS events (P=0.029). On the other hand, the combined probabilities Chi square test produced a P<0.005 value that might indicate that some SOS function other than UmuDC might have played a role in contributing to the elevated TLS events. Among the candidate functions the role of Pol II in -G frameshift mutagenesis in the SmaG3-AAF sequence context has been ruled out by *in vivo* (Napolitano et al., 2000) and *in vitro* primer extension studies (Becherel and Fuchs, 2001). Pol IV that was originally shown to be involved in untargeted mutagenesis (see section 1.5.3.3.1) is likely to have a role in TLS past chemically modified base adducts. Such a lesion targeted mutagenic function for Pol IV has been shown for a BaP

modified guanine residue (Napolitano et al. 2000). According to Napolitano et al. (2000) the Pol IV plays as important a role as Pol V in promoting -G mutations in SmaG3-BaP target. In the following year Shen et al. (2002) showed that Pol IV but not Pol V was involved in insertion of a correct C residue across from the BaP modified guanine. Taken together with the Napolitano et al. (2000) observation, this might suggest that -G mutagenic readthrough past BaP guanine adduct required concerted actions of both Pol IV and Pol V, and that Pol IV might be involved in the correct C insertion step.

Considering the similar bulky nature of the BaP and AAF adducts it is tempting to consider the possible involvement of Pol IV in mutagenic synthesis past the AAF adduct. Although, in an *in vivo* study Napolitano et al. (2000) were not able to show a significant reduction in the frequency of -G mutagenic TLS past the SmaG3-AAF lesion in a *dinB* mutant strain, failure to detect a role for Pol IV might be due to at least two reasons. First, the study utilized strains that were only mutated for a single function, i.e. *dinB* or *umuDC* genes. Considering the demonstrated involvement of Pol V (*umuDC* genes product) in AAF induced -G mutations, detection of the low frequency Pol IV dependent frameshifts would be difficult. Second, in general the use of double stranded vectors carrying a single AAF adduct makes the detectability of the mutations even more difficult since the majority of the adducts can be bypassed by damage avoidance mechanisms resulting in much lower TLS frequencies.

The observations by Napolitano et al. (2000) do not necessarily exclude the intrinsic capacity of Pol IV in low frequency TLS past the G-AAF adduct. Indeed, *in vitro* studies not only showed that the Pol IV was able to insert primarily a correct C residue across from the G-AAF adduct but also caused other nucleotides to be misincorporated with a

preference order of A>T>G (Suzuki et al., 2001). In addition, induction of targeted -1 and -2 frameshifts (Suzuki et al., 2001) is an indication of the likelihood of involvement of Pol IV in *in vivo* mutagenic TLS past AAF adduct. It is noteworthy that like many *in vitro* primer extension studies Suzuki et al. (2001) used an excess molar ratio (50-fold) of the Pol IV over primer terminus. Considering the high heterology index value of Pol IV (HI=12.84, see also Table 1.1), the *dinB* gene would be derepressed at early stages after the SOS inducing signal is received and the protein would accumulate up to 2500 copies per cell when it is fully derepressed (Table 1.1). These conditions would provide sufficient intracellular pools of the polymerase per single copy of the ss-vector carrying a single AAF adduct that is transformed into the cell. Considering all these facts it seems reasonable to speculate that the Pol IV might have a role in TLS past the SmaG3-AAF adduct in JM103 Δ *umuC* strain, however, further experimentation is required to assess such a possibility.

4.4.4. Base substitution TLS events past SmaG3-AAF adduct is dependent on SOS regulated Pol V and Pol RI functions: The SmaG3-AAF induced base substitution TLS events comprise a relatively low proportion of the observed mutagenic events, i.e. 10- to 40-fold less than frameshifts, which account for 0.24%-2.6% of total bypasses in all cellular backgrounds regardless of their SOS induction state (Tables 3.2, 3.3, 3.5). The considerably lower occurrence of base substitutions as compared to frameshifts is a reflection of the importance of the reiterated nature of the target sequence that makes it a suitable hotspot for frameshift mutation. In addition the relatively low frequency of base substitution is probably due to the preferential tendency of the polymerases to incorporate primarily a correct C residue across from the G-AAF adduct. Overall, in all cellular

backgrounds misincorporation of purine nucleotides is preferred over T misinsertion, which is consistent with other reported observations (Bichara and Fuchs, 1985; Melchior *et al.*, 1994; see also section 1.6). Insertion of a T across from G3-AAF is the least frequent event and in some cases appeared to be undetectable. In the wild type JM103 background insertion of mismatching bases follows the order of A>G>T in an SOS⁻ background resulting in G→T, G→C and G→A mutations, respectively. Induction of the SOS functions resulted in an overall increase in TLS frequencies ranging from 2- to 6-fold across different classes of base substitutions with a preference given for purine nucleotide insertion (Table 3.2). In the absence of the functional Pol V the base substitution mutations were reduced by a factor of 8-40 in a JM103 Δ *umuC* strain regardless of the state of the cells for SOS induction (Table 3.3). This observation together with the SOS inducible nature of these mutations in the wild type strain suggested that Pol V was required for base misinsertions opposite the G-AAF adduct and their extension. In addition, expression of Pol RI from pGW1700 plasmid not only restored the base substitutions but also caused the TLS frequencies to exceed those that were measured in the wild type strain by reaching 63- to 240-fold higher level than *umuC* strain (Table 3.5). In a similar manner to Pol V the Pol RI showed the least preference for T insertion, and purine base incorporation across from the G-AAF adduct was preferred (Table 3.5), which is in agreement with other reports for prokaryotic and eukaryotic cells (Bichara and Fuchs, 1985; Melchior *et al.*, 1994; see also section 1.6). In Pol RI expressing cells the lack of extensive elevation of mutagenic TLS events in general, and especially base substitution events following SOS induction as compared to SOS⁻ cells, might be the result of the saturation of the intracellular pool of functional Pol

RI due to the ease of the proteolytic processing of MucA protein. Incorporation of any base other than C would result in formation of a mismatched replicational terminus. The relative levels of instability of these mismatched termini and different efficiencies of PolV versus Pol RI in promoting extension from these termini explains the variable frequency of base substitution mutations in different cellular backgrounds.

Several studies in prokaryotes and eukaryotes have shown the insertion of guanine and adenine nucleotides across from dG-AAF and dG-AF adducts that result in frequent induction of G→T and G→C transversions (Bichara and Fuchs, 1985; Carothers *et al.*, 1989 and 1993; Melchior *et al.*, 1994; Roy and Fuchs, 1994; Shibutani *et al.*, 1998). The NMR structure of dG-AF mismatched to G and A residues revealed a base displaced structure in which the dG-AF showed a *syn* glycosidic alignment, however, no base pairing was observed between the modified G and incoming G or A residues (Norman *et al.*, 1989; Abuaf *et al.*, 1995). It is likely that similar local DNA structural alterations will be induced by the G-AAF adduct, however, insight into the details of such changes will require structural studies. Whether selection and insertion of any of these mismatching bases require a transient hydrogen bond formation with the dG-AAF residue and/or the sugar-phosphate backbone, the role of the polymerases in facilitating the process would need further structural studies of a ternary complex formed between the polymerase/template and the primer strand that carries these mismatches.

4.4.5. TLS events past SmaG1-AAF adduct are modulated by SOS regulated error prone polymerases and the position of the AAF adduct in *SmaI* sequence context:

Compared to G3-AAF, the G1-AAF adduct exhibits extremely weak mutagenic capability. Earlier studies using an excision repair mutant JM103*uvrA* strain and double

stranded vectors that carried a *SmaI* site with a single AAF modification on either G1, G2 or G3 bases, measured the relative contributions of each of these modifications to the induced targeted -G frameshifts in contiguous G repeats and semi targeted -C frameshifts in the reiterated C residues located upstream from the modified G on the template strand (Lambert et al., 1992a). In comparison to G3-AAF, the frequency of targeted -G frameshifts induced by the G1-AAF modification was reduced by 218-fold to 0.007%, while -C semi-targeted frameshifts were induced by 74- and 132-fold higher efficiency than targeted -G frameshifts when the AAF was located on G1 and G2, respectively (Lambert et al., 1992a). These observations suggested that the position of the the AAF modification is an additional factor that modulates the nature of the TLS events past the G-AAF adduct. Since -G1 events were induced rarely by the G1-AAF adduct, therefore we were interested in measuring the most likely TLS events past a G1-AAF adduct, i.e. error free, G1→T and -C in Pol RI background. In JM103/*umuC* that expresses Pol RI the TLS events in a single stranded vector that carried the SmaG1-AAF modification were primarily error free (99%); the frequency of error free TLS was increased by 5-fold when the SOS regulon was induced (Table 3.6). The semi targeted -C frameshifts and G→T transversions were similarly induced by 10- and 5-fold, respectively, upon SOS induction (Table 3.6), which is an indication of the possible involvement of Pol RI in facilitating the TLS past G1-AAF adduct. The G→T base substitution events were less frequent at G1-AAF by 4- to 10-fold as compared to G3-AAF (Tables 3.6 and 3.5). This suggested that the sequences flanking the DNA adduct influence the extension of primer/template termini. Occurrence of the semi-targeted frameshifts are an indication of the ability of the G1-AAF to promote strand slippage in upstream runs of C. Taking the

observations obtained in this work and those of Lambert et al. (1992) together, the following model is proposed to explain the cellular TLS past the SmaG1-AAF adduct. A Umu-like polymerase e.g. Pol RI inserts a correct C residue at high frequency across from G1-AAF adduct. Extension from the C terminus and successive incorporation of the complementary nucleotides results in stabilization of the primer strand and the most frequent error free TLS outcome. However, G insertion following C:dG₁-AAF might result in formation of a subpopulation of an isomeric SMI conformation through slippage of the G terminus by one base and base pairing with the 5'-C₂ (Fig. 4.3). Such a slippage event is believed to be initiated and facilitated by the instability imposed on the terminal G:C₁ by the lateral transition of the instability that was created originally by the C:dG₁-AAF pair resulting in an increased frequency of slippage (Woodsen and Crothers, 1987; Kalnik et al., 1989; Lambert et al., 1992). Extension from the slipped intermediate results in a low frequency -C₁ semi targeted frameshift outcome (Fig. 4.3).

Alternatively, insertion of an A or a G nucleotide opposite dG1-AAF adduct can be promoted at a considerably lower frequency as compared to C insertion. Extension from these mispaired termini account for the low frequency G→T and -G₁ frameshifted TLS events. Unlike G→T TLS that requires a Pol RI dependent one step extension from A:dG1-AAF, the extension from the G terminus requires formation of a misaligned slipped intermediate in order to result in -G₁ frameshift. Such a SMI can be formed following the G misincorporation opposite the G1-AAF and its slippage to misalign with the C residue immediately 5' to the G1-AAF that leaves the adducted G in an extrahelical bulge (Fig. 4.3). Extension from the misaligned terminus followed by the second round replication of the nascent strand will fix the mutation.

Figure 4.3: TLS events past the SmaG1-AAF adduct. Possible mechanisms involved in induction of error free, semi targeted -C and targeted -G1, G→T and G→C are shown schematically. The G1 residue containing an AAF modification is shown in red. Incorporated bases across from the G1-AAF are shown in blue (See section 4.5.5 for details).

4.4.6. Unlike error free TLS the error prone -2 frameshift TLS events past NarG3-AAF adduct are dependent on an SOS regulated polymerase in addition to Pol V and Pol RI: TLS events past the NarG3-AAF adduct, which are carried on the single stranded circular vector pLHIINarG3-AAF, were studied using the wild type JM103 strain, the JM103 Δ umuC mutant strain and its derivatives that express Pol V or Pol RI *in trans* from the pDC322 and pGW1700 plasmids, respectively. In the wild type JM103 strain error free TLS is the major bypass event that accounts for 65-70% of total translesion synthesis past the NarG3-AAF adduct regardless of the state of the cells for SOS induction (Table 3.7). Induction of SOS functions resulted in about a 5-fold increase in error free TLS frequency ($P=0.003$), suggesting that an SOS inducible function(s) is involved in readthrough past the AAF adduct (Table 3.7). Elimination of Pol V activity in the JM103 Δ umuC strain resulted in about a 2.2-fold reduction in error free TLS past the same adduct, suggesting that Pol V contributed to the increased frequency of error free bypass that was observed in the wild type strain. Further evidence for involvement of Pol V in promoting error free TLS in the *NarI* sequence came from our observation of a 2- to 8-fold increased error free TLS in JM103 Δ umuC strain when Pol V was expressed *in trans* from the pDC322 plasmid as compared to JM103 Δ umuC strain that was mock transformed with pBR322 plasmid (Table 3.8). In order to see whether a similar effect could be promoted by other UmuDC like polymerases, the error free TLS events were measured in JM103 Δ umuC expressing the Pol RI from pGW1700 plasmid, which resulted in a 5- to 7-fold increase in error free TLS past NarG3-AAF adduct as compared to JM103 Δ umuC (Table 3.10). In both cases the error free TLS accounted for the majority of the bypass events regardless of SOS induction reaching to

70-80% of total bypass events. Induction of the SOS regulon in Pol V and Pol RI expressing JM103 Δ *umuC* strain resulted in 3.8-fold (P=0.003) and 1.6-fold (P=0.002) increases in error free bypasses that confirmed the SOS inducible nature of these polymerases. Taking all these observations together, it is concluded that Pol V and/or Pol RI are the major polymerases that are primarily involved in error free synthesis past the NarG3-AAF adduct.

Although inactivation of the *umuC* gene in the JM103 strain reduced the error free TLS past the NarG3-AAF adduct, induction of the SOS regulon resulted in at least a 2-fold increase in error free TLS frequency which appeared to be moderately significant (P=0.02). However, interestingly, in the absence of Pol V the -CG frameshifted TLS was elevated significantly by 6- to 11-fold in SOS⁻ and SOS⁺ backgrounds, respectively, as compared to wild type JM103 strain (Table 3.8). This observation is consistent with the hypothesis that there is at least one additional SOS inducible function other than Pol V and Pol RI that is capable of promoting efficient error prone synthesis past the NarG3-AAF adduct; however, this function may not be as efficient as Pol V or Pol RI in inducing error free bypass. The nature of the SOS inducible function has recently been investigated by *in vivo* and *in vitro* studies (Napolitano et al., 2000; Becherel and Fuchs, 2001). It was shown that in Δ *polB* mutant *E.coli* (lacking Pol II) the -CG frameshifts at the *NarI* target were reduced by 95-fold (Napolitano et al., 2000). In addition, in an *in vitro* primer extension study using a template that carried the NarG3-AAF adduct, the robust capacity of Pol II to synthesize past the modified G and produce a synthesis product that was short by 2 bases was demonstrated (Becherel and Fuchs, 2001). Based on our results and the other supporting *in vivo* and *in vitro* evidence, and considering the

great correlation of 5-fold increase in copy number of Pol II (Table 1.1) with a 5-fold elevation of -2 frameshifts in a *umuC* cellular background in response to activation of the SOS regulon (Table 3.8), and since the involvement of the Pol IV has already been ruled out (Napolitano et al., 2000), it seems plausible to consider Pol II as the major polymerase involved in the -2 frameshift pathway with a limited capacity to promote SOS inducible error free bypasses within this DNA sequence context (Table 3.8).

What role might Pol V and Pol RI have in -2 pathway? Expression of Pol V and Pol RI from pDC322 or pGW1700 plasmids, resulted in 4- to 10-fold and 2- to 11-fold reduction in -2 frameshifts in Pol V and Pol RI expressing *ΔumuDC* strain, respectively, with a concomitant increase in error free bypasses at 2-8 and 5-7 higher frequencies (Tables 3.9 and 3.10). Although, this antagonistic effect of Pol V and Pol RI with Pol II is responsible for the shift of the -2 mutagenic TLS in favor of more prevalent error free bypass events, it does not necessarily rule out the possible role of the Umu-like polymerases in the -2 pathway. Considering that the overall copy number of these functions, even when encoded chromosomally, supercedes the copy number of the Pol II (Table 1.1), and the inherent capacities of both Pol V and Pol RI for error free TLS past the NarG₃-AAF adduct, one might consider a limited minor role for these polymerases in the -2 pathway. In the absence of a direct *in vitro* assay, this notion is supported by indirect *in vivo* data. For example, the UV dose dependent induction of -2 frameshifts in NarG₃-AAF target in a *ΔumuDC* background showed lower TLS at all dosages with a plateau at 25% TLS frequency, while in a wild type strain the TLS frequency reached the plateau at 40% (Fuchs et al., 2001), which was congruent with another report that showed that in the *ΔumuDC* background the -2 frameshifts were reduced by about 1.6-fold as

compared to the wild type strain (Napolitano et al., 2000), suggesting a minor role for Pol V in the -2 pathway. All together, these observations are consistent with our findings upon which the following TLS model can be proposed.

The primer strand becomes accessible to either of Pol V, Pol RI or Pol II upon dissociation of the Pol III at AAF modified guanosine residue. The fate and nature of the TLS event is determined by the conformation of the lesion terminus and the competitive access to this terminus by any of the available polymerases. Upon insertion of a correct C residue across from the G-AAF adduct, a non-slipped lesion terminus is formed that comprises the major population of the primer termini (Fig. 4.4). Extension from the non-slipped lesion terminus is preferentially performed by either of the Pol V or Pol RI resulting in error free bypass events that accounts for the majority of the TLS events when these polymerases are expressed in the cells. However, extension from this terminus by Pol II is less frequent due to either limited affinity of Pol II to this terminus or its limited extension capacity, or both. Spontaneous lesion terminus isomerization leading to formation of a misaligned slipped primer strand, which as the basis for -2 frameshift events, might occur by one or two Watson-Crick base pairing of the slipped priming strand with the bases upstream of the lesion site on the template strand, leading to minor and major slipped mutagenic intermediate structures, respectively (Fig. 4.4). The major SMI is formed by slippage of the 3'-CG-5' bases in the priming strand and base pairing with 5'-G₂C-3' bases immediately 5' to the modified G on the template strand, while the minor SMI is formed by slippage of the C residue and formation of one Watson-Crick base pair with a G residue at the penultimate position 5' to modified G. The major and minor SMI result in formation of either of 5'-G₃C-3' or 5'-CG₃-3' bulged regions on

Fig 4.4: Non-mutagenic TLS event past an AAF modified guanine residue (G_3^*) in NarI site, results from incorporation of a cytosine residue across from the AAF modified G and extension thereafter. While the -2 mutagenic pathway is promoted by incorporation of a C residue across from the modified G and formation of two possible slipped mutagenic intermediates that are in equilibrium with the non-slipped form (Milhe *et al.*, 1996). Extension from the minor SMI is probably promoted by PolV or PolRI while extension from major SMI is preferred by PolII.

template strand, respectively (Milhe *et al.*, 1996). Extension from either of these SMI, results in a -2 frameshift mutation (Fig. 4.4) that might require different polymerases. In contrast to Pol II the Pol V/Pol RI dependent slipped primer extensions might be favored from the minor SMI rather than major SMI. Binding of Pol V and Pol RI to the major SMI will also result in competitive blocking of access of Pol II to its favored lesion terminus, the major SMI. This might explain the antagonistic effect of the Pol V and Pol RI against Pol II in reducing the -2 frameshifts. However, testing the preferential extension from minor SMI by Pol V and Pol RI awaits *in vitro* primer extension studies. Alternatively, one might consider the possibility that the minor SMI is not preferred by either polymerase; however, Pol V and Pol RI could exhibit less preference for extension from the major SMI in comparison to Pol II while the extension from the non-slipped terminus is preferred by Pol V and Pol RI. In this case the antagonistic effect of the Pol V and Pol RI against Pol II in promoting -2 frameshifts and the preference for error free TLS events in the presence of Pol V/Pol RI, would be the sole effect of the competitive access of the different polymerases to the lesion termini.

4.5. General model for mutagenesis and concluding remarks: Currently in the literature it is well established that cells, either prokaryotic or eukaryotic, harbor multiple numbers of specialized polymerases in addition to major replicative polymerases. These polymerases transiently take over synthesizing DNA and manage the replication fork when normal replication comes to a halt or is interrupted due to an alteration of the bases by physical or chemical means, or when the continuity of the template bases is disturbed. The sequence of events that results in restoration of normal replication by involvement and mediation of the error prone polymerases is called translesion synthesis, which is

distinct from other cellular mechanisms in having a capacity to synthesize past damaged bases with a certain risk of causing mutations. In *E.coli* the members of error prone polymerases include at least three out of five known polymerases that are damage inducible; these polymerases vary in their specificities and capacities to promote TLS past DNA damages of different natures but exhibit a certain substrate overlap. Considering their mutagenic capacities one would realize that these error prone polymerases should have restricted access to the replication fork; in that case, such access should be as short as possible in order to avoid excessive mutagenesis. On the other hand, considering the generally overlapping substrate specificities of Y-class of polymerases one may ask what determines which polymerase or polymerases deal with certain DNA damage that resides in the replication fork? It is also logical to ask, why over billions of years of evolutionary time scale, nature evolved polymerases that are prone to error rather than evolving a sophisticated ideal machinery that could precisely copy the DNA at the damaged sites without having the trouble of being stalled? Obviously, the time was not a constraint and the number of the evolutionary experiments that nature has conducted has been infinite and still continues. It is possible that such a sophisticated machinery has proven useless in an evolutionary sense. The reason would be that organisms tend to survive more successfully in any milieu and at any time scale if they equip themselves with resources and capacities that enable them to function between the two extremes of physiological death and genetic death. While highly accurate polymerases with a capacity to bypass any replicational road blocks would ultimately reduce and prevent genetic diversification and drive the organisms toward genetic death, an uncontrolled high mutation rate would knock out essential gene products and cause physiological death. The evolution of more

intricate regulatory mechanisms can, on one hand, offer the cells a short term benefit of TLS over the replicational impediments and may ultimately increase cellular survival in a real time challenging environment, and on the other hand may enhance the ability of the organism to survive over the long evolutionary time scale. The former has been best reflected by our data where it was shown that the SOS polymerases were able to increase the cellular survival by contributing primarily to the error free TLS events and less frequently to mutagenic processes.

How do the cells hand over the management of replication in the vicinity of DNA damage from a high fidelity replicative enzyme to a less accurate polymerase, and how do they monitor the activity of the error prone enzyme in order to assure its transient function? Although, the inducible nature of these enzymes in *E.coli* would impose a transcriptional control on the available copies of these enzymes per cell, reliance on mass effect and diffusion to the site of action, may be over simplistic if considered as the only mechanism at work. The highly organized heirarcheal temporal activation of the SOS inducible functions, and the pertinent signalling circuitry at transcriptional and post translational levels, has already been discussed in section 1.5.3.1.1. It is also possible that such a delicate control would also rely heavily on higher order protein-protein and protein-DNA interactions possibly at all levels. Indeed, formation of such multi protein complexes has been shown by Graham Walker's lab and suggests the necessity of interactions between the elements of the SOS regulon, including error prone polymerases, and the subunits of Pol III. If so, then access of the error prone polymerases to the damage site would not necessarily rely on a passive diffusion along the DNA but could be part of a highly coordinated and regulated process. The first step in polymerase switching

would require removal of the Pol III that has already been stalled in front of the damaged base. Whether Pol III simply falls off the stalling site after several futile attempts to incorporate a base opposite the damaged base, or is removed actively by an outcompeting incoming polymerase, is not relevant to the final outcome. However, it has been shown that all five *E.coli* polymerases including the SOS controlled polymerases do have intrinsic affinity to the β sliding clamp, the Pol III processivity factor (Stukenberg et al., 1991), which is deemed to be essential for targeting the incoming DNA polymerase to the primer terminus (Bonner et al., 1992; Sutton et al., 2001a; Lopez de Saro and O'Donnell, 2001; Lenne-Samuel et al., 2002). Lenne-Samuel et al. (2002) have shown that the Pol IV- β interaction is important for trafficking the polymerase to the site of DNA damage; this might suggest a similar role for β with other error prone polymerases.

In the case of Pol V it has been shown that UmuD₂C, and UmuD'₂C interact, albeit with varying affinities, with the same face of the β sliding clamp that γ clamp loader and α catalytic subunits do (Sutton et al., 2001a). This would mean that interaction of Pol V with the β clamp would antagonize and possibly eliminate its interaction with α and γ subunits of Pol III (Fig. 4.5). Interaction of the UmuD₂C with β halts the DNA replication and progression of normal cellular growth, while interaction of β with UmuD'₂C, the active Pol V, results in TLS. Besides the general role of the β clamp in recruiting error prone polymerases to the DNA damage site, the RecA*-ssDNA nucleofilament that is formed ahead of the halted replicative Pol III (see section 1.5.3.1.1), serves to localize UmuD' in the vicinity of the DNA damage and hence its increased access to the β clamp. Interaction between Pol V and β not only is necessary for competitive installation of Pol V at the damage site, but also, as *in vitro* kinetic

Fig. 4.5: A general mutagenesis model. A) Pol III stalls in front of a dG.AAF adduct. B) RecA nucleofilament is formed on template ssDNA upstream to adduct. C) Pol IV binds to β and Pol III falls off the template. D) SSB joins the complex. E) bidirectional disassembly of RecA filament. F) Pol V loses contact with RecA and dissociates from template. G) Pol III replaces Pol V and resumes DNA replication.

studies have shown, such an interaction activates the catalytic function of Pol V by at least 3-fold by lowering the $K_{m_{\text{dNTP}}}$ without affecting the V_{max} (Pham et al., 2001a). Correct positioning of Pol V at the 3' end of the primer strand is mediated by the protein-protein contacts that are made between Pol V and the 3'-end of the RecA* nucleofilament (Frank et al., 1993; Frank et al., 2000). In addition, physical contact between Pol V and RecA as well as SSB further stimulates the polymerase function of the Pol V by 340- and 1040-fold, respectively (Pham et al., 2001a).

The Pol V mediated distributive translesion synthesis past the template lesion is justified by the length of the RecA* nucleofilament that lies ahead of the polymerase, which ensures the transient stay of the Pol V on the template and hence prevents untargeted mutations (Fig 4.5). This is achieved by bidirectional disassembly of the RecA filament that results in removal of the RecA from the template a few bases beyond the DNA lesion. The normal 5'→3' disassembly is achieved in an ATP dependent manner while the ATP independent 3'→5' disassembly is catalyzed by the combined actions of Pol V and SSB during DNA synthesis while the Pol V association with β/γ complex maintains its contact with the 3'-tip of the nucleofilament. As a result of the concomittent bidirectional disassembly of RecA the short lived limited synthesis past the template lesion is ensured. At the end of Pol V-catalysed TLS, it is likely that Pol V is dissociated from the template leaving the 3' end of the primer strand accessible to the Pol III subunits where it can be reassembled and processive replication can resume.

The model described above addresses two fundamental aspects of the error prone polymerase dependent TLS. First, it considers a targeting mechanism for trafficking the error prone enzyme to the template lesion site and switching it for Pol III. Second, it

considers a mechanism for the transient nature of Pol V catalysed lesion bypass synthesis. However, it does not offer a mechanism for the involvement of multiple polymerases and switching one for the other. It is possible that Pol V, together with other error prone polymerases, are part of a higher order multiprotein polymerase cassette. If so, then the model should be refined further by characterizing all pertinent protein-DNA and protein-protein complexes and assigning roles and tasks to them.

5. List of References

- Abuaf P, Hingerty BE, Broyde S and Grunberger D, (1995). Solution structure of the N-(deoxyguanosine-8-yl)aminofluorene adduct opposite deoxyinosine and deoxyguanosine in DNA by NMR and computational characterization. *Chem Res Toxicol*, 8:369-378.
- Ali JA, Lohman TM, (1997). Kinetic measurement of the step size of DNA unwinding by *Escherichia coli* UvrD helicase. *Science*, 275:377-380.
- Allaben WT, Weeks CE, Weis CC, Burger GT, King CM., (1982). Rat mammary gland carcinogenesis after local injection of N-hydroxy-N-acyl-2-aminofluorenes: relationship to metabolic activation. *Carcinogenesis*, 3:233-240.
- Amundsen SK, Taylor AF, Smith GR.A (2002) Domain of RecC Required for Assembly of the Regulatory RecD Subunit Into the *Escherichia coli* RecBCD Holoenzyme. *Genetics*, 161:483-492.
- Aoufouchi S, Flatter E, Dahan A, Faili A, Bertocci B, Storck S, Delbos F, Cocea L, Gupta N, Weill JC, Reynaud CA., (2000). Two novel human and mouse DNA polymerases of the polX family. *Nucleic Acids Res.*, 28:3684-3693.
- Arey J., Zielinska B., Harger W.P., Atkinson R. and A.M. Winer, (1988). The contribution of nitrofluoranthenes and nitropyrenes to the mutagenic activity of ambient particulate organic matter collected in southern California. *Mut.Res.*207: 45-51.
- Arinc E, Adali O, Gencler-Ozkan AM., (2000). Induction of N-nitrosodimethylamine metabolism in liver and lung by in vivo pyridine treatments of rabbits. *Arch Toxicol.*, 74:329-334.
- Arnott S, Chandrasekaran R, Birdsall DL, Leslie AG, Ratliff RL, (1981). Left-handed DNA helices. *Nature*, 283:743-745.
- Asokan P, Das M, Rosenkranz HS, Bickers DR, Mukhtar H.(1985). Topically applied nitropyrenes are potent inducers of cutaneous and hepatic monooxygenases. *Biochem Biophys Res Commun.*, 129:134-140.
- Autret S, Levine A, Holland IB, Seror SJ., (1997). Cell cycle checkpoints in bacteria. *Biochimie.*, 79:549-554.
- Babu SR, Lakshmi VM, Hsu FF, Zenser TV, Davis BB., (1992). Role of N-glucuronidation in benzidine-induced bladder cancer in dog. *Carcinogenesis*, 13:1235-1240.
- Babu SR, Lakshmi VM, Hsu FF, Zenser TV, Davis BB., (1995). Glucuronidation of N-hydroxy metabolites of N-acetylbenzidine. *Carcinogenesis*, 16:3069-3074.
- Ban C, Yang W., (1998). Structural basis for MutH activation in *E.coli* mismatch repair and relationship of MutH to restriction endonucleases. *EMBO J.*, 17:1526-1534.
- Banerjee S.K., Christensen R.B., Lawrence C.W. and Leclerc J.E., (1988). Frequency and spectrum of mutations produced by a single cis-syn thymine-thymine cyclobutane dimer in a single-stranded vector. *Proc.Natl.Acad.Sci USA*, 85:8141-8145.
- Baranov PV, Gesteland RF, Atkins JF, (2002). Recoding: translational bifurcations in gene expression. *Gene.*, 286:187-201.

- Baranov PV, Gurvich OL, Fayet O, Prere MF, Miller WA, Gesteland RF, Atkins JF, Giddings MC, (2001). RECODE: a database of frameshifting, bypassing and codon redefinition utilized for gene expression. *Nucleic Acids Res.*, 29:264-267.
- Bartsch H, Malaveille C, Stich HF, Miller EC and Miller JA.(1977). Comparative electrophilicity, mutagenicity, DNA repair induction activity, and carcinogenicity of some N- and O-acyl derivatives of N-hydroxy-2-aminofluorene. *Cancer Res.*, 37:1468-1475.
- Bartsch H., (1981). Metabolic activation of aromatic amines and azo dyes. *IARC Sci Publ.*, 40:13-30.
- Bassett E, Vaisman A, Tropea KA, McCall CM, Masutani C, Hanaoka F, Chaney SG. , (2002). Frameshifts and deletions during in vitro translesion synthesis past Pt-DNA adducts by DNA polymerases beta and eta. *DNA Repair (Amst.)*, 1:1003-1016.
- Battista JR, Donnelly CE, Ohta T, Walker GC. (1990a). The SOS response and induced mutagenesis. *Prog Clin Biol Res.*, 340:169-178.
- Battista JR, Ohta T, Nohmi T, Sun W, Walker GC., (1990b). Dominant negative umuD mutations decreasing RecA-mediated cleavage suggest roles for intact UmuD in modulation of SOS mutagenesis. *Proc Natl Acad Sci U S A.*, 87:7190-7194.
- Batty DP, Wood RD., (2000). Damage recognition in nucleotide excision repair of DNA. *Gene.*, 241:193-204.
- Beamand JA, Barton PT, Tredger JM, Price Rj and Lake BG., (1998). Effect of some cooked food mutagens on unscheduled DNA synthesis in cultured precision-cut rat, mouse and human liver slices. *Food Chem Toxicol.*, 36:455-466.
- Becherel OJ, Fuchs RP., (2001). Mechanism of DNA polymerase II-mediated frameshift mutagenesis. *Proc Natl Acad Sci U S A*, 98:8566-8571.
- Begg, K. J., S. J. Dewar, and W. D. Donachie. 1995. A new Escherichia coli cell division gene. *ftsK*. *J. Bacteriol.* 177:6211-6222.
- Beije B and Moller, (1988) 2-Nitrofluorene and related compounds: prevalence and biological effects. *Mutat Res.* 196:177-209.
- Beland F.A., Beranek D.T., Dooly K.L., Heflich R.H., Kadlubar F.F., 1983. Arylamine-DNA adducts in vitro and in vivo: their role in bacterial mutagenesis and urinary bladder carcinogenesis. *Environ. Health. Perspect.*, 49: 125-134.
- Beland F.A. and Kadlubar F.F., (1985). Formation and persistence of arylamine DNA adducts in vivo *Environ. Health. Perspect.*, 62: 19-30.
- Beland FA, Beranek DT, Dooley KL, Heflich RH, Kadlubar FF., (1983). Arylamine-DNA adducts in vitro and in vivo: their role in bacterial mutagenesis and urinary bladder carcinogenesis. *Environ Health Perspect.*, 49:125-134.
- Beland FA, Kadlubar FF., (1985). Formation and persistence of arylamine DNA adducts in vivo. *Environ Health Perspect.*, 62:19-30.
- Beland FA, Melchior WB Jr, Mourato LL, Santos MA and Marques MM., (1997). Arylamine-DNA adduct conformation in relation to mutagenesis. *Mutat Res.*, 376:13-19.
- Belguise-Valladier P, Maki H, Sekiguchi M, Fuchs RP., (1994). Effect of single DNA lesions on in vitro replication with DNA polymerase III holoenzyme. Comparison with other polymerases. *J Mol Biol.*, 236:151-164.

- Beranek DT, White GL, Heflich RH, Beland FA., (1982). Aminofluorene-DNA adduct formation in *Salmonella typhimurium* exposed to the carcinogen N-hydroxy-2-acetylaminofluorene. *Proc Natl Acad Sci U S A.*, 79:5175-5178.
- Berry DF, Francis AJ, Bollag JM., (1987). Microbial metabolism of homocyclic and heterocyclic aromatic compounds under anaerobic conditions. *Microbiol Rev.*, 51:43-59.
- Bessho T, Nitta N, Negishi K, Hayatsu H. (1992). Blockage of polymerase-catalyzed DNA chain elongation by chemically modified cytosine residues in templates and the release of blockage for readthrough. *Nucleic Acids Res.*, 20:4213-4220.
- Bichara M, Fuchs RP., (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. *J Mol Biol.*, 183:341-351.
- Bitsch A, Hadjiolov N, Klohn PC, Bergmann O, Zwirner-Baier I and Neumann HG, (2000). Dose response of early effects related to tumor promotion of 2-acetylaminofluorene., *Toxicol Sci.*, 55:44-51.
- Bitsch A, Klohn PC, Hadjiolov N, Bergmann O, Neumann HG., (1999). New insights into carcinogenesis of the classical model arylamine 2-acetylaminofluorene. *Cancer Lett.*, 143:223-227.
- Blackwell LJ, Martik D, Bjornson KP, Bjornson ES, Modrich P., (1998). Nucleotide-promoted release of hMutS α from heteroduplex DNA is consistent with an ATP-dependent translocation mechanism. *J Biol Chem.*, 273:32055-32062.
- Bonner CA, Stukenberg PT, Rajagopalan M, Eritja R, O'Donnell M, McEntee K, Echols H, Goodman MF., (1992). Processive DNA synthesis by DNA polymerase II mediated by DNA polymerase III accessory proteins. *J Biol Chem.*, 267:11431-11438.
- Boshoff HI, Reed MB, Barry CE 3rd, Mizrahi V., (2003). DnaE2 polymerase contributes to in vivo survival and the emergence of drug resistance in *Mycobacterium tuberculosis*. *Cell.*, 113:183-193.
- Boudsocq F, Campbell M, Devoret R, Bailone A., (1997). Quantitation of the inhibition of Hfr x F- recombination by the mutagenesis complex UmuD'C. *J Mol Biol.*, 270:201-211.
- Boudsocq F, Ling H, Yang W, Woodgate R., (2002). Structure-based interpretation of missense mutations in Y-family DNA polymerases and their implications for polymerase function and lesion bypass. *DNA Repair (Amst).*, 1:343-358.
- Bouziane M, Miao F, Ye N, Holmquist G, Chyzak G, O'Connor TR., (1998). Repair of DNA alkylation damage. *Acta Biochim Pol.*, 45:191-202.
- Brent R, Ptashne M., (1981). Mechanism of action of the *lexA* gene product. *Proc Natl Acad Sci U S A.*, 78:4204-4208.
- Bridges BA, Mottershead RP., (1976). Mutagenic DNA repair in *Escherichia coli*. III. Requirement for a function of DNA polymerase III in ultraviolet-light mutagenesis. *Mol Gen Genet.*, 144:53-58.
- Bridges BA, Woodgate R., (1984). Mutagenic repair in *Escherichia coli*. X. The *umuC* gene product may be required for replication past pyrimidine dimers but not for the coding error in UV-mutagenesis. *Mol Gen Genet.*, 196:364-366.
- Bridges BA, Woodgate R., (1985a). The two-step model of bacterial UV mutagenesis. *Mutat Res.*, 150:133-139.

- Bridges BA, Woodgate R., (1985b). Mutagenic repair in *Escherichia coli*: products of the *recA* gene and of the *umuD* and *umuC* genes act at different steps in UV-induced mutagenesis. *Proc Natl Acad Sci U S A.*, 82:4193-4197.
- Brotcorne-Lannoye A, Maenhaut-Michel G., (1986). Role of RecA protein in untargeted UV mutagenesis of bacteriophage lambda: evidence for the requirement for the *dinB* gene. *Proc Natl Acad Sci U S A.*, 83:3904-3908.
- Bruck I, Goodman MF, O'Donnell M., (2003). The essential C family DnaE polymerase is error-prone and efficient at lesion bypass. *J Biol Chem.*, 278:44361-44368.
- Bruck I, Woodgate R, McEntee K, Goodman MF., (1996). Purification of a soluble UmuD'C complex from *Escherichia coli*. Cooperative binding of UmuD'C to single-stranded DNA. *J Biol Chem.*, 271:10767-10774.
- Bryant DW, McCalla DR, Luetschik P, Quilliam MA, McCarry BE., (1984). Metabolism of 1,8-dinitropyrene by *Salmonella typhimurium*. *Chem Biol Interact.*, 49:351-368.
- Burckhardt SE, Woodgate R, Scheuermann RH, Echols H., (1988). UmuD mutagenesis protein of *Escherichia coli*: overproduction, purification, and cleavage by RecA. *Proc Natl Acad Sci U S A.*, 85:1811-1815.
- Burnouf D, Fuchs RP. (1985). Construction of frameshift mutation hot spots within the tetracycline resistance gene of pBR322. *Biochimie*, 67:385-389.
- Burnouf D, Koehl P, Fuchs RP, (1989). Single adduct mutagenesis: strong effect of the position of a single acetylaminofluorene adduct within a mutation hot spot. *Proc Natl Acad Sci U S A.*, 86:4147-4151.
- Butler MA, Iwasaki M, Guengerich FP, Kadlubar FF., (1989). Human cytochrome P-450PA (P-450IA2), the phenacetin O-deethylase, is primarily responsible for the hepatic 3-demethylation of caffeine and N-oxidation of carcinogenic arylamines. *Proc Natl Acad Sci U S A.*, 86:7696-7700.
- Carothers AM, Steigerwalt RW, Urlaub G, Chasin LA, Grunberger D, (1989). DNA base changes and RNA levels in N-acetoxy-2-acetylaminofluorene-induced dihydrofolate reductase mutants of Chinese hamster ovary cells *J Mol Biol*, 208:417-428.
- Carothers AM, Urlaub G, Mucha J, Yuan W, Chasin LA, Grunberger D, (1993). A mutational hot spot induced by N-hydroxy-aminofluorene in dihydrofolate reductase mutants of Chinese hamster ovary cells. *Carcinogenesis*, 14:2181-2184.
- Carothers AM, Urlaub G, Mucha J, Yuan W, Chasin LA, Grunberger D., (1993). A mutational hot spot induced by N-hydroxy-aminofluorene in dihydrofolate reductase mutants of Chinese hamster ovary cells. *Carcinogenesis.*, 14:2181-2184.
- Casarett LJ, Doull J, Klaassen CD and Amdur MO, (1986). Casarett and Doull's toxicology: The basic science of poisons. 3rd ed. , Macmillan, New York.
- Cavalieri E, Rogan E, Higginbotham S, Cremonesi P and Salmasi S. (1988). Tumor-initiating activity in mouse skin and carcinogenicity in rat mammary gland of fluorinated derivatives of benzo[a]pyrene and 3-methylcholanthrene. *J Cancer Res. Clin. Oncol.*, 114:16-22.

- Chen HW, Randle DE, Gabbidon M, Julin DA. (1998). Functions of the ATP hydrolysis subunits (RecB and RecD) in the nuclease reactions catalyzed by the RecBCD enzyme from *Escherichia coli*. *J Mol Biol.*, 278:89-104.
- Chiapperino D, Kroth H, Kramarczuk IH, Sayer JM, Masutani C, Hanaoka F, Jerina DM, Cheh AM., (2002). Preferential misincorporation of purine nucleotides by human DNA polymerase ϵ opposite benzo[a]pyrene 7,8-diol 9,10-epoxide deoxyguanosine adducts. *J Biol Chem.*, 277:11765-11771.
- Cho BP and Beland FA, Marques MM., (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.
- Cho BP and Zhou L., (1999). Probing the conformational heterogeneity of the acetylaminofluorene-modified 2'-deoxyguanosine and DNA by 19F NMR spectroscopy. *Biochemistry.*, 38:7572-7583.
- Christmann M, Tomicic MT, Roos WP, Kaina B., (2003). Mechanisms of human DNA repair: an update. *Toxicology.*, 193:3-34.
- Chung JG, Levy GN, Weber WW., (1993). Distribution of 2-aminofluorene and p-aminobenzoic acid N-acetyltransferase activity in tissues of C57BL/6J rapid and B6.A-NatS slow acetylator congenic mice. *Drug Metab Dispos.*, 21:1057-1063.
- Clark AJ, Sandler SJ., (1994). Homologous genetic recombination: the pieces begin to fall into place. *Crit Rev Microbiol.*, 20:125-142.
- Coulondre, C., J. H. Miller, P. J. Farabaugh, and W. Gilbert. (1978). Molecular basis of base substitution hot spots in *Escherichia coli*. *Nature*, 274:775-780.
- Courcelle J, Carswell-Crompton C, Hanawalt PC., (1997). recF and recR are required for the resumption of replication at DNA replication forks in *Escherichia coli*. *Proc Natl Acad Sci U S A.*, 94:3714-3719.
- Courcelle J, Crowley DJ, Hanawalt PC., (1999). Recovery of DNA replication in UV-irradiated *Escherichia coli* requires both excision repair and recF protein function. *J Bacteriol.*, 181:916-922.
- Courcelle J, Khodursky A, Peter B, Brown PO, Hanawalt PC., (2001). Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli*. *Genetics.*, 158:41-64.
- Cox MM., (2000). Recombinational DNA repair in bacteria and the RecA protein. *Prog Nucleic Acid Res Mol Biol.*, 63:311-366.
- Cramer J.W., Miller J.A. and Miller E.C., (1960). N-hydroxylation: A new metabolic reaction observed in the rat with the carcinogen 2-acetylaminofluorene. *J.Biol.Chem.*, 235:885-888.
- Crowley, D. J. and P. C. Hanawalt, 1998 Induction of the SOS response increases the efficiency of global nucleotide excision repair of cyclobutane pyrimidine dimers, but not 6-4 photoproducts, in UV-irradiated *Escherichia coli*. *J. Bacteriol.* 180:3345-3352.
- Danielson PB, Foster JL, Cooper SK, Fogleman JC., (1999). Diversity of expressed cytochrome P450 genes in the adult Mediterranean Fruit Fly, *Ceratitis capitata*. *Insect Mol Biol.* 8:149-159.
- D'Ari R (1985). The SOS system. *Biochimie.*, 67:343-347.
- D'Ari R, Huisman O., (1982). DNA replication and indirect induction of the SOS response in *Escherichia coli*. *Biochimie.*, 64:623-627.

- Doudney CO., (1990). DNA-replication recovery inhibition and subsequent reinitiation in UV-radiation-damaged *E. coli*: a strategy for survival. *Mutat Res.*, 243:179-186.
- Douhal A, Kim SK, Zewail AH., (1995). Femtosecond molecular dynamics of tautomerization in model base pairs. *Nature.*, 378:260-263.
- Duvauchelle JB, Blanco L, Fuchs RP, Cordonnier AM., (2002). Human DNA polymerase mu (Pol mu) exhibits an unusual replication slippage ability at AAF lesion. *Nucleic Acids Res.*, 30:2061-2067.
- Ebina Y, Takahara Y, Kishi F, Nakazawa A, Brent R., (1983). LexA protein is a repressor of the colicin E1 gene. *J Biol Chem.*, 258:13258-13261.
- El-Bayoumy K, Lavoie EJ, Hecht SS, Fow EA and Hoffmann D., (1981a). The influence of methyl substitution of the mutagenicity of nitronaphthalenes and nitrobiphenyls. *Mutat Res.*, 81:143-153.
- El-Bayoumy K, LaVoie EJ, Tulley-Freiler L and Hecht SS., (1981b). Effects of ortho-methyl substituents on the mutagenicity of aminobiphenyls and aminonaphthalenes. *Mutat Res.*, 90:345-354.
- El-Bayoumy K, Rivenson A, Johnson B, DiBello J, Little P, Hecht SS. (1988). Comparative tumorigenicity of 1-nitropyrene, 1-nitrosopyrene, and 1-aminopyrene administered by gavage to Sprague-Dawley rats. *Cancer Res.*, 48:4256-4260.
- El-Bayoumy K, Rivenson A, Upadhyaya P, Chae YH, Hecht SS. (1993) Induction of mammary cancer by 6-nitrochrysene in female CD rats. *Cancer Res.*, 53:3719-3722.
- Evans FE and Levine RA., (1986). Conformation and configuration at the central amine nitrogen of a nucleotide adduct of the carcinogen 2-(acetylaminofluorene as studied by ¹³C and ¹⁵N NMR spectroscopy. *J Biomol Struct Dyn.*, 3:923-934.
- Evans FE, Levine RA., (1988). NMR study of stacking interactions and conformational adjustments in the dinucleotide-carcinogen adduct 2'-deoxycytidylyl-(3→5)-2'-deoxy-8-(N-fluoren-2-ylac etamido) guanosine. *Biochemistry.*, 27:3046-3055.
- Falnes PO, Johansen RF, Seeberg E., (2002). AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. *Nature*, 419:178-182.
- Ferentz AE, Walker GC, Wagner G., (2001). Converting a DNA damage checkpoint effector (UmuD2C) into a lesion bypass polymerase (UmuD'2C). *EMBO J.*, 20:4287-4298.
- Fernandez De Henestrosa AR, Ogi T, Aoyagi S, Chafin D, Hayes JJ, Ohmori H, Woodgate R. (2000). Identification of additional genes belonging to the LexA regulon in *Escherichia coli*. *Mol Microbiol.*, 35:1560-1572.
- Firoozan M, Grant CM, Duarte JA, Tuite MF., (1991). Quantitation of readthrough of termination codons in yeast using a novel gene fusion assay. *Yeast.*, 7:173-183.
- Frank EG, Cheng N, Do CC, Cerritelli ME, Bruck I, Goodman MF, Egelman EH, Woodgate R, Steven AC., (2000). Visualization of two binding sites for the *Escherichia coli* UmuD'(2)C complex (DNA pol V) on RecA-ssDNA filaments. *J Mol Biol.*, 297:585-597.
- Frank EG, Hauser J, Levine AS, Woodgate R., (1993). Targeting of the UmuD, UmuD', and MucA' mutagenesis proteins to DNA by RecA protein. *Proc Natl Acad Sci U S A.*, 90:8169-8173.

- Frank, E.G., D.G. Ennis, M. Gonzalez, A.S. Levine, and R. Woodgate. (1996a). Regulation of SOS mutagenesis by proteolysis. *Proc. Natl. Acad. Sci.* 93: 10291-10296
- Frank, E.G., M. Gonzalez, D.G. Ennis, A.S. Levine, and R. Woodgate. (1996b). In vivo stability of the Umu mutagenesis proteins: A major role for RecA. *J. Bacteriol.* 178: 3550-3556
- Freitag N, McEntee K. (1989). "Activated"-RecA protein affinity chromatography of LexA repressor and other SOS-regulated proteins. *Proc. Natl. Acad. Sci. USA*, 86:8363-8367.
- Friedberg EC, Fischhaber PL, Kisker C., (2001a). Error-prone DNA polymerases: novel structures and the benefits of infidelity. *Cell.*, 107:9-12.
- Friedberg EC. (1996) Relationships between DNA repair and transcription. *Annu Rev Biochem.*, 65:15-42.
- Friedberg EC., (2001b). Why do cells have multiple error-prone DNA polymerases? *Environ Mol Mutagen.*, 38:105-110.
- Friedberg, E. C., G. C. Walker, and W. Siede, (1995). "DNA repair and mutagenesis" ASM Press, Washington, D.C.
- Fuchs R.P.P. and Daune M., (1972). Physical studies on deoxyribonucleic acid after covalent binding of a carcinogen. *Biochemistry*, 11:2659-2666.
- Fuchs RP, Koffel-Schwartz N, Pelet S, Janel-Bintz R, Napolitano R, Becherel OJ, Broschard TH, Burnouf DY, Wagner J., (2001). DNA polymerases II and V mediate respectively mutagenic (-2 frameshift) and error-free bypass of a single N-2-acetylaminofluorene adduct. *Biochem Soc Trans.*, 29:191-195.
- Fuchs RP, Napolitano RL., (1998). Inactivation of DNA proofreading obviates the need for SOS induction in frameshift mutagenesis. *Proc Natl Acad Sci U S A.*, 95:13114-13119.
- Fuchs RP, Schwartz N, Daune MP., (1981). Hot spots of frameshift mutations induced by the ultimate carcinogen N-acetoxy-N-2-acetylaminofluorene. *Nature.*, 294:657-659.
- Fuchs RP., (1984). DNA binding spectrum of the carcinogen N-acetoxy-N-2-acetylaminofluorene significantly differs from the mutation spectrum. *J Mol Biol.*, 177:173-180.
- Fulco AJ, (1991). P450BM-3 and other inducible bacterial P450 cytochromes: biochemistry and regulation. *Annu Rev Pharmacol Toxicol*, 31:177-203.
- Garcia A, Lambert IB and Fuchs RP., (1993). DNA adduct-induced stabilization of slipped frameshift intermediates within repetitive sequences: implications for mutagenesis. *Proc. Natl. Acad. Sci. (USA)*, 90:5989-5993.
- Gelardi A, Morini F, Dusatti F, Penco S, Ferro M. (2001), Induction by xenobiotics of phase I and phase II enzyme activities in the human keratinocyte cell line NCTC 2544. *Toxicol In Vitro.*, 15:701-11.
- Gerlach VL, Aravind L, Gotway G, Schultz RA, Koonin EV, Friedberg EC. 1999. Human and mouse homologs of Escherichia coli DinB (DNA polymerase IV), members of the UmuC/DinB superfamily. *Proc. Natl. Acad. Sci. USA*, 96:11922-11927.
- Gerlyng P., Grotmol T., P.O. Seglen. (1994). Effect of 4-acetylaminofluorene and other tumour promoters on hepatocellular growth and binucleation. *Carcinogenesis*, 15: 371-379.

- Gesteland RF, Atkins JF, (1996). Recoding: dynamic reprogramming of translation. *Annu Rev Biochem.*, 65:741-768.
- Glatt H., (1997). Sulfation and sulfotransferases 4: bioactivation of mutagens via sulfation. *FASEB J.*, 11:314-321.
- Goksoyr A, Husoy AM., (1998). Immunochemical approaches to studies of CYP1A localization and induction by xenobiotics in fish. *EXS.*, 86:165-202.
- Goldsmith M, Sarov-Blat L, Livneh Z., (2000). Plasmid-encoded MucB protein is a DNA polymerase (pol RI) specialized for lesion bypass in the presence of MucA', RecA, and SSB. *Proc Natl Acad Sci U S A.*, 97:11227-112231.
- Gonzalez M, Rasulova F, Maurizi MR, Woodgate R., (2000). Subunit-specific degradation of the UmuD/D' heterodimer by the ClpXP protease: the role of trans recognition in UmuD' stability. *EMBO J.*, 19:5251-5258.
- Gonzalez M, Woodgate R., (2002). The "tale" of UmuD and its role in SOS mutagenesis. *Bioessays.*, 24:141-148.
- Gonzalez MC, Marteau C, Franchi J, Migliore-Samour D. (2001), Cytochrome P450 4A11 expression in human keratinocytes: effects of ultraviolet irradiation. *Br J Dermatol.*, 145:749-757.
- Gonzalez, M., E.G. Frank, J.P. McDonald, A.S. Levine, and R. Woodgate. 1998. Structural insights into the regulation of SOS mutagenesis. *Acta Biochim. Polon.* 45: 163-172
- Goodman MF, Tippin B., (2000a). Sloppier copier DNA polymerases involved in genome repair. *Curr Opin Genet Dev.*, 10:162-168.
- Goodman MF, Tippin B., (2000b). The expanding polymerase universe. *Nat Rev Mol Cell Biol.*, 1:101-109.
- Goodman MF., (2002). Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu Rev Biochem.*, 71:17-50.
- Gradia, S., Subramanian, D., Wilson, T., Acharya, S., Makhov, A., Griffith, J., and Fishel, R. (1999) hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Mol. Cell*, 3: 255-261.
- Green MD, Tephly TR., (1998). Glucuronidation of amine substrates by purified and expressed UDP-glucuronosyltransferase proteins. *Drug Metab Dispos.*, 26:860-867.
- Grossman L, Thiagalingam S., (1993). Nucleotide excision repair, a tracking mechanism in search of damage. *J Biol Chem.*, 268:16871-16874.
- Grunberger D, Nelson JH, Cantor CR, Weinstein IB., (1970). Coding and conformational properties of oligonucleotides modified with the carcinogen N-2-acetylaminofluorene. *Proc. Natl. Acad. Sci. U S A.*, 66:488-494.
- Gupta RP, Abou-Donia MB., (1998). Cytochrome P450 enzymes in chickens: characteristics and induction by xenobiotics. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol.*, 121:73-83.
- Hagensee ME, Timme TL, Bryan SK, Moses RE., (1978). DNA polymerase III of *Escherichia coli* is required for UV and ethyl methanesulfonate mutagenesis. *Proc Natl Acad Sci U S A.*, 84:4195-41999.
- Hall MC, Matson SW, (1999). The *Escherichia coli* MutL protein physically interacts with MutH and stimulates the MutH-associated endonuclease activity. *J Biol Chem.*, 274:1306-1312.

- Hanawalt PC. (1966). The U.V. sensitivity of bacteria: its relation to the DNA replication cycle. *Photochem Photobiol.*, 5:1-12.
- Hauser J, Levine AS, Ennis DG, Chumakov KM, Woodgate R., (1992). The enhanced mutagenic potential of the MucAB proteins correlates with the highly efficient processing of the MucA protein. *J Bacteriol.*, 174:6844-6851.
- Havener JM, McElhinny SA, Bassett E, Gauger M, Ramsden DA, Chaney SG., (2003). Translesion synthesis past platinum DNA adducts by human DNA polymerase μ . *Biochemistry.*, 42:1777-1788.
- Heelis PF, Kim ST, Okamura T, Sancar A., (1993). The photo repair of pyrimidine dimers by DNA photolyase and model systems. *J Photochem Photobiol B.*, 17:219-228.
- Heflich RH, Neft RE., (1994). Genetic toxicity of 2-acetylaminofluorene, 2-aminofluorene and some of their metabolites and model metabolites. *Mutat Res.*, 318:73-114.
- Hein DW. (1988 a). Genetic polymorphism and cancer susceptibility: evidence concerning acetyltransferases and cancer of the urinary bladder. *Bioessays.*, 9:200-204.
- Hein DW. (1988 b). Acetylator genotype and arylamine-induced carcinogenesis. *Biochim Biophys Acta.*, 948:37-66.
- Higashitani N, Higashitani A, Horiuchi K., (1995). SOS induction in *Escherichia coli* by single-stranded DNA of mutant filamentous phage: monitoring by cleavage of LexA repressor. *J Bacteriol.*, 177:3610-3612.
- Higashitani N, Higashitani A, Horiuchi K., (1995). SOS induction in *Escherichia coli* by single-stranded DNA of mutant filamentous phage: monitoring by cleavage of LexA repressor. *J Bacteriol.*, 177:3610-3612.
- Higashitani N, Higashitani A, Roth A, Horiuchi K., (1992). SOS induction in *Escherichia coli* by infection with mutant filamentous phage that are defective in initiation of complementary-strand DNA synthesis. *J Bacteriol.*, 174:1612-1618.
- Higson FK., (1992 a) Microbial degradation of nitroaromatic compounds. *Adv Appl Microbiol.*, 37:1-19.
- Higson FK., (1992 b). Microbial degradation of biphenyl and its derivatives. *Adv Appl Microbiol.*, 37:135-164.
- Hingerty B.E. and Broyde S., (1986). Energy minimized structures of cacinogen-DNA adducts: 2-Acetylaminofluorene and 2-Aminofluorene. *Journal of biomolecular Structure and Dynamics* 4: 365-371.
- Hoeljmackers JH., (2001). Genome maintenance mechanisms for preventing cancer. *Nature.*, 411:366-374.
- Hoffmann GR, Fuchs RP, (1997). Mechanisms of frameshift mutations: insight from aromatic amines. *Chem Res Toxicol.*, 10:347-459.
- Holland IB, Jones C., (1985). The role of the FtsZ protein (SfiB) in UV-induced division inhibition and in the normal *Escherichia coli* cell division cycle. *Ann Inst Pasteur Microbiol.*, 136:165-171.
- Holmes J Jr, Clark S, Modrich P., (1990). Strand-specific mismatch correction in nuclear extracts of human and *Drosophila melanogaster* cell lines. *Proc Natl Acad Sci U S A.*, 87:5837-5841.

- Howard PC, McCoy EC, Rosenkranz HS., (1987). Sequential and differing nitroreductive pathways for mutagenic nitropyrenes in *Salmonella typhimurium*. *Mutagenesis.*, 2:431-432.
- Howard-Flanders P, Rupp WD, (1972). Recombinational repair in UV-irradiated *Escherichia coli*. *Johns Hopkins Med J.*, 1:212-25.
- Hsieh SE, Lo HH, Chung JG., (1998). The characteristics of arylamine N-Acetyltransferase in *Pseudomonas aeruginosa*. *Curr Microbiol.*, 36:353-360.
- Huisman, O. & D'Ari, R.,(1981). An inducible DNA replication-cell division coupling mechanism in *E. coli*. *Nature*, 290: 797-799.
- Ide F, Ishikawa T, Takayama S., (1981). Detection of chemical carcinogens by assay of unscheduled DNA synthesis in rat tracheal epithelium in short-term organ culture. *J. Cancer Res. Clin. Oncol.*, 102:115-26.
- Imadia K., Lee MS, Wang CY and King CM, (1991a). Carcinogenicity of dinitropyrenes in the weanling female CD rat. *Carcinogenesis*, 12:1187-1191.
- Imaida K, Hirose M, Tay L, Lee MS, Wang CY and King CM., (1991b). Comparative carcinogenicities of 1-, 2-, and 4-nitropyrene and structurally related compounds in the female CD rat. *Cancer Res.*, 51:2902-2907.
- Jackson AL, Loeb LA, (2001). The contribution of endogenous sources of DNA damage to the multiple mutations in cancer. *Mutat Res.*, 477:7-21.
- Jakoby WB, Sekura RD, Lyon ES, Marcus CJ and Wang J, (1980). In *Enzymatic basis of detoxication* (Jakoby WB, ed.) Vol.2, Toronto Academic Press.
- Jez JM, Flynn TG, Penning TM., (1997). A nomenclature system for the aldo-keto reductase superfamily. *Adv Exp Med Biol.*, 414:579-600.
- Johansson M.A., Jägestad M., (1994). Occurrence of mutagenic/carcinogenic heterocyclic amines in meat and fish products, including pan residues, prepared under domestic conditions. *Carcinogenesis*, 15: 1511-1518.
- Johnson RE, Kondratyck CM, Prakash S, Prakash L., (1999b). hRAD30 mutations in the variant form of xeroderma pigmentosum. *Science.*, 285:263-265.
- Johnson RE, Prakash S, Prakash L.,(1999a). Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Poleta. *Science.*, 283:1001-1004.
- Johnson RE, Washington MT, Haracska L, Prakash S, Prakash L., (2000a). Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature.*, 406:1015-1019.
- Johnson RE, Washington MT, Prakash S, Prakash L., (2000b). Fidelity of human DNA polymerase eta. *J Biol Chem.*, 275:7447-7450.
- Jonczyk, P., and A. Nowicka., (1996). Specific in vivo protein-protein interactions between *Escherichia coli* SOS mutagenesis proteins. *J. Bacteriol.* 178:2580-2585.
- Junop MS, Obmolova G, Rausch K, Hsieh P, Yang W. (2001). Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair. *Mol Cell.*, 7:1-12
- Kadlubar FF, Miller JA, Miller EC., (1977). Hepatic microsomal N-glucuronidation and nucleic acid binding of N-hydroxy arylamines in relation to urinary bladder carcinogenesis. *Cancer Res.*, 37:805-814.

- Kadlubar FF., (1994). Biochemical individuality and its implications for drug and carcinogen metabolism: recent insights from acetyltransferase and cytochrome P4501A2 phenotyping and genotyping in humans. *Drug Metab Rev.*, 26:37-46.
- Kalnik MW, Norman DG, Zagorski MG, Swann PF, Patel DJ. (1989). Conformational transitions in cytidine bulge-containing deoxytridecanucleotide duplexes: extra cytidine equilibrates between looped out (low temperature) and stacked (elevated temperature) conformations in solution. *Biochemistry.*, 28:294-303.
- Kannouche P, Fernandez de Henestrosa AR, Coull B, Vidal AE, Gray C, Zicha D, Woodgate R, Lehmann AR., (2003). Localization of DNA polymerases η and ι to the replication machinery is tightly co-ordinated in human cells. *EMBO J.*, 22:1223-1233.
- Karow ML, Rogers EJ, Lovett PS, Piggot PJ., (1998). Suppression of TGA mutations in the *Bacillus subtilis* spoIIR gene by prfB mutations. *J Bacteriol.*, 180:4166-4170.
- Karpinsky GE, McCoy EC, Rosenkranz HS, Mermelstein R. (1982). The chemical activation of non-mutagenic nitrated polycyclic aromatic hydrocarbons to mutagens. *Mut. Res.*, 92:29-37.
- Kato R, Yamazoe Y., (1994). Metabolic activation of N-hydroxylated metabolites of carcinogenic and mutagenic arylamines and arlamides by esterification. *Drug Metab Rev.*, 26:413-429.
- Kato R., (1986). Metabolic activation of mutagenic heterocyclic aromatic amines from protein pyrolysates. *Crit Rev Toxicol.*, 16:307-348.
- Kato T, Shinoura Y., (1977). Isolation and characterization of mutants of *Escherichia coli* deficient in induction of mutations by ultraviolet light. *Mol Gen Genet.*, 156:121-131.
- Kedderis GL, Miwa GT. The metabolic activation of nitroheterocyclic therapeutic agents. *Drug Metab Rev.*, 19:33-62.
- Keen NT, Tamaki S, Kobayashi D, Trollinger D., (1988). Improved broad-host-range plasmids for DNA cloning in gram-negative bacteria. *Gene.*, 70:191-197.
- Kenyon CJ, Walker GC., (1980). DNA-damaging agents stimulate gene expression at specific loci in *Escherichia coli*. *Proc Natl Acad Sci U S A.*, 77:2819-2823.
- Khidhir MA, Casaregola S, Holland IB, (1985). Mechanism of transient inhibition of DNA synthesis in ultraviolet-irradiated *E. coli*: inhibition is independent of recA whilst recovery requires RecA protein itself and an additional, inducible SOS function. *Mol Gen Genet.*, 199:133-140.
- Kim SR, Maenhaut-Michel G, Yamada M, Yamamoto Y, Matsui K, Sofuni T, Nohmi T, Ohmori H., (1997). Multiple pathways for SOS-induced mutagenesis in *Escherichia coli*: an overexpression of dinB/dinP results in strongly enhancing mutagenesis in the absence of any exogenous treatment to damage DNA. *Proc Natl Acad Sci USA.*, 94:13792-13797.
- Kim SR, Matsui K, Yamada M, Gruz P, Nohmi T. (2001). Roles of chromosomal and episomal dinB genes encoding DNA pol IV in targeted and untargeted mutagenesis in *Escherichia coli*. *Mol Genet Genomics.*, 266:207-215.
- King CM, Phillips B., (1968). Enzyme-catalyzed reactions of the carcinogen N-hydroxy-2-fluorenylacetamide with nucleic acid. *Science.*, 159:1351-1353.

- King CM, Traub NR, Lortz ZM, Thissen MR., (1979). Metabolic activation of arylhydroxamic acids by N-O-acyltransferase of rat mammary gland. *Cancer Res.*, 39:3369-3372.
- Kleman M, Overvik E, Mason G, Gustafsson JA.(1990). Effects of the food mutagens MeIQx and PhIP on the expression of cytochrome P450IA proteins in various tissues of male and female rats. *Carcinogenesis.*, 11:2185-2189.
- Koehl P., Burnouf D. and Fuchs R.P.P. (1989a). Construction of plasmids containing a unique acetylaminofluorene adduct located within a mutation hot spot. *J.Mol.Biol.*, 207: 355-364.
- Koehl P., Valladier P., Lefevre JF. and Fuchs RPP, (1989b). Strong structural effect of the position of a single acetylaminofluorene adduct within a mutation hot spot. *Nucleic Acids Res.* 17: 9531-9541.
- Koffel-Schwartz N, Fuchs RPP, (1989). Genetic control of AAF-induced mutagenesis at alternating GC sequences: an additional role for RecA. *Mol Gen Genet*, 215:306-311.
- Koffel-Schwartz N, Verdier JM, Bichara M, Freund AM, Daune MP, Fuchs RP, (1984). Carcinogen-induced mutation spectrum in wild-type, *uvrA* and *umuC* strains of *Escherichia coli*. Strain specificity and mutation-prone sequences. *J Mol Biol*, 177:33-51.
- Kogoma T., (1997). Stable DNA replication: interplay between DNA replication, homologous recombination, and transcription. *Microbiol Mol Biol Rev.*, 61:212-238.
- Koh WS, Yang KH, Jeong TC, Delany B and Kaminski NE., (1995). 2-Acetylaminofluorene inhibits the activation of immune responses by blocking cell cycle progression at G1 phase. *Arch. Toxicol.*, 69:350-356.
- Kokoska RJ, Bebenek K, Boudsocq F, Woodgate R, Kunkel TA., (2002). Low fidelity DNA synthesis by a γ family DNA polymerase due to misalignment in the active site. *J Biol Chem.*, 277:19633-19638.
- Kowalczykowski SC. (1991). Biochemistry of genetic recombination: energetics and mechanism of DNA strand exchange. *Annu Rev Biophys Biophys Chem.*, 20:539-575.
- Kowalczykowski SC., (2000). Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem Sci.*, 25:156-165.
- Kriek E (1974).Carcinogenesis by aromatic amines. *Biochim Biophys Acta.* 355:177-203.
- Kulaeva OI, Koonin EV, McDonald JP, Randall SK, Rabinovich N, Connaughton JF, Levine AS, Woodgate R., (1996). Identification of a DinB/UmuC homolog in the archeon *Sulfolobus solfataricus*.*Mutat Res.*, 357:245-253.
- Kunz BA, Straffon AF, Vonarx EJ., (2000). DNA damage-induced mutation: tolerance via translesion synthesis. *Mutat Res.*, 451:169-185.
- Kuzminov A., (1999). Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. *Microbiol Mol Biol Rev.* 63:751-813.
- Lambert IB, Carroll C, Laycock N, Duval L, Whiteway J, Lawford I, Turner G, Booth R, Douville S, Nokhbeh MR. (1998), The mutational specificity of 1-nitroso-6-nitropyrene in the *lacI* gene of *Escherichia coli* strains deficient in nucleotide excision repair. *Mutagenesis*, 13:9-18.

- Lambert IB, Carroll C, Laycock N, Koziarz J, Lawford I, Duval L, Turner G, Booth R, Douville S, Whiteway J, Nokhbeh MR. (2001), Cellular determinants of the mutational specificity of 1-nitroso-6-nitropyrene and 1-nitroso-8-nitropyrene in the lacI gene of Escherichia coli. *Mutat Res.*, 484:19-48.
- Lambert IB, Gordon AJ, Bryant DW, Glickman BW, McCalla DR. (1991) The action of 1-nitroso-8-nitropyrene in Escherichia coli: DNA adduct formation and mutational consequences in the absence of nucleotide excision-repair. *Carcinogenesis.*, 12:879-884.
- Lambert IB, Gordon AJ, Glickman BW, McCalla DR., (1992b). The influence of local DNA sequence and DNA repair background on the mutational specificity of 1-nitroso-8-nitropyrene in Escherichia coli: inferences for mutagenic mechanisms. *Genetics.*, 132:911-927.
- Lambert IB, Napolitano RL and Fuchs RPP, (1992a). Carcinogen-induced frameshift mutagenesis in repetitive sequences. *Proc.Natl.Acad.Sci.USA*, 89:1310-1314.
- Lamers MH, Perrakis A, Enzlin JH, Winterwerp HH, de Wind N, Sixma TK. (2000). The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature*, 407:711-717.
- Landi MT, Sinha R, Lang NP, Kadlubar FF. (1999). Chapter 16. Human cytochrome P4501A2. *IARC Sci Publ.*, 148:173-195.
- Lawrence CW, Banerjee SK, Borden A, LeClerc JE., (1990). T-T cyclobutane dimers are misinstructive, rather than non-instructive, mutagenic lesions. *Mol Gen Genet.*, 222:166-168.
- Lawrence CW, Borden A, Banerjee SK, LeClerc JE., (1990). Mutation frequency and spectrum resulting from a single abasic site in a single-stranded vector. *Nucleic Acids Res.*, 18:2153-2157.
- Le Chatelier E, Becherel OJ, d'Alencon E, Canceill D, Ehrlich SD, Fuchs RP, Janniere L., (2004) Involvement of DnaE, the second replicative DNA polymerase from *Bacillus subtilis*, in DNA mutagenesis. *J Biol Chem.*, 279:1757-1767.
- LeClerc JE, Borden A, Lawrence CW., (1991). The thymine-thymine pyrimidine-pyrimidone(6-4) ultraviolet light photoproduct is highly mutagenic and specifically induces 3' thymine-to-cytosine transitions in Escherichia coli. *Proc Natl Acad Sci U S A.*, 88:9685-9689.
- Lee M, and Yang KH, (2000). 2-Acetylaminofluorene suppresses immune response through the inhibition of nuclear factor-kappaB activation during the early stage of B cell development. *Toxicol Lett.*, 114:173-180.
- Lenne-Samuel N, Wagner J, Etienne H, Fuchs RP., (2002). The processivity factor beta controls DNA polymerase IV traffic during spontaneous mutagenesis and translesion synthesis in vivo. *EMBO Rep.*, 3:45-49.
- Levy GN, Chung JG, Weber WW, (1994). 2-Aminofluorene metabolism and DNA adduct formation by mononuclear leukocytes from rapid and slow acetylator mouse strains. *Carcinogenesis.*, 15:353-357.
- Lewis AJ, Otake Y, Walle UK, Walle T., (2000). Sulphonation of N-hydroxy-2-acetylaminofluorene by human dehydroepiandrosterone sulphotransferase. *Xenobiotica.*, 30:253-261.

- Lewis LK, Harlow GR, Gregg-Jolly LA, Mount DW., (1994). Identification of high affinity binding sites for LexA which define new DNA damage-inducible genes in *Escherichia coli*. *J Mol Biol.*, 241:507-523.
- Li Y, Baldauf S, Lim EK, Bowles DJ., (2001). Phylogenetic analysis of the UDP-glycosyltransferase multigene family of *Arabidopsis thaliana*. *J Biol Chem.*, 276:4338-4343.
- Lindahl T, Karran P, Wood RD. (1997). DNA excision repair pathways. *Curr Opin Genet Dev.*, 7:158-169.
- Lindahl T, Sedgwick B, Sekiguchi M, Nakabeppu Y. (1988). Regulation and expression of the adaptive response to alkylating agents. *Annu Rev Biochem.*, 57:133-57.
- Lindeman B, Skarpen E, Oksvold MP and Huitfeldt HS, (2000) The carcinogen 2-acetylaminofluorene inhibits activation and nuclear accumulation of cyclin-dependent kinase 2 in growth-induced rat liver. *Mol Carcinogenesis.*, 27:190-199.
- Lindsley JE, Fuchs RP., (1994). Use of single-turnover kinetics to study bulky adduct bypass by T7 DNA polymerase. *Biochemistry.*, 33:764-772.
- Ling H, Boudsocq F, Plosky BS, Woodgate R, Yang W., (2004). Replication of a cis-syn thymine dimer at atomic resolution. *Nature.* 424:1083-1087.
- Ling H, Boudsocq F, Woodgate R, Yang W., (2001). Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. *Cell.*, 107:91-102.
- Little JW., (1984). Autodigestion of *lexA* and phage lambda repressors. *Proc Natl Acad Sci USA.*, 81(5):1375-1379.
- Little JW., (1991). Mechanism of specific LexA cleavage: autodigestion and the role of RecA coprotease. *Biochimie.*, 73:411-421.
- Little JW., (1993). LexA cleavage and other self-processing reactions. *J. Bacteriol.*, 175:4943-4950.
- Loechler EL, (1996). The role of adduct site-specific mutagenesis in understanding how carcinogen-DNA adducts cause mutations: perspective, prospects and problems. *Carcinogenesis*, 17:895-902.
- Löfroth G., Hefner E., Alfheim I. and Møller M., (1980). Mutagenic activity in photocopies. *Science.* 209:1037-1039.
- Longo V, Amato G, Salvetti A, Gervasi PG., (2000). Heterogenous effects of anthraquinones on drug-metabolizing enzymes in the liver and small intestine of rat. *Chem Biol Interact.*, 126:63-77.
- Lopez de Saro FJ, O'Donnell M., (2001). Interaction of the beta sliding clamp with MutS, ligase, and DNA polymerase I. *Proc Natl Acad Sci U S A.*, 98:8376-8380.
- Lu FM, Chak KF., (1996). Two overlapping SOS-boxes in *ColE* operons are responsible for the viability of cells harboring the *Col* plasmid. *Mol Gen Genet.*, 251:407-411.
- Luceri F, Pieraccini G, Moneti G. and Dolara P., (1993). Primary aromatic amines from side-stream cigarette smoke are common contaminants of indoor air. *Toxicol Ind Health.*, 9:405-413.
- Luo Y, Pfuetzner RA, Mosimann S, Paetzel M, Frey EA, Cherney M, Kim B, Little JW, Strynadka NC., (2001). Crystal structure of LexA: a conformational switch for regulation of self-cleavage. *Cell.*, 106:585-594.

- Macintyre G, Atwood CV, Cupples CG, (2001). Lowering S-adenosylmethionine levels in *Escherichia coli* modulates C-to-T transition mutations. *J Bacteriol.*, 183:921-927.
- Mackenzie PI, Owens IS, Burchell B, Bock KW, Bairoch A, Belanger A, Fournel-Gigleux S, Green M, Hum DW, Iyanagi T, Lancet D, Louisot P, Magdalou J, Chowdhury JR, Ritter JK, Schachter H, Tephly TR, Tipton KF, Nebert DW., (1997). The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics.*, 7:255-269.
- Maenhaut-Michel G, Janel-Bintz R and Fuchs RP., (1992). A umuDC-independent SOS pathway for frameshift mutagenesis. *Mol Gen Genet.*, 235:373-380.
- Maenhaut-Michel G, Janel-Bintz R and Fuchs RP., (1992). A umuDC-independent SOS pathway for frameshift mutagenesis. *Mol Gen Genet.*, 235:373-380.
- Malejka-Giganti D, Niehans GA, Reichert MA, Bennett KK and Bliss RL. (1999). Potent carcinogenicity of 2,7-dinitrofluorene, an environmental pollutant, for the mammary gland of female Sprague-Dawley rats. *Carcinogenesis.* 20:2017-2023.
- Margison GP, Santibanez-Koref MF., (2002). O6-alkylguanine-DNA alkyltransferase: role in carcinogenesis and chemotherapy. *Bioessays.*, 24:255-266.
- Marians KJ. (2000). PriA: at the crossroads of DNA replication and recombination. *Prog Nucleic Acid Res Mol Biol.*, 63:39-67.
- Marians KJ., (2000). PriA-directed replication fork restart in *Escherichia coli*. *Trends Biochem Sci.*, 25:185-189.
- Marnett LJ, (2000). Oxyradicals and DNA damage., *Carcinogenesis.*, 21(3):361-370.
- Marsh, L., and G. C. Walker. (1985). Cold sensitivity induced by overproduction of UmuDC in *Escherichia coli*. *J. Bacteriol.* 162:155-161.
- Marti TM, Kunz C, Fleck O. (2002). DNA mismatch repair and mutation avoidance pathways. *J Cell Physiol.*, 191:28-41.
- Masutani C, Kusumoto R, Iwai S, Hanaoka F., (2000). Mechanisms of accurate translesion synthesis by human DNA polymerase η . *EMBO J.*, 19:3100-3109.
- Masutani C, Kusumoto R, Yamada A, Dohmae N, Yokoi M, Yuasa M, Araki M, Iwai S, Takio K, Hanaoka F, (1999). The XPV (Xeroderma pigmentosum variant) gene encodes human DNA polymerase η . *Nature* 399:700-704.
- Mathews D.H., Sabina J., Zuker M., and D. H. Turner, (1999). Expanded Sequence Dependence of Thermodynamic Parameters Improves Prediction of RNA Secondary Structure. *Journal of Molecular Biology*, 288:911-940.
- Matsushita H. and Iida Y. (1986). Application of capillary gas chromatography to environmental analysis. *J. High Resolut. Chromatogr. Chromatogr. Commun.*, 9:708-711.
- McClary JA, Witney F, Geisselsoder J., (1989) Efficient site-directed in vitro mutagenesis using phagemid vectors. *Biotechniques.*, 7:282-289.
- McCoy EC, Rosenkranz EJ, Rosenkranz HS, Mermelstein R. (1981). Nitrated fluorene derivatives are potent frameshift mutagens. *Mutat. Res.* 90:11-20.
- McCoy EC, Rosenkranz HS, Mermelstein R., (1981). Evidence for the existence of a family of bacterial nitroreductases capable of activating nitrated polycyclics to mutagens., *Environ Mutagen.*, 3:421-427.

- McDonald JP, Levine AS, Woodgate R., (1997). The *Saccharomyces cerevisiae* RAD30 gene, a homologue of *Escherichia coli* dinB and umuC, is DNA damage inducible and functions in a novel error-free postreplication repair mechanism. *Genetics.*, 147:1557-1568.
- McDonald JP, Raptic-Otrin V, Epstein JA, Broughton BC, Wang X, Lehmann AR, Wolgemuth DJ, Woodgate R., (1999). Novel human and mouse homologs of *Saccharomyces cerevisiae* DNA polymerase eta. *Genomics.*, 60:20-30.
- McDonald JP, Tissier A, Frank EG, Iwai S, Hanaoka F, Woodgate R. , (2001). DNA polymerase iota and related rad30-like enzymes. *Philos Trans R Soc Lond B Biol Sci.*, 356:53-60.
- Mechanic LE, Frankel BA, Matson SW, (2000). *Escherichia coli* MutL loads DNA helicase II onto DNA. *J Biol Chem.*, 275:38337-38346.
- Melchior WB Jr, Marques MM, Beland FA, (1994). Mutations induced by aromatic amine DNA adducts in pBR322. *Carcinogenesis*, 15:889-899.
- Melchior WB Jr, Marques MM, Beland FA., (1994). Mutations induced by aromatic amine DNA adducts in pBR322. *Carcinogenesis.*, 15:889-899.
- Menzel R, Bogaert T, Achazi R., (2001), A Systematic Gene Expression Screen of *Caenorhabditis elegans* Cytochrome P450 Genes Reveals CYP35 as Strongly Xenobiotic Inducible. *Arch Biochem Biophys.*, 395:158-68.
- Messier F, Lu C, Andrews P, McCarry BE, Quilliam MA, McCalla DR., (1981). Metabolism of 1-nitropyrene and formation of DNA adducts in *Salmonella typhimurium*. *Carcinogenesis.*, 2:1007-1011.
- Meunier V, Bourrie M, Julian B, Marti E, Guillou F, Berger Y, Fabre G. Expression and induction of CYP1A1/1A2, CYP2A6 and CYP3A4 in primary cultures of human hepatocytes: a 10-year follow-up. *Xenobiotica.*, 30:589-607.
- Milhe C, Fuchs RP, Lefevre JF., (1996). NMR data show that the carcinogen N-2-acetylaminofluorene stabilises an intermediate of -2 frameshift mutagenesis in a region of high mutation frequency. *Eur J Biochem.*, 235:120-127.
- Miller E.C., Miller J.A. and Enemoto M., (1964). The comparative carcinogenesis of 2-acetylaminofluorene and its N-hydroxy metabolite in mice, hamsters and guinea pigs. *Cancer Res.* 24:2018-2032.
- Miller E.C., Miller J.A. and Hartman H.A., (1961). N-hydroxy-2-acetylaminofluorene: a metabolite of 2-acetylaminofluorene with increased carcinogenic activity in the rat. *Cancer Res.* 21:815-824.
- Miller EC, Miller JA, Boberg EW, Delclos KB, Lai CC, Fennell TR, Wiseman RW, Liem A.,(1985). Sulfuric acid esters as ultimate electrophilic and carcinogenic metabolites of some alkenylbenzenes and aromatic amines in mouse liver. *Carcinog Compr Surv.*, 10:93-107.
- Miller EC, Miller JA., (1981). Searches for ultimate chemical carcinogens and their reactions with cellular macromolecules. *Cancer.*, 47:2327-2345.
- Miller EC., (1978). Some current perspectives on chemical carcinogenesis in humans and experimental animals: Presidential Address. *Cancer Res.*, 38:1479-1496.
- Miller JA and Miller EC. (1983). Some historical aspects of N-aryl carcinogens and their metabolic activation. *Environ Health Perspect* 49:3-12.
- Miller JA, Miller EC. (1969). The metabolic activation of carcinogenic aromatic amines and amides. *Prog Exp Tumor Res.*, 11:273-301.

- Miller JA, Sandine RB, Miller EC and Rusch HP, (1955). The carcinogenicity of compounds related to 2-acetylaminofluorene II. Variations in the bridges and the 2-substituent. *Cancer Res.*, 22:1002-1014.
- Modrich P, Lahue R. (1996). Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu Rev Biochem.*, 65:101-33.
- Modrich P. (1989). Methyl-directed DNA mismatch correction. *J Biol Chem.*, 264:6597-6600.
- Modrich P. (1991). Mechanisms and biological effects of mismatch repair. *Annu Rev Genet.*, 25:229-53.
- Moller L, Lax I. and Eriksson LC, (1993). Nitrated polycyclic aromatic hydrocarbons: a risk assessment for the urban citizen. *Environ Health Perspect.* 1993 Oct;101 Suppl 3:309-315.
- Moore P, Strauss BS., (1979). Sites of inhibition of in vitro DNA synthesis in carcinogen- and UV-treated phi X174 DNA. *Nature.*, 278:664-666.
- Moore PD, Bose KK, Rabkin SD, Strauss BS., (1981). Sites of termination of in vitro DNA synthesis on ultraviolet- and N-acetylaminofluorene-treated phi X174 templates by prokaryotic and eukaryotic DNA polymerases. *Proc Natl Acad Sci U S A.*, 78:110-114.
- Morel Y, Mermoud N, Barouki R., (1999). An autoregulatory loop controlling CYP1A1 gene expression: role of H(2)O(2) and NFI. *Mol Cell Biol.*, 19:6825-6832.
- Mount DW, Low KB, Edmiston SJ., (1972). Dominant mutations (lex) in *Escherichia coli* K-12 which affect radiation sensitivity and frequency of ultraviolet light-induced mutations. *J Bacteriol.*, 112:886-893.
- Mudgett JS, Buckholt M, Taylor WD., (1991). Ultraviolet light-induced plasmid-chromosome recombination in *Escherichia coli*: the role of recB and recF. *Gene*, 97:131-136.
- Mukherjee A, Cao C, Lutkenhaus J., (1998). Inhibition of FtsZ polymerization by SulA, an inhibitor of septation in *Escherichia coli*. *Proc Natl Acad Sci U S A.*, 95:2885-2890.
- Mulder GJ, (1981). In sulfate metabolism and sulfate conjugation; Proceedings of an international workshop held at Noordwijkehout, the Netherlands, London, Eng. Taylor Press.
- Murli S, Opperman T, Smith BT, Walker GC., (2000). A role for the umuDC gene products of *Escherichia coli* in increasing resistance to DNA damage in stationary phase by inhibiting the transition to exponential growth. *J Bacteriol.*, 182:1127-1135.
- Murli S, Walker GC. (1993). SOS mutagenesis. *Curr Opin Genet Dev.*, 3:719-725.
- Napolitano R, Janel-Bintz R, Wagner J, Fuchs RP. (2000). All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis. *EMBO J.*, 19:6259-6265.
- Napolitano RL, Fuchs RP., (1997). New strategy for the construction of single-stranded plasmids with single mutagenic lesions. *Chem Res Toxicol.*, 10(6):667-671.
- Napolitano RL, Lambert IB and Fuchs RP., (1994). DNA sequence determinants of carcinogen-induced frameshift mutagenesis. *Biochemistry*, 33:1311-1315.
- Napolitano RL, Lambert IB and Fuchs RP., (1997). SOS factors involved in translesion synthesis. *Proc Natl Acad Sci U S A*, 94:5733-5738.

- Narasimhan V, Bryan AM, (1976). Temperature-induced perturbations in the circular dichroic spectrum of the synthetic polymer poly[d(G-C)]. *Biochim. Biophys. Acta.*, 435:433-437.
- Nelson DR, Kamataki T, Waxman DJ, Guengerich FP, Estabrook RW, Feyereisen R, Gonzalez FJ, Coon MJ, Gunsalus IC, Gotoh O, (1993). The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. *DNA Cell Biol.*, 12:1-51.
- Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR, Gotoh O, Coon MJ, Estabrook RW, Gunsalus IC, Nebert DW., (1996). P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics.*, 6:1-42.
- Nelson JR, Lawrence CW, Hinkle DC., (1996a). Deoxycytidyl transferase activity of yeast REV1 protein. *Nature.*, 382:729-731.
- Nelson JR, Lawrence CW, Hinkle DC., (1996b) Thymine-thymine dimer bypass by yeast DNA polymerase zeta. *Science.*, 272:1646-1649.
- Nilsson M, Ryden-Aulin M., (2003). Glutamine is incorporated at the nonsense codons UAG and UAA in a suppressor-free *Escherichia coli* strain. *Biochim Biophys Acta.*, 1627:1-6.
- Nohmi T, Battista JR, Dodson LA, Walker GC., (1988). RecA-mediated cleavage activates UmuD for mutagenesis: mechanistic relationship between transcriptional derepression and posttranslational activation. *Proc Natl Acad Sci U S A.*, 85:1816-1820.
- Norman D, Abuaf P, Hingerty BE, Live D, Grunberger D, Broyde S, Patel DJ., (1989). NMR and computational characterization of the N-(deoxyguanosin-8-yl)aminofluorene adduct [(AF)G] opposite adenosine in DNA: (AF)G[syn].A[anti] pair formation and its pH dependence. *Biochemistry.*, 28:7462-7476.
- Nussbaum M, Fiala ES, Kulkarni B, El-Bayoumy K and Weisburger JH., (1983). In vivo metabolism of 3,2'-dimethyl-4-aminobiphenyl (DMAB) bearing on its organotropism in the Syrian golden hamster and the F344 rat. *Environ Health Perspect.*, 49:223-231.
- O'Handley SF, Sanford D.G., Xu R., Lester C.C., Hingerty B.E., Broyde S. and Krugh T.R., (1992) Structure of an acetylaminofluorene (AAF) modified DNA oligomer, Structure and Function, volume I: Nucleic Acids, pp 137-145, Adenine Press, NY.
- O'Handley SF, Sanford DG, Xu R, Lester CC, Hingerty BE, Broyde S, Krugh TR, (1993). Structural characterization of an N-acetyl-2-aminofluorene (AAF) modified DNA oligomer by NMR, energy minimization, and molecular dynamics. *Biochemistry*, 32:2481-2497.
- Obmolova G, Ban C, Hsieh P, Yang W., (2000). Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature.*, 407:703-710.
- Ogi T, Kato T Jr, Kato T, Ohmori H. 1999. Mutation enhancement by DINB1, a mammalian homologue of the *Escherichia coli* mutagenesis protein dinB. *Genes Cells* 4:607-618.
- Ohashi E, Bebenek K, Matsuda T, Feaver WJ, Gerlach VL, Friedberg EC, Ohmori H, Kunkel TA., (2000a). Fidelity and processivity of DNA synthesis by DNA

- polymerase kappa, the product of the human DINB1 gene. *J Biol Chem.*, 275:39678-39684.
- Ohashi E, Ogi T, Kusumoto R, Iwai S, Masutani C, Hanaoka F, Ohmori H., (2000b). Error-prone bypass of certain DNA lesions by the human DNA polymerase kappa. *Genes Dev.*, 14:1589-1594.
- Ohmori H, Friedberg EC, Fuchs RP, Goodman MF, Hanaoka F, Hinkle D, Kunkel TA, Lawrence CW, Livneh Z, Nohmi T, Prakash L, Prakash S, Todo T, Walker GC, Wang Z, Woodgate R., (2001). The Y-family of DNA polymerases. *Mol Cell.*, 8:7-8.
- Ohmori H, Hatada E, Qiao Y, Tsuji M, Fukuda R., (1995). dinP, a new gene in *Escherichia coli*, whose product shows similarities to UmuC and its homologues. *Mutat Res.*, 347:1-7.
- Okinaka R.T., Nicolas J.W., Whaley T.W., Strinse G.F., 1984. Phototransformation of 2-aminofluorene into N-oxidized mutagens. *Carcinogenesis*, 5:1741-1743.
- Opperman T, Murli S, Smith BT, Walker GC., (1999). A model for a umuDC-dependent prokaryotic DNA damage checkpoint. *Proc Natl Acad Sci U S A.*, 96:9218-9223.
- Paquette SM, Bak S, Feyereisen R., (2000). Intron-exon organization and phylogeny in a large superfamily, the paralogous cytochrome P450 genes of *Arabidopsis thaliana*. *DNA Cell Biol.*, 19:307-317.
- Parker SL, Tong T, Bolden S, Wingo PA. (1996). Cancer statistics, 1996. *CA Cancer J Clin.*, 46:5-27.
- Patrianakos C. and Hoffmann D., (1979). Chemical studies on tobacco smoke. LXIV. On the analysis of aromatic amines in cigarette smoke. *J. Anal.Toxicol.*, 3:150-154.
- Pegg AE, Byers TL. (1992). Repair of DNA containing O6-alkylguanine. *FASEB J.*, 6:2302-2310.
- Pegg AE., (2000). Repair of O(6)-alkylguanine by alkyltransferases. *Mutat Res.*, 462:83-100.
- Perrino, F. W., S. Harvey, and S. M. McNeill. (1999). Two functional domains of the epsilon subunit of DNA polymerase III. *Biochemistry* 38:16001-16009.
- Petit MA, Bedale W, Osipiuk J, Lu C, Rajagopalan M, McInerney P, Goodman MF, Echols H., (1994) Sequential folding of UmuC by the Hsp70 and Hsp60 chaperone complexes of *Escherichia coli*. *J Biol Chem.*, 269:23824-23829.
- Pham P, Bertram JG, O'Donnell M, Woodgate R, Goodman MF., (2001a). A model for SOS-lesion-targeted mutations in *Escherichia coli*. *Nature.*, 409:366-370.
- Pham P, Rangarajan S, Woodgate R, Goodman MF. (2001b). Roles of DNA polymerases V and II in SOS-induced error-prone and error-free repair in *Escherichia coli*. *Proc Natl Acad Sci U S A.*, 98:8350-8354.
- Pohl FM, Jovin TM, (1972). Salt-induced co-operative conformational change of a synthetic DNA: equilibrium and kinetic studies with poly (dG-dC). *J Mol. Biol.*, 67:375-396.
- Prakash S, Prakash L. (2000). Nucleotide excision repair in yeast. *Mutat Res.*, 451:13-24.
- Qiu Z, Goodman MF. (1997). The *Escherichia coli* polB locus is identical to dinA, the structural gene for DNA polymerase II. Characterization of Pol II purified from a polB mutant. *J Biol Chem.*, 272:8611-8617.

- Radman M. (1975). SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis. *Basic Life Sci.*, 5:355-367.
- Rajagopalan M, Lu C, Woodgate R, O'Donnell M, Goodman MF, Echols H., (1992). Activity of the purified mutagenesis proteins UmuC, UmuD', and RecA in replicative bypass of an abasic DNA lesion by DNA polymerase III. *Proc Natl Acad Sci U S A.*, 89:10777-10781.
- Rajagopalan M, Lu C, Woodgate R, O'Donnell M, Goodman MF, Echols H., (1992). Activity of the purified mutagenesis proteins UmuC, UmuD', and RecA in replicative bypass of an abasic DNA lesion by DNA polymerase III. *Proc Natl Acad Sci U S A.*, 89:10777-107781.
- Rangarajan S, Woodgate R, Goodman MF, (2002) Replication restart in UV-irradiated *Escherichia coli* involving pols II, III, V, PriA, RecA and RecFOR proteins. *Mol Microbiol.*, 43:617-628.
- Rangarajan S, Woodgate R, Goodman MF., (1991). A phenotype for enigmatic DNA polymerase II: a pivotal role for pol II in replication restart in UV-irradiated *Escherichia coli*. *Proc Natl Acad Sci U S A.*, 96:9224-9229.
- Rattray AJ, Strathern JN., (2003). Error-prone DNA polymerases: when making a mistake is the only way to get ahead. *Annu Rev Genet.*, 37:31-66.
- Rechkoblit O, Zhang Y, Guo D, Wang Z, Amin S, Krzeminsky J, Louneva N, Geacintov NE., (2002). trans-Lesion synthesis past bulky benzo[a]pyrene diol epoxide N2-dG and N6-dA lesions catalyzed by DNA bypass polymerases. *J Biol Chem.*, 277:30488-30494.
- Reuven NB, Arad G, Maor-Shoshani A, Livneh Z., (1999). The mutagenesis protein UmuC is a DNA polymerase activated by UmuD', RecA, and SSB and is specialized for translesion replication. *J Biol Chem.*, 274:31763-31766.
- Reuven NB, Tomer G, Livneh Z., (1998). The mutagenesis proteins UmuD' and UmuC prevent lethal frameshifts while increasing base substitution mutations. *Mol. Cell* 2:191-199.
- Robison K, McGuire AM, Church GM., (1998). A comprehensive library of DNA-binding site matrices for 55 proteins applied to the complete *Escherichia coli* K-12 genome. *J Mol Biol.*, 284:241-254.
- Roland K.L., Smith M.H., Rupley J.A. and Little J.W., (1992). In vitro analysis of mutant LexA proteins with an increased rate of specific cleavage. *J. Mol. Biol.*, 228:395-408.
- Rosenkranz EJ, McCoy EC, Mermelstein R, Rosenkranz HS., (1982). Evidence for the existence of distinct nitroreductases in *Salmonella typhimurium*: roles in mutagenesis., *Carcinogenesis.*, 3:121-123.
- Rosenkranz HS and Mermelstein R. (1983). Mutagenicity and genotoxicity of nitroarenes. All nitro-containing chemicals were not created equal. *Mutat Res.* 114:217-267.
- Rosenkranz HS, McCoy EC, Sanders DR, Butler M, Kiriazides DK, Mermelstein R. (1980). Nitropyrenes: isolation, identification, and reduction of mutagenic impurities in carbon black and toners. *Science.*, 209:1039-1043.
- Ross JA, Leavitt SA, (1998). Induction of mutations by 2-acetylaminofluorene in lacI transgenic B6C3F1 mouse liver. *Mutagenesis*, 13:173-179.

- Rothman N, Talaska G, Hayes RB, Bhatnagar VK, Bell DA, Lakshmi VM, Kashyap SK, Dosemeci M, Kashyap R, Hsu FF, Jaeger M, Hirvonen A, Parikh DJ, Davis BB, Zenser TV., (1997). Acidic urine pH is associated with elevated levels of free urinary benzidine and N-acetylbenzidine and urothelial cell DNA adducts in exposed workers. *Cancer Epidemiol Biomarkers Prev.*, 6:1039-1042.
- Roush AA, Suarez M, Friedberg EC, Radman M, Siede W., (1998). Deletion of the *Saccharomyces cerevisiae* gene RAD30 encoding an *Escherichia coli* DinB homolog confers UV radiation sensitivity and altered mutability. *Mol Gen Genet.*, 257:686-692.
- Roy A. and Fuchs RPP, (1994). The mutational spectrum induced in *S.cerevisiae* by the carcinogen N-2-acetylaminofluorene. *Mol.Gen.Genet.*, 245:69-77.
- Rupp WD, Howard-Flanders P. (1968). Discontinuities in the DNA synthesized in an excision-defective strain of *Escherichia coli* following ultraviolet irradiation. *J Mol Biol.*, 28:291-304.
- Rupp WD, Howard-Flanders P. (1968). Discontinuities in the DNA synthesized in an excision-defective strain of *Escherichia coli* following ultraviolet irradiation. *J Mol Biol.*, 31:291-304.
- Sage E, Leng M, (1981). Conformational changes of poly(dG-dC) . poly(dG-dC) modified by the carcinogen N-acetoxy-N-acetyl-2-aminofluorene. *Nucleic Acids Res.*, 9:1241-1250.
- Sakai H, Park SS, Kikkawa Y., (1992). Differential oxidase activity of hepatic and pulmonary microsomal cytochrome P-450 isozymes after treatment with cytochrome P-450 inducers. *Biochem Biophys Res Commun.*, 187:1262-1269.
- Sancar A, Tang MS. (1993). Nucleotide excision repair. *Photochem Photobiol.*, 57:905-921.
- Sancar A. (1994). Structure and function of DNA photolyase. *Biochemistry.*, 33:2-9.
- Sancar A., (1995). Excision repair in mammalian cells. *J Biol Chem.*, 270:15915-15918.
- Sancar A., (1996). DNA excision repair. *Annu Rev Biochem.*, 65:43-81
- Sancar GB, (1990). DNA photolyases: physical properties, action mechanism, and roles in dark repair. *Mutat Res.*, 236:147-160.
- Sancar GB, (2000). Enzymatic photoreactivation: 50 years and counting. *Mutat Res.*, 451:25-37.
- Sandler SJ, Marians KJ, (2000). Role of PriA in replication fork reactivation in *Escherichia coli*. *J Bacteriol.*, 182:9-13.
- Santella R.M., Grunberger D., Broyde S., Hingerty B.E., (1981b). Z-DNA conformation of N-2-acetylaminofluorene modified poly(dG-dC) determined by reactivity with anti cytidine antibody and minimized potential energy calculations. *Nucleic Acids Res.*, 9:5459-5467.
- Santella R.M., Grunberger D., Weinstein I.B., Rich A., (1981a). Induction of the Z conformation in poly(dG-dC).poly(dG-dC) by binding of N-2-acetylaminofluorene to guanine residues. *Proc Natl Acad Sci U S A*, 78:1451-1455
- Sassanfar M, Dosanjh MK, Essigmann JM, Samson L. (1991). Relative efficiencies of the bacterial, yeast, and human DNA methyltransferases for the repair of O6-methylguanine and O4-methylthymine. Suggestive evidence for O4-

- methylthymine repair by eukaryotic methyltransferases. *J Biol Chem.*, 266:2767-2771.
- Sassanfar M, Roberts JW. (1990). Nature of the SOS-inducing signal in *Escherichia coli*. The involvement of DNA replication. *J Mol Biol.*, 212:79-96.
- Schaaper RM, Koffel-Schwartz N and Fuchs RP., (1990). N-acetoxy-N-acetyl-2-aminofluorene-induced mutagenesis in the *lacI* gene of *Escherichia coli*. *Carcinogenesis*, 11:1087-1095.
- Schnarr M, Oertel-Buchheit P, Kazmaier M, Granger-Schnarr M., (1991). DNA binding properties of the LexA repressor. *Biochimie.*, 73:423-431.
- Schofield MJ, Nayak S, Scott TH, Du C, Hsieh P., (2001). Interaction of *Escherichia coli* MutS and MutL at a DNA mismatch. *J Biol Chem.*, 276:28291-28299.
- Schuetzle D. (1983). Sampling of vehicle emissions for chemical analysis and biological testing. *Environ Health Perspect.* 47:65-80.
- Sedgwick SG. (1986). Inducible DNA repair in microbes. *Microbiol Sci.* 3:76-83.
- Selby CP, Sancar A., (1993). Transcription-repair coupling and mutation frequency decline. *J Bacteriol.*, 175:7509-7514.
- Selby CP, Witkin EM, Sancar A. (1991). *Escherichia coli* *mfd* mutant deficient in "mutation frequency decline" lacks strand-specific repair: in vitro complementation with purified coupling factor. *Proc Natl Acad Sci U S A.*, 88:11574-11578.
- Servant L, Cazaux C, Bieth A, Iwai S, Hanaoka F, Hoffmann JS., (2002). A role for DNA polymerase beta in mutagenic UV lesion bypass. *J Biol Chem.*, 277:50046-50053.
- Shan Q, Bork JM, Webb BL, Inman RB, Cox, MM. (1997). RecA protein filaments: end-dependent dissociation from ssDNA and stabilization by RecO and RecR proteins. *J Mol Biol.*, 265:519-540.
- Shapiro R, Sidawi D, Miao YS, Hingerty BE, Schmidt KE, Moskowitz J, Broyde S. (1994). Conformation of amine-modified DNA: 2-aminofluorene- and 2-acetylaminofluorene-modified deoxydinucleoside monophosphates with all possible nearest neighbors. A comparison of search and optimization methods. *Chem Res Toxicol*, 7:239-253.
- Shen X, Sayer JM, Kroth H, Ponten I, O'Donnell M, Woodgate R, Jerina DM, Goodman MF., (2002). Efficiency and accuracy of SOS-induced DNA polymerases replicating benzo[a]pyrene-7,8-diol 9,10-epoxide A and G adducts. *J Biol Chem.*, 277:5265-5274.
- Shevell DE, Friedman BM, Walker GC. (1990). Resistance to alkylation damage in *Escherichia coli*: role of the Ada protein in induction of the adaptive response. *Mutat Res.*, 233:53-72.
- Shibutani S, Suzuki N, Grollman AP., (1998). Mutagenic specificity of (acetylamino)fluorene-derived DNA adducts in mammalian cells. *Biochemistry*, 37:12034-12041.
- Shinagawa H, Iwasaki H, Kato T, Nakata A., (1988). RecA protein-dependent cleavage of UmuD protein and SOS mutagenesis. *Proc Natl Acad Sci U S A.*, 85:1806-1810.

- Shinagawa H. (1996). SOS response as an adaptive response to DNA damage in prokaryotes. *EXS.*, 77:221-235.
- Silvian LF, Toth EA, Pham P, Goodman MF, Ellenberger T., (2001) Crystal structure of a DinB family error-prone DNA polymerase from *Sulfolobus solfataricus*. *Nat Struct Biol.*, 8:984-989.
- Sim E, Payton M, Noble M, Minchin R., (2000). An update on genetic, structural and functional studies of arylamine N-acetyltransferases in eucaryotes and procaryotes. *Hum Mol Genet.*, 9:2435-2441.
- Sinha D. and Dao .TL., (1974). A direct mechanism of mammary carcinogenesis induced by 7,12-dimethyl-benz(a)anthracene. *J Natl Cancer Inst.*, 53:841-846.
- Smith BT, Walker GC. 1998. Mutagenesis and more: umuDC and the *Escherichia coli* SOS response. *Genetics*, 148:1599–1610.
- Smith CA, Wang M, Jiang N, Che L, Zhao X, Taylor JS., (1996). Mutation spectra of M13 vectors containing site-specific Cis-Syn, Trans-Syn-I, (6-4), and Dewar pyrimidone photoproducts of thymidyl-(3'-->5')-thymidine in *Escherichia coli* under SOS conditions. *Biochemistry.*, 35:4146-4154.
- Solution structure of the N-(deoxyguanosin-8-yl)-1-aminopyrene ([AP]dG) adduct opposite dA in a DNA duplex. *Biochemistry*, 38:10843-10854.
- Sommer S, Bailone A, Devoret R., (1993). The appearance of the UmuD'C protein complex in *Escherichia coli* switches repair from homologous recombination to SOS mutagenesis. *Mol Microbiol.*, 10:963-971.
- Sommer S, Boudsocq F, Devoret R, Bailone A. (1998). Specific RecA amino acid changes affect RecA-UmuD'C interaction. *Mol Microbiol.*, 28:281-291.
- Spampinato C, Modrich P.(2000). The MutL ATPase is required for mismatch repair. *J Biol Chem.*, 275:9863-9869.
- Sary A, Kannouche P, Lehmann AR, Sarasin A., (2003). Role of DNA polymerase eta in the UV mutation spectrum in human cells. *J Biol Chem.*, 278:18767-18775.
- Steinborn G., (1978). Uvm mutants of *Escherichia coli* K12 deficient in UV mutagenesis. I. Isolation of uvm mutants and their phenotypical characterization in DNA repair and mutagenesis. *Mol Gen Genet.*, 165:87-93.
- Streisinger G, Okada Y, Emrich J, Newton J, Tsugita A, Terzaghi E, Inouye, (1966). M.Frameshift mutations and the genetic code. *Cold Spring Harb Symp Quant Biol.*, 31:77-84.
- Streisinger G, Owen J. (1985). Mechanisms of spontaneous and induced frameshift mutation in bacteriophage T4. *Genetics.*, 109:633-659.
- Stukenberg PT, Studwell-Vaughan PS, O'Donnell M., (1991). Mechanism of the sliding beta-clamp of DNA polymerase III holoenzyme. *J Biol Chem.*, 266:11328-11334.
- Su SS, Lahue RS, Au KG, Modrich P. (1988). Mismatch specificity of methyl-directed DNA mismatch correction in vitro. *J Biol Chem.*, 263:6829-6835.
- Sugimura T., (1985). Carcinogenicity of mutagenic heterocyclic amines formed during the cooking process. *Mutation. Res.*, 150:33-41.
- Sugimura T., (1986). Studies on environmental chemical carcinogenesis in Japan. *Science.*, 233:312-318.

- Sutton MD, Farrow MF, Burton BM, Walker GC., (2001a). Genetic interactions between the *Escherichia coli* umuDC gene products and the beta processivity clamp of the replicative DNA polymerase. *J Bacteriol.*, 183:2897-2909
- Sutton MD, Opperman T, Walker GC., (1999). The *Escherichia coli* SOS mutagenesis proteins UmuD and UmuD' interact physically with the replicative DNA polymerase. *Proc Natl Acad Sci U S A*, 96:12373-12378.
- Sutton MD, Smith BT, Godoy VG, Walker GC. (2000). The SOS response: recent insights into umuDC-dependent mutagenesis and DNA damage tolerance. *Annu Rev Genet.*, 34:479-497.
- Sutton MD, Walker GC., (2001). Managing DNA polymerases: coordinating DNA replication, DNA repair, and DNA recombination. *Proc Natl Acad Sci U S A.*, 98:8342-8349.
- Sutton, M. D., and G. C. Walker., (2001c). umuDC-mediated cold sensitivity is a manifestation of functions of the UmuD₂C complex involved in a DNA damage checkpoint control. *J. Bacteriol.* 183:1215-1224.
- Sutton, M. D., S. Murli, T. Opperman, C. Klein, and G. C. Walker., (2001b). umuDC-dnaQ interaction and its implications for cell cycle regulation and SOS mutagenesis in *Escherichia coli*. *J. Bacteriol.*, 183:1085-1089.
- Suzuki N, Ohashi E, Hayashi K, Ohmori H, Grollman AP, Shibutani S., (2001). Translesional synthesis past acetylaminofluorene-derived DNA adducts catalyzed by human DNA polymerase kappa and *Escherichia coli* DNA polymerase IV. *Biochemistry.*, 40:15176-15183.
- Tadmor Y, Ascarelli-Goell R, Skaliter R, Livneh Z., (1992). Overproduction of the beta subunit of DNA polymerase III holoenzyme reduces UV mutagenesis in *Escherichia coli*. *J Bacteriol.*, 174:2517-2524.
- Taft-Benz, S. A., and R. M. Schaaper. (1999). The C-terminal domain of dnaQ contains the polymerase binding site. *J. Bacteriol.*, 181:2963-2965.
- Tagliaferri F, Teodori L, Valente MG, Stipa F, Cucina A, Gohde W, Colettii D, Alo P and Stipa S., (2000). In vitro proliferation and in vivo malignancy of cell lines simultaneously derived from a chemically-induced heterogeneous rat mammary tumor. *In Vitro Cell Dev Biol Anim.*, 36:163-166.
- Tang M and Lieberman MW., (1983). Quantification of adducts formed in DNA treated with N-acetoxy-2-acetylaminofluorene or N-hydroxy-2-aminofluorene: comparison of trifluoroacetic acid and enzymatic degradation. *Carcinogenesis.*, 4:1001-1006.
- Tang M, Bruck I, Eritja R, Turner J, Frank EG, et al., (1998). Biochemical basis of SOS-induced mutagenesis in *Escherichia coli*: reconstitution of in vitro lesion bypass dependent on the UmuD'₂C mutagenic complex and RecA protein. *Proc. Natl. Acad. Sci. USA*, 95:9755-9760.
- Tang M, Pham P, Shen X, Taylor J-S, O'Donnell M, (2000). Roles of *E. coli* DNA polymerase IV and V in lesion-targeted and untargeted SOS mutagenesis. *Nature*, 404:1014-1018.
- Tang M, Shen X, Frank EG, O'Donnell M, Woodgate R, Goodman MF., (1999). UmuD'₂C is an error-prone DNA polymerase, *Escherichia coli* pol V. *Proc. Natl. Acad. Sci. USA*, 96:8919-8924.

- Tippin B, Pham P, Goodman MF., (2004) Error-prone replication for better or worse. *Trends Microbiol.* 2004 Jun;12(6):288-295.
- Tissier A, Frank EG, McDonald JP, Iwai S, Hanaoka F, Woodgate R., (2000a). Misinsertion and bypass of thymine-thymine dimers by human DNA polymerase ϵ . *EMBO J.*, 19:5259-5266.
- Tissier A, McDonald JP, Frank EG, Woodgate R., (2000b). Pol ϵ , a remarkably error-prone human DNA polymerase. *Genes Dev.*, 14:1642-1650.
- Tocher JH., (1997). Reductive activation of nitroheterocyclic compounds. *Gen. Pharmacol.*, 28:485-487.
- Tokiwa H. and Ohnishi Y., (1986). Mutagenicity and carcinogenicity of nitroarenes and their sources in the environment. *Crit. Rev. Toxicol.*, 17:23-70.
- Townson SM, Boreham PF, Upcroft P, Upcroft JA., (1994). Resistance to the nitroheterocyclic drugs. *Acta Trop.*, 56:173-194.
- Trautwein C, Will M, Kubicka S, Rakemann T, Flemming P and Manns MP (1999). 2-acetaminofluorene blocks cell cycle progression after hepatectomy by p21 induction and lack of cyclin E expression. *Oncogene.*, 18:6443-6453.
- Trewick SC, Henshaw TF, Hausinger RP, Lindahl T, Sedgwick B., (2002). Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. *Nature*, 419:174-178.
- Trincao J, Johnson RE, Escalante CR, Prakash S, Prakash L, Aggarwal AK., (2001). Structure of the catalytic core of *S. cerevisiae* DNA polymerase ϵ : implications for translesion DNA synthesis. *Mol Cell.*, 8:417-426.
- Vaisman A, Masutani C, Hanaoka F, Chaney SG., (2000). Efficient translesion replication past oxaliplatin and cisplatin GpG adducts by human DNA polymerase ϵ . *Biochemistry.*, 39:4575-4580.
- van der Meer JR, de Vos WM, Harayama S, Zehnder AJ., (1992). Molecular mechanisms of genetic adaptation to xenobiotic compounds. *Microbiol Rev.*, 56:677-694.
- Van Houten B. (1990). Nucleotide excision repair in *Escherichia coli*. *Microbiol Rev.* 54:18-51.
- Vandewiele D, Borden A, O'Grady PI, Woodgate R, Lawrence CW., (1998). Efficient translesion replication in the absence of *Escherichia coli* Umu proteins and 3'-5' exonuclease proofreading function. *Proc Natl Acad Sci U S A.*, 95:15519-21554.
- Vineis P, Bartsch H, Caporaso N, Harrington AM, Kadlubar FF, Landi MT, Malaveille C, Shields PG, Skipper P, Talaska G., (1994). Genetically based N-acetyltransferase metabolic polymorphism and low-level environmental exposure to carcinogens. *Nature.*, 369:154-156.
- Vineis P, Caporaso N, Tannenbaum SR, Skipper PL, Glogowski J, Bartsch H, Coda M, Talaska G, Kadlubar F., (1990). Acetylation phenotype, carcinogen-hemoglobin adducts, and cigarette smoking. *Cancer Res.*, 50:3002-3004.
- Viswanathan M, Burdett V, Baitinger C, Modrich P, Lovett ST., (2001). Redundant exonuclease involvement in *Escherichia coli* methyl-directed mismatch repair. *J Biol Chem.*, 276:31053-31058.
- Viswanathan M, Lovett ST., (1999). Exonuclease X of *Escherichia coli*. A novel 3'-5' DNase and DnaQ superfamily member involved in DNA repair. *J Biol Chem.*, 274:30094-300100.

- Volkert MR. (1988). Adaptive response of *Escherichia coli* to alkylation damage. *Environ Mol Mutagen.*, 11:241-255.
- Wagner J, Etienne H, Janel-Bintz R, Fuchs RP., (2002). Genetics of mutagenesis in *E. coli*: various combinations of translesion polymerases (Pol II, IV and V) deal with lesion/sequence context diversity. *DNA Repair (Amst).*, 1:159-167.
- Wagner J, Gruz P, Kim SR, Yamada M, Matsui K, et al.1999. The *dinB* gene encodes a novel *E. coli* DNA polymerase, DNA pol IV, involved in mutagenesis. *Mol. Cell* 4:281-286.
- Walker GC, Dobson PP., (1979). Mutagenesis and repair deficiencies of *Escherichia coli* *umuC* mutants are suppressed by the plasmid pKM101. *Mol Gen Genet.*, 172:17-24.
- Walker GC, Marsh L, Dodson L. (1985). Cellular responses to DNA damage. *Environ Health Perspect.*, 62:115-117.
- Walker GC. (1985). Inducible DNA repair systems. *Annu Rev Biochem.*, 54:425-457.
- Walker GC. (1995). SOS-regulated proteins in translesion DNA synthesis and mutagenesis. *Trends Biochem Sci.*, 20:4164-4220.
- Wang AH, Quigley GJ, Kolpak FJ, Crawford JL, van Boom JH, van der Marel G, Rich A, (1979). Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature*, 282:680-686.
- Wang CY, Zukowski K, Lee MS, Imaida K. (1987). Production of urothelial tumors in the heterotopic bladder of rats by instillation of N-glucuronosyl or N-acetyl derivatives of N-hydroxy-2-aminofluorene. *Cancer Res.*, 47:3406-3409.
- Wang Z., (2001). Translesion synthesis by the UmuC family of DNA polymerases. *Mutat Res.*, 486:59-70.
- Wang, L., and J. Lutkenhaus., (1998). FtsK is an essential cell division protein that is localized to the septum and induced as part of the SOS response. *Mol. Microbiol.* 29:731-740.
- Warnecke D, Erdmann R, Fahl A, Hube B, Muller F, Zank T, Zahringer U, Heinz E., (1999). Cloning and functional expression of UGT genes encoding sterol glucosyltransferases from *Saccharomyces cerevisiae*, *Candida albicans*, *Pichia pastoris*, and *Dictyostelium discoideum*. *J Biol Chem.*, 274:13048-13059.
- Washington MT, Wolfle WT, Spratt TE, Prakash L, Prakash S., (2003). Yeast DNA polymerase eta makes functional contacts with the DNA minor groove only at the incoming nucleoside triphosphate. *Proc Natl Acad Sci U S A.*, 100:5113-5118.
- Watabe T, Ishizuka T, Isobe M, Ozawa N.,(1982). A 7-hydroxymethyl sulfate ester as an active metabolite of 7,12-dimethylbenz[alpha]anthracene. *Science.*, 215:403-405.
- Webb BL, Cox MM, Inman RB., (1997). Recombinational DNA repair: the RecF and RecR proteins limit the extension of RecA filaments beyond single-strand DNA gaps. *Cell*, 91:347-356.
- Weisburger E.K. and Weisburger J.H., (1958). Chemistry, carcinogenicity and metabolism of 2-fluorenamine and related compounds. *Adv.Cancer.Res.*, 5:331-431.
- West SC. (1996). The RuvABC proteins and Holliday junction processing in *Escherichia coli*. *J Bacteriol.*, 178:1237-1241.

- Westra JG and Visser A., (1979). Quantitative analysis of N-(guanine-8-yl)-N-acetyl-2-aminofluorene and N-(guanine-8-yl)-2-aminofluorene in modified DNA by hydrolysis in trifluoroacetic acid and high pressure liquid chromatography. *Cancer Lett.*, 8:155-162.
- Williams GM, Jeffrey AM., (2000). Oxidative DNA damage: endogenous and chemically induced. *Regul Toxicol Pharmacol.*, 32:283-292.
- Wilson R.H., DeEds F. and Cox A.J., Jr. (1941). The toxicity and carcinogenic activity of 2-acetaminofluorene. *Cancer Res.*, 1:595-608.
- Witkin EM, Roegner-Maniscalco V, Sweasy JB, McCall JO. (1987). Recovery from ultraviolet light-induced inhibition of DNA synthesis requires umuDC gene products in recA718 mutant strains but not in recA+ strains of *Escherichia coli*. *Proc Natl Acad Sci U S A.*, 84:6805-6809.
- Witkin EM. (1976). Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. *Bacteriol Rev.*, 40:869-907.
- Witkin EM. (1981). Functions induced by damaged DNA: summary. *Prog Nucleic Acid Res Mol Biol.*, 26:2472-2550.
- Witkin EM. (1989). Ultraviolet mutagenesis and the SOS response in *Escherichia coli*: a personal perspective. *Environ Mol Mutagen.* 14 Suppl 16:30-34.
- Wood RD, (1996). DNA repair in eukaryotes. *Annu Rev Biochem.*, 65:135-167
- Woodgate R, Bridges BA, Herrera G, Blanco M., (1987) Mutagenic DNA repair in *Escherichia coli*. XIII. Proofreading exonuclease of DNA polymerase III holoenzyme is not operational during UV mutagenesis. *Mutat Res.*, 183:31-37.
- Woodgate R, Ennis DG., (1991). Levels of chromosomally encoded Umu proteins and requirements for in vivo UmuD cleavage. *Mol Gen Genet.*, 229:10-16.
- Woodgate R, Levine AS. (1996). Damage inducible mutagenesis: recent insights into the activities of the Umu family of mutagenesis proteins. *Cancer Surv.* 28:117-140.
- Woodgate R., (1999). A plethora of lesion-replicating DNA polymerases. *Genes Dev.*, 13:2191-2195.
- Woodgate R., (2001). Evolution of the two-step model for UV-mutagenesis. *Mutat Res.*, 485:83-92.
- Woodgate, R., M. Rajagopalan, C. Lu, and H. Echols. (1989). UmuC mutagenesis protein of *Escherichia coli*: purification and interaction with UmuD and UmuD'. *Proc. Natl. Acad. Sci. USA* 86:7301-7305.
- Woodside AM, Guengerich FP. (2002). Misincorporation and stalling at O(6)-methylguanine and O(6)-benzylguanine: evidence for inactive polymerase complexes. *Biochemistry.*, 41:1039-1050.
- Woodson SA, Crothers DM., (1987). Proton nuclear magnetic resonance studies on bulge-containing DNA oligonucleotides from a mutational hot-spot sequence. *Biochemistry.*, 26:904-912.
- Yamada A, Masutani C, Iwai S, Hanaoka F., (2000). Complementation of defective translesion synthesis and UV light sensitivity in xeroderma pigmentosum variant cells by human and mouse DNA polymerase eta. *Nucleic Acids Res.*, 28:2473-2480.
- Yang W., (2000). Structure and function of mismatch repair proteins., *Mutat Res.*, 460:245-456

- Yasuda T, Morimatsu K, Horii T, Nagata T, Ohmori H., (1998). Inhibition of *Escherichia coli* RecA coprotease activities by DinI. *EMBO J.*, 17:3207-3216.
- Yasuda T, Morimatsu K, Kato R, Usukura J, Takahashi M, Ohmori H., (2001). Physical interactions between DinI and RecA nucleoprotein filament for the regulation of SOS mutagenesis. *EMBO J.*, 20:1192-1202.
- Yebra, M. J., and A. S. Bhagwat. (1995). A cytosine methyltransferase converts 5-methylcytosine in DNA to thymine. *Biochemistry*, 34:14572-14557.
- Yu X, Egelman EH. (1993). The LexA repressor binds within the deep helical groove of the activated RecA filament. *J. Mol. Biol.* 231:29-40.
- Yuan F, Zhang Y, Rajpal DK, Wu X, Guo D, Wang M, Taylor JS, Wang Z., (2000). Specificity of DNA lesion bypass by the yeast DNA polymerase ϵ . *J Biol Chem.*, 275:8233-8239.
- Zenno S, Kobori T, Tanokura M, Saigo K., (1998). Purification and characterization of NfrA1, a *Bacillus subtilis* nitro/flavin reductase capable of interacting with the bacterial luciferase. *Biosci Biotechnol Biochem.*, 62:1978-1987.
- Zenno S, Koike H, Kumar AN, Jayaraman R, Tanokura M, Saigo K., (1996a). Biochemical characterization of NfsA, the *Escherichia coli* major nitroreductase exhibiting a high amino acid sequence homology to Frp, a *Vibrio harveyi* flavin oxidoreductase., *J Bacteriol.*, 178:4508-4514.
- Zenno S, Koike H, Tanokura M, Saigo K., (1996b). Gene cloning, purification, and characterization of NfsB, a minor oxygen-insensitive nitroreductase from *Escherichia coli*, similar in biochemical properties to FRase I, the major flavin reductase in *Vibrio fischeri*. *J Biochem.*, 120:736-744.
- Zenser TV, Lakshmi VM, Davis BB., (1998). N-glucuronidation of benzidine and its metabolites. Role in bladder cancer. *Drug Metab Dispos.*, 26:856-859.
- Zhang QY, Wikoff J, Dunbar D, Kaminsky L., (1996). Characterization of rat small intestinal cytochrome P450 composition and inducibility. *Drug Metab Dispos.*, 24:322-328.
- Zhang Y, Wu X, Guo D, Rechkoblit O, Taylor JS, Geacintov NE, Wang Z., (2002a). Lesion bypass activities of human DNA polymerase μ . *J Biol Chem.*, 277:44582-44587.
- Zhang Y, Wu X, Rechkoblit O, Geacintov NE, Taylor JS, Wang Z., (2002b). Response of human REV1 to different DNA damage: preferential dCMP insertion opposite the lesion. *Nucleic Acids Res.*, 30:1630-1638.
- Zhang Y, Wu X, Yuan F, Xie Z, Wang Z., (2001a) Highly frequent frameshift DNA synthesis by human DNA polymerase μ . *Mol Cell Biol.*, 21:7995-8006.
- Zhang Y, Yuan F, Wu X, Rechkoblit O, Taylor JS, Geacintov NE, Wang Z., (2000d). Error-prone lesion bypass by human DNA polymerase ϵ . *Nucleic Acids Res.*, 28:4717-4724.
- Zhang Y, Yuan F, Wu X, Taylor JS, Wang Z., (2001b). Response of human DNA polymerase ι to DNA lesions. *Nucleic Acids Res.*, 29:928-935.
- Zhang Y, Yuan F, Wu X, Wang M, Rechkoblit O, Taylor JS, Geacintov NE, Wang Z., (2000a). Error-free and error-prone lesion bypass by human DNA polymerase κ in vitro. *Nucleic Acids Res.*, 28:4138-4146.

- Zhang Y, Yuan F, Wu X, Wang Z., (2000b). Preferential incorporation of G opposite template T by the low-fidelity human DNA polymerase iota. *Mol Cell Biol.*, 20:7099-7108.
- Zhang Y, Yuan F, Xin H, Wu X, Rajpal DK, Yang D, Wang Z., (2000c). Human DNA polymerase kappa synthesizes DNA with extraordinarily low fidelity. *Nucleic Acids Res.*, 28:4147-4156.
- Zhou BL, Pata JD, Steitz TA., (2001). Crystal structure of a DinB lesion bypass DNA polymerase catalytic fragment reveals a classic polymerase catalytic domain. *Mol Cell.*, 8:427-437.