The *supF* assay for understanding
DNA adduct-induced mutagenesis: traditional
application and development of a
site-specific version

Thesis submitted for the degree of

Doctor of Philosophy

At the University of Leicester

By

Evagelos Liapis BSc (Hons) (London)

Department of Biochemistry

University of Leicester

November 2007

UMI Number: U495651

UMI

Dissertation Publishing

ProQuest

Abstract

The *supF* assay for understanding DNA adduct-induced mutagenesis: traditional application and development of a site-specific version

by Evagelos Liapis

Understanding the origin of mutations is central to understanding the aetiology of cancer. Several mutagenesis assay systems have been developed to measure the potency and mutagenic specificity of genotoxins and to study mutational mechanisms. A popular mutagenesis system is the *supF* assay, which employs the pSP189 double stranded shuttle vector plasmid that can be treated with carcinogens *in vitro* and transfected into human cells. Following replication, recovered plasmids can be screened in indicator bacteria for induced mutations in the *supF* gene.

A primary aim of this study was to establish and utilise the *supF* assay to investigate the mutagenicity of the cancer drug tamoxifen in target endometrial cells. In particular, the *supF* assay was employed to ascertain the mutations caused by two model reactive intermediates of tamoxifen, α-acetoxytamoxifen and 4-hydroxytamoxifen quinone methide (4-OHtamQM) in methylated pSP189 plasmid. The plasmid was methylated *in vitro* in order to allow application of the LwPy53 algorithm which enables *in silico* prediction of the G→T mutation distribution along the human *p53* gene, using α-acetoxytamoxifen and 4-OHtamQM-induced mutation data from the *supF* assay.

Relative mutation frequencies increased proportionally with the adduct level for α-acetoxytamoxifen, up to ~18 times the background frequency, while 4-OHtamQM failed to raise the mutation frequency above that of the untreated control. The majority of mutations in α-acetoxytamoxifen-treated plasmid were GC →TA transversions, while GC→AT transitions predominated in both 4-OHtamQM-treated plasmid and the untreated control. Based on the *p53* G→T transversion predictions, α-acetoxytamoxifen induced damage resulted in two hotspots at positions 244 and 273; these mutations might be expected to occur in tamoxifen associated endometrial tumours if tamoxifen adducts play a role in cancer development.

The second part of the thesis was devoted to the development and validation of a novel site-specific assay in *E.coli* MBM7070 and human cells to assess the mutagenicity of individual adducts, rather than compounds. The assay, which is a site-specific version of the *supF* assay, was validated and utilized to investigate the mutagenic potential and types of mutations caused by individual O^6-MeG adducts situated in intact double stranded or gapped plasmids, within two different sequence contexts, in both *E.coli* and human cells. The methods successfully established can be easily adapted to enable studies of any stable adduct that can be synthesized and incorporated into the plasmid.

To my parents,

Angelos & Eleni Liapi

Acknowledgements

Publications

Published Abstracts

Liapis, E.; Pearson, H.; Farmer, P. B.; Brown, K. Development of a site-specific assay for investigating the mutagenicity of DNA adducts in human cells. *United Kingdom Environmental Mutagen Society 28th Annual meeting, Bradford, UK. Mutagenesis* (2005) 20, 484, #51.

Liapis, E.; Lewis, P.D.; McLuckie, K. I. E.; Farmer, P. B.; Brown, K. Mutagenicity of tamoxifen DNA adducts in human endometrial cells and in silico prediction of *p53* mutation hotspots. *United Kingdom Environmental Mutagen Society 30th Annual meeting, Cardiff, UK. Mutagenesis* (2007) 22, 446, #57.

Liapis, E.; Farmer, P. B.; Brown, K. Development of a site-specific mutagenesis assay for the investigation of DNA adducts in human cells. *United Kingdom Environmental Mutagen Society 30th Annual meeting, Cardiff, UK. Mutagenesis* (2007) 22, 447 #58.

Contents

CHAPTER 1

Introduction

1.1 CANCER ETIOLOGY

Cancer is a collection of diseases that share uncontrolled growth as a common feature [1]. It is caused by environmental and inherited events that transform a normal cell to a malignant cell through a sequence of molecular changes. A majority of cancers are sporadic and some 1-5% are due to single-gene dominant traits [2, 3].

Each year an estimated 10.9 million people worldwide are diagnosed with cancer [4]. As shown in Table 1.1, most of these people (45%) are in Asia, because of the size of its population.

New cases of Cancer, 2002	Number	% of Total
Africa	649,800	6
Asia	4,876,900	45
Europe	2,820,800	26
Latin America & the Caribbean	833,000	8
Northern America	1,570,500	14
Oceania	111,500	1
Developed Countries	5,016,100	46
Developing Countries	5,827,500	54
World	10,862,500	100

Table 1.1 Cancer incidence for the regions of the world, 2002 estimates [4]

In 1981 Doll and Peto published a seminal review of trends and causes of cancer [5]. Based on the epidemiological observation that migrants tend to acquire the cancer rates of their new country, they concluded that differences in cancer rates can be attributed, in part, to environmental factors such as smoking, diet, reproductive behaviour, infection and occupational exposures (Figure 1.1). There were many criticisms of the Doll and Peto studies, such as that they placed too much emphasis on lifestyle factors-for example, smoking-with too little emphasis on involuntary exposures, such as environmental

3

carcinogens [6]. Nevertheless, in the past 25 years, many epidemiological studies have confirmed the contribution of exogenous factors to the etiology of cancers [7] and have expanded the list of factors to include obesity and lack of physical activity.
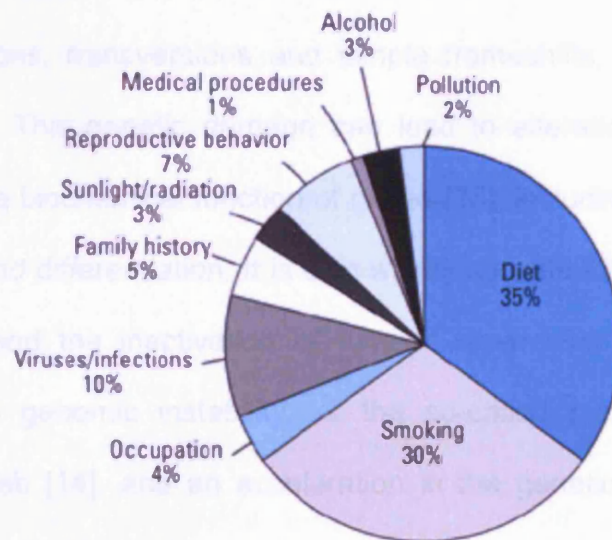
Figure 1.1 Pie Chart of the most common causes of cancer [5]

There is now substantial experimental evidence to support the multistage model of carcinogenesis (Figure 1.2) [8-10]. According to the clonal evolution theory of cancer proposed by Nowell [10] all cancers arise by expansion of cell numbers from a single cell.
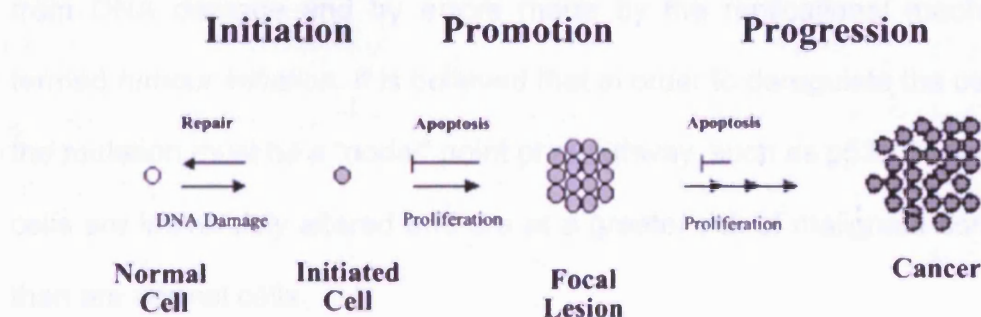
Figure 1.2 Multistage Model of Carcinogenesis

4

If a single abnormal cell is to give rise to a tumour, it must pass on its abnormality to its progeny: the aberration has to be heritable [11]. As cancers develop, they often acquire an increasing number of genetic alterations. At the chromosomal level, aneuploidy, recombination, translocation, large insertions and deletions are commonly observed. At the DNA level, point mutations, such as transitions, transversions and simple frameshifts, have also been recognised [12]. This genetic damage can lead to alterations in either the expression or the biochemical function of genes [13], including those involved in proliferation and differentiation. It is also widely accepted that the activation of oncogenes, and the inactivation of tumour suppressor and DNA-repair genes, leads to genomic instability, or the so-called *mutator phenotype*, proposed by Loeb [14], and an acceleration in the genetic changes taking place. The number of genes involved in neoplastic development is not known with certainty. Most colorectal cancers have three or more altered genes [15, 16] and estimates of as many as 10 or more mutational changes have been proposed to occur in adult human cancers [17].

The life history of a cancer can be divided into three stages: initiation, promotion, and progression (Figure 1.2) [1]. Within the conceptual framework of multistage carcinogenesis, the initial genetic change (mutation) that results from DNA damage and by errors made by the replicational machinery is termed *tumour initiation*. It is believed that in order to deregulate the cell cycle, the mutation must hit a "nodal" point of a pathway, such as p53. Thus, initiated cells are irreversibly altered and are at a greater risk of malignant conversion than are normal cells.

Tumour promotion comprises the selective clonal expansion of initiated cells and the epigenetic effects of tumour promoters facilitate this process. It is caused by a continuous disturbance of cellular signal transduction that results in an overstimulation of metabolic pathways along which mediators of cell proliferation and inflammation as well as genotoxic by-products are generated [18]. The end products of tumour promotion are generally benign lesions or foci of preneoplastic cells. These cells are more vulnerable to tumourigenesis because they now represent a larger, more rapidly proliferating target population for the further action of chemical carcinogens, oncogenic viruses, and other cofactors [19].

The progression of benign tumours to malignant cancers is a phase in carcinogenesis clearly distinct from promotion [20]. Unlike benign tumours, malignant tumours undergo a series of changes as they develop from a premalignant to a malignant state. This process is known as *tumour progression* (the development of cancer cells exhibiting permanent, irreversible, qualitative and heritable changes in one or more characteristics not previously manifested in a cancer cell population). An important consequence of tumour progression is the development of tumour heterogeneity.

A first problem in understanding the etiology of a cancer is to discover whether the heritable aberration is due to a genetic change, that is, an alteration in the cell's DNA sequence, or to an *epigenetic* change, that is, any change in a phenotype that does not result from an alteration in the DNA sequence [11].

1.1.1 Epigenetic Mechanisms of Carcinogenesis

Epigenetic mechanisms regulating gene expression include signal transduction pathways, DNA methylation, chromatin remodelling, translational control and post-translation modifications [21, 22]. In sharp contrast with genetic modifications, epigenetic modifications are highly dynamic and reversible. The best way to characterise the epigenetics of cancer is as an example of dysregulation, including global hypomethylation, individual gene hypermethylation, chromatin modification, and altered genomic imprinting.

Methylation of DNA at the 5-position of cytosine is important in the regulation of gene expression and is one possible mechanism for the heritable change in cancer cells [23]. Overall hypomethylation takes place predominantly in repetitive and endoparasitic DNA sequences and has been linked to the generation of chromosomal instability. On the other hand, hypermethylation occurs in the CpG islands located in the promoters of certain tumour-suppressor genes, such as p16^{INK4a}, BRCA1 or hMLH, leading to gene silencing [24]. Abnormal hypermethylation of the regulatory region of genes could initiate and/or reinforce their inactivation either by directly interfering with the binding of transcription factors or by the generation of inactive chromatin. On the other hand, growth inducing genes, such as oncogenes, can become overexpressed as a result of hypomethylation or gene amplification [25]. Changes in DNA methylation patterns can also render certain genomic regions unstable by altering the chromatin structure, leading to deletions, inversions and chromosomal losses [26].

An area that has attracted more recent attention in cancer epigenetics is genome imprinting, a genetic phenomenon by which certain genes are expressed in a parent of origin specific manner. Normal imprinting, mediated by both DNA methylation and histone modifications, ensures paternally determined, heritable transcriptional expression of one gene allele and repression of the other [27]. Several genes undergo loss of imprinting (LOI) in cancers, such that both alleles are expressed in the tumour [28]. Finally, chromatin remodelling, mediated by post translational modifications (such as acetylation, deacetylation, methylation, ubiquitination and phosphorylation) of tails of histones may enhance or repress the transcription of certain genes [29].

1.1.2 Genetic mechanisms of carcinogenesis

There are, however, good reasons to think that the vast majority of cancers are initiated by genetic changes. Genetic changes can be classified as gene mutations, chromosome rearrangements, gene amplification, and aneuploidy. The origin of the somatic mutation theory is generally credited to Theodor Boveri [30] and there is now considerable evidence to support it, as discussed in [31].

First, cells of a variety of cancers can be shown to have a shared abnormality in their DNA sequence that distinguishes them from the normal cells surrounding the tumour. Point mutations have been observed to activate proto-oncogenes and to inactivate tumour-suppressor genes in certain cancers [32]. Chromosome rearrangements of oncogenes are also well documented. Gene amplification as well as numerical chromosome changes

are important in a number of different cancers [31]. Second, many of the agents known to give rise to cancer also cause genetic changes. Therefore, chemicals that induce any of the four distinct types of genetic events above could heritably alter a critical target gene necessary for neoplastic development. These observations support the use of mutagenicity assays in the evaluation of carcinogenic risks of chemicals to humans. Finally, the conclusion that somatic mutations underlie cancer is supported by studies of people who inherit a strong susceptibility to the disease. For example, people with the disease *xeroderma pigmentosum* have defects in the cellular system that repairs DNA damage induced by UV light, and they experience a hugely increased incidence of skin cancers [11]. Other examples of recognised syndromes of cancer with a hereditary component are Li-Fraumeni syndrome, familial adenomatous polyposis and retinoblastoma, which are initiated by inherited mutations in the p53, APC and RB-1 genes respectively [33-35] .

1.2 DNA DAMAGE

DNA replication has a remarkable fidelity, estimated to produce just 1×10^{-9}-10^{-10} mutations/nucleotide [36]. However, when the replication machinery encounters a DNA lesion, this fidelity becomes largely irrelevant, because the lesion may cause a physical barrier to the progression of replication and/or the coding information may be distorted. A perplexing diversity of lesions arises in DNA from three main causes. First, base alterations can arise from the inherent instability of specific chemical bonds that constitute the normal chemistry of nucleotides under physiological conditions of temperature and pH. These spontaneous alterations in the chemistry of DNA bases include tautomeric shifts (transient rearrangement of bonding), deamination and loss

of bases [37]. Second, DNA damage can result from reactive oxygen species (ROS) and other reactive metabolites that are produced during endogenous metabolic processes. Finally, environmental agents such as ultraviolet (UV) radiation and numerous genotoxic chemicals cause alterations in DNA structure, which if left unrepaired, may lead to mutations that enhance cancer risk [38]. These lesions produce structural and dynamic changes in DNA, which impair its function. Failure to correctly repair DNA damage can result in mutations, cancer and death.

1.2.1. Endogenous DNA damage

Endogenous DNA damage occurs at a high frequency compared with exogenous damage and the types of damage produced by normal cellular processes is similar to those caused by some environmental agents [39]. In fact it has been estimated that even without excessive exposure to DNA damaging agents, each day no less than 20,000 lesions are formed per cell in DNA from endogenous sources [40]. Spontaneous mutations are those that arise in the absence of known mutagen treatment and they account for the "background rate" of mutation.

1.2.1.1 Tautomeric Shifts

Tautomers are organic compounds that are interconvertible by a chemical reaction called tautomerization. The reaction results in the formal migration of a hydrogen atom or proton, accompanied by a switch of a single bond and adjacent double bond. The concept of tautomers that are interconvertible by tautomerizations is called tautomerism. Tautomerism is a special case of structural isomerism and can play an important role in non-canonical base

10

pairing in DNA and RNA molecules [41]. Nucleotide substitution can arise due to the phenomenon of tautomerism. Each base may exist as two possible alternative structures that interconvert. Such structural isomers that exist in dynamic equilibrium are known as tautomers. These tautomers exhibit different hydrogen-binding capabilities, which can lead to mispairing during replication and the introduction of errors into the DNA. Thymine has enol and keto isomers (Figure 1.3).
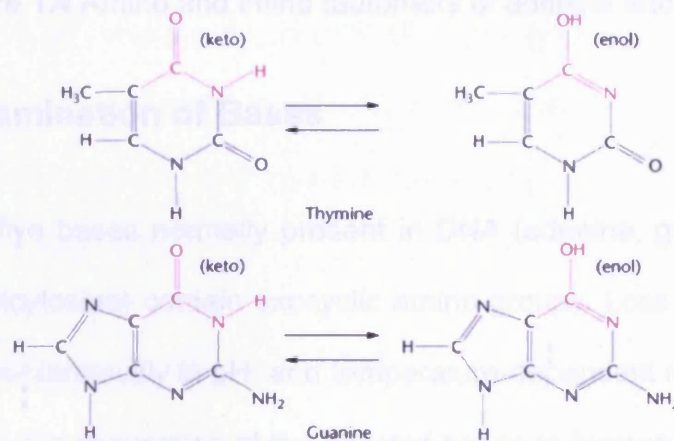
Figure 1.3 Keto and enol tautomers of thymine and guanine

The common keto form pairs with adenine, but the rare enol form base pairs with guanine. Guanine also has enol and keto tautomers. In this case the rare enol-guanine base pairs with thymine rather than cytosine. Similarly, adenine equilibrates between common amino and rare imino tautomers (Figure 1.4). The rare imino-adenine base pairs with cytosine instead of thymine. While cytosine also has amino and imino tautomers, they both pair with adenine. However, in general, the cell is able to limit the mutation rate due to tautomeric changes by proofreading and by the action of DNA repair systems [42].
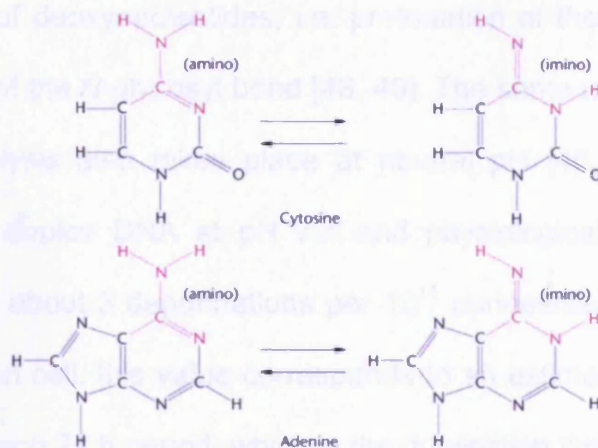
Figure 1.4 Amino and imino tautomers of adenine and cytosine

1.2.1.2 Deamination of Bases

Four of the five bases normally present in DNA (adenine, guanine, cytosine and 5-methylcytosine) contain exocyclic amino groups. Loss of these groups can occur spontaneously in pH- and temperature-dependent reactions of DNA and results in the conversion of the affected bases to hypoxanthine, xanthine, uracil and thymine respectively [43]. Failure to repair these deaminated bases before replication can give rise to mutations, since the loss of the amino group changes the pairing properties of the base [44].

1.2.1.3 Loss of Bases - Depurination and Depyrimidation

Bases can be lost from nucleic acids by cleavage of the N-glycosidic bond, leaving the sugar-chain intact, to produce an abasic site. The loss of purines and pyrimidines from DNA has been most extensively studied at acidic pH, although depurination and depyrimidation can also occur at appreciable rates at neutral or alkaline pH [45-47]. The mechanism of hydrolytic DNA depurination at acidic pH is thought to be the same as that established for

acid hydrolysis of deoxynucleotides, i.e. protonation of the base followed by direct cleavage of the N-glycosyl bond [48, 49]. The same mechanism of acid-catalysed hydrolysis also takes place at neutral pH [46, 50]. The rate of depurination of duplex DNA at pH 7.4 and physiological ionic strength is calculated to be about 3 depurinations per 10^{11} purines/second at 37°C [48]. For a mammalian cell, this value corresponds to an estimated loss of 10,000 purines during each 24 h period, which is the duplication time for many human cells [46]. Pyrimidine nucleosides are considerably more stable than purine nucleosides with respect to the N-glycosyl linkage of the bases to deoxyribose. The mechanism of depyrimidation is the same as for depurination, but cytosine and thymine are lost at rates only 1/20 of that for adenine or guanine [51].

1.2.1.4 Oxidative Damage to DNA

"Oxidative stress" is a term associated with both enhanced production of reactive oxygen species (ROS) and reduced efficacy of protection by antioxidant enzymes or low molecular weight antioxidants. Oxidation of cellular macromolecules occurs at significant frequencies in aerobic organisms due to byproducts of normal metabolism and the immune response. Furthermore, the oxidation of both mitochondrial and nuclear DNA has been implicated in human disease and aging. It has been generally estimated that about 1-3% of the oxygen consumed in humans is converted to ROS [52, 53]. These ROS include superoxide (O_2^-), hydrogen peroxide (H_2O_2), hydroxyl radicals ($^.OH$) and singlet oxygen ($^1O_2$) which can oxidise DNA, leading to several types of DNA damage, namely base modifications, sugar damage, strand breaks, abasic sites and DNA-protein crosslinks [54].

13

ROS also attack other cellular components such as lipids to generate reactive intermediates capable of binding to DNA, giving rise to alkylated lesions, including etheno- and propano adducts [55, 56]. Lipid hydroperoxides are the initial products of unsaturated fatty acid oxidation, but they are relatively short lived. However, they can react with metals to produce epoxides and aldehydes. The major aldehyde products of lipid peroxidation are crotonaldehyde, acrolein, 4-hydroxynonenal (HNE) and malondialdehyde (MDA). These reactive substances damage DNA by the formation of exocyclic adducts. MDA appears to be the most mutagenic product of lipid peroxidation, whereas HNE is the most toxic [57]. The short-lived ·OH radical is the most reactive of the primary ROS and can cause severe damage to DNA, protein and membrane lipids [58, 59]. Although less reactive, H_2O_2 is also believed to cause damage, either directly or indirectly, through further reduction to OH· in the Fenton reaction [60], which can be summarised as:

$$H_2O_2 + Fe^{2+} \rightarrow \cdot OH + OH^- + Fe^{3+}$$

Similarly, free superoxide radical is also not very reactive with DNA, however, it dismutates either spontaneously or through enzymatic processing by superoxide dismutases to produce H_2O_2. Whilst the Fenton reaction typically involves Fe^{2+} it can be catalysed by other transition metal ions such as Cu^{2+} [61]. In fact, the carcinogenic effect of certain metal salts has been attributed to their involvement in the Fenton reaction [62]. In the presence of reducing agents such as ascorbate [63] together with NADH [64] and superoxide, Fe^{3+} will recycle back to the active Fe^{2+} form:

$$O_2^- + Fe^{3+} \rightarrow O_2 + Fe^{2+}$$
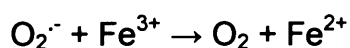
The two main modes of DNA attack by ROS are (i) addition to the double bonds of DNA bases and (ii) hydrogen abstraction from the deoxyribose sugar units in DNA. Over 80 products of DNA base damage caused by ROS are known [65]. One of the most widely studied oxidative DNA lesions is 7, 8-dihydro-8-oxo-2'-deoxyguanosine (8-oxodG). In DNA it readily assumes a *syn* conformation which can form a stable Hoogsteen mispair with dA, resulting in GC→TA transversion mutations following replication [66-68]. This lesion represents at least 5% of all oxidative lesions formed in DNA [69]. Steady state levels of this adduct determined by various techniques in human cells ranged over three orders of magnitude, depending largely on which assay is used to measure it [70]. A recent interlaboratory study has analysed 8-oxodG in the DNA of lymphocytes isolated from venous blood from healthy young male volunteers in several European countries [71]. The median concentration of 8-oxodG in lymphocyte DNA, calculated from the mean values of each group of subjects as determined by HPLC, was 4.24 per 10^6 guanines, while the median concentration of formamidopyrimidine DNA glycosylase (FPG) - sensitive sites, measured with the comet assay, was 0.34 per 10^6 guanines. In addition, identical samples of HeLa cells were supplied to all participants as a reference standard. The median values for 8-oxodGuo in HeLa cells were 2.78 per 10^6 guanines (by HPLC) and 0.50 per 10^6 guanines (by enzymic methods, such as the comet assay, alkaline unwinding or alkaline elution). The main conclusion of this study was that the true background level of base oxidation in DNA is orders of magnitude lower than has often be claimed in the past.

The majority of cells possess defence mechanisms against the potential harmful effects of ROS, such as superoxide dismutase, which converts its

substrate superoxide, into oxygen and hydrogen peroxide, and catalase, which converts hydrogen peroxide into water and oxygen. Additionally, there are various DNA repair processes, such as the base excision repair (BER) pathway and strand break repair enzymes, which remove these lesions to maintain the integrity of the genome [72].

1.2.2 Environmental DNA Damage

Epidemiological studies suggest that in developed societies exogenous factors are involved in about 75-80% of cancer cases [5, 73]. The notion that the environment and lifestyle have the principal role in the causation of sporadic cancer is also supported by analyses of cancer cases in cohort twins [74] and by analyses of family-cancer databases [75].

1.2.2.1. Chemical Carcinogens

1.2.2.1.1 Classification of Chemical Carcinogens

Chemical carcinogens can be broadly classified as "genotoxic" and "non-genotoxic". The term genotoxic is used to specifically describe carcinogens that damage DNA. The mechanism of non-genotoxic carcinogens (described above) or "epigenetic carcinogens" does not involve reaction with DNA but these chemicals display other properties underlying an increase in neoplasms, such as promoting activity [76].

However, the categorisation of carcinogens according to their mechanism of action is not straightforward, as it is evident that in some cases they can exert multiple effects via several, rather than a single, mechanism. For example, some chemicals that do not react directly with DNA, such as topoisomerase

inhibitors or inhibitors of the spindle apparatus or associated motor proteins, can result in genetic changes through indirect mechanisms including chromosome rearrangements and aneuploidy [77]. Additionally, evidence exists that metabolically activated estrogens can produce several types of DNA damage *in vitro* and in rodent models, forming estrogen-DNA adducts [78], and inducing oxidative damage indirectly through the production of free radicals [79]. It has been also demonstrated that several synthetic estrogens such as mestranol, 17α-ethinyl estradiol (EE) and diethylstilbestrol (DES) can act as promoters of hepatocarcinogenesis in rats of both sexes [80, 81]. On the other hand, DNA-reactive chemicals, such as methylating agents, may cause heritable epigenetic changes by altering gene methylation or gene expression [82] as well as induce mutations. Therefore, any attempt to classify chemical carcinogens must consider the complexity and the multitude of possible mechanisms.

1.2.2.1.2 Non-genotoxic carcinogens

Non-genotoxic or non-DNA-reactive carcinogens exert their effects through processes that do not involve direct covalent binding of the chemical or its metabolites to DNA [83]. Their biochemical modes of action are diverse, encompassing sustained cytotoxicity and disruption of cell-cycle checkpoints [84], cytochrome P450 enzyme induction [85], interference of repair enzymes [86], blocking apoptosis [87] and stimulation of oxidative stress [88]. In contrast to DNA-binding mutagens, non-DNA reactive carcinogens are considered to exhibit threshold concentration-effect response curves [89].

Peroxisome proliferators (PP) are a large class of structurally diverse industrial and pharmaceutical chemicals including plasticizers, herbicides and

hypolipidemic drugs. As their name implies, PPs cause a dramatic increase in the size and number of peroxisomes in exposed animals or cells. Peroxisomes are subcellular organelles found in most plant and animal cells that perform a wide variety of metabolic functions, including H_2O_2-derived respiration, β-oxidation of fatty acids and cholesterol metabolism [90]. Prolonged exposure to PPs causes an increased incidence of liver tumours in male and female mice and rats [91]. They are not considered to be genotoxic agents because they are almost uniformly negative in assays that measure covalent binding to DNA, production of DNA adducts and short-term tests for mutagenicity [92]. Although a few adverse effects on DNA (e.g. chromosomal aberrations) have been reported [93-95], these results may be due to the perturbation of processes that maintain the integrity of DNA rather than direct DNA damage [96]. Two hypotheses have been proposed to account for PP-induced hepatocarcinogenesis in rodents: an increase in indirect DNA damage through induction of oxidative stress and alteration of hepatocyte growth control by enhanced cell proliferation or decreased apoptosis [97].

Oxidative stress can also be triggered by certain chemicals that produce ROS, such as paraquat and potassium bromate [98, 99]. Other exogenous sources of ROS are tobacco smoke, transition metals, ethanol, redox cycling compounds or physical irradiation by multiple sources. In principle, chemicals that give rise to excess ROS production and lipid peroxidation will cause different types of toxicity, including genotoxicity and cell death.

Hormones represent another class of important human carcinogens. Hormone-related cancers share a common causal agent: both endogenous and exogenous hormones. For instance, administration of 17 β-estradiol (E_2)

to mice raises the incidence of mammary, pituitary, uterine, cervical, vaginal, testicular, lymphoid and bone tumours [100]. Several lines of evidence suggest that estrogens are epigenetic carcinogens acting via a promoting effect related to stimulation of proliferation of estrogen-responsive cells. This hypothesis is also supported by experimental observations of tumour-promoting effects of estrogens on carcinogen-initiated mammary [101], liver [102] and uterine [103] tumours in rodent models. Thus, use of anti-hormone therapy, such as tamoxifen, to interrupt this hormonal stimulus, could be expected to slow the process of progression until hormone independence occurs late in the pathway [104].

Compounds with an aneugenic (i.e. they cause aneuploidy) but not mutagenic effect may be characterised as carcinogenic only at high, toxic doses [105]. Aneuploidy in somatic human cells arises whenever the number of chromosomes deviates from 46. Effects such as aneuploidy, chromosome loss and non-disjunction have received particular attention, because of their potential involvement in the carcinogenic process [106, 107]. Nonetheless, currently it remains controversial whether aneuploidy is causal of cancers or simply reflects consequential changes in cells after they have become transformed [108, 109]. Colcemid and vinblastine induce aneuploidy by modifying the number of spindle fibres which regulate the segregation of chromosomes during mitosis and meiosis. It is expected that aneuploidy or polyploidy are induced only if most fibres are damaged, because of their redundancy in living cells [110].

1.2.2.1.3 DNA Reactive Carcinogens

Humans are inevitably exposed to over 25,000 toxic compounds that are potentially or demonstrably mutagenic in lower organisms [111]. These mutagens have a variety of sources, e.g. reactive or potentially reactive compounds found in the diet, automotive exhaust emissions, cigarette smoke and medicinal drugs. For a proportion of these chemicals a link between exposure and the development of human cancer has been established. In many cases, this link has been deduced from epidemiological studies of cancer incidence in workers exposed to specific chemicals in an industrial setting. For example, workers exposed to chromate in the chromate production industry and in the manufacture of pigments exhibit an elevated risk of lung cancer [112, 113].

Chemical reaction with DNA is a basis for mutagenicity and accordingly, genotoxic carcinogens account for the majority of chemical carcinogens [114]. The covalent interaction product between the specific chemical and DNA, known as a DNA adduct, is a precursor lesion, that if not repaired, may result in a mutation [115]. A genotoxic carcinogen may be an intrinsically reactive compound that can covalently bind to DNA or in most cases it requires prior metabolic activation, often through the P450 mono-oxygenase pathway [115]. The ultimate reactive species of genotoxic carcinogens are, in general, electrophiles; compounds that have an electron deficient atom which leads them to attack atoms that are electron rich (nucleophiles) [116]. The double helix has a number of N and O atoms that present targets for electrophiles. These atoms exist not only in the base, but also in the phosphodiester backbone.

In general, the stronger electrophiles display a greater range of nucleophilic targets. The site of DNA reaction can be predicted, to some extent, by the electrophilicity of the activated species and the 3-dimensional shape of the molecule [115]. However, at the moment it is difficult to predict the exact spectrum of damage induced by any given electrophile *in vivo* because of the complex interactions between the double helix and the chemical and higher order structures such as histones [37]. Modification of DNA by an electrophilic carcinogen can result in a number of consequences; these events maybe dependent on adduct structure, at what stage of the cell cycle the damage is induced, where the adduct is formed (sequence context) and the type of repair process involved in responding to the damage [116]. DNA adduct formation can induce a variety of effects influencing the cell cycle and DNA integrity, namely i) blockage or impediment of DNA replication, leading to arrested cell division or chromosomal aberrations, ii) hydrolysis, leading to abasic sites which may cause mutations and iii) misincorporation of bases during replication, leading to mutations [37]. Genotoxic carcinogens can be categorised according to structure and include alkylating agents, aralkylating agents, and arylamines [72, 117].

Alkylating agents

Humans are exposed to alkylating compounds produced both endogenously and present in the environment, or diet [118]. Carcinogens that transfer alkyl residues to DNA include the nitrosamines, aliphatic epoxides, aflatoxins, lactones, nitrosoureas, mustards, haloalkanes, alkyl triazenes and sultones [117]. Sources of endogenous alkylation include nitrosating agents and lipoperoxidation [119]. However, the doses of exogenous alkylating agents,

such as cytostatic drugs in cancer therapy, often far exceed those humans are

exposed to from endogenous sources [120].

Although numerous potential sites for alkylation have been identified in all four

bases, not all of them have equal reactivity and different types of agents have

different preferential binding sites. The most common sites for alkylation are

shown in Figure 1.5.
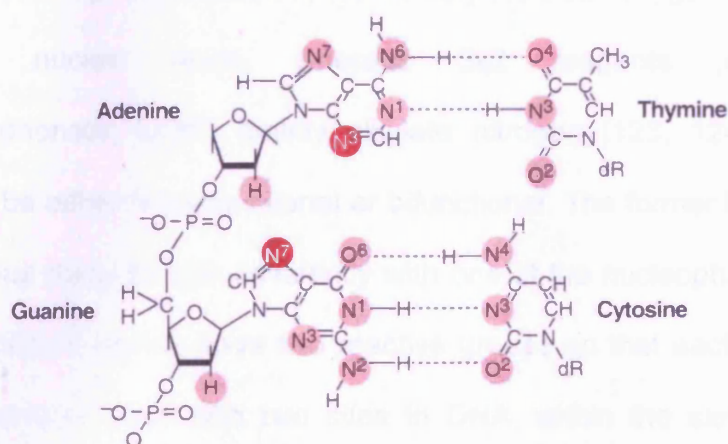
Figure 1.5 Nucleophilic centres in DNA that are the most reactive to alkylating

agents [37].

Although the reaction sites tend to be the same, the percentage of

modification at each position depends on the specific alkylating agent. In

general, it seems that alkylating agents that are not particularly ionic in nature

(e.g. dimethyl-nitrosamine or methyl methane-sulphonate) tend to react more

extensively with the more-nucleophilic ring nitrogen atoms, whereas those that

have greater ionic character (e.g. *N*-ethyl-*N*-nitrosourea) show greater

preference for reaction at the oxygen atoms in DNA, such as the O^6 position

of guanine and the phosphodiester groups of the DNA backbone [121, 122].

Mechanisms of alkylation are separated into two pathways, either S_N1 or S_N2, based on the kinetics of the alkylation reaction. Unimolecular nucleophilic substitution involves the formation of an intermediate carbonium ion, which then rapidly attacks DNA and other molecules (first-order kinetics). Therefore, the rate-limiting step in this type of reaction is formation of the carbonium ion. By contrast, S_N2 alkylating agents generate the reactive group upon direct contact with the DNA (second-order kinetics). Alkylating agents of the S_N1 type (e.g. N-methyl-N-nitrosourea, MNU) alkylate both oxygen and nitrogen atoms in nucleic acids, whereas S_N2 reagents (e.g. methyl methanesulphonate, MMS) mainly alkylate nitrogen [123, 124]. Alkylating agents can be either monofunctional or bifunctional. The former have a single reactive group that interacts covalently with one of the nucleophilic centers in DNA. Bifunctional agents have two reactive groups so that each molecule is potentially able to react with two sites in DNA, within the same strand or across two complementary strands.

Examples of alkylating agents found in environmental sources include alkyl sulfonates, N-nitroso compounds, triazenes and nitrogen mustards. N-nitroso compounds, especially nitrosamines, formed in tobacco smoke may represent a major environmental source of exposure to alkylating agents for humans. After metabolic activation they give rise to alkylated DNA adducts, such as O^6-meG and larger pyridyloxobutyl DNA adducts [125]. There is strong evidence that the specific nitrosamines found in tobacco are causative factors in human cancers of the lung, esophagus, pancreas and oral cavity [125].

Aralkylating agents are chemicals that transfer aralkyl groups to nucleotides, to form DNA adducts. Carcinogens of this category include the large group of

polycyclic aromatic hydrocarbons (PAH), aflatoxins, alkenyl benzenes, pyrrolizidine alkaloids and those nitroaromatics that are activated through the dihydrodiol epoxide mechanism [117].

PAHs are products of incomplete pyrolysis of organic matter and are widespread in the environment [130]. It is believed that the largest concentrations of PAHs are inhaled by smokers [131, 132]. Therefore, smoking-related PAHs are strongly implicated as being responsible for the initiation and development of lung cancer. Activated diol epoxides of PAHs form bulky covalent DNA adducts, primarily at the exocyclic N^2 position of guanine residues and the N^6 position of adenines [130]. An ultimate metabolite of the well studied PAH benzo[a]pyrene, is benzopyrene diolepoxide (BPDE). A strong correlation has been found between preferential sites of BPDE adduct formation in the human *p53* gene and lung cancer mutational hotspots, providing an etiological link between BPDE exposure and human lung cancer initiation [133].

The aflatoxins are naturally occurring mycotoxins found on foods such as corn and peanuts and they have been demonstrated to induce liver cancer in many experimental animal systems and also in humans based on experimental data and epidemiological studies in human populations [126, 127]. The four naturally occurring aflatoxins (B1, B2, G1 and G2) are highly substituted coumarins with a fused dihydrofuran configuration. Their carcinogenic potency decreases in the order B1>G1>B2>G2 [128]. Aflatoxins are known to be activated to their ultimate carcinogenic species by the cytochrome P450 mono-oxygenase system and pathways of aflatoxin metabolism include epoxidation, which leads to covalent DNA binding. The activated metabolites

preferentially bind to the N7 position of guanine, giving rise to the major aflatoxin-N7-guanine adduct (AFB1-N^7-gua). One possible pathway through which aflatoxin exerts its carcinogenicity and synergistic interaction with hepatitis B virus (HBV) infection in hepatocellular carcinoma (HCC) induction, is through its capacity to induce GC→TA transversions as a predominant mutation. One of the most striking examples of a molecular fingerprint in the human p53 gene is a characteristic G→T transversion at the third base of codon 249 which is observed in people with liver cancer from regions of high aflatoxin exposure [129].

Arylamines

The most heterogeneous groups of carcinogens are those that transfer an arylamine residue to DNA. This group includes the aromatic amines and amides, the aminoazo dyes, the nitroaromatics and the heterocyclic aromatic amines. Arylamines are widely used as dye intermediates, especially for azo dyes, pigments and optical brighteners, as intermediates for photographic chemicals, pharmaceuticals and agricultural chemicals, in polyurethanes, and as antioxidants in polymers [134]. Several arylamines are mutagenic and carcinogenic [135]. In humans, benzidine, 2-naphthylamine and 4-ABP have been classified by the IARC as group 1 human carcinogens [134]. It has been estimated that occupational exposures to aromatic amines explain up to 25 % of bladder cancers occurring in some regions of Western countries [136].

The metabolism of arylamines has been studied intensively and ring oxidation, N-glucuronidation, N-acetylation and N-oxidation are the major metabolic pathways of arylamines in mammals [137]. DNA adducts of arylamines can be formed through several metabolic intermediates, including N-

hydroxyarylamines, *N*-hydroxyarylamides, *O*-glucuronides and a variety of reactive arylamine esters such as *N*-sulphonyloxy, *N*-acetoxy and *N*-polyoxy derivatives. The major adducts are typically formed through linkage of the chemicals N-atom to the C8 position and amino groups of guanine. These sites are usually unaffected by alkylating agents [117].

1.2.2.2. Ionising Radiation

The carcinogenic potential of ionising radiation was recognised soon after Roentgen's discovery of X-rays in 1895 [138]. The principal sources of external background exposure are cosmic radiation and radionuclides naturally occurring on the planet.

Studies in several animal models [139], developed to investigate radiation-induced carcinogenesis, have shown radiation to be a "universal carcinogen" in that it will induce cancer in most tissues of most species at all ages, including the fetus. The types of cancer induced by radiation are of the same histological type as occur spontaneously, but their distribution may differ. Studies in *in vitro* cellular systems [140, 141] concluded that radiation is a relatively weak carcinogen and mutagen compared with certain chemical agents such as the PAHs.

Ionising radiation damages DNA through both direct and indirect effects. Direct damage results from the absorption of the radiation energy by DNA (which is considered to include the first layer of tightly bound water), leading to ionisation of bases and sugars [142]. It has been suggested that the direct action of ionising radiation may lead to the ejection of an electron from the unsaturated C5 or C6 position and that the resulting cation radical may

further react with a hydroxyl ion [142]. Indirect effects occur when species formed by radiation in water or other surrounding molecules react with cellular DNA. The predominance of water in biological systems means ROS formed by irradiation of water are the major sources of indirect damage to DNA [142, 143]. Thus, direct and indirect radiation effects may result in identical reactive intermediates that cause damage to DNA bases and strand breaks [72].

1.2.2.3. Ultraviolet Radiation

The ultraviolet portion of the solar spectrum is undoubtedly the major factor in the induction of skin cancer. This is well established for squamous and basal cell cancers together with malignant melanoma [144, 145]. The role of the sun in inducing these cancers was initially suggested by a number of clinical observations and has been subsequently confirmed by epidemiologic studies [146]. In relation to DNA damage, ultraviolet light can damage DNA by both direct and indirect mechanisms.

DNA is the most prominent cellular chromophore for the absorption of UV-B [147]. The absorption of energy causes excitation of a single electron to a higher, less stable, energy level. This reactive intermediate can then degrade or react with a variety of organic molecules in an oxygen-independent manner. This leads mostly to the formation of dimeric photoproducts at bipyrimidine sites in a strong sequence dependent manner. The most frequent lesions induced by UV-B and UV-C radiation are *cis-syn* cyclobutane pyrimidine dimers (CPDs) and pyrimidine-pyrimidone (6-4) photoproducts [(6-4) PPs]. Several minor photoproducts, such as the pyrimidine hydrates and thymine glycols are also formed [72]. Furthermore, UV-B can cause limited

numbers of strand breaks and single base deletions have also been detected [148]. Evidence also exists for the induction of 8-OHdG following UV-B exposure [149].

DNA is only a weak chromophore in the UV-A region. Instead, other organic molecules, referred to as photosensitizers, absorb photons of UV-A energy and may subsequently transfer this energy indirectly to DNA via reactive oxygen intermediates or radicals [148].

1.3 DNA REPAIR AND REPLICATION IN CANCER

Evolution has created a variety of sophisticated DNA repair systems to deal with most of the insults within the cell's vital genetic information. However, if these repair processes are saturated or defective, persistent lesions may interfere with DNA replication and eventually give rise to mutations [150]. Defective DNA repair generally causes cancer, developmental abnormalities and early aging [72]. At least three main partly overlapping damage repair pathways operate in mammals: direct damage reversal, excision repair systems and DNA double strand break (DSB) repair. Moreover, an additional replicational pathway, known as damage tolerance or postreplication repair, has been recently discovered [151].

1.3.1 Direct Damage Reversal

The direct reversal of DNA damage is by far the simplest repair mechanism described to date and consists of polypeptide chains, with enzymatic properties, binding to the damaged base and restoring the DNA genome to its normal state in a single-reaction step [152]. The major polypeptides involved

in this pathway are: i) DNA photolyase, which is the enzyme responsible for removing cyclobutane dimers from DNA in a light-dependent process termed photoreactivation [153] ii) DNA-alkyltransferases, which catalyse the efficient transfer of alkyl groups, ranging in size from methyl to benzyl groups, from the O^6 of guanine to the enzyme's active-site cysteine residue [154] , and iii) the ABH2 and ABH3 dioxygenases which catalyse the oxidation and release of the methyl group from 1-methyladenine and 3-methylcytosine [155]. Placental mammals (including humans), however, lack the proteins that catalyse photoreversal and, thus, can repair only alkylation damage [155]. DNA photolyases have been purified and characterised from a number of other organisms, including *E.coli*, the yeast *S.cerevisiae* and the fruit fly *Drosophila melanogaster* [156].

1.3.2 Double Strand Break (DSB) Repair

DSBs arise from ionizing radiation or X-rays, free radicals, chemicals and during replication of single strand breaks. These potentially cytotoxic lesions are repaired by the DSB repair system, which consists of the homologous recombination and the non homologous end joining (NHEJ) repair modes.

During DSB repair via homologous recombination, missing DNA is restored using the intact homologous sequence provided by the sister chromatid. This pathway requires that the cell be diploid for the region that requires repair. In vertebrate organisms, this pathway is restricted to the late S and G2 phases of the cell cycle [157]. As the name implies, NHEJ is a process of direct DNA end-to-end fusion that does not require strand exchanges or the availability of homologous DNA [158, 159]. In vertebrate organisms, this pathway is active

throughout the cell cycle [157], therefore, NHEJ is thought to repair the majority of DSBs.

1.3.3 Excision Repair Pathways

Most base modifications in the cellular genome are repaired by the excision repair pathways: base excision repair (BER), nucleotide excision repair (NER) and mismatch repair (MMR). These pathways all involve recognition of DNA damage, removal of a single-stranded section of DNA that contains the damage, repair replication across the gap and restoration of DNA continuity by ligation of the repaired strand [155]. The only substantial difference between the three excision repair pathways is that MMR uses exonucleases to excise a long (up to a 1,000 bp [160]), single-stranded region containing the damage, starting from one incision in the damaged strand. By contrast, BER and NER use endonucleases to incise the strand very near to, or at the site of abasic sites or damage for excision of the damaged fragment [155].

The MMR pathway corrects mismatches in DNA that arise primarily from replicative errors and recombination between heterologous regions of DNA [161, 162]. Approximately 15% of colorectal cancers develop through mechanisms whereby this form of DNA maintenance is lost, leading to 100-1,000-fold increases in error rates during replication [163].

The high prevalence of endogenous DNA damage corresponds with highly efficient repair of such damage, necessary for cell survival [164, 165]. BER is initiated by a class of DNA repair enzymes known as DNA glycosylases [166]. These enzymes recognise only a particular type of base damage, a particular inappropriate base, or a particular mispairing. However, some of these

enzymes recognise more than one type of damage base. DNA glycosylases are either monofunctional or bifunctional. Monofunctional glycosylases remove the mispaired base only, leaving an intact *apurinic, apyrimidinic* or *abasic* (AP) site. Bifunctional glycosylases, in addition to the glycosylase activity have a lyase acivity that cleaves DNA 3' of the abasic site. AP sites can also result from depurination or depyrimidination of DNA following spontaneous hydrolysis of N-glycosyl bonds. The repair of AP sites requires further enzymatic processing and is initiated by a second class of BER enzymes called AP endonucleases, which specifically recognise these sites in duplex DNA. AP endonucleases produce incisions or nicks in duplex DNA by hydrolysis of the phosphodiester bond immediately 5' to the AP site. Hydrolytic cleavage of the phosphodiester bond generates a 5' terminal deoxyribose-phosphate residue that is removed by yet another class of enzymes, including exonucleases as well as enzymes with deoxyribophosphodiesterase (dRPase) activity. The sequential action of these three classes of enzymes results in the generation of single nucleotide gap in the DNA during BER. The missing nucleotides are replaced by DNA synthesis and covalently joined to the parental DNA by DNA ligation.

NER consists of a ubiquitous "cut and patch" mechanism, in which the damage is excised as part of a short oligonucleotide. This is in contrast to BER, in which the chemically modified moiety is excised in base form. The NER is responsible for removal of large helix-distorting DNA adducts induced by UV light and bulky electrophilic chemicals in mammalian cells. This is highlighted by the fact that patients with the rare disease xeroderma pigmentosum (XP), who are defective in NER and have a compromised ability

to repair UV induced damage, have a 2000-fold increased incidence of sunlight-induced skin cancer [167].

1.3.4 Damage Tolerance

Despite the robust excision repair systems that remove DNA lesions before replication, cells also possess efficient strategies for tolerating lesions persisting at the replication fork during DNA synthesis. A group of mechanistically different processes, which are collectively referred to as damage tolerance strategies, can counteract this situation and facilitate replication on damaged templates. Two basic strategies of damage tolerance can be distinguished: i) damage avoidance (DA) or gap-filling homologous recombination repair (GFRR) and ii) translesion synthesis (TLS).

In DA the replication machinery avoids directly copying the damaged template. Two models for damage avoidance have been proposed (Figure 1.6).
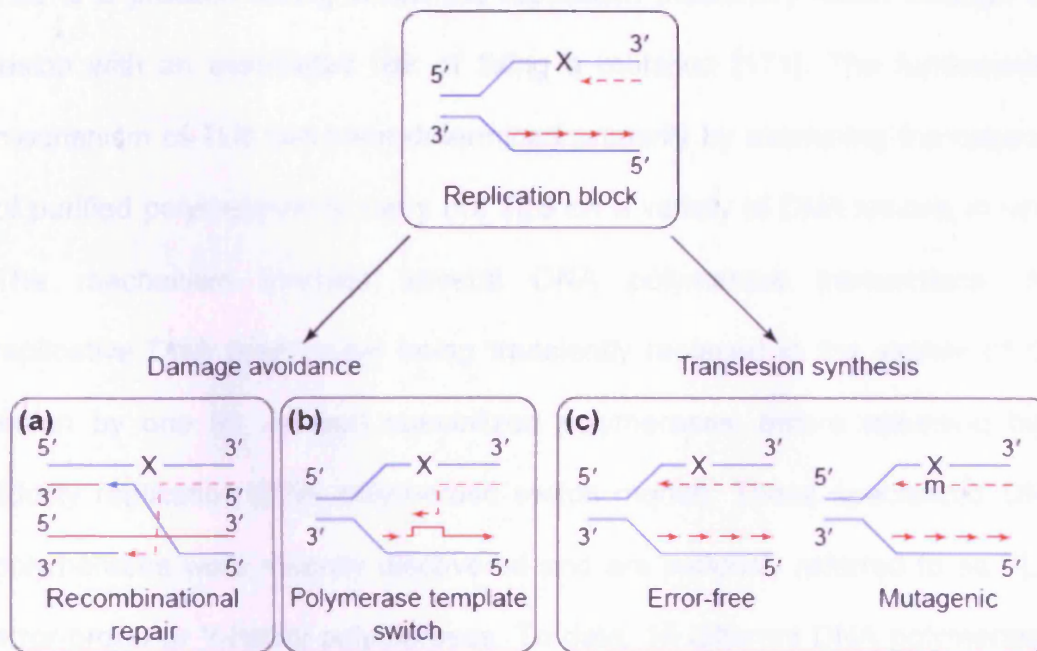
Figure 1.6 Damage tolerance strategies. Proposed DA mechanisms comprise (a) recombinational repair and (b) DNA polymerase strand-switching and (c) translesion synthesis [168].

The first is *recombinational repair*, in which a gap in the newly synthesized daughter strand created by dissociation of the replicating polymerase is repaired by a recombination mechanism involving the sister chromatid [168]. In the second model, *DNA polymerase template switching*, the DNA polymerase switches temporarily from the damaged parental template to the undamaged newly synthesized strand of the sister chromatid before returning to the parental template downstream of the lesion [169]. Both DA models are essentially error-free. Although our knowledge about the specific molecular mechanisms underlying these processes is elementary, sufficient evidence exists to suggest that DA is the major damage tolerance strategy in *E. coli* and *S. cerevisiae* [170].

TLS is a process during which the replication machinery reads through the lesion with an associated risk of fixing a mutation [171]. The fundamental mechanism of TLS has been determined primarily by examining the capacity of purified polymerases to carry out TLS on a variety of DNA lesions *in vitro*. The mechanism involves several DNA polymerase transactions, the replicative DNA polymerase being transiently replaced in the vicinity of the lesion by one (or several) specialized polymerases, before resuming high fidelity replication (DNA polymerase switch model). These specialized DNA polymerases were recently discovered and are variously referred to as TLS, error-prone or Y-family polymerases. To date, 16 different DNA polymerases have been identified in eukaryotes, of which 12 appear to have specialized functions in DNA repair or chromosome stability [172]. The four founding members of the Y-superfamily of DNA polymerases are pol η, pol ι, pol κ and pol Rev1 [173], even though Y family translesion polymerases are continuously being identified. Structurally, Y family polymerases have a much more open catalytic site that can accommodate damaged bases and/or tolerate mispairing of the incoming nucleotide [173]. However, the same properties that allow lesion bypass can also result in reduced stringency. As a result some of these polymerases have a much lower fidelity on undamaged templates [173]. In fact, when copying undamaged DNA they manifest error rates two to four orders of magnitude greater than those of replicative polymerases [174]. As the inserted base might be incorrect, TLS is a potentially mutagenic event. In fact, TLS is responsible for most mutations caused by DNA damaging agents, such as sunlight and a variety of chemical carcinogens [175, 176]. TLS therefore permits the cell to "tolerate" the presence of DNA damage as a means of avoiding potentially fatal replicative

risk, at the expense of replicational accuracy. Tolerance of DNA damage via TLS may be responsible for a significant fraction of damage-induced mutations linked to cancer and other genetic diseases [177].

1.4 MUTAGENESIS

Understanding the origin of mutations is central to understanding the aetiology of cancer. Not only is mutation linked to cancer and birth defects, but is also implicated in a number of other human diseases ranging from atherosclerosis to diabetes [88]. Therefore, DNA mutations are directly implicated in human health. The process by which mutations are generated is referred to as mutagenesis. It is important to note that damage to DNA itself is not a mutagenic event. DNA replication and subsequent cell division is necessary to convert chemical damage to an inheritable change in DNA that we call a mutation. Thus, proliferation is a vital factor in the formation of mutations and in the expansion of clones of cells bearing these mutations [178].

1.4.1 Mutational spectra

In general, a mutation spectrum is a distribution of mutation frequencies along nucleotide sequences and is compiled by the analysis of a large number of mutant target sequences [179]. Mutations in target sequences are revealed by specific experimental test-systems or, in cases of disease causing genes, by clinical studies. A mutation spectrum can include substitutions, deletions, insertions, inversions, transpositions, or other complex mutations. Mutation frequencies vary along a nucleotide sequence and positions with an exceptionally high mutation frequency are termed mutational "hotspots" [180-

182]. A mutation hotspot should be assignable as a hotspot with $\geq$ 0.95 probability [179]. Mutational hotspots can provide molecular evidence of the importance of a specific environmental carcinogen in the development of a particular tumour. For example, there is a strong correlation between mutational hotspots in the *p53* gene of lung tumours from smokers and preferential sites of adduct formation by PAHs found in tobacco smoke [183]. Thus, the mutational spectrum found in certain genes may, in some instances, serve as a fingerprint of the causative agent [184]. The basis of the uneven mutation frequency across a gene lies in two categories of factors: structural (e.g. DNA sequence context effects and adduct structure) and biological (e.g. differences in DNA polymerases and cell type) [185].

1.4.2 Factors affecting the mutagenic potential of a genotoxic chemical

1.4.2.1 Effect of DNA sequence context

DNA sequence context can affect both base substitution and frameshift mutational pathways. Regarding base substitution mutagenesis, the most extensive studies have been carried out with a variety of adducts formed by polycyclic aromatic hydrocarbons [186-190]. Sequence context also affects the mutagenic specificity of the PhIP-C8-dG adduct and the mutation frequency (MF) of O^6-meG [191]. There are several DNA context features that can influence DNA damage and repair such as homonucleotide runs, specific mutable motifs, potential zDNA, direct and inverted repeats and minisatellites [192, 193]. At this time it is not entirely clear as to how sequence in general can affect mutation frequency or over what distance sequence

determinants might be active [194]. In most cases, only local nucleotide sequence context was studied. In many cases, mutational hotspots emerge due to neighbouring nucleotides [195, 196]. However, several studies indicate that nucleotides that are distal, as well as those that are proximal to the adduct, are capable of influencing both the mutation frequency and the distribution of base substitution mutations [194, 197-200]. The contribution of sequence context has been extensively studied in the case of frameshift mutagenesis. For example, it was shown that the C8-dG adduct of the bulky carcinogen 2-acetylaminofluorene (2AAF-C8-dG) induces predominantly -1 frameshifts when located in dG runs, but -2 frameshifts when located at G2 in $NarI$-like sites (5'-$N_aG_1CG_2cN_b$) [201-203]. Notably, MF can vary dramatically (by 30- to 50-fold) depending on the base at N_b, while N_a has little influence [204].

1.4.2.2 Adduct structure

Different mutagens/carcinogens induce different patterns of mutations, implying that the structure of different adducts must result in the induction of different mutations. This conclusion was substantiated by the finding that relatively subtle changes in adduct structure can lead to significant changes in the patterns of mutagenesis. Perhaps the first evidence for this phenomenon was obtained by comparing the mutagenesis of 2-aminofluorene (2AF) and 2-acetylaminofluorene (2AAF) [205, 206]. While 2AF induces predominantly base substitution mutations, the presence of the acetyl group results in 2AAF principally inducing frameshift mutations. It is clear that this difference is due to mutational differences caused by the dominant adducts, 2AF-C8-dG and 2AAF-C8-dG [202, 207]. This difference is likely to be due to the bulky acetyl

group in 2AAF-C8-dG causing steric interference with the sugar attached to the modified base, thereby leading to a greater tendency to adopt the syn conformation compared to 2AF-C8-dG [208]. Other examples involve situations where an adduct can exist in different stereochemistries, changes of which can affect the pattern of mutagenesis [186, 187, 209].

1.4.2.3 Biological factors

The complex nature of adduct mutagenesis can be further influenced by biological factors. For example, different DNA polymerases can process adducts differently, as was demonstrated in a number of in vitro studies [68, 210-212]. The situation seems to be even more complex in vivo, as the discovery of TLS polymerases implied that there is significant potential for complexity in how a cell might cope with the bypass of DNA damage [213]. Furthermore, the SOS damage response in E.coli is generally believed to enhance survival from DNA damage, although with increased mutagenesis [156]. Another important factor could be the rate of DNA repair. DNA repair rates vary for transcribed and non-transcribed strands of the same gene [214-217]. Finally, it is likely that other high-level features of gene and chromatin structure also have influence on the mutagenic processing and repair of DNA adducts [192].

1.4.2.4 Mechanisms of mutagenesis

A good mechanistic model for frameshift hotspots is the "Streisinger slippage model" which is thought to involve a conformational change to a Streisinger slippage intermediate, whereby the dC incorporated opposite the adduct pairs with a downstream dG moiety in the template, followed by additional DNA

38

synthesis. The notion that frameshift mutagenesis involves a slipped intermediate is also supported by physical studies (thermal stability measurements, chemical probes and by NMR) [218, 219] and predictions can be made with some confidence as to where mutations might occur. However, in the case of base substitution hotspots, it is difficult to predict their location in a gene for a given mutagen, as little is known about the mechanisms by which mutagens induce base substitution mutations [220].

Possible mechanisms of substitution mutagenesis include mis-informational mechanisms, such as pathways whereby the carcinogen moiety of an adduct might cause the base to tautomerize, ionize, rotate or wobble, as well as non-informational mechanisms [221]. A lesion is classified as "non-informational" if it cannot base pair via hydrogen bonding [185]. Some adducts principally induce a single kind of mutation, while others can induce many different types of mutations [185]. Although the mechanism by which an adduct can induce different kinds of mutations is still a matter of debate, it is possible that a single DNA adduct may adopt several different conformations in DNA and each may potentially induce a different type of mutation [185]. Physical studies, using techniques such as 3D-NMR of carcinogen DNA adducts have yielded considerable information about the structures and conformations adopted in DNA. These studies, however, provide limited insight into what is actually occurring mechanistically during mutagenesis [185].

1.4.3 Experimental Mutagenesis Systems

In environmental toxicology, mutagenesis assay systems have been developed to measure potency (the strength of the mutagen) and mutagenic

specificity (the types of mutations caused by the mutagen). In addition, systems have been developed for the investigation of mutational mechanisms. The mutagenic and carcinogenic potential of many chemicals can be confirmed by animal bioassays and by any number of *in vitro* assays of mutagenic activity (e.g. the Ames assay). Predictions of the carcinogenic potential of a much larger number of environmental agents for which epidemiological data are lacking have been made by extrapolation from animal bioassays and mutagenesis data [222]. Nevertheless, despite the availability of large amounts of such data, it is usually not possible to identify the specific environmental exposures responsible for the majority of human cancers. This is partly because of the small number of confirmed human carcinogens and partly due to the large variety and mixtures of agents to which individuals are exposed. In order to provide the missing link between exposure and cancer, investigators have sought to identify specific signatures of environmental agents that can be identified in the exposed population. One potential signature is the mutational specificity of genotoxins.

The majority of assay systems employ marker genes, the mutagenesis of which can be scored by gain or loss of function. Analysis of genes, such as *p53*, in the tumour itself cannot alone provide much information about the carcinogenic exposure because mutations in the tumour are selected and clonally expanded during tumour development. Also, tumour cells appear to exhibit elevated levels of spontaneous mutagenesis. For application in human populations, ideally a mutagenesis target should be small, easily accessible, easily selected and easily sequenced. Phenotypic selection restricts the mutation spectrum to detectable nucleotides in which a mutation leads to phenotypic changes. Forward mutagenesis assays can detect various

changes that lead to an observable loss of activity of a gene product. Base substitutions, frameshifts, deletions and insertions can be monitored at many sites, permitting the role of DNA sequence context on the mutagenic outcome to be ascertained. In contrast, in reverse mutation assays, an activity that has been lost due to a mutant allele is regained, either by back mutation to the original wild-type sequence or by other mutations that restore the selected phenotype (pseudo wild-type). Generally, forward assays can score a wider variety of mutations than reverse assays.

A widely used reversion system is the Ames *Salmonella* test. This system is based on the reversion of histidine auxotrophs of *Salmonella typhimurium*. Several different *S. typhimurium* strains, together with the agent to be tested, are plated on minimum medium plates containing a limited amount of histidine. This amount of histidine allows the bacteria to undergo several generations of growth so that DNA lesions can be processed into mutations and the mutant phenotype can be expressed. After the histidine has been exhausted, His$^+$ revertants grow as distinct colonies, each one representing an independent mutational event [37], while non-mutated bacteria do not grow in the absence of histidine. The Ames test is a very effective tool for determining the relative mutagenicity of compounds and has a range of 5 orders of magnitude. Despite the power of this system and the ease of use, the Ames test and similar reversion assay systems cannot be used to develop mutation spectra for compounds of interest.

To determine the nature of mutational changes caused by a particular mutagen and also the site specificity of those changes, a large and unbiased collection of mutations is required. In general, the number of mutational

41

events that can cause a given mutant allele to revert is quite limited, so that all reversion systems are biased for the detection of a few specific events at a few specific sites. In principle, a much more unbiased collection of mutants can be obtained by using a forward mutation system [37]. A forward mutation spectrum in a reasonably large RNA or protein-coding gene usually possesses many (more than 50) detectable sites [223]. Examples of reporter genes that have been used as targets in forward mutational systems are shown in Table 1.2.

Gene	Function encoded
lacI+	Repressor of lac operon
lacZ+	β-Galactosidase
lacZ'	α-Complementing fragment of β-Galactosidase
λcI+	λ repressor
galK+	Galactose kinase
supF	Tyrosine suppressor tRNA
tet+	Tetracycline resistance
HPRT	Hypoxanthine-guanine phosphoribosyltransferase
APRT	Adenine phosphoribosyltransferase
gpt+	Guanine-hypoxanthine phosphoribosyltransferase
HSVαtk+	HSV thymidine kinase

Table 1.2 Examples of genes used as targets in forward mutational systems [37].

Reporter genes have been utilised for the study of mutagenesis in three main approaches; (i) as endogenous genes, (ii) in vector-based systems and (iii) in transgenic animals.

The most widely employed systems for studying mutational specificity in human cells are the hypoxanthine phosphoribosyl transferase (HPRT) system [224] and the supF [225] system. In the HPRT assay, this endogenous human gene is used as the mutagenesis target and mutations can be analysed either

in lymphoblasts from exposed individuals or in culture cells treated with the agent of interest.

1.4.3.1 Endogenous Genes

Naturally occurring chromosomal reporter genes have many advantages for mutagenesis studies. They have a native chromatin structure and are present in all cell types. Forward mutation assays have been developed in which the loss of activity of the target gene by mutation confers a selective growth advantage on the cell in the presence of an appropriate drug. There is a requirement for the gene to be haploid, and this is realised by using X chromosome-linked genes such as HPRT or cells hemizygous for the marker gene such as *APRT* or *DHFR*. The *HPRT* gene has been the most actively studied locus in investigations of mutational agents. It controls the enzyme hypoxanthine-guanine phosphoribosyltransferase that plays a role in the purine salvage pathway. This enzyme also phosphoribosylates purine analogues such as 6-thioguanine (6-TG), 6-mercaptopurine and 8-azaguanine (8-AG) - a necessary step for their cytotoxicity [226]. Resistance to these analogues provides a highly efficient selective system for *HPRT* mutant cells, allowing them to grow while wild-type cells are killed [227].

There are two principal limitations associated with the use of endogenous markers. Induction of mutations with mutagens necessarily involves treatment of cells as well, with the triggering of an associated stress response [228]. Therefore, all mutagenesis occurs in the context of a cell responding to stress and that can influence the mutagenic outcome [229, 230]. Another limitation of the system is in the large size of the target gene; the gene in genomic DNA is

approximately 44kb in length [231] and the coding region is 654bp [232]. A large group of mutants would have to be screened to allow meaningful analysis and generation of mutation spectra. Data from experiments with the *HPRT* locus have been reviewed by Cariello and Skopek [233].

1.4.3.2 Shuttle vectors

Until the 1980s mutational analysis systems were used to analyse mutagenesis in prokaryotes and lower eukaryotes but then it became feasible to develop systems that would allow mutagenesis studies in mammalian cells. These systems, known as *shuttle vectors*, consist of a target gene that is used to monitor mutagenesis and additional sequences that permit replication and selection in either mammalian cells or bacteria. There are three categories of vector-based reporter systems for mutation detection and analysis. They are distinguished by the extra- or intrachromosomal location of the reporter gene and by the technology used to identify and recover mutations.

1.4.3.2.1 Extrachromosomal Transient Vectors

Transient extrachromomal vectors replicate to high copy number (10^4 to 10^5 copies per cell) in host cells over a few days' period. DNA is recovered from the mammalian cells and then shuttled back into the appropriate indicator bacteria. The sequences permitting replication in mammalian cells are usually derived from viruses, such as SV40 and polyoma, while the sequences permitting replication in bacterial cells are usually derived from a plasmid. The *lacI* and *supF* genes were the first reporter genes to be used extensively [225, 234], with the latter gene currently in widespread use. One of the advantages of these vectors is that plasmid can be treated with mutagen under controlled

conditions *in vitro* prior to transfection into mammalian cells. Transient shuttle vectors are also useful for examining mutagenesis in various mammalian cell lines with repair deficiencies [235-237]. However, they also introduce the possibility that replication of the extrachromosomal DNA differs in some respect from that of genomic DNA and that this will influence the nature of the mutagenesis that occurs [37].

1.4.3.2.2 Extrachromosomal Stable Episomal Vectors

These vectors are maintained as stable episomes in cells in culture; i.e., the vector replicates autonomously rather than becoming integrated into the cell's genome. Those in current use rely on information from Epstein-Barr virus and bovine papillomavirus replication systems. These vectors are present in approximately 10 to 100 copies per cell nucleus and replicate synchronously with the host cell. The first of these vectors carried thymidine kinase as a marker gene [238], while a separate group constructed a similar version with the lacI gene as a marker [239]. Generally, cells carrying the vector are treated with a mutagen of choice and several alkylators, such as *N*-methyl-*N*-nitrosourea and *N*-ethyl-*N*-nitrosourea have been tested [240, 241]. A severe limitation of episomal vectors is their limited sensitivity because of the low copy number of plasmid DNA in mammalian cells.

1.4.3.3 Mutagenesis Reporters in Transgenic Animals

Transgenic mouse mutation assays afford the opportunity to evaluate mutation frequency and specificity in individual tissues of the whole animal. The most widely used system is the Big Blue that utilises the bacteriophage lambda-based *lacI* gene and was developed by Stratagene [242]. Mutational

data has been generated for a variety of mutagens in different tissues and different routes of administration. A similar mouse system carrying the supF gene has also been developed [243]. The advantage of both systems is that mutational data generated by them can be compared directly with the large databases available from studies in *E.coli* or mammalian cells.

1.4.4 The supF assay

In the *supF* forward mutation assay, developed by Seidman and co-workers [244], the *supF* target gene is located in a shuttle vector plasmid that can replicate in both mammalian and *E.coli* cells. One of the advantages of these vectors is that the plasmid can be treated under control conditions *in vitro* prior to transfection into mammalian cells. It is worth mentioning that the vector is preferentially treated with an active form of the chemical *in vitro*, so there must be an appropriate reactive metabolite available. The supF system has been used to analyse the mutational specificity of over 40 different agents [245], including UV, BPDE, and aflatoxin. In each case the mutational specificity is presented as a "mutation spectrum", that is the position and type of mutation in the *supF* gene sequence in each of a collection of independent mutants. In some cases, the mutation spectrum has been compared to the distribution of initial damage in the *supF* gene [246]. In other cases, the spectrum for the same agent has been compared in the presence or absence of specific DNA repair pathways [247-249].
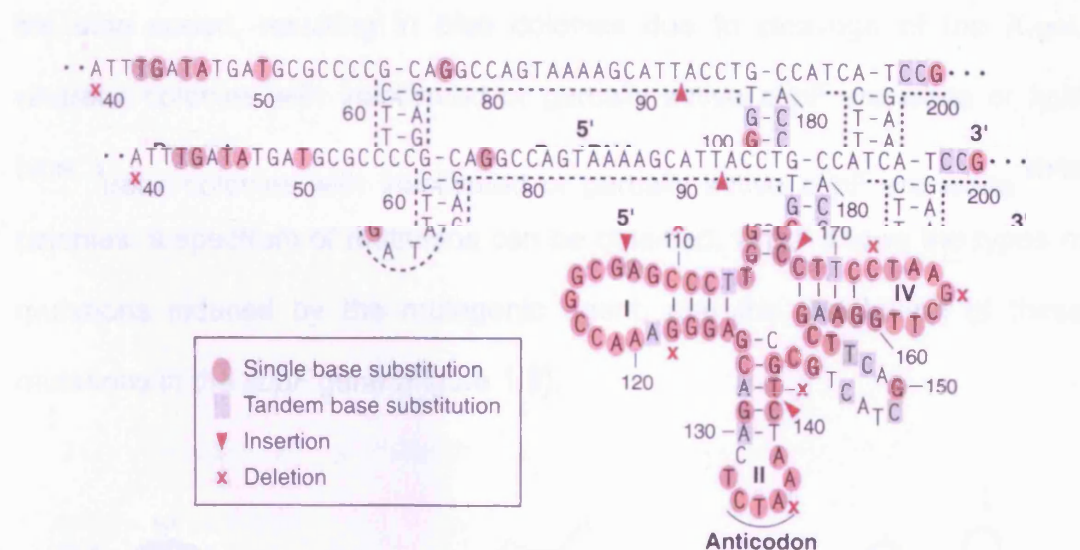
Figure 1.7 Hypothetical secondary structure of a single-stranded DNA containing the *supF* tRNA gene sequence and showing the location of mutations that inactivate *supF* function. (*Adapted from Kraemer and Seidman [244]*)

The *supF* gene encodes an amber suppressor tRNA (Figure 1.7). The presence of the stop codon within a coding sequence will lead to the production of a truncated, inactive protein. Active *supF* allows for the incorporation of tyrosine at the amber chain termination codons (UAG). Hence, the *supF* tRNA can suppress an amber mutation and allow production of a full length, active protein. After the plasmid is replicated in mammalian cells and recovered, the most common method for identifying *supF* mutants involves introduction of the shuttle vector into a bacterial indicator strain of *E.coli* (MBM7070) carrying an amber mutation in the *lacZ* gene, which encodes for β-galactosidase (the enzyme responsible for cleaving the disaccharide lactose to glucose and galactose) [250]. Therefore, this strain of *E.coli* cannot convert lactose to glucose and galactose. When the transformed bacteria are grown on plates containing the lactose analogue X-gal, the presence of plasmid with non-mutated, functional *supF* allows read-through of

47

the stop codon, resulting in blue colonies due to cleavage of the X-gal, whereas colonies with inactivated or partially active *supF* are white or light blue. When mutated *supF* gene DNA is sequenced from a number of white colonies, a spectrum of mutations can be obtained, which shows the types of mutations induced by the mutagenic agent, and the distribution of these mutations in the *supF* gene (Figure 1.8).
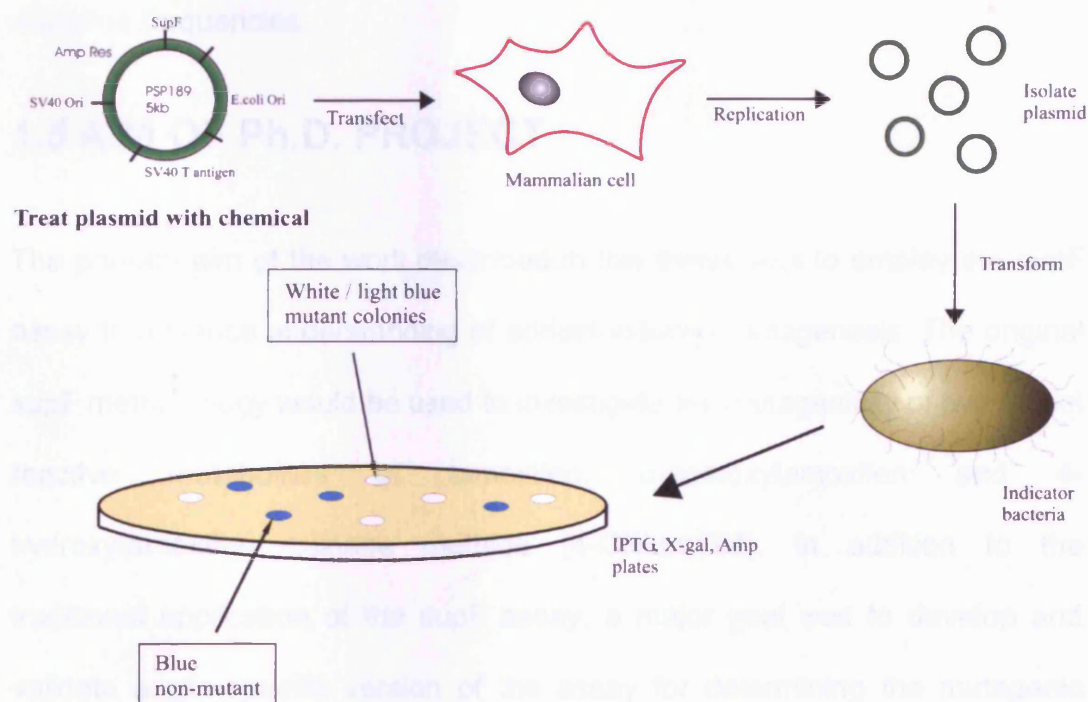
Figure 1.8 Outline of the *supF* assay method.

One of the advantages of the *supF* reporter gene is its extreme sensitivity to mutagenic inactivation. Because tRNAs require a precise structure to be functional, tRNA genes are very sensitive to the substitution mutations, much more so than protein coding genes. In the latter, sequence changes that do not change the amino acid sequence, or those amino acid substitutions that do not affect the function of the protein, will not be scored. In the *supF* gene

all six possible base substitutions, insertions and deletions can be detected [244]. Small insertions or deletions have also been found that inactivate *supF* function. The tRNA sequence itself or its promoter can be mutated. Therefore, the selection bias associated with protein coding mutagenesis marker genes is largely avoided. Besides providing a reliable target for studying the mutational specificity of a compound, the supF assay can give an indication of relative mutagenicity of different chemicals, by enabling comparisons of mutation frequencies.

1.5 AIM OF Ph.D. PROJECT

The primary aim of the work described in this thesis was to employ the supF assay to advance understanding of adduct-induced mutagenesis. The original supF methodology would be used to investigate the mutagenicity of two model reactive metabolites of tamoxifen, α-acetoxytamoxifen and 4-hydroxytamoxifen quinone methide (4-OHtamQM). In addition to the traditional application of the supF assay, a major goal was to develop and validate a site-specific version of the assay for determining the mutagenic potential of individual DNA adducts, in different sequence contexts within a double stranded plasmid, in both *E.coli* and human cells. The assay should be novel in that it should permit positioning of the DNA adducts anywhere on the *supF* gene, including studies of multiple lesions, and allow easy identification of the types of mutations induced by these adducts. The assay should also provide an indication of the relative mutagenicity for a given DNA adduct. The assay was validated using O^6-MeG as an easily obtainable model DNA adduct. However, the methods can be adapted to the study of any stable adduct that can be synthesized and incorporated into the plasmid.

CHAPTER 2

Mutagenicity of tamoxifen DNA adducts in human endometrial cells and *in silico* prediction of *p53* mutation hotspots

2.1 INTRODUCTION

2.1.1 Tamoxifen and breast cancer

Breast cancer (BC) is the most common form of malignancy in women throughout the world. In 2005, breast cancer caused 502,000 deaths (7% of cancer deaths; almost 1% of all deaths) worldwide [251]. Each year almost 44,100 cases of breast cancer are diagnosed in the UK, of which 99.2% are female cases and around 0.8% are male. Breast cancer causes more than 12,500 deaths each year in the UK [252]. If current BC rates stay constant, a female born today has 1 in 8 chance of developing BC sometime during her life. Despite the increase in its incidence, mortality related to BC is decreasing in many western countries. Data from the USA and UK show that between 1987 and 1997 there was a substantial reduction in breast cancer mortality, both in middle age (approximately 25%) and to a lesser extent in old age [253].This is primarily due to increased awareness and x-ray mammography screening, as well as multidisciplinary treatment. However, more than 60% of BC patients will ultimately die from distant disease recurrence. The mainstay of BC treatment is surgery when the tumour is localized, with possible adjuvant hormonal therapy, chemotherapy and/or radiotherapy. More recently small molecules and monoclonal antibodies that target growth factor receptors and their downstream signaling pathways have entered in clinical trials, with the aim to minimize toxicity related side-effects. Patients with oestrogen receptor positive tumors (approximately 60-65% of primary breast cancers [254]) will typically receive hormonal therapy after chemotherapy is completed to deprive cancer cells of oestrogen that causes cell proliferation. Typical

hormonal treatments include selective oestrogen receptor modulators (SERMs) (such as tamoxifen and raloxifene), aromatase inhibitors (e.g. Arimidex) and pituitary downregulators (e.g. Zoladex).

Tamoxifen [*trans*- (*Z*)-1-[4-[2-(dimethylamino) ethoxy] phenyl]-1,2-diphenyl-1-butene] (ICI 46 474) is a non-steroidal SERM that has become the treatment of choice for women diagnosed with all stages of hormone-responsive breast cancer (both oestrogen receptor (ER) and/or progesterone receptor positive) [255]. It is also the most widely prescribed cancer chemotherapeutic in the world [256]. Since its introduction in 1971 it has been used successfully to treat many millions of women [257]. Treatment results in an extension of disease-free survival [258] and a decrease in recurrence rates of breast cancer by 47% [257]. It was also found to have a chemopreventative role in women at high risk of developing BC, as a 49% decrease in the incidence of invasive breast cancer was observed in the tamoxifen arm versus placebo in the NSABP P-1 prevention trial, sponsored by the National Cancer Institute (NCI) [259]. Based on these findings, tamoxifen was approved by the United States Food and Drug Administration (FDA) for use as a chemopreventative agent in women at high-risk for breast cancer.

Paradoxically, cancer chemotherapeutic agents themselves may be carcinogenic. Several epidemiological as well as randomized, prospective trials have established an increased risk of endometrial cancer in association with prolonged tamoxifen treatment, with relative risks ranging from 2.53 to 7.5 [259-264]. This led to the International Agency for Research on Cancer (IARC) classifying tamoxifen as a group 1 (human) carcinogen [265]. The

underlying mechanism of carcinogenicity for tamoxifen remains elusive and many properties of the drug have been implicated, including its genotoxic, oestrogenic and epigenetic effects. Findings such as these raised concerns about the widespread preventative use of tamoxifen by healthy women. They have also led to significant debate over whether its carcinogenic effects are due to its oestrogenic activity through the ER and/or its genotoxic potential.

2.1.2 Metabolism of tamoxifen

For environmental carcinogens, the most difficult parameters to assess are exposure and internal dose. In contrast, these factors are usually well characterized for therapeutic agents. The metabolism and metabolic activation of tamoxifen has been widely studied and it is now well established that several Phase I and Phase II enzymes are involved in the process.

Phase I metabolism of tamoxifen has been demonstrated to occur at several positions on the molecule [266-268]. The principal sites of Phase I metabolism *in vivo,* are the nitrogen atom of the side chain (N-oxidation and demethylation) and the 4-position (hydroxylation). These pathways give rise to tamoxifen *N*-oxide (TAM *N*-oxide), *N*-desmethyltamoxifen (*N*-desmethylTAM) and 4-hydroxytamoxifen (4-OHTAM) metabolites. Another minor metabolic route includes the hydroxylation of the α-position of the ethyl side chain to form α-hydroxytamoxifen (α-OHTAM). Although α-OHTAM is a minor metabolite, accounting for ~0.1% of the administered tamoxifen dose in rats [269], it is believed to be responsible for the major tamoxifen DNA adduct seen in rats, mice and humans [270]. Notably, when α-OHTAM was administered to rats, it was found to bind to DNA at up to 50-fold higher levels

than an equal concentration of tamoxifen [271]. Results from *in vitro* studies suggest that tamoxifen is initially activated to α-OHTAM by cytochrome P450 3A4 in humans [267, 272], but the involvement of other isoforms has also been proposed [273, 274]. In addition, the ability of minor and/or extrahepatic P450s to catalyze tamoxifen α-hydroxylation remains to be assessed fully.

Phase II metabolic activation of α-OHTAM, with and without concomitant *N*-demethylation, is mediated by a sulphotransferase, specifically by an isoform of hydroxysteroid sulphotransferase (ST2A2 in rat [275], SULT2A1 in human [276]) to generate reactive esters. A subsequent loss of sulphate from these esters results in short-lived reactive carbocation intermediates which react preferentially with exocyclic amino groups of guanine and adenine [270, 277]. Moreover, activation of α-OHTAM metabolites by acetyltransferases has been proposed [272, 278], but not substantiated [279].

Using α-acetoxytamoxifen or the sulphate ester of α-OHTAM as model reactive metabolites, the major reaction products between activated α-OHTAM and dG residues in DNA *in vitro* are α-(deoxyguanosin-N^2–yl)tamoxifen (dG-N^2-TAM) adducts in which the α-carbon of tamoxifen is covalently linked to the exocyclic amino group of dG [270, 277, 280] (Figure 2.1). Since the carbocation arising from these esters readily undergoes rotation about the central bond, both *cis* and *trans* forms are produced and once formed they are stable and do not interconvert. Additionally, as the α-carbon of tamoxifen is chiral, each geometric isomer exists as two diastereoisomers (fr-1 & fr-2 for the *trans* isomer and fr-3 & fr-4 for the *cis*). These adducts have been shown to co-elute with the major adduct peaks

formed in the livers of rats administered tamoxifen [277, 281]. A minor tamoxifen deoxyadenosine adduct has also been identified in which linkage is through the amino group [270].

It remains to be ascertained whether these metabolic pathways of bio-activation to a reactive electrophilic species are relevant to humans. Metabolism of tamoxifen in humans appears to be qualitatively similar to metabolism in rodents [282] and α-OHTAM is a detectable metabolite in the plasma of women on long-term tamoxifen therapy [283]. Although there have been some case reports of acute liver toxicity of tamoxifen [284], there have been no findings of increased liver cancer among tamoxifen-treated women [262, 285]. To explain this apparent resistance to the hepatocarcinogenic effects of tamoxifen in humans compared to rats, two main mechanisms have been proposed. First, there is evidence that α-OHTAM is a poorer substrate for sulphotransferase in humans than in rats [286]. Second, glucuronidation, which is likely to be a detoxification (inactivation) pathway, seems to predominate over sulphation in incubations of α-OHTAM with human liver microsomal fractions [287].

Figure 2.1 Formation of tamoxifen-DNA adducts in rat liver via α-hydroxylation of tamoxifen metabolites. Adapted from [279]

2.1.3 Carcinogenicity of tamoxifen in rodents

Tamoxifen induces hepatocellular carcinomas in rats [288-290] and numerous tamoxifen-DNA adducts have been detected in rat liver, formed in a dose-related manner [277, 291, 292]. The incidence of tumours in a lifetime exposure study [289] was similar in male and female rats and showed a dose-dependent increase over the range of 5-35 mg/kg/day. This is 10-50-fold higher than the dose used in humans (i.e., 0.8 mg/kg/day for a 50-kg woman). However, lifetime exposure is not necessary for the induction of liver tumours. Following exposure to dietary tamoxifen for only 3 months, after discontinuation of treatment, five out of 15 rats developed liver tumours after 20 months without any subsequent promotion, whereas 12 of 14 animals had liver tumours when phenobarbitone was used as a promoter [293]. It is also established that tamoxifen is not able to act as tumour initiator if administered as a single dose [293], although it can promote liver cancer initiated by *N*-nitrosodiethylamine [294].

In contrast to rats, mouse hepatocarcinogenicity did not occur either in the presence or absence of phenobarbital promotion [295]. Instead, tumours of the testis are induced in males, and of the ovaries in females [296]. Furthermore, when tamoxifen is administered neonatally, tumours develop in the uterus [297]. Finally, when administered transplacentally to mice, tamoxifen causes a high incidence of hyperplasia in the reproductive tract, and a lower incidence of tumours in offspring [298].

2.1.4 Carcinogenicity of tamoxifen in women receiving tamoxifen therapy

In recent years, results indicate not only an increased incidence of endometrial cancer in women taking tamoxifen but also increased mortality from the disease [299], implying that the tamoxifen-induced neoplasms are in some way pathologically different from spontaneous tumours, not associated with use of the drug. Other studies have raised the concern that tamoxifen-related endometrial cancers may behave more aggressively compared with those found in patients with unopposed estrogen [299-302] and that endometrial thickening is evident in a much larger percentage of women [303].

2.1.5 Genotoxicity of tamoxifen

dG-N^2-TAM and α-(N^2-deoxyguanosinyl)-*N*-desmethyltamoxifen (dG-N^2-*N*-desTAM) account for the major adducts in the liver of rats treated with tamoxifen [277, 291, 292, 304, 305] and have been detected in several tissues including the reproductive organs of monkeys treated with tamoxifen [306, 307]. Hepatic tamoxifen-DNA adduct levels were relatively low in tamoxifen-exposed mice compared to equivalently treated rats [295, 308]. In addition, α-(N^2- deoxyguanosinyl) tamoxifen *N*-oxide (dG-N^2-TAM *N*-oxide) was detected as a minor adduct in the liver of mice treated with tamoxifen. The *trans-* and *cis-* forms of dG-N^2-TAM *N*-oxide accounted for 7.2% and 0.7%, respectively, of the total tamoxifen DNA adducts observed [308]. Formation of these adducts may be due to high levels of N-oxidation by flavin-containing monooxygenase in the mouse liver [309]. Finally, α-(N^2-

deoxyguanosinyl)-4-hydroxytamoxifen (dG-N^2-4-OHTAM) is formed when 4-OHTAM quinone methide (4-OHtamQM) [310], produced by oxidation or by horseradish peroxidase activation of 4-OHTAM *in vitro*, reacts with dG *in vitro* [311]. However, this adduct was not found in the liver of rats treated with tamoxifen, α-OHTAM, or 4-OHTAM [281, 312], although the 4-OHTAM adduct potentially coeluted with an adduct seen only at trace levels in the liver of a tamoxifen treated rat.

While the formation of tamoxifen-DNA adducts in rat liver is well established, evidence for the formation of tamoxifen adducts in extrahepatic tissues is more controversial. In two studies, a low level of adduct formation in the kidney was reported [291, 313], although other studies have produced mostly negative results [302, 314]. Several studies in which the rat uterus was investigated reported negative findings [291, 312-316], whereas one study has reported low levels [317]. In contrast, a more sensitive accelerator mass spectrometry study found [^{14}C]-tamoxifen to be bound to DNA in liver, intestine, reproductive tract, spleen, lung and kidney of rats dosed orally with the compound [318].

The situation regarding tamoxifen-DNA adducts in humans is still controversial. Some research groups have detected tamoxifen-DNA adducts in endometrial tissues of breast cancer patients treated with tamoxifen, using ^{32}P-postlabeling analysis [319-321] and accelerator mass spectrometry [322] while other groups could not detect any adducts in endometrial tissues [323-325] and lymphocytes of tamoxifen-treated women [326, 327]. However, in a recent study [328], low level of tamoxifen-DNA adducts were detected in

lymphocytes from a small proportion of breast cancer patients receiving tamoxifen. *In vitro* experiments have also provided conflicting evidence concerning the ability of human endometrial cells to metabolically activate tamoxifen to DNA binding products [323, 329-332].

2.1.6 Estrogenic effects of tamoxifen

It is not clearly understood how SERMs, such as tamoxifen and raloxifene, can be antioestrogenic in some cells and oestrogenic in others. Current evidence suggests that oestrogen (E2) antagonism in the breast results from the ability of tamoxifen to inhibit co-activator binding and promote co-repressor binding to the ligand-binding domain, thus converting the ER into an active repressor in this tissue [333]. To date, more than 20 regulatory proteins have been identified that bind to ERα and ERβ and modulate their function, each acting either as a positive (co-activator) or negative transcriptional regulator (co-repressors) [334]. Recent studies [335] reported that in endometrial cells, tamoxifen increased the recruitment of the co-activator SRC-1 to the ERα target promoter, which enhanced its oestrogen-like actions in this tissue, while raloxifene, which lacks oestrogen agonist activity in the endometrium, had no effect. In recent years, alternative mechanisms of SERM action have been proposed. For example, it has been proposed that the SERM-receptor complex may activate transcription by tethering to promoters that do not contain the classical oestrogen response elements (EREs) through protein-protein interactions with other DNA-bound transcription factors [336, 337]. Another idea is that the ratio of ERα to ERβ is important for expressing the oestrogen-like effects of antioestrogens.

Interestingly, both tamoxifen and raloxifene function as pure antagonists when acting through ERβ on genes containing EREs but can function as partial agonists when acting on them through ERα [338]. Finally, there is a growing body of evidence to suggest that steroid hormones can initiate signalling from the exterior of the cell. Nongenomic effects have been reported for many classes of steroids including oestrogens [339].

2.1.7 Epigenetic effects of tamoxifen

It has been recently demonstrated that genotoxic carcinogens, in addition to exerting genotoxic effects, often cause epigenetic alterations [340]. Furthermore, it is becoming increasingly evident that these induced epigenetic changes may play a role in carcinogenesis [340-344]. However, it has not yet been established whether epigenetic changes induced by carcinogens play a causative role in carcinogenesis or are merely a consequence of the transformed state [340, 345]. A recent study has presented evidence that tamoxifen can induce epigenetic alterations in the livers of rats exposed to tamoxifen-containing diet. The epigenetic alterations include a substantial decrease in cytosine DNA methylation, and a decreased histone H4 lysine 20 trimethylation in liver tissue accompanied by altered activity and expression of maintenance and *de novo* DNA methyltransferases. These alterations were specific for liver tissue only and were unaffected in non-target tissues, such as pancreas, spleen and mammary glands.

2.1.8 Mutagenicity of tamoxifen-DNA adducts

The mutagenic potential of dG-N^2-TAM, one of the major tamoxifen DNA adducts formed in rat liver and detected in DNA from treated women, was established using site-specific mutagenesis in a single-strand vector propagated in simian kidney (COS-7) cells [346]. dG-N^2-TAM adducts promoted primarily G→T transversions, along with small numbers of G→A transitions (Table 2.1).

	dG-N^2-TAM			
Mutation	trans-1	trans-2	cis-1	cis-2
G → T	1.1%	9.6%	10.9%	12.3%
G → A	1.5	2.8	1.7	1.7
G → C	0.7	0	0.8	0
Total	3.3	12.4	13.4	14

Table 2.1 Mutagenic Potential of single dG-N^2-TAM adducts in mammalian cells. Adapted from [346].

The mutagenic specificities were similar to those observed in primer extension reactions catalyzed by mammalian DNA polymerases on dG-N^2-TAM-modified DNA templates [347] and in the liver DNA of *MacI* transgenic rats [348, 349]. In addition, adducts formed by α-acetoxytamoxifen (which yields the same dG-N^2-TAM adducts formed in rats after treatment with tamoxifen) preferentially induced GC→TA transversions in the pSP189 *supF* gene when replicated in human kidney cells [350] or GM00637 human fibroblasts [351]. Finally, the role of pol η and pol κ in the mutagenic processing of tamoxifen-DNA adducts has been recently investigated in an *in vitro* study [352]. With

pol η, a small amount of direct dAMP incorporation was observed along with deletions, indicating that G→T transversions and deletions were produced. Pol κ promoted small amounts of dTMP incorporation, indicating that G→A mutations occurred. These observations were also supported by steady-state kinetic studies.

Bulky DNA adducts, such as those formed by tamoxifen are typically removed by the NER system. Using reconstituted mammalian and human NER enzymes *in vitro*, it was shown that tamoxifen-DNA adducts are removed slowly from DNA [353]. The repair rate of hepatic tamoxifen-DNA adducts has been determined in rats using ^{32}P-postlabeling/PAGE and ^{32}P-postlabeling/HPLC analyses. The adduct half-life was approximately 25 days and among the four different diastereoisomers, the trans form (fr-1) was removed much more slowly than the other forms [354]. In addition, in a separate study [351], the mutation frequency resulting from replication of α-acetoxytamoxifen-induced damage in NER-deficient cells (XPA) was higher than that observed with NER-proficient cells, suggesting that NER plays a significant role in the removal of tamoxifen-DNA adducts.

2.1.9 The LwPy53 algorithm

Ideally, determining the association between a specific carcinogen and cancer type would involve knowing not just the types of mutation that would arise in a reporter gene but also the actual carcinogen-specific mutation spectrum in a biologically important gene such as *p53*. Such carcinogen-specific mutation distributions along a cancer gene in combination with knowledge of preferential adduct distribution and exposure would allow for greater certainty

when predicting the cause of tumourigenesis. The LwPy53 algorithm [355] ultilizes mutation data at the level of the dinucleotide from the *supF* gene to predict chemically induced hot-spots along the human *p53* gene. The algorithm incorporates additional fixed parameters to represent environmental factors that contribute to the overall shape of the mutation spectrum: absolute deviation from mean DNA curvature and nucleosome positioning. Both factors can be assumed to contribute to the regional accessibility of a mutagen to a DNA sequence and the rate of DNA repair of adducts. These factors are derived from information concerning predicted *p53* chromatin structure [356] and deviation from the predicted average *p53* DNA curvature.

The algorithm was recently used to predict the *p53* benzo(a)pyrene diolepoxide (BPDE) mutation distribution based on *supF* mutation data, and the resulting spectrum was remarkably similar to the *p53* mutation pattern observed in lung cancer of smokers [355]. The resulting predicted mutation distribution reveals strong mutation hotspots at codons 157, 248 and 273 that correlate with known BPDE adduct hotspots within *p53* [357]. The predicted spectra were closer to the *in vivo* mutation spectra of smokers with lung cancer when *supF* data generated using CpG methylated pSP189 plasmid DNA was used compared to data obtained using unmethylated plasmid, which is employed in the conventional assay. This finding lends further support to the idea that methylated CpG sites are a preferential target for certain carcinogens [358].

2.1.10 Aims

The aim of the work presented in this chapter is to investigate the mutagenicity of the tamoxifen-DNA adducts dG-N^2-TAM and dG-N^2-4-OHTAM. These adducts were formed using the model reactive intermediates α-acetoxytamoxifen and 4-hydroxytamoxifen quinone methide (4-OHtamQM) (Figure 2.2). The *supF* assay was employed to ascertain the mutations caused by these adducts in human endometrial cells. In order to utilise the LwPy53 algorithm the plasmid was methylated *in vitro* at CpG sites prior to treatment with α-acetoxytamoxifen or 4-OHtamQM. Since the plasmid is treated *in vitro*, aliquots of the treated DNA can be analysed for adduct quantification in parallel to the mutation assay; A previously developed HPLC-^{32}P-postlabelling method [281] was used for this purpose.

α-Acetoxytamoxifen

Deoxyguanosine

4-Hydroxytamoxifen quinone methide

α-(N^2-deoxyguanosinyl)-tamoxifen

α-(N^2-deoxyguanosinyl)-4-hydroxytamoxifen
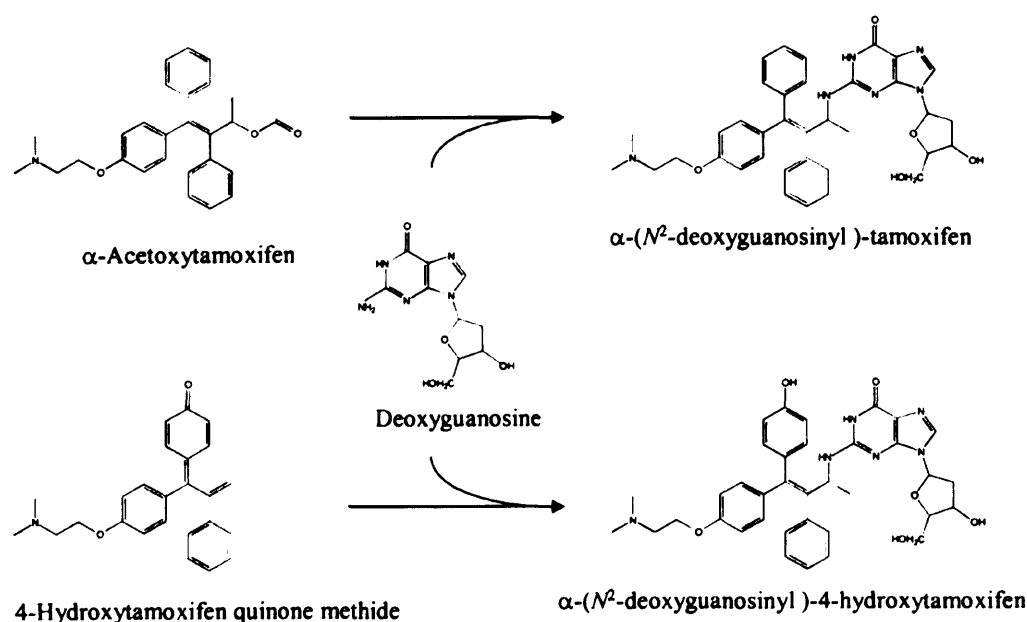
Figure 2.2 Reaction of α-acetoxytamoxifen and 4-hydroxytamoxifen quinone methide with deoxyguanosine yields the major dG-N^2-TAM and the minor dG-N^2-4-OHTAM adducts respectively. Adapted from [350].

2.2 MATERIALS AND METHODS

Materials

All chemicals were from Sigma-Aldrich, Poole, Dorset, UK unless otherwise stated.

2.2.1 Human cell lines

Human endometrial adenocarcinoma (Ishikawa) cells (oestrogen receptor negative) were cultured from cells provided by Dr Ian White (University of Leicester). Ishikawa cells were derived from a well differentiated human endometrial adenocarcinoma from a 39 year old woman that lacked tamoxifen treatment [359, 360]. Ishikawa cells were grown in DMEM/F12, phenol red-free (Invitrogen, Carlsbad, CA, USA) medium supplemented with 10% fetal calf serum (Life Technologies Ltd, Paisley, UK) and 2mM GlutaMAX at 37°C in 5% CO_2 in air.

2.2.2 Shuttle vector plasmid and bacterial strain

The plasmid pSP189 containing the *supF* gene and *E.coli* strain MBM7070 {F⁻ , lacZamCA7020, lacγ1, hsd R⁻, ara D139, Δ(araABC-leu) 7679, gal U, gal K, rps L, thi (derivative of MC1061) [361]} were gifts from Dr. M. Seidman (National Institute of Aging, NIH, Baltimore, MD, USA).

Methods

2.2.3 In vitro methylation of recombinant plasmid with CpG methyltransferase (M.SssI)

Dried pSP189 plasmid (50 µg) was methylated *in vitro* using 50 Units of CpG methyltransferase (NEB), 5 µl 32mM S-adenosylmethionine (SAM) (NEB) and 25 µl of NEBuffer 2 (10x) (NEB) in a 250 µl reaction which was incubated for 2 h at 37°C. Methylated pSP189 plasmid was purified using Sureclean (Bioline, Bath, UK) prior to reaction with α-acetoxytamoxifen or 4-OHtamQM.

2.2.4 Digestion of methylated and unmethylated (control) pSP189 plasmid with HpaII restriction endonuclease

In order to confirm successful methylation of the plasmid, aliquots (2 µg) of methylated or unmethylated pSP189 (control) were digested with 50 Units of HpaII (NEB) and 6 µl NEBuffer 1 (10x) in a 60 µl reaction for 1 h at 37°C. HpaII will not cut sites that have been methylated by SssI methylase.

2.2.5 Acetylation of α-hydroxytamoxifen

The *trans* isomer of α-hydroxytamoxifen was synthesised using the method described by Foster and co-workers [362] by Dr Robert Britton. *Trans*-α-Acetoxytamoxifen was prepared from *trans* α-hydroxytamoxifen using the published method [270]. Briefly, α-hydroxytamoxifen (100 mg, 260 µmol) was washed with 3×5 ml pyridine (followed by rotary evaporation each time). α-hydroxytamoxifen was redissolved in anhydrous pyridine (2 ml) and

anhydrous acetic anhydride (100 μl, 100 μmol) was added. After 18 h at 37°C

the reaction was stopped by addition of diethyl ether (50 ml) and washed with

water (3×30 ml). The organic layer was dried (Na_2SO_4) and concentrated

then the product, trans-α-acetoxytamoxifen, was separated from unreacted

starting materials by column chromatography (2% methanol in

dichloromethane). Mass analysis of the product revealed a protonated

molecular ion $(M+H)^+$ with m/z 430, confirming the molecular formula

$C_{28}H_{31}NO_3$ (429.230). Fractions containing purified α-acetoxytamoxifen were

combined and dried by rotary evaporation, then a known mass of the residue

was re-dissolved in ethanol to produce a stock solution for reaction with

pSP189 plasmid.

2.2.6   Treatment   of   pSP189   plasmid   with   α-acetoxytamoxifen

Aliquots of methylated pSP189 plasmid (200 μg in 200 μl tris-EDTA buffer, pH

8.0) were treated with varying doses of α-acetoxytamoxifen in 100 μl ethanol

(10 μM, 25 μM and 50 μM final concentrations) by Dr Keith McLuckie.

Samples were prepared in duplicate, along with a single ethanol treated

control (200 μg plasmid only) and incubated overnight at 37°C. Unreacted

compound was then extracted with water saturated ethyl acetate (4 × 400 μl),

using a needle and syringe to remove the organic layer. The DNA was

precipitated with 2 M sodium acetate (10 μl) and ice-cold ethanol (800 μl). The

DNA samples were pelleted by centrifugation using a bench top centrifuge

(Hettich, Germany) the supernatant was removed and pellets washed with

70% ice-cold ethanol, followed by 100% ice-cold ethanol. Pellets were re-

dissolved in 400 µl sterile tissue culture grade water. The yield of DNA was calculated from the optical density at 260 nm according to the following relationship:

DNA concentration = A_{260} × Dilution × 50 (constant for dsDNA)

The plasmid DNA was stored at -80°C until transfection into cells was performed.

2.2.7 Treatment of pSP189 plasmid with 4-OHtamQM

4-Hydroxytamoxifen (6.52 mg) was activated to its quinone methide by Dr Keith McLuckie, by stirring for 30 min with silver (II) oxide (42.92 mg) in 1.8 ml dry chloroform (Liehr, *et al.*, 1983). The reaction mixture was filtered and dried to a yellow-brown residue under nitrogen. Assuming 100% conversion, this was dissolved in 50:50 ethanol:acetonitrile, and appropriate volumes were added to 100 µg pSP189 plasmid in 500 µl H_2O. Ethanol:acetonitrile (50:50) was then added to adjust the total volume to 625 µl and produce final concentrations of 0, 50, 100 and 250 µM 4-OHtamQM. After incubation at 37°C for 18 h the unreacted 4-OHtamQM was extracted by 5 × 400 µl diethyl ether. Plasmid DNA was precipitated with 3M sodium acetate / ice-cold ethanol as described above and re-dissolved in 200 µl sterile tissue culture grade water. The yield of DNA was calculated from the optical density at 260 nm. The plasmid DNA was stored at -80°C until transfection into cells was performed.

2.2.8   HPLC-^{32}P-Postlabelling   analysis   of   α-acetoxytamoxifen   and   4-OHtamQM   treated   pSP189   plasmid DNA

The HPLC-^{32}P-postlabelling analysis (described in sections 2.2.8.1-2.2.9.3) was conducted by Robert Crookston.

2.2.8.1 Digestion

The assay method used is outlined in Figure 2.3. Aliquots of α-acetoxytamoxifen or 4-OHtamQM treated pSP189 plasmid and untreated plasmid (negative control) (10 µg of each) were dried by vacuum centrifugation in 1.5 ml ultra-centrifuge tubes. To each tube was added 6.25 µl digestion mixture. The digestion mixture contained 175 mU micrococcal nuclease (Boehringer, Berkshire, UK), 3 mU calf spleen phosphodiesterase, 1 µl SSCC buffer (100 mM sodium succinate and 50 mM calcium chloride) and sterile water to make up to 6.25 µl. Samples were incubated overnight (17 h) at 37°C.

2.2.8.2 Nuclease P1 treatment

After overnight digestion 5 µl (4 µg) was removed from the digestion mixture and transferred to a separate tube. The remaining 1µg was retained for later use (Section 2.2.8.4). To enrich the samples for adducts nuclease P1 (9 µg) was added to each tube and incubated at 37°C for 1 h. Tris HCl (10 mM, pH 7.6, 2.4 µl) was added to stop the incubation.

^{32}P-Postlabelling Assay

A G C T G-Tam

5'   3'

Chemically modified DNA

↓ Micrococcal nuclease
   Spleen phosphodiesterase

A   G   C   T   G-Tam

HO   HO   HO   HO   HO
   P   P   P   P   P

↓ Nuclease P1 enhancement

G-Tam

HO
      P
Adducted nucleotide

↓ ^{32}P-ATP
   T4 polynucleotide kinase

G-Tam

^{32}P   HPLC separation + online
      P   radiochemical detection
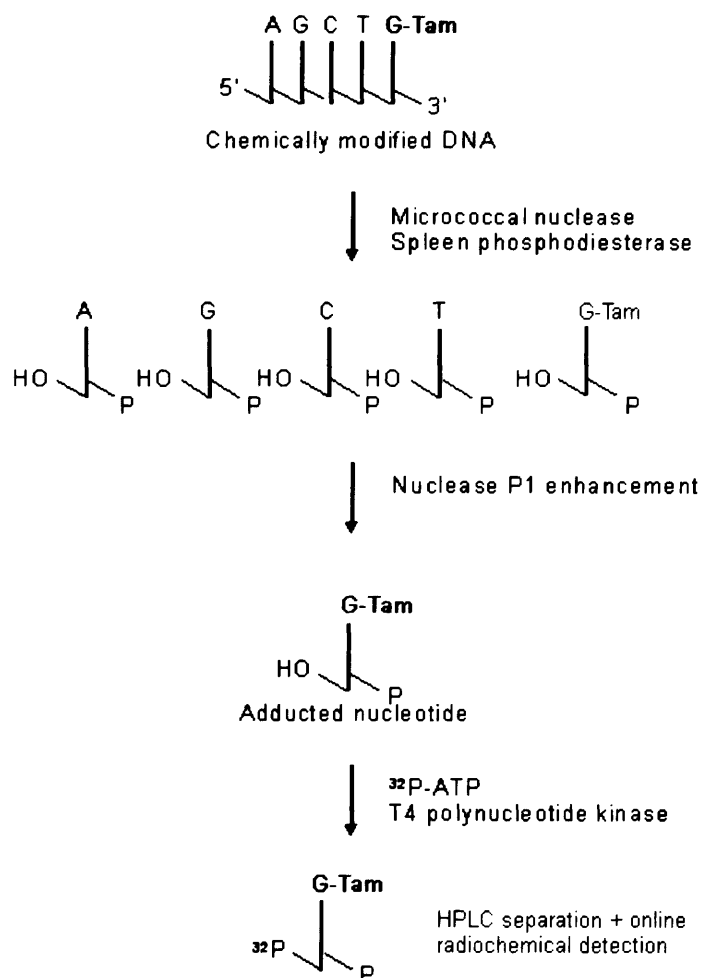
Figure 2.3 Outline of the ^{32}P-postlabelling assay for quantitation of tamoxifen DNA adducts

2.2.8.3 Radiolabelling of adducted nucleotides

T4 Polynucleotide kinase (Roche Diagnostics Ltd., Lewes, East Sussex, UK) (0.5 µl) and labelling buffer (2 µl) (containing 200 mM Tris HCl, 100 mM magnesium chloride, 100 mM DTT, 10 mM spermidine, filter sterilised) was

added to each nuclease P1 treated sample. The samples were ^{32}P-radiolabelled by addition of ATP (50 µCi [γ-^{32}P]ATP, Amersham Life Sciences Ltd, Little Chalfont, Bucks., UK) and incubated at 37°C for 1 h, then an aliquot from each tube was removed and diluted (1:20). This was used to analyse the efficiency of the nuclease P1 enrichment step. The remainder of each sample was stored at -20°C until HPLC analysis.

2.2.8.4 Labelling normal nucleotide samples

The 1µg samples remaining after the digestion step (Section 2.2.8.1) were diluted by a factor of 4800 and ^{32}P-postlabelled as above (Section 2.2.8.3) with 20 µCi [γ-^{32}P] ATP. The samples were split into two aliquots: one aliquot was incubated with apyrase (80 mU) for 30 min at 37°C (+apyrase), whilst the other was incubated without enzyme (-apyrase). The aliquots were diluted with sterile water (final volume 100 µl).

2.2.8.5 Efficiency of labelling and nuclease P1 digestion

+Apyrase, –apyrase and nuclease P1 efficiency samples were spotted (2 µl) onto polyethyleneimine (PEI)-cellulose thin layer chromatography (TLC) plates (Merck Eurolab Ltd., Poole, Dorset UK), and developed in 0.12 M sodium phosphate, pH 6.8 for approximately 2 h. The plates were dried, wrapped in Saran wrap and scanned using a phosphorimager.

2.2.9 HPLC analysis of ^{32}P-postlabelled α-acetoxytamoxifen or 4-OHtamQM treated pSP189 plasmid DNA digests

2.2.9.1 HPLC instrumentation

HPLC analysis of radiolabelled samples was performed using a Varian HPLC system consisting of a 210 Prostar solvent delivery module and 410 autosampler (Varian Ltd., Walton on Thames, Surrey, UK) coupled to an on-line radioactivity detector (β-ram, Lablogic, Sheffield, UK) fitted with a solid phase cell (500 µl). Data collection was performed using Laura, an MS Windows package (Lablogic). DNA adduct separation was performed using a Hypersil C$_{18}$ BDS (5 µm, 250x4.6 mm) reverse-phase column (Thermo Electron Corporation, Runcorn, UK).

2.2.9.2 HPLC conditions

HPLC analysis was performed at ambient temperature according to the published method [281]. Solvent A was 2 M ammonium formate, pH 4.0. This was prepared by dissolving ammonium formate in water and adjusting the pH to 4.0 with 98% formic acid. Solvent B was a mixture of acetonitrile and methanol 6:1 (v/v). Optimum separation was achieved using a system of 80% A, 20% B for 40 minutes followed by a linear gradient of 20-45% B for 25 minutes at a flow rate of 1 ml/min.

2.2.9.3 Adduct calculations

Known amounts of [Y-^{32}P]-ATP were injected onto the HPLC, a peak (eluting at 3 min) was collected and the radioactivity measured in disintegrations per min (d.p.m.) by scintillation counting. The ATP peak area was plotted against d.p.m. to give a standard curve. HPLC peak areas corresponding to tamoxifen adducts were measured and values applied to this standard curve to give values in d.p.m. Relative adduct labelling (RAL) was then calculated by the method of Reddy and Randerath (1986) based on the specific activity (SA) of the [γ-^{32}P]ATP (expressed as d.p.m. per pmol) and the amount of DNA used (1μg DNA = 3240 pmol deoxyribonucleoside 3'-monophosphates [dNp])

RAL = d.p.m. in adduct peak / (SA [Y-^{32}P]-ATP × pmol dNp used for analysis)

2.2.10 Fugene-6 mediated transfection of Ishikawa cells with α-acetoxytamoxifen and 4-hydroxytamoxifen quinone methide (4-OHtamQM) treated pSP189 plasmid DNA

Ishikawa cells (grown at 37°C, 5% CO_2, 90% confluent) were split 1 in 10 and approximately 2×10^6 cells were plated in 9 cm transfection plates in 10 ml DMEM/F12, phenol red-free medium supplemented with 10% fetal calf serum (Life Technologies Ltd, Paisley, UK) and 2mM GlutaMAX and grown for 48 h (37°C, 5% CO_2). Cells were 40-60% confluent prior to transfection. Three hours before transfection, used medium was aspirated and fresh medium added (10 ml). Fugene-6 (Roche Diagnostics Ltd, Lewes, East Sussex, UK) transfection reagent (30 μl) was slowly added to foetal calf serum free

DMEM/F12, phenol red-free medium (500 µl). The tube was gently tapped to mix the contents. Plasmid DNA (10 µg) was added to the Fugene solution and incubated at room temperature for 45 min. The aliquot of DNA was carefully added dropwise to each transfection plate and swirled gently to mix. Plates were then returned to the incubator (37°C, 5% CO_2) and 24 h after transfection, used media was aspirated and replaced with fresh media (10 ml). The plasmid was reclaimed from the Ishikawa cells between 45-48 h after transfection using a plasmid mini kit (Qiagen, Crawley, UK).

2.2.11 Isolation of pSP189 plasmid from Ishikawa cells

Media was aspirated from the transfection plates. Cells were scraped, using disposable plastic cell scrapers (Fisher Scientific, Loughborough, UK) into the residual volume of medium on each plate then pipeted into 1.5 ml micro centrifuge tubes. The plates were washed by scraping remaining cells into buffer P1 (300 µl, Qiagen, Crawley, UK). The Qiagen mini plasmid preparation protocol was followed for the remaining steps. Briefly, this involves alkaline cell lysis and separation, plus washing of plasmid DNA on an anion exchange column, followed by elution and ethanol precipitation of purified DNA. Yeast tRNA (Invitrogen, Paisley, UK) was used as a co-precipitant to enhance plasmid yield. Briefly, the eluted plasmid was mixed with 100% ethanol (700 µl) and yeast tRNA (30 µl, 1 µg/µl) and incubated at -20°C for 20 min. Subsequently, plasmid DNA was precipitated by centrifuging at 14,000 g for 30 min in a microcentrifuge. The supernatant was removed and the pellet was washed once with 70% ethanol. Syringes and fine needles were used to remove supernatants to prevent loss of the DNA pellets. DNA

pellets were air-dried and dissolved in 43 µl tris-EDTA buffer, pH 8.0, overnight (16 h). Reclaimed plasmid was stored at -20°C.

2.2.12 Digestion of pSP189 plasmid with Dpn1 restriction enzyme

Recovered plasmid was incubated with 2 units Dpn1 restriction enzyme (New England Biolabs (UK) Ltd., Hitchin, UK) for 3 h at 37°C in NEBuffer 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM dithiothreitol) in a total volume of 50 µl. Restriction enzyme was inactivated by heat (80°C) for 20 min, reconstituted in 20 µl TE buffer and stored at -20°C prior to transformation of electrocompetent *E.coli* (see section 2.2.1.7). Treatment of plasmid with the Dpn1 restriction enzyme, after it has been recovered from human cells, will digest any plasmid molecules which have not been produced by replication, i.e. any original plasmid containing the DNA adducts. Dpn1 does this by recognising methylated adenosines in the sequence 5'-GA^TC-3' in DNA and catalysing a double stranded cut on the DNA strand 3' to the adenine. All plasmid molecules that have been replicated in the human cells will have lost their methylated adenosines, as only bacterial cells possess the appropriate methylating enzymes, and will therefore remain in covalently closed circular form.

2.2.13 Preparation of plates for bacterial screening assay

Ampicillin (800 µl, 50 mg/ml), X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactose, 600 µl, 50 mg/ml) and IPTG (isopropyl β-D-thiogalactoside, 200 µl,

50 mg/ml) were added to molten agar (400 ml, 50°C). Plates were poured, 200 ml per 525 cm^2 square bioassay dish or 25 ml per 9 cm diameter round plate, and stored overnight in darkness to cure. Plates were then stored at 4°C until needed and warmed in a 37°C incubator prior to use, to dry any excess moisture.

2.2.14 Production of E.coli MBM7070 glycerol stocks

Luria Bertoni broth (LB broth, 12 ml) was inoculated with a 1 ml aliquot of *E.coli* MBM7070 (15% glycerol in LB broth). The culture was incubated for approximately 4 h in an orbital shaking incubator (Gallenkamp, Germany) at 37°C and 250 rpm. The starter culture was further used to inoculate a 500 ml flask containing 125 ml LB broth. This was grown for 16 h at 37°C and 250 rpm. The next morning, 36 ml of culture from each flask were combined with 36 ml of 30% sterile glycerol in LB broth. These were aliquoted and frozen at -80°C for future use.

2.2.15 Plasmid preparation

LB broth (2x12 ml) containing 100 µg/ml ampicillin was inoculated with 2 x 1 ml aliquots of *E.coli* MBM7070 (15% glycerol in LB broth) containing the pSP189 plasmid. The culture was incubated for approximately 4 h in an orbital shaking incubator at 37°C and 250 rpm. These starter cultures were further used to inoculate four 2 l flasks each containing 500 ml LB broth (containing 100 µg/ml ampicillin). These were grown at 37°C and 250 rpm. After 16 h growth, glycerol stocks were made as above and the plasmid was prepared

from the culture using a Qiagen Endofree Mega plasmid preparation kit (Qiagen, Crawley, UK) according to the manufacturers' protocol. Briefly, this involves alkaline cell lysis and separation of plasmid DNA on an anion exchange column, followed by ethanol precipitation of purified DNA. The yield of DNA was calculated from the optical density at 260 nm as shown below.

DNA concentration = A_{260} × Dilution × 50 (constant for dsDNA)

The yield of plasmid would generally be approximately 1 mg per 500 ml overnight culture. The plasmid DNA was stored at -80° C in 100 μg aliquots in TE Buffer pH 7.4).

2.2.16 Production of electrocompetent E.coli MBM7070

LB broth (2 x 12 ml) was inoculated with a 1 ml aliquot of MBM7070 E.coli (15% glycerol in LB). The cultures were incubated for 16 h (overnight) in an orbital shaking incubator at 37°C and 250 rpm. The next morning the starter cultures were used to inoculate two 2 l flasks each containing 500 ml LB broth. These were incubated in an orbital shaking incubator (Gallenkamp, Germany) at 37°C and 250 rpm until an optical density of 0.5-0.7 at 600 nm was reached. The cells were chilled on ice for 15 minutes then transferred to four 250 ml polypropylene copolymer centrifuge bottles. The tubes were centrifuged at 4200 rpm and 4°C in a Beckman J2-21m/E centrifuge using a JA-14 rotor (Beckman, Palo Alto, Ca., USA). The supernatant was removed and the pellets resuspended in 250 ml ice-cold sterile water. The cells were then centrifuged as before followed by another wash cycle (250 ml water). The pellets were then resuspended in 50 ml ice-cold 10% (v/v) glycerol

solution. The cells were centrifuged at 4200 rpm, 4°C for 10 min. The supernatant was carefully removed and the pellets re-suspended in an equivalent volume of 10% (v/v) glycerol. The cells were stored at -80°C in 100 µl aliquots, ready for use.

2.2.17 High-efficiency transformation of E.coli MBM7070 by electroporation

pSP189 plasmid DNA (3 µl, recovered from human cells) was transferred to an aliquot of electrocompetent cells (100 µl, defrosted on ice) and mixed gently before being incubated on ice for 1 min. The DNA-cell mixture was then transferred to a Geneflow 0.2 cm gap electroporation cuvette (Geneflow Ltd, Staffordshire, UK) pre-cooled on ice. The DNA-cell mixture was electroporated at 2.5 kV, 25 µF and 200 Ω using Biorad Gene Pulser apparatus (time constant 4.2 ms) (Biorad, Hercules, CA, USA). Immediately after electroporation 1 ml SOC medium was added. The mixture was transferred to a 15 ml sterile polypropylene tube and incubated for 30-45 minutes at 37°C and 250 rpm in an orbital shaker to allow expression of the ampicillin resistance phenotype.

2.2.18 Screening for mutant colonies

Transformed MBM7070 cells were plated out on 245x245mm plates (VWR International, Hunter Boulevard, Leicestershire, UK) and grown overnight in darkness at 37°C. Colonies were counted using a high resolution sorcerer image analysis/colony counting system (Perceptive Instruments Ltd, Haverhill,

Suffolk). The plate images were analysed using an automatic image analysis system (Sorcerer v2.2, Perceptive Instruments Ltd). A counting marker pen was used to count mutant colonies. Any white colonies (mutants) were noted and a mutation frequency was calculated.

Mutation Frequency = number of white colonies (mutant) / (number of blue colonies (WT) + number of white colonies (mutant))

Any white looking colonies were picked using a flamed wire loop and streaked out to single colonies on fresh 9 cm plates (25 µg/ml IPTG, 75 µg/ml X-gal, and 100 µg/ml ampicillin). Plates were grown in darkness for 16 h at 37°C. Single white colonies were re-streaked until it was possible to isolate single mutants (Figure 2.4).
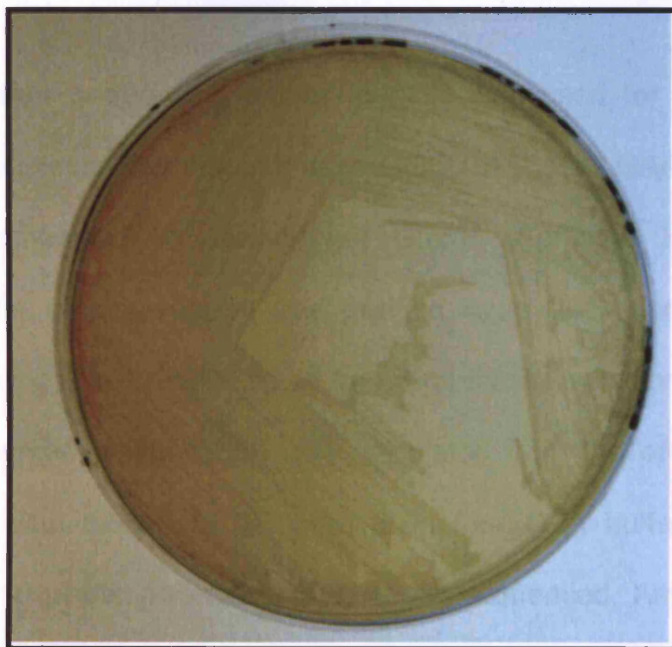
Figure 2.4 Illustrates a streak plate of white MBM7070 *E.coli* colonies containing mutant pSP189 plasmid.

2.2.19 Plasmid sequencing

Plasmid derived from mutant (white) colonies was amplified with the TempliPhi™ DNA Sequencing Template Amplification Kit (GE Healthcare, Amersham Place, Buckinghamshire, UK) before DNA sequencing. Templiphi generates microgram quantities of template DNA utilizing bacteriophage Phi29 DNA polymerase and rolling circle amplification. Briefly, a maximum of 1 μl of the template to be amplified is added to 5 μl of sample buffer. The sample is heated to 95°C for three minutes to gently lyse the bacteria and release the plasmid DNA. The sample is cooled and combined with 5 μl of reaction buffer and 0.2 μl of enzyme mix and incubated at 30°C for 4–18 h. At the end of this incubation, the DNA polymerase is inactivated by heating at 65°C for 10 min. The tube should now contain approximately 1–1.5 μg of amplified DNA that can be used directly in a DNA sequencing reaction.

A forward primer sequencing primer (psp2F) was used for all sequencing reactions. The primer had the sequence GGC GAC ACG GAA ATG TTG AA and was synthesised by biomers.net GmbH (Söflinger Str. 100, Ulm, Germany). The oligonucleotide was purified by a 20% denaturing PAGE electrophoresis gel and reconstituted in 1.5 ml sterile water prior to use. The sequencing primer concentration was calculated from the optical density at 260 nm and was diluted to 0.8 pmol/μl in tris-EDTA buffer, pH 8.0 and submitted at 10 μl per plasmid sample to be sequenced. Amplified plasmid DNA samples and sequencing primer were sent to the Protein and Nucleic Acid Chemistry Laboratory (PNACL, University of Leicester, Leicester, UK) for DNA sequencing. DNA Sequencing was carried out on an Applied Biosystems

Model 3730 DNA sequencer. The pSP189 shuttle vector contains an 8 base

"signature sequence" giving 4^8 (65,536) possible unique sequences [363]. Any

mutants with a duplicated "signature" were excluded from further analysis.

2.2.20 Utilization of the LwPy53 algorithm

In order to utilize the LwPy53 algorithm, mutation data from the *supF* assay

had to be converted into an appropriate format. Mutation data was used from

three *supF* spectra: (i) the α-acetoxytamoxifen 25μM+50μM combined

spectrum, (ii) the 4-OHtamQM combined spectrum and (iii) the combined

control spectrum. All base substitutions, along with point deletions and

insertions at each *supF* position were recorded in an Excel spreadsheet (for

an example see appendix 3). The output of the algorithm is a score that

represents the probability of mutation at each nucleotide of the *p53* gene

(exons 5, 7 and 8). Each score was converted into a percentage per

spectrum, by dividing the score by the sum of all scores and multiplying by

100.

2.3 RESULTS

The mutagenicity of the major tamoxifen DNA adduct, induced by treatment

with α-acetoxytamoxifen, and one potential minor tamoxifen DNA adduct,

generated by reaction with 4-OHtamQM, was compared using the *supF*

forward mutation assay and a relevant target cell line, human endometrial

adenocarcinoma (Ishikawa) cells. In order to utilise the LwPy53 algorithm, the

pSP189 plasmid was methylated *in vitro* at CpG sites prior to treatment with

α-acetoxytamoxifen or 4-OHtamQM. CpG methylation of the plasmid was confirmed by HpaII digestion.

Treated plasmid was transfected into Ishikawa cells where replication occurred then recovered plasmid was used to transform MBM7070 indicator *E. coli*. Any plasmids containing a mutation in the *supF* gene grew as white colonies while non-mutant wild type plasmids grew as blue colonies. Mutant colonies were picked from agar plates and added to the Templiphi reaction to amplify their plasmids and these were subsequently sequenced to determine the mutation(s) in the *supF* gene region of the plasmid. The LwPy53 algorithm was then applied to exons 5, 7 and 8 of *p53* using spontaneous, α-acetoxytamoxifen and 4-OHtamQM induced mutation data from *supF* to obtain a predicted human *p53* spectrum that might be expected to be present in endometrial tumours of tamoxifen treated women, if tamoxifen DNA adduct formation plays a role in mutation induction in humans.

2.3.1 Hpa II Digestion of methylated and umethylated pSP189

Aliquots of CpG methylated and unmethylated (control) pSP189 plasmid were digested with the HpaII restriction endonuclease in order to confirm successful methylation by CpG methyltransferase and the products of the reaction were loaded on a 1% agarose gel (Figure 2.5). HpaII has the following recognition sequence:

$$5'...C^{\triangledown}CGG...3'$$

$$3'...GGC_{\triangle}C...5'$$

However, Hpall will not cut sites that have been methylated by CpG methyltransferase or Hpall Methyltransferase. As a result, methylated plasmid is protected from Hpall digestion.
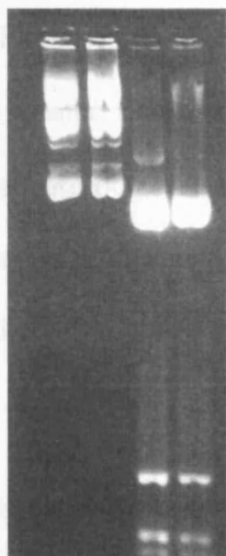
Figure 2.5 Hpall digestion of pSP189. Photograph of a 1% preparative agarose gel showing the Hpall-digested products. CpG methylated pSP189 was digested with 50 units (lane 1) or 100 units (lane 2) of Hpall. Umethylated pSP189 (control) was digested with 50 units or 100 units of Hpall. (lanes 3 and 4 respectively).

2.3.2 Tamoxifen-DNA adduct quantification by ^{32}P-postlabelling

2.3.2.1 DNA adduct quantification by ^{32}P-postlabelling of α-acetoxytamoxifen treated pSP189 plasmid

Shuttle vector plasmid pSP189 was modified by reaction with α-acetoxytamoxifen and analysed by ^{32}P-postlabelling to quantify adduct levels formed at different doses. A positive control of liver DNA from a Wistar Han rat which had been treated with tamoxifen in the diet for six months (40 mg/kg) was also analysed in parallel. Figure 2.6B shows that as reported previously [281, 350], α-acetoxytamoxifen treated plasmid gave one major peak that co-eluted with the dG-N^2-tamoxifen adduct detected in DNA from a tamoxifen treated rat (Figure 2.6D). An additional minor peak is also observed eluting just prior to the major adduct, which, based on retention time, is the N-demethylated dG-N^2-tamoxifen adduct [364].

The relationship between α-acetoxytamoxifen dose and DNA adduct numbers is illustrated in Figure 2.7A. It is clear from the graph that the number of DNA adducts increased linearly as the dose of α-acetoxytamoxifen increased from 10 μM to 50 μM. The degree of pSP189 modification by α-acetoxytamoxifen equates to a level of 1.8, 3.0 or 3.8 adducts per plasmid (4952 basepairs), for the 10, 25 and 50 μM doses respectively (Table 4.2), which is equivalent to a range of between 174-376 adducts per 10^6 nucleotides. The same batch of plasmid that was analysed by ^{32}P-postlabelling was used for all subsequent *supF* assay investigations in Ishikawa cells.
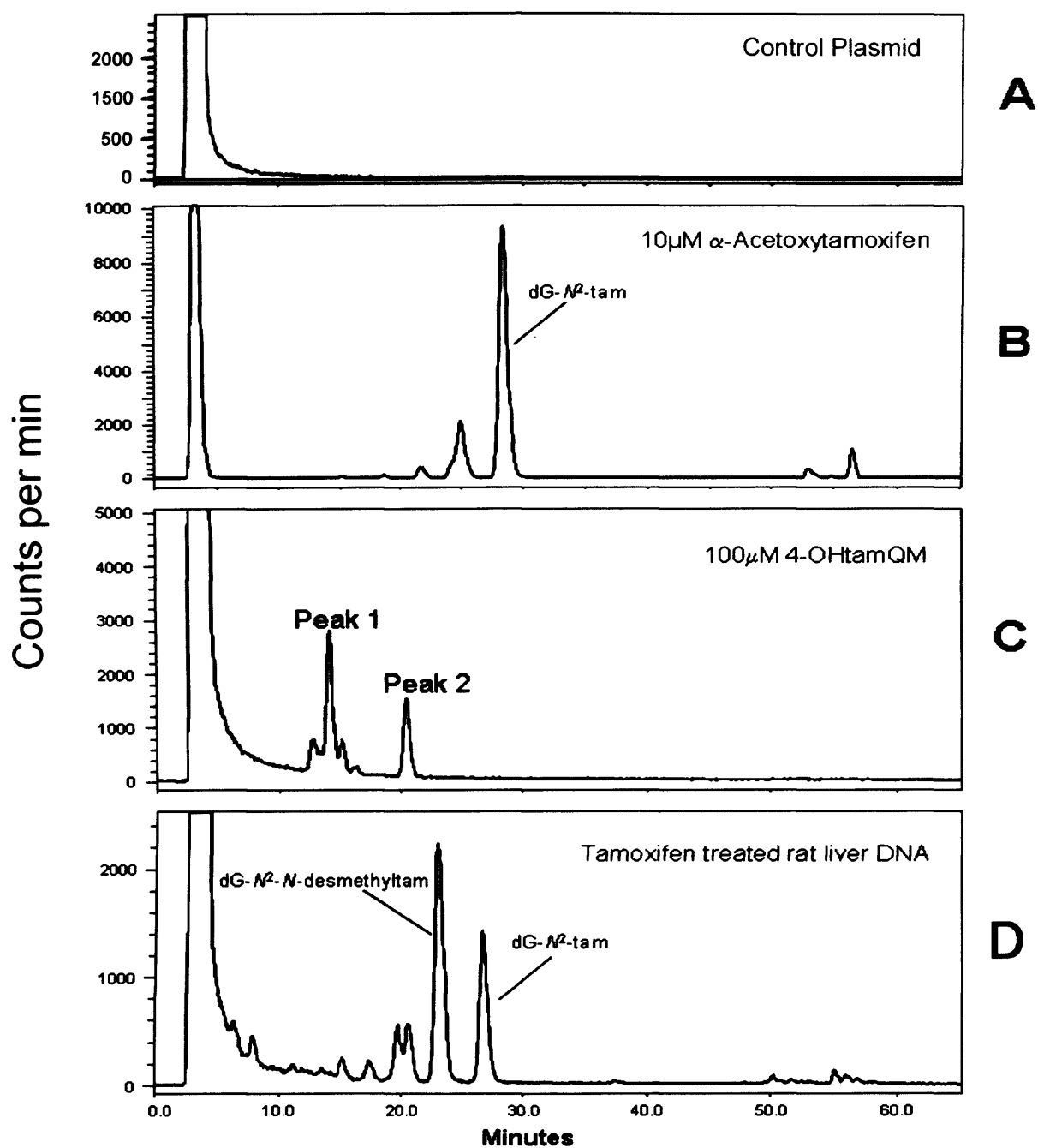
Figure 2.6 Representative radioactive HPLC chromatograms of ^{32}P-labelled digests from control (A), α-acetoxytamoxifen (B), 4-OHtamQM (C) treated plasmids and liver DNA from a tamoxifen treated rat (D). HPLC analysis was conducted by Robert Crookston.
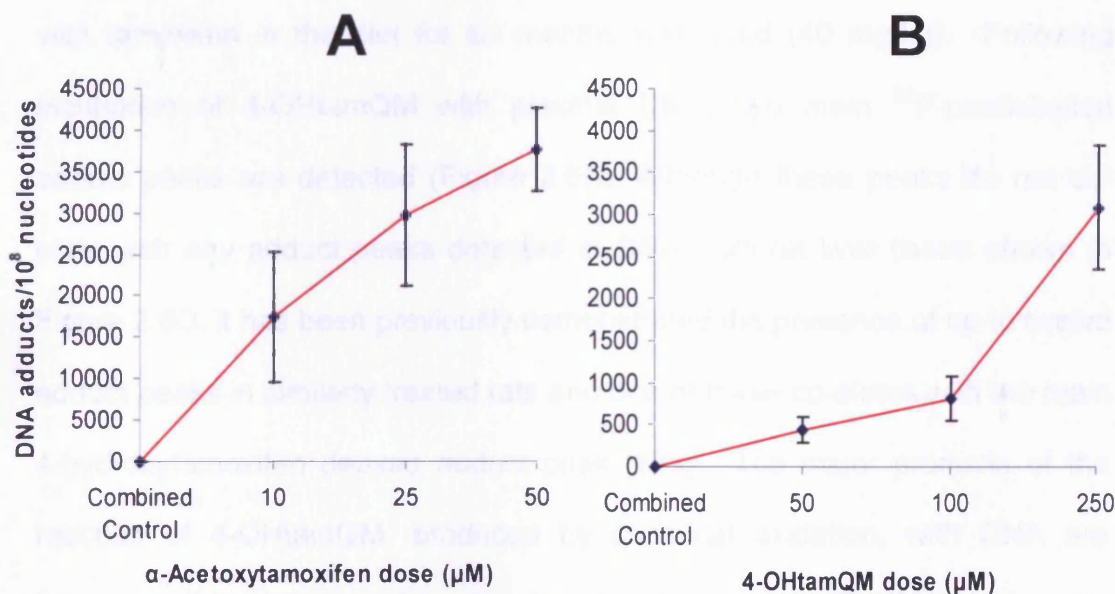
Figure 2.7 Dose-related increase in DNA adduct formation induced by α-acetoxytamoxifen (A) and 4-OHtamQM (B) determined by HPLC-^{32}P-postlabelling analysis. Data points represent the average of duplicate measurements of two plasmid samples per dose. Error bars represent standard deviation. Combined control represents unadducted methylated pSP189. The raw data used to construct this chart is tabulated in appendices 1 and 2. Analysis was conducted by Robert Crookston.

2.3.2.2 DNA adduct quantification by ^{32}P-postlabelling of 4-OHtamQM treated pSP189 plasmid

Shuttle vector plasmid pSP189 was modified by reaction with 4-OHtamQM and analysed by ^{32}P-postlabelling to quantify adduct levels at different doses. Shuttle vector plasmid pSP189 was modified by reaction with 4-OHtamQM and analysed by ^{32}P-postlabelling to quantify adduct levels at different doses.

A positive control of liver DNA from a Wistar Han rat which had been treated with tamoxifen in the diet for six months was used (40 mg/kg). Following incubation of 4-OHtamQM with plasmid DNA, two main ^{32}P-postlabelled adduct peaks are detected (Figure 2.6C). Although these peaks do not co-elute with any adduct peaks detected in DNA from rat liver tissue shown in Figure 2.6D, it has been previously demonstrated the presence of up to twelve adduct peaks in similarly treated rats and one of these co-elutes with the main 4-hydroxytamoxifen derived adduct peak [281]. The major products of the reaction of 4-OHtamQM, produced by chemical oxidation, with DNA are known to be isomers of a 4-hydroxylated form of dG-N^2-tamoxifen (α-(deoxyguanosin-N^2-yl)-4-hydroxytamoxifen) [311]. This reactive intermediate can also be generated enzymatically and it has been shown that incubation of 4-hydroxytamoxifen with horseradish peroxidase yields one major ^{32}P-postlabelled adduct peak, which corresponds to peak 1 observed in this study [281, 365].

This difference in adduct profile is probably due to the different methods used to activate 4-hydroxytamoxifen. The additional adduct, which is present at low levels with peroxidase activation, may be an isomer of dG-N^2-4-hydroxytam or may be an as yet unidentified adduct. Higher molar concentrations of 4-OHtamQM were used in the incubations compared to α-acetoxytamoxifen, as the former is known to generate lower levels of adducts [365]. However, despite the lower adduct levels, these doses have previously been shown to induce an increase in M.F. greater than α-acetoxytamoxifen which is why they were considered appropriate in this study. The relationship between 4-OHtamQM dose and DNA adduct numbers is illustrated in Figure 2.7B. A

dose-dependent increase in adduct formation was observed which is consistent with previous reports [350]. In the present study, 4-OHtamQM (50 μM) induced a 95-fold lower level of DNA adducts than an equimolar dose of α-acetoxytamoxifen. The three treatments with 4-OHtamQM (50, 100 and 250 μM) induced approximately 0.04, 0.08 and 0.3 adducts per plasmid (Table 4.5), which is equivalent to a range of between 4- 31 adducts per 10^6 nucleotides.

2.3.3 Part A: Mutagenicity of plasmid treated with α-acetoxytamoxifen replicated in Ad293 cells

2.3.3.1 Mutation frequency in *supF* gene

Treatment of pSP189 plasmid with α-acetoxytamoxifen induced an increase in mutation frequency for all doses relative both to the solvent (50% ethanol:50%

Table 2.2. Mutation frequency induced by α-acetoxytamoxifen treated pSP189 replicated in Ishikawa cells

Treatment	Colonies Screened	Number of Mutants	Mutation Frequency[a]	Adduct Number[b] (± S.D.)
Solvent Control	78487	39	4,9	0 ± 0
Water Control	65910	48	7,2	0 ± 0
Combined Control	144397	87	6,0	
10 μM	192820	297	15,4	174,0 ± 78
25 μM	38861	200	51,4	297,0 ± 85
50 μM	14813	134	90,4	376,4 ± 61

[a] Mutation frequency per 10^4 colonies
[b] Adduct number per 10^6 nucleotides

acetonitrile) and water (and acetonitrile) control, as illustrated in Table 2.2.

The mutation frequency of control plasmid treated with water only was similar to the mutation frequency of plasmid incubated with the solvent used to dissolve the α-acetoxytamoxifen. Furthermore, the mutation spectra induced

by these two types of control were also very similar (Figure 2.12). Therefore, it can be assumed that the presence of ethanol and acetonitrile had no effect on the spontaneous mutagenesis of the plasmid. In order to increase the pool of mutated colonies for subsequent sequence analysis data from the two types of control were combined (combined control). The mutation frequency of the combined control sample was similar to that observed in previous studies from this laboratory using different cell lines (3.3-6.6 mutants/10^4 colonies in Ad293 cells [350] and 7.06-12.27/10^4 in human fibroblasts [249]. Previous studies from other laboratories have reported lower values of background mutation frequency in Ad293 cells, ranging from 0.1-1.8/10^5 [366] to 1.3/10^4 colonies [367]. As illustrated in Figure 2.8 and Table 2.2, treatment with α-acetoxytamoxifen induced a linear increase in the mutation frequency relative to the solvent control (3.1-fold for 10 µM, 10.4-fold for 25 µM and 18.4-fold for 50 µM).
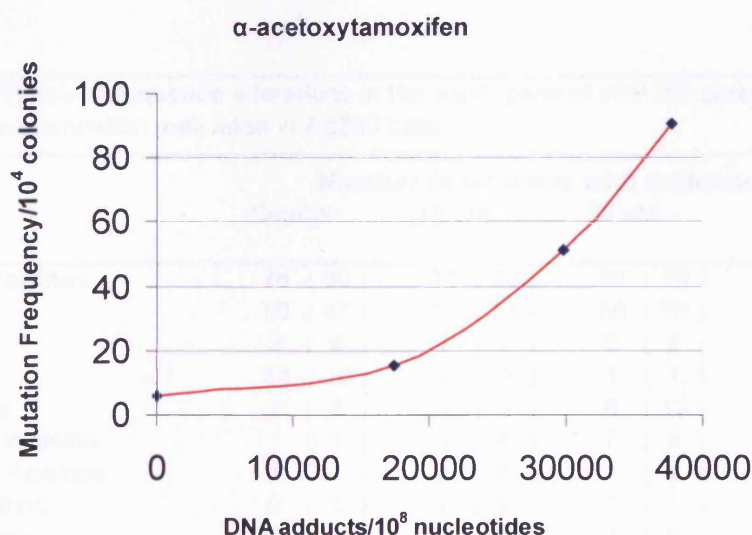
Figure 2.8 Mutation frequency vs. DNA adduct level detected in α-acetoxytamoxifen treated plasmid.

2.3.3.2 Types of mutation found in *supF* gene

The types of mutations induced by α-acetoxytamoxifen are illustrated in Table 2.3. For all sequenced plasmids, including controls, the predominant mutations are in the form of base substitutions (ranging from 60% - 92% of all mutations seen). When these are subdivided into single base, tandem (i.e., two or three adjacent point mutations) or multiple mutations, the single base substitutions are most common (between 47% and 80%) and the proportions increase in a dose related manner for treated plasmid compared to control. Tandem substitutions range from 0% to 2% and multiple mutations accounted for between 1% and 12%. The proportion of frameshift mutations ranges from 2% to 12%. Large deletions are common in control plasmid, occurring in 39% of all cases, whereas the incidence decreases with increasing α-acetoxytamoxifen dose. Therefore, the mutation profile changes with α-acetoxytamoxifen treatment compared to spontaneous mutations.

Table 2.3. Types of sequence alterations in the *supF* gene of pSP189 plasmids treated with α-acetoxytamoxifen replicated in Ad293 cells

Mutations	Number of plasmids with mutations (%)			
	Control	10 μM	25 μM	50 μM
Base substitutions	**76** (60)	**35** (63)	**59** (79)	**70** (92)
Single	60 (47)	31 (55)	56 (75)	61 (80)
Tandem	3 (2)	1 (2)	2 (2)	0 (0)
Multiple	13 (10)	3 (5)	1 (1)	9 (12)
Frameshifts	**2** (2)	**5** (9)	**9** (12)	**5** (7)
Single base deletion	1 (1)	2 (4)	7 (9)	4 (5)
Single base insertion	1 (1)	2 (4)	1 (1)	1 (1)
Large insertions	0 (0)	1 (2)	1 (1)	0 (0)
Large deletions	49 (39)	16 (29)	7 (9)	1 (1)
Total plasmids sequenced	**127** (100)	**56** (100)	**75** (100)	**76** (100)

Table 2.4 shows the types of single base substitutions induced by α-acetoxytamoxifen. At all dose levels, excluding control, transversions are the preferred mutation (46-79% of all substitutions). In the combined control plasmid the major mutations are GC→AT transitions and GC→TA transversions (accounting for 54% and 23% of all substitutions respectively) with GC→CG transversions contributing 20% and AT→TA transversions 3% of base substitutions. At the 10 μM α-acetoxytamoxifen dose the major mutations are the GC→TA transversions and GC→AT transitions (each accounting for 32% of all substitutions), followed by GC→CG (29%), AT→CG and AT→GC (3% each).

Table 2.4. Types of single base and tandem substitution mutations in *supF* gene of pSP189 plasmids treated with α-acetoxytamoxifen replicated in Ishikawa cells

Mutations	Number of plasmids with mutations (%)			
	Combined Control	10 μM	25 μM	50 μM
Transversions	**28** (46)	**20** (65)	**38** (68)	**49** (79)
GC→TA	14 (23)	10 (32)	27 (48)	38 (61)
GC→CG	12 (20)	9 (29)	6 (11)	7 (11)
AT→TA	2 (3)	0 (0)	0 (0)	1 (2)
AT→CG	0 (0)	1 (3)	5 (9)	3 (5)
Transitions	**33** (54)	**11** (35)	**18** (32)	**13** (21)
GC→AT	33 (54)	10 (32)	18 (32)	12 (19)
AT→GC	0 (0)	1 (3)	0 (0)	1 (2)
Total single base substitutions	**61** (100)	**31** (100)	**56** (100)	**62** (100)

At the two higher doses there is a pronounced preference for the induction of GC→TA transversions (48% at 25 μM and 61% at 50 μM). At the 25 μM dose the other major substitution is the GC→AT transition (32%) with GC→CG and AT→CG transversions being relatively uncommon (11% and 9% respectively). After treatment with the highest dose (50 μM) the second most common

substitution is the GC→AT transition (19%) as was the case for the 25 μM dose, followed by GC→CG transversion (11%), AT→CG transversion (5%), AT-TA transversions and AT-GC transitions (2% each).

Figure 2.9 shows a comparison of the different proportions of base substitution mutations induced in the *supF* gene by the various α-acetoxytamoxifen treatments or in the combined control. As the dose increases, the proportion of GC→TA transversions increases linearly and the proportion of GC→AT transitions decreases. It is clear from the graphical representation that transitions are the major substitution type in the control, while transversions dominate in treated samples.
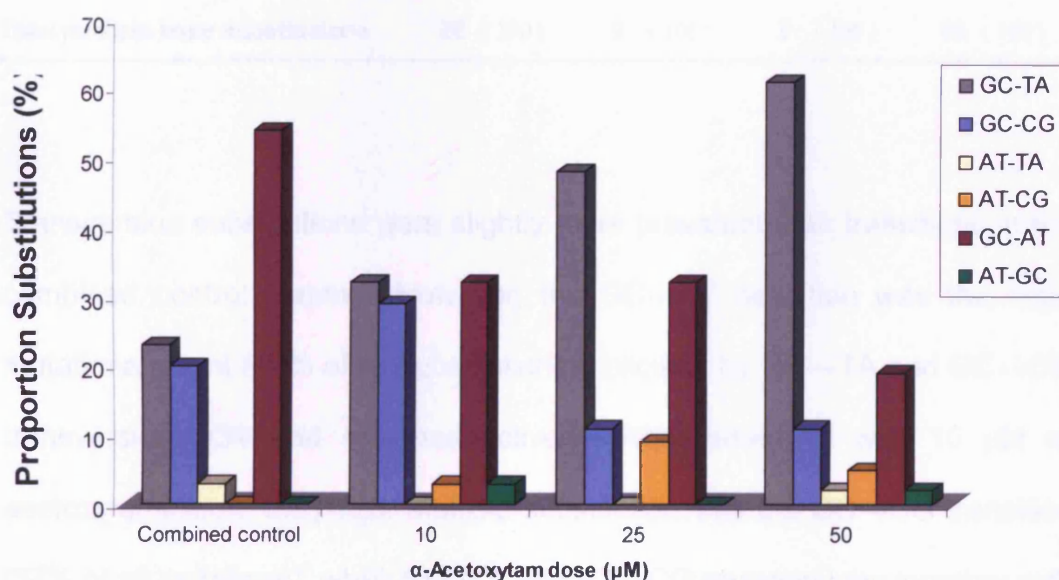
Figure 2.9 Bar chart illustrating the proportion of different single base and tandem substitution mutations seen in combined control plasmid and in plasmid treated with α-acetoxytamoxifen (10, 25 and 50 μM).

Multiple substitution mutations were much less frequent compared to single substitutions (Tables 2.3 and 2.5). Transversions were more common than transitions in general, although the number of multiple mutations in the treated plasmids was too small to draw definitive conclusions.

Table 2.5. Types of multiple base substitution mutations in the *supF* gene of pSP189 plasmids treated with α-acetoxytamoxifen replicated in Ishikawa cells

Mutations	Number of plasmids with mutations (%)			
	Combined Control	10 µM	25 µM	50 µM
Transversions	**15** (54)	**8** (89)	**2** (100)	**9** (47)
GC→TA	11 (39)	0 (0)	1 (50)	8 (42)
GC→CG	4 (14)	6 (67)	0 (0)	1 (5)
AT→TA	0 (0)	1 (11)	0 (0)	0 (0)
AT→CG	0 (0)	1 (11)	1 (50)	0 (0)
Transitions	**13** (46)	**1** (11)	**0** (0)	**10** (53)
GC→AT	13 (46)	0 (0)	0 (0)	10 (53)
AT→GC	0 (0)	1 (11)	0 (0)	0 (0)
Total multiple base substitutions	**28** (100)	**9** (100)	**2** (100)	**19** (100)

Transversion substitutions were slightly more prevalent than transitions in the combined control plasmid. However, the GC→AT transition was the main mutational event (46% of all substitutions), followed by GC→TA and GC→CG transversions (39 and 14% respectively). After treatment with 10 µM α-acetoxytamoxifen, the major multiple substitution was the GC→CG transition (67% of all mutations), while AT→TA and AT→CG transversions together with AT→GC transitions accounted for 11% each. Dosing the plasmid with 25 µM α-acetoxytamoxifen only yielded one multiple mutation; this contained two transversions, a GC→TA and an AT→CG substitution. The 50 µM treatment induced multiple substitutions primarily in the form of GC→AT transitions

(53%), closely followed by GC→TA transitions (42%). The remaining substitution was a GC→CG transversion (5%).

Figure 2.10 illustrates graphically different proportions of multiple base substitution mutations induced in the *supF* gene by the various α-acetoxytamoxifen treatments, and through spontaneous means. Whilst there are no clear trends, the same types of substitutions (mainly GC→TA and GC→AT) were present at similar proportions in the control and highest dose (50 μM) plasmid. However, there were very few multiple substitutions at the 10 and 25 μM doses and as a result no firm conclusions can be drawn about these treatments.
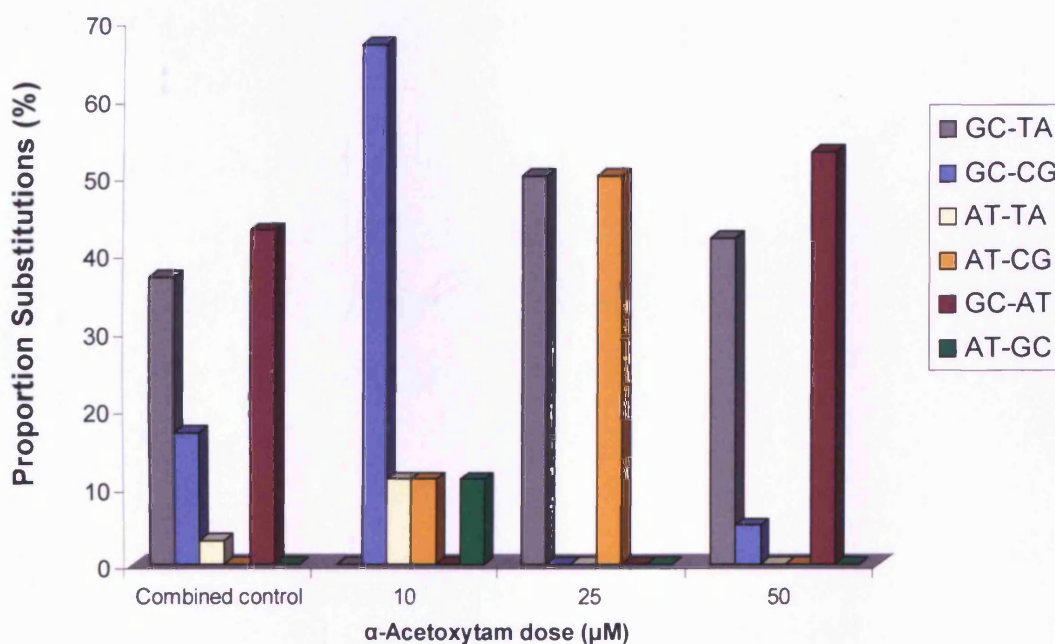
Figure 2.10 Bar chart illustrating the proportion of different multiple base substitution mutations in the *supF* gene found in untreated control mutants and in those induced by treatment with α-acetoxytamoxifen (10, 25 and 50 μM).

The proportion of frameshift mutations was slightly increased with α-acetoxytamoxifen treatment, compared to control, although there was no dose response effect (Table 4.3). Insertions were uncommon in general. Of the control plasmids sequenced 1% had single base deletions, whereas at 10, 25 and 50 μM α-acetoxytamoxifen single deletions occurred in 4%, 9% and 5% of cases of the *supF* gene respectively. Interestingly, a large deletion spanning bases 97→267 was present in 40% of the control plasmids sequenced. The frequency of this large deletion was decreased in a linear fashion with increased α-acetoxytamoxifen doses, and accounted for 29, 9 and 1% of all mutations at 10, 25 and 50 μM respectively. The deleted sequences are illustrated in Figure 2.11.
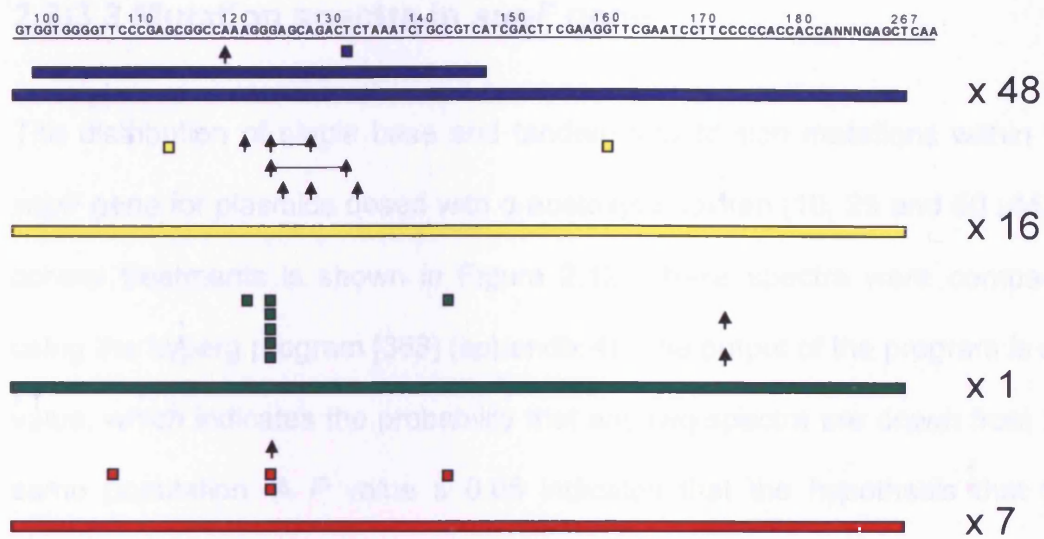
Key: Deletions: ■ Control ■ 10µM ■ 25µM ■ 50µM

↑ Denotes the following insertions:

Control; A @119

10µM; G @128, as part of a multiple mutation, A @121,

G @124 and G @128, as part of a multiple mutation

G @124 and G @132, G @133, TC @125

25 µM; T @172, CCC @172

50 µM; G @124

Figure 2.11 Deletions and insertions in the *supF* gene seen in spontaneous mutants and in those induced by treatment with α-acetoxytamoxifen (10, 25 and 50 µM). Numbers on the right indicate the number of plasmid sequences with the same large deletion of bases 97→267.

2.3.3.3 Mutation spectra in *supF* gene

The distribution of single base and tandem substitution mutations within the

supF gene for plasmids dosed with α-acetoxytamoxifen (10, 25 and 50 μM) or

control treatments is shown in Figure 2.12. These spectra were compared

using the hyperg program [368] (appendix 4). The output of the program is a *P*

value, which indicates the probability that any two spectra are drawn from the

same population. A *P* value $\leq$ 0.05 indicates that the hypothesis that the

spectra are the same can be rejected. When the solvent and water control

spectra are compared, they are found to be statistically similar to each other

($P=0.096$). Furthermore, the 10 μM α-acetoxytamoxifen treatment did not

induce a spectrum significantly different from the combined control ($P=0.133$).

In contrast, treatment of plasmid with the intermediate (25 μM) and high (50

μM) doses resulted in spectra significantly different from the combined control

($P < 0.05$). Since both spectra were distinct from the control single and

tandem substitutions from these two α-acetoxytamoxifen doses were

combined in one spectrum (as shown in Figure 2.13) for hotspot analysis and

utilization of the LwPy53 algorithm. Mutations are not generally distributed

randomly, but concentrated at one or more sites, known as hotspots. A

hotspot was defined as a site where the number of mutations observed was 4-

fold or more, greater than the number expected for a random (Poisson)

distribution. Single and multiple base substitutions have been presented on

separate spectra because it has been suggested that these multiple mutations

arise through a different mechanism to single base mutations [369].

Solvent Control

```
100        110        120        130        140        150        160        170        180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
        A    T                AC   G  C    AA      T         A    GA  T    A              AA
        A                            C    G       T                        A
                                     T    G       A
                                     A            A
                                     A            *
                                     *
```

Water Control

```
100        110        120        130        140        150        160        170        180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
        A AA    TAA    A  A   A T     A  A      T    A                    TC    T    T
        A A                          A  G      T    A                    T          A
                                     A  G      T    A
                                     A  G      A
                                     A  *      A
                                     A         *
                                     A
                                     A
                                     C
                                     *
```

10μM α-Acetoxytamoxifen

```
100        110        120        130        140        150        160        170        180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
        A A  A       T     G     C    A T  GG G   T T  T      A A  T    T        TT    A
        C                            T                                  C        AT
        C                            A
                                     A
                                     A
                                     C
                                     C
                                     C
                                     *
```

25μM α-Acetoxytamoxifen

```
100        110        120        130        140        150        160        170        180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
        TA    A    T CA  G    T    AT T    A      T      T  T      CT      TT    T    TT  G      CA
        A    G   A  T         T    A  T    G             A         CA      AT         TA         C
             T                T    T       G                       C       AA         A
                             T    T                                       A          A
                             T    A                                       A
                             *    C                                       A
                                  *                                       C
                                                                          *
```

50μM α-Acetoxytamoxifen

```
100        110        120        130        140        150        160        170        180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
      T   A T     A T   A  C  TTT AA  T    A       AGT    A    T  C  C   TT   AT    TT         ACA
          G       A T   A    TT    A T    A                       A  TT    T    A
                  T         T       T    G                        AT    T
                  A                 T                             AT    C
                  C                                               AT    C
                  *                                               A*    *
                                                                  C
                                                                  *
```

Figure 2.12 Mutation spectra depicting single base and tandem substitutions induced in the *supF* gene after treatment with α-acetoxytamoxifen (10, 25 and 50 μM), solvent (ethanol:acetonitrile) or water. Red asterisks denote hotspots. A hotspot was defined as a site where the number of mutations observed was 4-fold or more, greater than the number expected for a random (Poisson) distribution. Tandem mutations are underlined.

25µM+50µM α-Acetoxytamoxifen Combined

```
                                    C
                                    C
                                    C
                                    A
                                    A                A
                                    A        *       A
                                    A        *       A
                                    A        G       A
                                    A        G       A
                                    A        G       T
                                    A        G       T
                                    A        G       T        A
 AAAA     A            C            A        A        T        A                          A       T              A
 AAAA     TTA     A   A    AAT    G T       AA       T       A  A          GA      T      A      TC       T    T    AA
 100       110        120         130       140          150             160            170          180
 GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
 TTAA   TA   T CA  A C TTT AT T   A        TGT   T T     TCTC   A   TT   AT    TT  G   ACA
  A   GG    AT A    TT  AA T       A       A     A A     CA    C   TT    T     TT        CA
      AT  G   TT   AT   A           A                    C         TT    T     TA        C
      A  T        T      G                                         AT    T     A
      T           T      G                                         AT    C     A
      T           T      G                                         AT    C     A
      A           T                                                AT
      C           A                                                AA
      *       *   T                                                A *
                  C                                                A
                  *                                                A
                                                                   A
                                                                   A
                                                                   A
                                                                   C
                                                                   C
                                                                   *
```
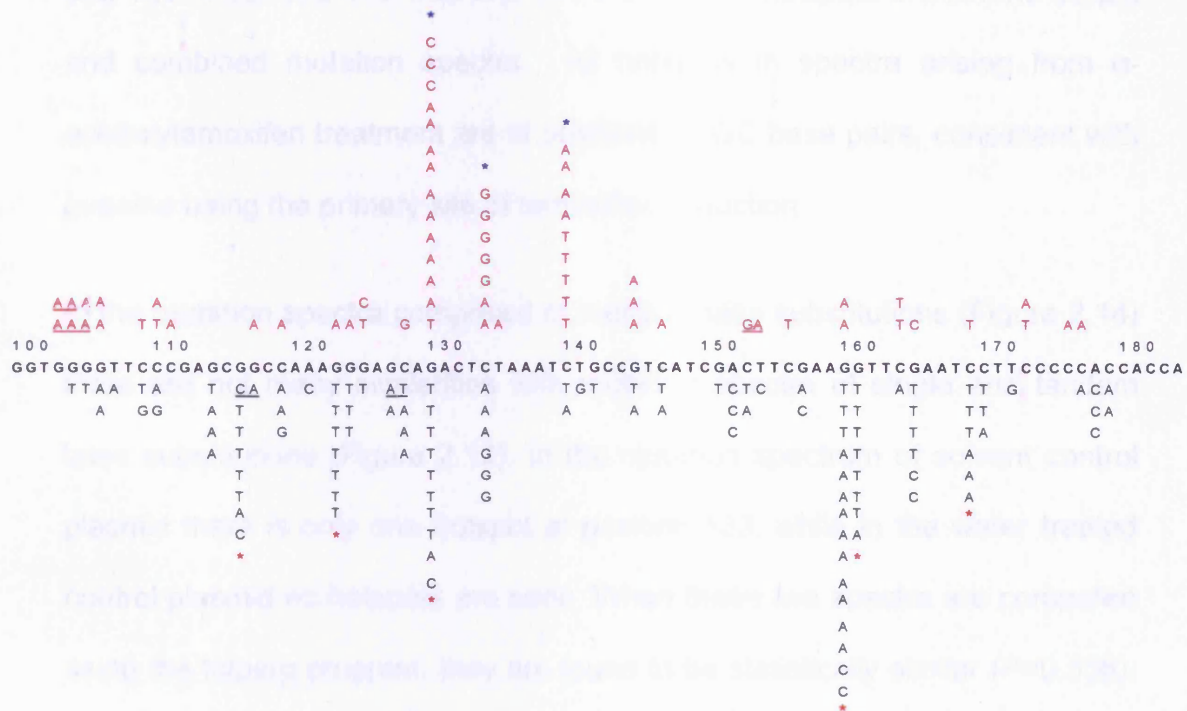
Figure 2.13 Combined mutation spectrum depicting single base and tandem substitutions in the *supF* gene induced after treatment with α-acetoxytamoxifen (25 and 50 µM). Single base and tandem substitutions induced following both control treatments are shown in red above the *supF* gene sequence. Hotspots are represented by asterisks. Tandem mutations are underlined.

In the mutation spectrum induced in control plasmid treated with solvent only, there are two hotspots at positions 129 and 139. Whilst in the water control, hotspots also exist at the sites, together with a third at position 133. At the lower dose of α-acetoxytamoxifen (10 µM) there is one hotspot at position 129, while at the middle dose (25 µM) there are three hotspots at positions 122, 129 and 159. At the higher dose (50 µM) there are four hotspots at positions 115, 159, 160 and 164. If the mutation spectra from the two top doses of α-acetoxytamoxifen treated plasmids (25 and 50 µM) are combined

(Figure 2.13) there are now six hotspots at positions 115, 122, 129, 159, 160 and 168. Positions 115, 159 and 160 are sites of hotspots in both the 50 μM and combined mutation spectra. All hotspots in spectra arising from α-acetoxytamoxifen treatment are at positions of GC base pairs, consistent with guanine being the primary site of tamoxifen adduction.

In the mutation spectra comprised of multiple base substitutions (Figure 2.14) there are not many similarities with mutation spectra of single and tandem base substitutions (Figure 2.12). In the mutation spectrum of solvent control plasmid there is only one hotspot at position 133, while in the water treated control plasmid no hotspots are seen. When these two spectra are compared using the hyperg program, they are found to be statistically similar (P=0.536). Likewise, multiple mutation spectra from all α-acetoxytamoxifen treated samples were also similar to the combined control (P>0.05). After treatment with 10 μM α-acetoxytamoxifen, a single mutation hotspot was induced at the GC basepair position 117 (Figure 2.14). There were not enough multiple substitutions to give a mutation spectrum for the 25 μM dose of α-acetoxytamoxifen but at the highest dose (50 μM) there was one hotspot at position 102, which occurred at a GC basepair. When all α-acetoxytamoxifen treatments are combined in one spectrum the same two hotspots at positions 102 and 117 are seen. On the other hand, in combined control plasmids, three hotspots (shown in red) are observed at positions 115, 133 and 139. There were sites for hotspots which corresponded to those seen in the single and tandem substitution mutation spectra (Figure 2.12). For the combined control plasmids, these sites were at positions 133 and 139, while for the lower α-acetoxytamoxifen dose (10 μM) there was one site at position 117.

Solvent Control

```
 100       110       120       130       140       150       160       170       180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
                   A  TT  G    T   T  A         T C  T    A    G
                   T     G    A                                A
                         A
                         *
```

Water Control

```
 100       110       120       130       140       150       160       170       180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
         T      TA    A AA A   T    T
         A                     T    T
```

10µM α-Acetoxytamoxifen

```
 100       110       120       130       140       150       160       170       180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
                C G       G  GGG                    A
                  G
                  G
                  *
```

25µM α-Acetoxytamoxifen

```
 100       110       120       130       140       150       160       170       180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
G T
```

50µM α-Acetoxytamoxifen

```
 100       110       120       130       140       150       160       170       180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
     T      T  TA   ATAT   T  T  A        A                   A
     T      T T     C  A
     A
     *
```

Combined α-Acetoxytamoxifen

```
                                  *
                                  G
                                  G
              *                   A        *
      A                    A A A  T        T                          A
      T              TA    TATT T      T T A     GA T C  T    A     G
 100       110       120       130       140       150       160       170       180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
     GTT    T   TACG    A TA GT GGT    T A      A         A     A
     T      T T     G      C  A  G
     A             G
     *             *
```

Figure 2.14 Mutation spectra depicting multiple substitutions induced after treatment with α-acetoxytamoxifen (10, 25 and 50 µM) or in solvent and water controls. Non-treated combined control multiple base substitutions are shown in red above the *supF* gene sequence of the combined α-acetoxytamoxifen spectrum. Asterisks denote hotspots. Tandem mutations are underlined.

2.3.4 Part B: Mutagenicity of plasmid treated with 4-OHtamQM and replicated in Ad293 cells

2.3.4.1 Mutation frequency in *supF* gene

Treatment of pSP189 plasmid with 4-OHtamQM did not induce a significant

increase in mutation frequency for all doses in comparison to the combined

control, as illustrated in Table 2.6 and Figure 2.15.

Table 2.6 Mutation frequency induced by 4-hydroxytamoxifen quinone methide replicated in Ishikawa cells

Treatment	Colonies Screened	Number of Mutants	Mutation Frequency[a]	Adduct Number[b] (± S.D.)
Combined Control	144397	87	6,0	0 ± 0
50 µM	126813	82	6,47	4,3 ± 0,8
100 µM	117465	113	9,6	8,2 ± 2,7
250 µM	131023	95	7,2	31,0 ± 6,4

[a] Mutation frequency per 10^4 colonies

[b] Adduct number per 10^6 nucleotides. Adduct values were determined from duplicate analysis of two plasmid samples/dose except for at 50µM where only 3 values were obtained.

Although the number of adducts increased proportionally with higher 4-

OHtamQM concentrations, the mutation frequency increased only slightly

compared to the untreated control and was not dose-related, as depicted in

Figure 2.15. The highest mutation frequency was induced by 100 µM 4-

OHtam ($9.6/10^4$ colonies), however this was only 1.96-fold higher than the

spontaneous mutation frequency in the combined control plasmid.
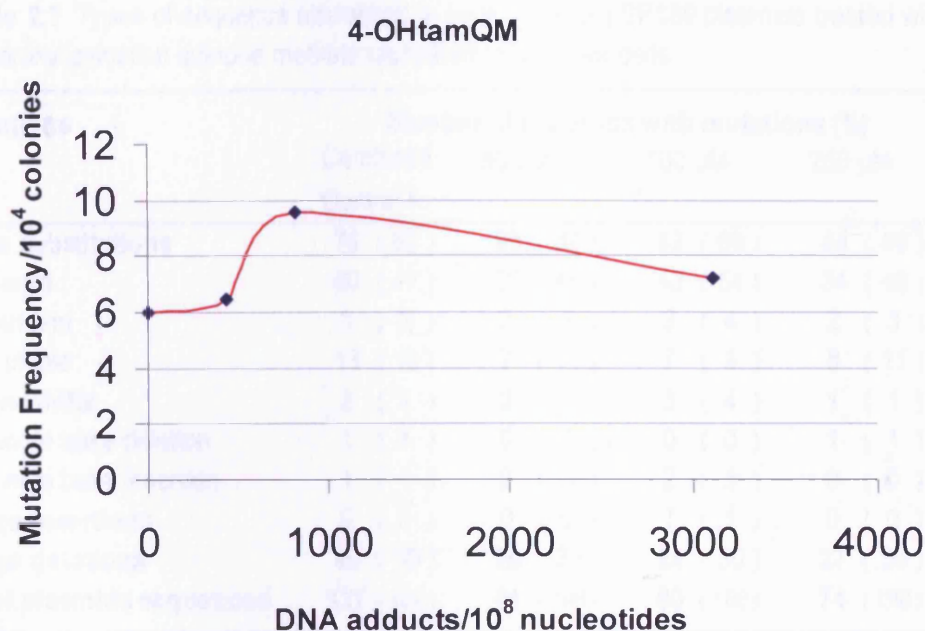
Figure 2.15 Relationship between DNA adduct level and mutation frequency for 4-OHtamQM treated plasmid.

2.3.4.2 Mutation types found in *supF* gene

Table 2.7 shows the types of sequence alterations in the *supF* gene of pSP189 plasmid treated with 4-OHtamQM. For all doses, including the combined control, base substitutions were the major type of mutation observed, accounting for between 60% and 66% of all plasmids sequenced. At all treatment concentrations, including the combined control, single and multiple base substitution mutations were most prevalent, while tandem substitutions were far less frequent.

Table 2.7 Types of sequence alterations in *supF* gene of pSP189 plasmids treated with 4-hydroxytamoxifen quinone methide replicated in Ishikawa cells

Mutations	Number of plasmids with mutations (%)			
	Combined Control [a]	50 µM	100 µM	250 µM
Base substitutions	76 (60)	39 (64)	53 (66)	44 (59)
Single	60 (47)	29 (48)	43 (54)	34 (46)
Tandem	3 (2)	3 (5)	3 (4)	2 (3)
Multiple	13 (10)	7 (11)	7 (9)	8 (11)
Frameshifts	2 (2)	2 (3)	3 (4)	1 (1)
Single base deletion	1 (1)	0 (0)	0 (0)	1 (1)
Single base insertion	1 (1)	2 (3)	2 (3)	0 (0)
Large insertions	0 (0)	0 (0)	1 (1)	0 (0)
Large deletions	49 (39)	20 (33)	24 (30)	27 (36)
Total plasmids sequenced	127 (100)	61 (100)	80 (100)	74 (100)

[a] Combined control data is the same as presented in Table 2.3

Table 2.8 illustrates the types of single base substitution mutations induced by the different treatments with 4-OHtamQM. Transitions were the major type of base substitution mutation in combined control, 50 and 100 µM treated plasmid samples, while in the 250 µM treated plasmid, the major base substitution type were transversions. When the plasmid was treated with 4-OHtamQM, there was a preference for transitions at GC base pairs. Of the transition mutations, GC→AT were the most prevalent, seen in between 47 and 52% of all single base substitution mutations, AT→GC transitions only appearing in three plasmids at the 100 µM treatment. Transversions were less common, with GC→TA transversions accounting for between 23% and 31%, GC→CG for between 14% and 21% and AT→TA for between 3% and 9%.

Table 2.8. Types of single base and tandem substitution mutations in *supF* gene
of pSP189 plasmids treated with 4-hydroxytamoxifen quinone methide
replicated in Ishikawa cells

Mutations	Number of plasmids with mutations (%)			
	Combined Control [a]	50 µM	100 µM	250 µM
Transversions	**28** (46)	**14** (48)	**17** (40)	**18** (53)
GC→TA	14 (23)	9 (31)	10 (23)	8 (24)
GC→CG	12 (20)	4 (14)	7 (16)	7 (21)
AT→TA	2 (3)	1 (3)	0 (0)	3 (9)
AT→CG	0 (0)	0 (0)	0 (0)	0 (0)
Transitions	**33** (54)	**15** (52)	**26** (60)	**16** (47)
GC→AT	33 (54)	15 (52)	23 (53)	16 (47)
AT→GC	0 (0)	0 (0)	3 (7)	0 (0)
Total single base substitutions	**61** (100)	**29** (100)	**43** (100)	**34** (100)

[a] Combined control data is the same as presented in Table 2.4

A graphical representation of the different single and tandem base substitution
mutation types induced by 4-OHtamQM relative to the combined control
mutations is illustrated in Figure 2.16. As the dose of 4-OHtamQM increases
the proportion of substitutions did not alter significantly compared to the
combined control.

Figure 2.16 Bar chart illustrating the proportion of different single base and tandem substitution mutations in the *supF* gene observed in combined control mutants and in those induced by treatment with 4-OHtamQM (50, 100 and 250 μM).

Table 2.9 shows the types of multiple substitution mutations induced in plasmids and reveals a preference for transitions in all treated samples, particularly at the highest dose, but not in the combined control. In addition, all transitions were of the GC→AT type in all plasmids, including the control. In the untreated combined control plasmids the proportion of transversions (54%) was slightly higher than that of transitions (46%), with GC→TA and GC→CG transversions accounting for 39% and 14% of all substitutions respectively and with GC→AT transitions accounting for 46%.

Table 2.9 Types of multiple base substitution mutations in *supF* gene of pSP189
plasmids treated with 4-hydroxytamoxifen quinone methide replicated in Ishikawa cells

Mutations	Number of plasmids with mutations (%)			
	Combined Control [a]	50 µM	100 µM	250 µM
Transversions	**15** (54)	**10** (48)	**8** (44)	**2** (8)
GC→TA	11 (39)	6 (29)	2 (11)	2 (8)
GC→CG	4 (14)	3 (14)	6 (33)	0 (0)
AT→TA	0 (0)	1 (5)	0 (0)	0 (0)
AT→CG	0 (0)	0 (0)	0 (0)	0 (0)
Transitions	**13** (46)	**11** (52)	**10** (56)	**24** (92)
GC→AT	13 (46)	11 (52)	10 (56)	24 (92)
AT→GC	0 (0)	0 (0)	0 (0)	0 (0)
Total multiple base substitutions	**28** (100)	**21** (100)	**18** (100)	**26** (100)

[a] Combined control data is the same as presented in Table 2.5

There was a slight increase in the percentage of transition substitutions in plasmids treated with the lower and intermediate doses (52% and 56% of all substitutions respectively) and a much more pronounced increase in the highest dose (92% of all substitutions). In contrast, there was a decrease in the proportion of transversion mutations that mirrored the increase in transitions as the dose of 4-OHtamQM increased. All of the transitions and transversions occurred at GC basepairs, in all plasmids (including the untreated control samples) except for one plasmid at the lower dose (50 µM), in which the mutation occurred at an AT basepair.
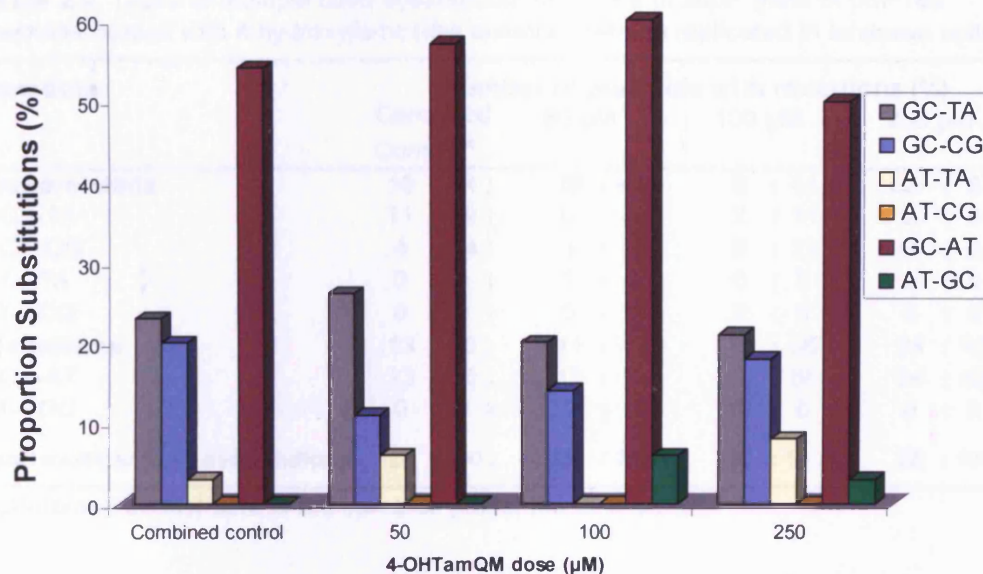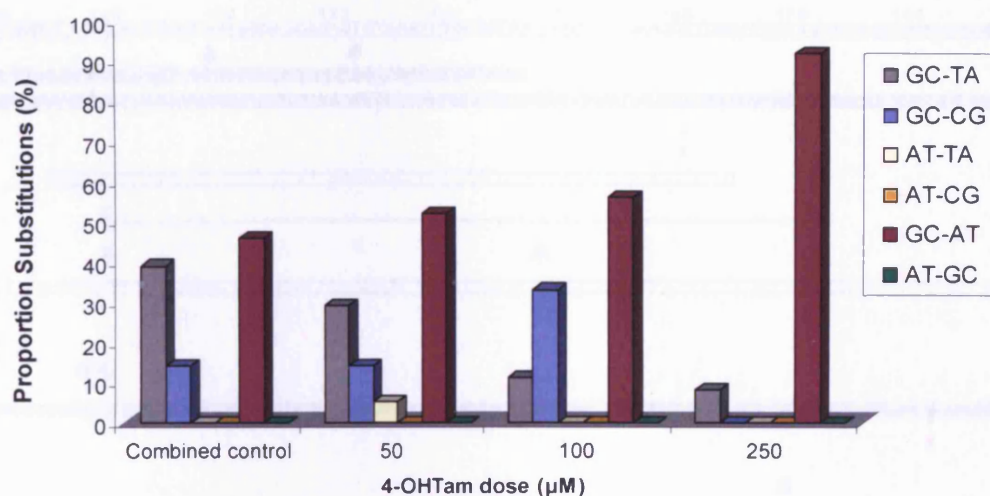
Figure 2.17 Bar chart illustrating the proportion of different multiple base substitution mutations in the *supF* gene observed in combined control mutants and those induced by treatment with 4-OHtamQM (50, 100 and 250 µM).

The proportion of multiple substitutions induced by 4-OHtamQM is illustrated graphically in Figure 2.17. No definite trends can be seen in relation to increasing dose. The only exception seems to be the highest dose (250 µM), in which there was a high preference for GC→AT transitions that was not evident at lower doses.

Frameshift mutations were relatively uncommon, only accounting for up to 4% of all mutations in the treated plasmid, and 2% in the combined control (Table 2.7). Deletions were more prevalent, especially large deletions where greater than two adjacent bases were lost. The proportion of large deletions was similar at all doses, including the combined control and ranged from 30% to 39%. Figure 2.18 illustrates the different insertions and deletions, including their sequence and position in the *supF* gene.
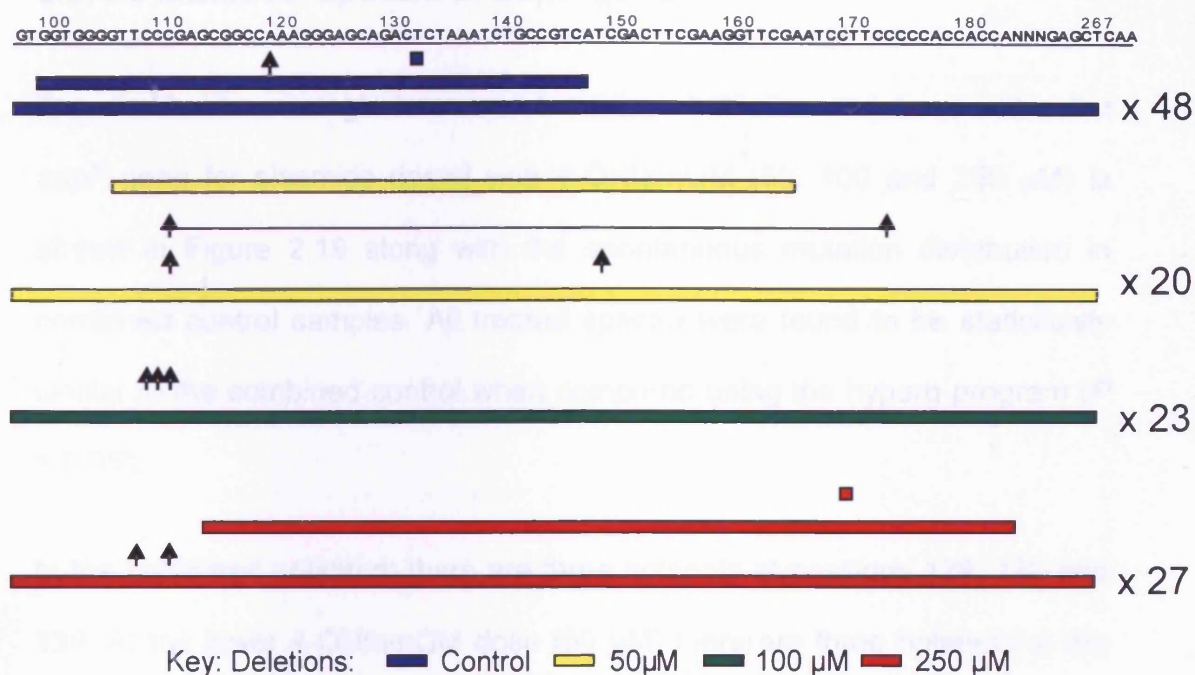
Figure 2.18 Deletions and insertions in the *supF* gene found in combined control mutants and in those induced by treatment with 4-OHtamQM (50, 100 and 250 µM). Numbers on the right indicate the number of plasmid sequences with the same large deletion of bases 97→267.

2.3.4.3 Mutation spectra in *supF* gene

The distribution of single base and tandem substitution mutations within the *supF* gene for plasmids dosed with 4-OHtamQM (50, 100 and 250 µM) is shown in Figure 2.19 along with the spontaneous mutation distribution in combined control samples. All treated spectra were found to be statistically similar to the combined control when compared using the hyperg program ($P > 0.05$).

In the combined spectrum there are three hotspots at positions 129, 133 and 139. At the lower 4-OHtamQM dose (50 µM) there are three hotspots at the same positions as in the control (129, 133 and 139) plus one hotspot at position 123. At the intermediate dose (100 µM) there are four hotspots, two of which are also seen in the combined control (129 and 139) plus two at positions 108 and 168. At the highest dose (250 µM) there are three hotspots that are also present in the combined control (positions 129, 133 and 139) and a new one at 124. If all three treatments are combined there are now five hotspots at positions 108, 129, 133, 139, and 168 but only two are not present in the control spectrum (Fig. 2.20). All mutation hotspots are at positions of GC basepairs.

50µM 4-OHtamQM

```
100       110       120       130       140       150       160       170       180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
     T     TT            T     T A   AT   A    A      A       TA  AT
                         T     T A   T                        TA
                         A     A G   T
                         *     A G   T
                         A     * T
                         C       A
                         *       G
                                 *
```

100µM 4-OHtamQM

```
100       110       120       130       140       150       160       170       180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
  AAT    TA    A      T   T G GC   T A  TCA  T TC T   C        TT  TT
         TA               A G   T A   A      A                 T   T
         TG               A G   T A                            T
         T                A     A                              T
         *                A     A                              *
                          A     *
                          C
                          *
```

250µM 4-OHtamQM

```
100       110       120       130       140       150       160       170       180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
 T  AT A   TTA      G T  TA   GA  G    T G            A   T A C   TT AA
    A               AA   T    G       A              A           T
                    A    A    G
                    A    A    *       *
                    *    *
```

Figure 2.19 Mutation spectra depicting single base and tandem substitutions induced in the *supF* gene observed in the combined untreated control or after treatment with 4-OHtamQM (50, 100 and 250 µM). Red asterisks denote hotspots. Tandem mutations are underlined.

Combined 4 -OHtamQM Treatments + Combined Control

```
                       *
                       C
                       C
                       C
                       A
                       A                 *
                       A      *          A
                       A      G          A
                       A      G          A
                       A      G          T
                       A      G          T
AAAA    A              C      A   A      T      A              A     T               A
AAAA    TTA    A A   AAT G T  AA      T       A A      GA T    A     TC    T     T  AA
100       110       120      130       140        150        160        170        180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
T  ATTA   TTA    A G T  TT   GA G AC     AT AG TCA   T TC TA   CT ATA   ATT ATT
   AT A   TT                 TA   T   A      T      A A   A A        TA    TT  T
     A    TA                 TA   T   G      T      A               C     T    T
          TA                 AA   T   G      T                      T     A
          TG                 AA   T   G      T                      T
          T                       A G G                             T
          *                       A G T                                      *
                                  A   T
                                  A G A
                                  A G A
                                  A   *
                                  A   A
                                  A   A
                                  A   G
                                  A   G
                                  A
                                  A   *
                                  A
                                  C
                                  C
                                  *
```
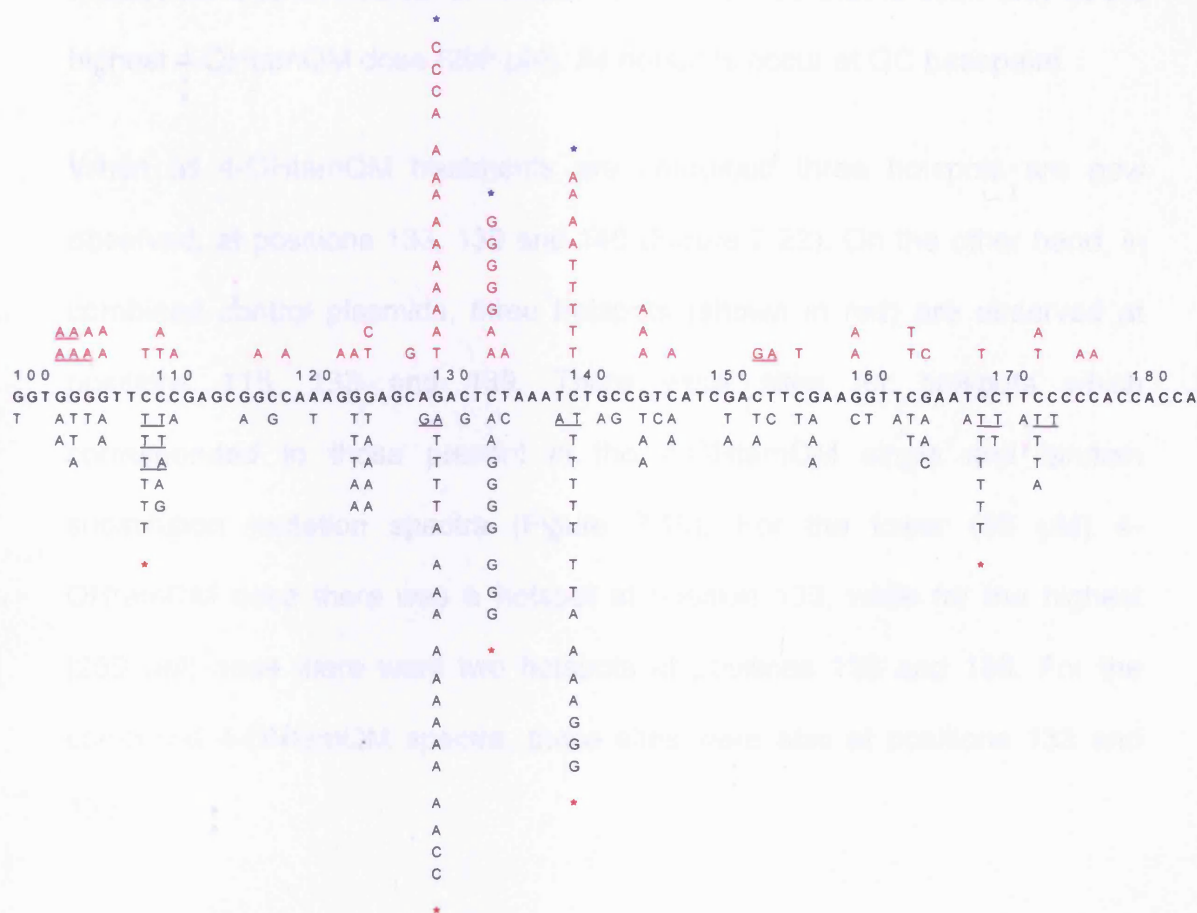
Figure 2.20 Mutation spectrum depicting single base and tandem substitutions induced in the *supF* gene after treatment with 4-OHtamQM (all doses combined). Non-treated combined control single base and tandem substitutions are shown in red above the gene sequence. Hotspots are represented by asterisks. Tandem mutations are underlined.

The distribution of multiple substitution mutations within the *supF* gene for plasmids dosed with 4-OHtamQM (50, 100 and 250 μM) or control treatments is shown in Figure 2.21. There is a single mutation hotspot in one control (solvent only) treated plasmid at position 133. This hotspot is also present in the multiple mutation spectrum induced by all 4-OHtamQM treatments.

Finally, there is an additional hotspot at position 139 that is seen only at the highest 4-OHtamQM dose (250 µM). All hotspots occur at GC basepairs.

When all 4-OHtamQM treatments are combined three hotspots are now observed, at positions 133, 139 and 146 (Figure 2.22). On the other hand, in combined control plasmids, three hotspots (shown in red) are observed at positions 115, 133 and 139. There were sites for hotspots which corresponded to those present in the 4-OHtamQM single and tandem substitution mutation spectra (Figure 2.19). For the lower (50 µM) 4-OHtamQM dose there was a hotspot at position 133, while for the highest (250 µM) dose there were two hotspots at positions 133 and 139. For the combined 4-OHtamQM spectra, these sites were also at positions 133 and 139.

Solvent Control

```
100        110        120        130        140        150        160        170        180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
                     A   T   T   G       T   T   A       T   C   T     A     G
                     T           G       A                                 A
                                 A
                                 *
```

Water Control

```
100        110        120        130        140        150        160        170        180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
           T         TA      A  AA  A     T           T
           A                              T
```

50μM 4-OHtamQM

```
100        110        120        130        140        150        160        170        180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
           T                 T    TT A  T    T   T   T         A   T   G       T
           A                 A    A   T     A       A
                                      G
                                      G
                                      *
```

100μM 4-OHtamQM

```
100        110        120        130        140        150        160        170        180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
           T                 T   G  T    T   A   A     T         T         T
                                 G       T       T                         T
                                 G
                                 G
                                 G
                                 G
                                 *
```

250μM 4-OHtamQM

```
100        110        120        130        140        150        160        170        180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
T          T      T    T   A    A  T    T      T  T           T   T       T
                                 T    T        A  T           T   T
                                 T    T                           T
                                 T    T
                                 T    T
                                 *    *
```

Figure 2.21 Mutation spectra in the *supF* gene depicting multiple substitutions induced after treatment with 4-OHtamQM (50, 100 and 250 μM), or water or solvent only. Asterisks denote hotspots.

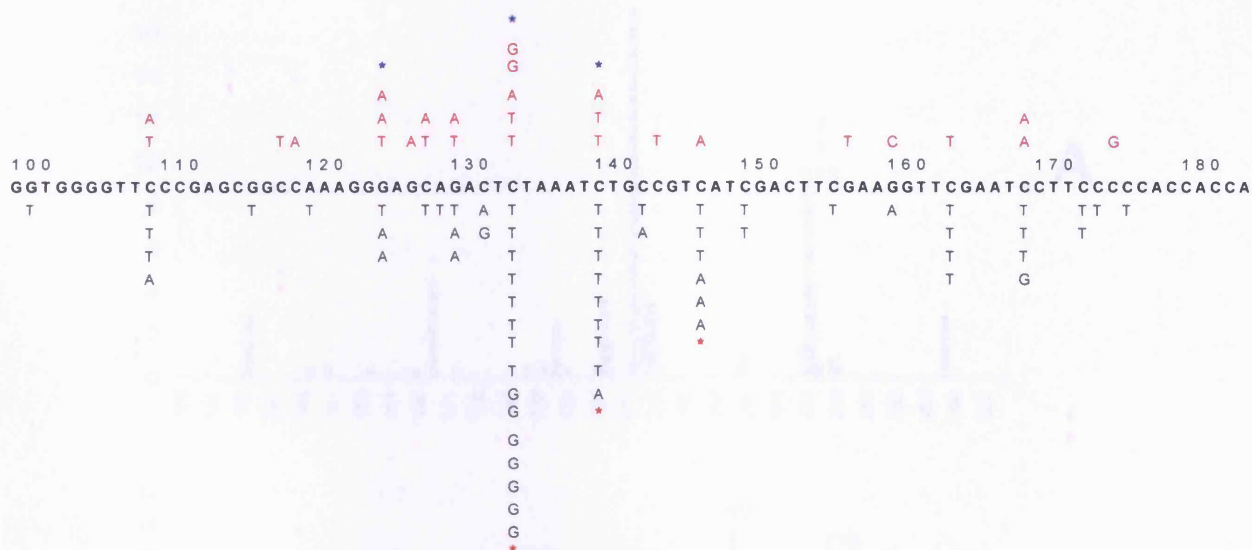Combined 4 -OHtamQM Treatments + Control

```
                                        *
                                        G
                                        G
                          *             A           *
                          A    A  A     A           A
               A          A    A A A    T           T                            A
               T     TA   T    A T T    T           T    T    A          T  C  T  A       G
100       110       120       130       140       150       160       170       180
GGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACCACCA
T          T       T    T     T    TTT A T      T    T    T         T  A  T     T    TT T
T                        A    A G T      T    A  T    T              T     T     T
T                        A    A   T      T       T                   T     T
A                                 T      T       A                   T     G
                                  T      T       A
                                  T      T       A
                                  T      T       *
                                  T      T
                                  G      A
                                  G      *
                                  G
                                  G
                                  G
                                  G
                                  *
```

Figure 2.22 Mutation spectra depicting multiple substitutions induced in the *supF* gene after treatment with 4-OHtamQM (50, 100 and 250 µM) or present in the combined control samples. Non-treated combined control multiple base substitutions are shown in red above the gene sequence. Asterisks denote hotspots.

2.3.5 Part C: In silico prediction of p53 G→T mutation spectra

The LwPy53 algorithm was applied to *supF* mutation data generated by treatment of methylated pSP189 with the two highest α-acetoxytamoxifen doses (25 and 50µM), all three 4-OHtamQM doses (50, 100 and 250 µM) or by spontaneous means (combined control). The resulting predicted mutation distributions are illustrated in Figures 2.23, 2.24 and 2.25.
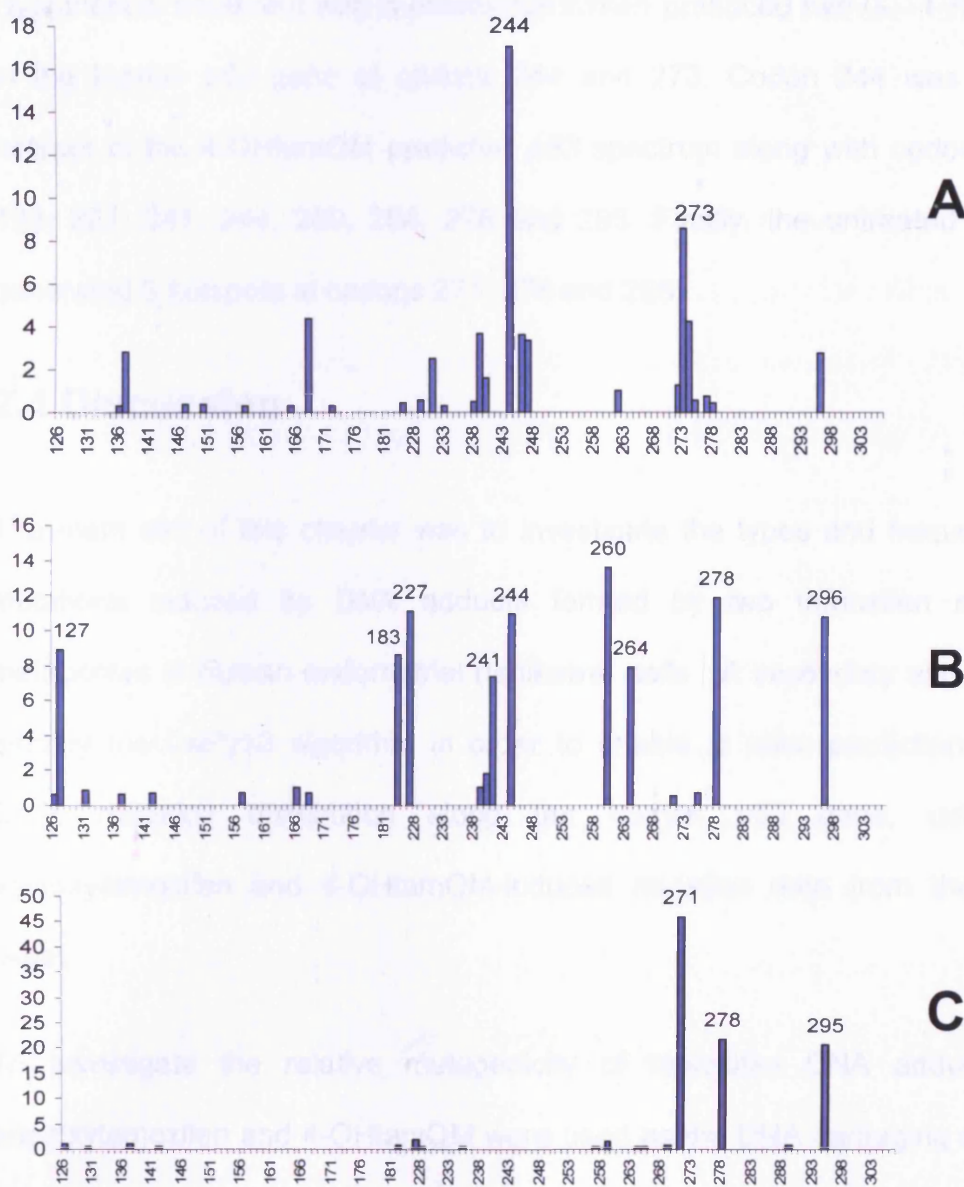
Figure 2.23 Predicted human *p53* G→T mutation spectrum for exons 5, 7 and 8 from data generated using methylated *supF* plasmid treated with α-acetoxytamoxifen (A), 4-OHtamQM (B) or untreated combined control (C). Hotspots were defined as sites containing 5% or more of the mutations in the spectrum. The vertical axes show the percentage of G→T mutations for each site relative to the overall number of G→T mutations for that particular spectrum.

Hypothetical treatment with α-acetoxytamoxifen produced two G→T hotspots in the human *p53* gene at codons 244 and 273. Codon 244 was also a hotspot in the 4-OHtamQM predicted *p53* spectrum along with codons 127, 183, 227, 241, 244, 260, 264, 278 and 296. Finally, the untreated control generated 3 hotspots at codons 271, 278 and 295.

2.4 Discussion

The main aim of this chapter was to investigate the types and frequency of mutations induced by DNA adducts formed by two tamoxifen reactive metabolites in human endometrial (Ishikawa) cells. A secondary aim was to employ the LwPy53 algorithm in order to enable *in silico* prediction of the G→T mutation distribution along the human *p53* gene, using α-acetoxytamoxifen and 4-OHtamQM-induced mutation data from the *supF* assay.

To investigate the relative mutagenicity of tamoxifen DNA adducts, α-acetoxytamoxifen and 4-OHtamQM were used as the DNA damaging agents. α-Acetoxytamoxifen is a model of the reactive tamoxifen sulphate ester metabolite [280], and forms the same α-(N^2-deoxyguanosinyl)-tamoxifen DNA adduct as seen in rats [277, 305], mice [295, 308], monkeys [306, 307] and women undergoing tamoxifen treatment [320, 321]. 4-Hydroxytamoxifen is one of the major metabolites of tamoxifen [370], which via oxidation to its quinone methide, is proposed to form one of the minor tamoxifen DNA adducts (α-(N^2-deoxyguanosinyl)-4-hydroxytamoxifen) in rat liver [311].

2.4.1 Part A: Mutagenicity of plasmid treated with α-acetoxytamoxifen replicated in Ishikawa cells

Replication of α-acetoxytamoxifen adducted plasmid in human endometrial (Ishikawa) cells resulted in an increase in mutation frequency above background control levels. The mutation frequency increased sharply at higher doses, which is in contrast to an earlier study from this laboratory carried out in *E-coli* that reported a lack of correlation between gross adduct number on the pLIZ lambda shuttle vector and mutagenicity in the *lacI* gene (Lowes, *et al.,* 1999).

In the present study, for the lower α-acetoxytamoxifen dose (10μM), there was a small relative increase (~3-fold) in mutation frequency above the untreated control. In contrast, mutation frequency was sharply increased with higher doses, suggesting that the ultimate biological effect of tamoxifen may depend not only on the number of dG-N^2-tamoxifen adducts originally formed, but also on processes such as cellular DNA repair, which may remove these lesions from the plasmids before DNA replication occurs. Higher damage levels may saturate the available repair mechanisms *in vitro*, resulting in a non-linear increase in mutation frequency. The mutation frequency of α-acetoxytamoxifen in this investigation was consistent with results from previous studies in our laboratory. In particular, treatment of plasmid with the highest α-acetoxytamoxifen dose (50 μM) induced about 22 x 10^{-4} mutant colonies on an adduct per plasmid basis. These results are comparable to results reported in Ad293 cells (~14 x 10^{-4}) [350] and GM00637 fibroblasts (~4.5 x 10^{-4}) [351] using the same concentration of α-acetoxytamoxifen (50

µM). Additionally, the adduct levels were comparable between the studies which suggests that the methods used are reproducible, providing added confidence in the results. Therefore, α-acetoxytamoxifen appears to be similarly mutagenic regardless of the cells the damage is replicated in, although this was not the case for 4-OHtamQM. Previously published *supF* studies using different bulky carcinogens have reported lower levels of mutagenicity. Specifically, when compared on an adduct per plasmid basis, α-acetoxytamoxifen (in the current study) was about 39 times more mutagenic than benzo[c]phenanthrene [371], 22 times more mutagenic than 1,6-dinitropyrene [372], and 9 to 18 times more mutagenic than BPDE [373, 374]. These results indicate that α-(N^2-deoxyguanosinyl)-tamoxifen is an extremely potent mutagenic lesion.

In the current study the prevalent substitution mutations in control plasmids were GC→AT transitions. These results are consistent with previously reported *supF* mutation spectra in human Ad293 cells [375, 376], lymphoblasts [377] and monkey kidney cells [378], in which the GC→AT transition was the preferred spontaneous mutation. Upon treatment with α-acetoxytamoxifen, GC→TA transversions become the major mutation type for α-acetoxytamoxifen single base substitutions and were induced in a dose-dependent manner. GC→TA transversions are the preferred mutations induced by dG-N^2-TAM adducts, as established by site-specific mutation studies in COS-7 cells [346], previous supF experiments [350, 351], primer extension reactions *in vitro* [347, 352] and from *in vivo* investigations in transgenic rats [348, 349]. Examination of the mutational specificity of polyaromatic hydrocarbons and other compounds which form DNA blocking

lesions, has suggested that the preferential incorporation of A opposite a bulky lesion allows polymerase extension of the inserted A (the "A rule") [379]. A straightforward explanation of the A rule is that in the presence of a bulky adduct or when a base is missing, some structural property of the template often allows A to be preferentially inserted [380, 381]. An alternative (and not exclusive possibility) is that polymerases preferentially bind adenine nucleotides and are therefore more likely to insert an A [382]. As well as misincorporation of A opposite the dG-N^2-tamoxifen adduct, misincorporation of G and, in some sequences, T was also shown to be possible, albeit at lower frequency [346, 347, 352]. The occurrence of these misincorporation events, would lead to induction of GC→CG and GC→AT mutations, respectively.

Single and multiple base substitutions have been presented on separate spectra because it has been suggested that these multiple mutations may arise through a different mechanism to single base mutations [369]. Multiple mutations are defined as two or more point mutations occurring at distinct positions of a gene (not in tandem). For all treated and untreated plasmid spectra, hotspots were at GC base pair sites presumably as a result of preferential adduct formation on deoxyguanosine, including spontaneous endogenous lesions.

In the case of single and tandem substitution spectra, position 129 was a hotspot in all α-acetoxytamoxifen and untreated control spectra, except for the top α-acetoxytamoxifen dose (50 μM) spectrum. The same position (129) was the only hotspot observed in a previously reported *supF* spontaneous

spectrum, after replication of pSP189 in Ad293 cells [376]. In the present

study, GC→AT transitions constituted 71% of the substitutions (observed at

this position) in the combined untreated control, while in the combined α-

acetoxytamoxifen spectrum, GC→TA transversions were predominant at this

site (80%). The difference observed in the mutational specificity of this hotspot

suggests that position 129 is mutated both spontaneously and due to the

presence of the dG-N^2-TAM DNA adducts induced by α-acetoxytamoxifen but

different nucleotides are incorporated opposite the lesion. Positions 133 and

139 were the two other hotpspots present in the combined control spectrum

but were not observed in any of the α-acetoxytamoxifen spectra. Therefore, it

is likely that mutations at these two positions were purely spontaneous. This

conclusion is further supported by a recent comparative analysis of

mammalian cell line spontaneous mutation spectra in *supF* [383], in which

hotspot C133 was the most frequent hotspot and hotspots G129 and C133

also occurred frequently.

Using the hyperg analysis program, only the mutation spectra from the two top

α-acetoxytamoxifen doses were found to be statistically different from the

combined control ($P<0.05$). When these spectra were combined into a single

spectrum (Figure 2.13) six hotspots were evident at positions 115, 123, 129,

159, 160 and 168. The hotspot at position 159 was present in both the

intermediate (25 μM) and high dose (50 μM) α-acetoxytamoxifen single

spectra. Positions 159 and 160 were also hotspots in the combined single and

tandem α-acetoxytamoxifen mutation spectrum of a previous *supF* study

carried out in this lab [350].

With regard to multiple base substitutions, three hotspots were present in the combined control spectrum at positions 124, 133 and 139 (Figure 2.14). Two of these hotspots (133 and 139) were also present in the single and tandem control mutation spectrum. Upon treatment with α-acetoxytamoxifen, two unique hotspots were induced at positions 102 and 117. However, multiple mutation spectra from α-acetoxytamoxifen treated plasmids failed to differentiate from the control when analysed by hyperg (P>0.05).

Frameshift mutations were relatively uncommon in both the control and α-acetoxytamoxifen treated samples. Nevertheless, there were 7 single base deletions and 3 single base insertions at position 124. Notably, a large deletion of bases 97→267 was present in 39% of the combined control mutants sequenced. The frequency of this large deletion decreased markedly upon treatment with α-acetoxytamoxifen and accounted for 29%, 9% and 1% of all mutations in the 10, 25 and 50 μM treated plasmids respectively as the mutation pattern became more characteristic of tamoxifen induced damage and was less influenced by the spontaneous alterations. The deletion was also observed in recent spontaneous *supF* spectra from another member of our lab (Tompkins et al, unpublished data). The nature of this background mutation is uncertain, although there is strong evidence that short direct repeats mediate deletions and duplications in DNA [384, 385]. Two possible mechanisms for these events are (i) recombination between short homologous repeats and (ii) DNA polymerase slippage between short repeated sequences. Short-homology-independent illegitimate recombination (SHIIR) [386] occurs spontaneously or is induced, by UV light or other DNA-damaging agents, and requires short regions of homology between

recombination sites. These regions usually contain 4–13 bp of homologous sequences [387, 388]. One such candidate site of sequence homology in the junction of the 171bp deletion observed in the present study is a 5bp sequence (GAGCT). Furthermore, since this short sequence is positioned next to the 8bp signature of the pSP189 plasmid, a potentially larger homologous sequence could be formed, up to 13bp.

2.4.2 Part B: Mutagenicity of plasmid treated with 4-OHtamQM replicated in Ishikawa cells

Treatment of methylated pSP189 with 4-OHtamQM failed to raise the mutation frequency considerably above the background control levels at all three doses. This finding is in sharp contrast to the mutation frequencies reported in a previous *supF* investigation of 4-OHtamQM in Ad293 cells by a member of our lab [350], in which 4-OHtamQM induced damage was found to be about 25 times more mutagenic than α-acetoxytamoxifen derived adducts on an adduct per plasmid basis. Interestingly, the adduct levels were remarkably similar between the two studies, indicating that the differences in induced mutagenicity were not due to differences in the extent of adduct formation.

Although the reason(s) behind this discrepancy are not known, several possible explanations could account for it. First, a major difference between the two investigations is that methylated plasmid was used in the current study, while plasmid was adducted with 4-OHtamQM in the absence of methylation in the previous study. It is possible therefore, that DNA

methylation at CpG sites could potentially interfere with the adduction reaction or the repair of induced damage in the cells. The adduction reaction could probably be affected in terms of adduct distribution but not in terms of overall adduct level (see comment above). Moreover, the types of adducts formed by 4-OHtamQM in methylated DNA might be different and possible crosslink formation induced by 4-hydroxytamoxifen derived adducts in unmethylated plasmid [350] is less likely to occur with CpG methylation. Earlier reports indicate that methylation of CpG sequences may affect not only binding of chemically reactive intermediates with DNA, but also the conformational properties of the resultant DNA adducts formed [389, 390]. Another important difference between the two studies is that plasmids adducted with 4-OHtamQM were replicated in different cell lines (Ishikawas vs Ad293). It is known that cell specific factors (e.g. different polymerase profiles) can affect the mutagenic processing of DNA adducts [391]. Finally, a third possibility is that reactive species (such as free radicals) generated from the 4-OHtamQM during the adduction of plasmid may have been more abundant in the previous study [350], leading to an overestimation of the mutagenic potency of 4-OHtamQM. 4-OHtamQM is believed to be a common reactive intermediate produced by microsomal, chemical and peroxidase activation of 4-OHtam and is responsible for the formation of identical 4-hydroxylated dG-N^2tam adducts in each system [392, 393]. Dimers of 4-hydroxytamoxifen formed via an undefined mechanism involving free radical generation have been detected upon incubation with both rat liver microsomes [393] or horseradish peroxidase [392]. Similar radical species may also arise as a consequence of chemical oxidation and could potentially cause other types of DNA damage

such as strand breaks, in addition to the 4-OHtam-dGN^2 adducts. However, the fact that the adduction reaction and synthesis of the 4-OHtamQM protocols were identical in each investigation makes this possibility unlikely.

The types of single-tandem and multiple substitutions generated by 4-OHtamQM treatment at all doses were in similar proportion to the types of mutations seen in the untreated control. In addition, the patterns of mutation distribution along the *supF* gene were virtually indistinguishable between the 4-OHtamQM treatments and the untreated control. Accordingly, all 4-OHtamQM mutation spectra (both single-tandem and multiple) were found to be statistically similar to the untreated control spectra when compared by Hyperg ($P>0.05$). Furthermore, the 4-OHtamQM and untreated control shared 3 hotspots in the combined single-tandem spectrum and 2 in the combined multiple spectrum. Altogether, these results suggest that treatment of methylated plasmid with 4-OHtamQM produced no real effect on the background spontaneous mutagenic specificity.

2.4.3 Part C: In silico prediction of p53 G→T mutation spectra

The *supF* assay is extremely valuable for quantifying the relative mutagenicity of different chemicals and also for providing a qualitative insight into the types and distribution of mutations in the *supF* gene. However, this is not a biologically important gene, just a marker for mutagenesis. Moreover, these experiments do not provide information on whether tamoxifen derived DNA

adducts cause mutations in human genes that might be involved in tamoxifen induced endometrial carcinogenesis.

Application of the LwPy53 algorithm represents an attempt to predict the mutational specificity of a mutagen in a biologically relevant gene, based on experimentally generated *supF* data. The *m*utation data from the *supF* assay experiments described in parts I and II were used as input data for the LwPy53 algorithm that predicts chemically-induced hotspots along the human *p53* gene. Single, tandem and multiple substitutions, along with single deletions and insertions were recorded for each nucleotide position of the *supF* gene. These mutation data were subsequently utilized by the LwPy53 algorithm to predict *in silico* the distribution of G→T transversions along exons 5, 7 and 8 of *p53*.

Based on the *p53* G→T transversion predictions, treatment with α-acetoxytamoxifen resulted in two hotspots at positions 244 and 273. Codon 273 is a unique hotspot, while hotspot 244 is also present in the 4-OHtamQM *p53* spectrum. Position 273 is a common hotspot for bulky adducts such as benzo[a]pyrene [394] and has been shown to reside within regions of slow DNA repair [395]. Furthermore, characteristic mutations in the *p53* gene have also been found in rat hepatocarcinomas induced by tamoxifen administration [396]. Therefore, it is plausible that the presence of this hotspot in *p53* genes of women with tamoxifen associated endometrial cancer could imply a bulky adduct, such as dG-N^2-TAM plays a role in mutation induction and cancer development. The extensive IARC TP53 mutation database [397] contains over 15 000 *p53* somatic mutations recorded for many different cancer types.

However, the record for endometrial cancer is incomplete, with little data available and there is no distinction between tamoxifen treated and control patients. A future study comparing the pattern of *p53* mutations in spontaneous and tamoxifen induced endometrial tumours could address these issues.

Interestingly, there was only one hotspot shared between the 4-OHtamQM and combined control *p53* spectra. One would expect a greater similarity between the two spectra, based on the *supF* results. However, LwPy53 predicts mutations based on different types of base substitution frequencies at different dinucleotides. Therefore, even if the *supF* spectra were not significantly different using the very stringent Hypergeometric test one would still get different spectra predicted from the control and 4-OHtamQM spectra because certain mutation types occur at certain dinucleotides. If the 4-OHtamQM spectrum is aligned with the control, one can see that the 4-OHtamQM has a larger proportion of G>T mutations at GG dinucleotides. Subsequently, when interpreting hotspots in the *p53* spectra one needs to look at the frequency of that mutation type in the *supF* spectrum and what dinucleotides they fall within. The same thinking should be applied to all predicted mutations. Mutation hotspots can only occur in the predicted *p53* spectrum if a mutation in that nucleotide leads to an amino acid change (i.e. selectable) and is exposed in a nucleosome (i.e. can be accessed by the mutagen). Also the probability of mutation is increased if the nucleotide lies in a very inflexible region estimated by DNA curvature. This is why hotspots are not seen in some places where just by sequence alone might have been expected.

2.4.4 Summary

The mechanisms through which tamoxifen causes endometrial cancer are not yet established but the potential contribution of tamoxifen-induced DNA damage is a subject of much interest. The overall aim of this work was to investigate the mutagenicity and importance of tamoxifen DNA adducts in human endometrial cells. Relative mutation frequencies increased proportionally with the adduct level for α-acetoxytamoxifen treated plasmid, up to ~18 times the background frequency. One could conclude that the α-acetoxytamoxifen metabolite of tamoxifen has the potential to cause serious mutagenic damage to DNA. However, 4-OHtamQM induced damage did not cause a dose-related increase in mutation frequency above that of the untreated control, which contrasts with previous studies from this laboratory [350, 365]. Additionally, only α-acetoxytamoxifen generated statistically different *supF* mutation spectra compared to the spontaneous pattern; the lack of mutagenicity demonstrated for 4-OHtamQM induced damage was consistent with a mutation spectrum similar to the spectrum observed in control untreated plasmid. Differences in the cellular background and/or methylation status of the plasmid could account for this disagreement. Finally, the LwPy53 algorithm was employed to predict G→T transversion hotspots in the human *p53* gene based on the experimental *supF* data. Codons 244 and 273 were the predicted G→T hotspots for the α-acetoxytamoxifen-treated plasmid. This suggests that if tamoxifen DNA adducts are directly involved in the initiation of human cancer then mutations may preferentially occur at these sites in endometrial tumours of treated women.

CHAPTER 3

Development of a novel site-specific

mutation assay in *E.coli*

3.1 INTRODUCTION

3.1.1. The O^6-methylguanine (O^6-MeG) DNA adduct

Many alkylating agents (described in Section 1.2.2.1.3) are known or suspected human carcinogens whose biological properties and mechanism of action have been extensively explored in cell and animal models over several decades [398]. Methylating agents are a type of alkylating agents that react with DNA by delivery of a methyl (-CH3) group and are capable of reacting with a number of nucleophilic sites on DNA (Figure 3.1). Methyl methane sulfonate (MMS), dimethyl sulfate (DMS), N-methyl-N-nitrosourea (MNU), and N-methyl-N-nitrosoguanidine (MNNG) have been used extensively as direct-acting methylating agents. Methylation damage from both endogenous [399-401] and exogenous [402] sources, as opposed to methylation by enzymatic reactions (e.g., 5-methylcytosine, N^6-methyladenine), can result in mutagenicity, genotoxicity and cytotoxicity [403]. A critical lesion responsible for these endpoints is O^6-MeG, which is induced in minor amounts (maximally 8% of total alkylation products) in the DNA of cells exposed to methylating agents [404]. Wide variations in levels of O^6-MeG have been detected in human DNA in several studies of various populations exposed to methylating agents (reviewed in [405, 406]). The levels of DNA damage observed in human studies as a result of exogenous exposures (noniatrogenic), such as tobacco smoke [407, 408] or nitrated-treated food [409, 410] is of the order of 1 adduct per 10^7-10^9 normal DNA bases, whereas that arising from endogenous exposures may potentially be several orders of magnitude higher [406].
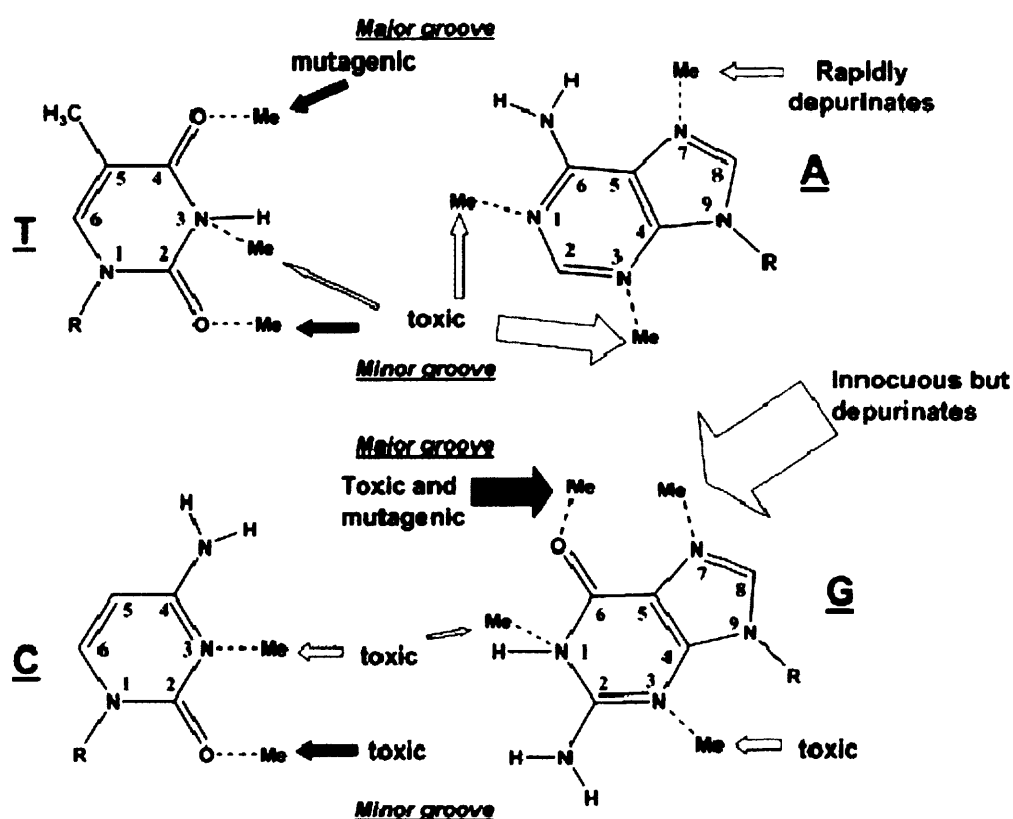
Figure 3.1 Potential sites of chemical methylation in double-strand DNA. The arrows point to each methyl adduct and whether the adduct is known to be predominantly toxic or mutagenic. The open arrows represent sites that are methylated by MMS, MNNG, and MNU. The filled arrows point to sites that are methylated by MNNG and MNU but not detectably by MMS. The size of the arrows roughly represents the relative proportion of adducts. Adapted from [411].

O^6-MeG can be responsible for a wide variety of biological effects, such as recombination and toxicity but the precise molecular mechanisms of these effects remain to be defined. However, the miscoding properties of this adduct are well established. The mutagenic capacity of O^6-MeG lies in its ability to direct the misincorporation of thymine during replication [412-417]. Although

cytosine is still the best Partner for O^6-MeG with regard to Watson-Crick base pairing, pairing with C or T both cause structural distortions in the helix [413, 418]. The frequency of inserting a T opposite this adducted base can be 7-fold higher than insertion of C and the difference is dependent on sequence context [419]. As a result, G→A transition mutations arise following two rounds of replication of DNA containing O^6-MeG. Furthermore, O^6-MeG does not cause an overt block to *in vitro* replication using *E.coli* Pol I [420], but it can slow the rate of polymerization [421]. Both prereplicative and postreplicative DNA repair systems are involved in the repair of O^6-MeG DNA adducts.

Prereplicative repair of O^6-MeG is performed by the DNA repair protein O^6-methylguanine-DNA-methyltransferase (MGMT). This protein contains an active cysteine residue which facilitates the removal of, and acts as a recipient for methyl groups from O^6-MeG in DNA [422]. The protein becomes inactivated during the process, as the demethylation reaction has a 1:1 stoichiometry and is irreversible. Therefore, the capacity of adapted cells for O^6-MeG repair is determined by the number of preformed MGMT molecules. A coordinated inducible adaptive response to DNA methylation is not a feature of mammalian cells. In cultured mammalian cell lines the MGMT levels are low and are simply depleted, rather than induced, by chronic dosage. In contrast, in bacteria, there is a several hundred-fold increase in the levels of MGMT in response to methylating agents [423]. It is generally thought that the repair of methylated DNA is mediated, for the most part, by MGMT. However, when the *E. coli* genome contains a high level of methylation damage a large proportion of O^6-MeG is repaired by the *uvr* nucleotide excision repair

pathway [424]. Despite this, the *uvr* repair pathway only provides a moderate amount of protection against methylation induced mutation.

Postreplicative repair of O^6-MeG lesions is mediated by the MMR system. The cytotoxic cellular responses to O^6-MeG turn out to be dependent upon a functional MMR repair system [425]. The major function of MMR is to correct replication errors that escape proofreading at the replication fork. In addition to alkyltransferase repair, mismatch repair is known to play a role in modulating the mutagenicity and genotoxicity of O^6-MeG residues. Defective mismatch repair is known to be associated with tolerance to the cytotoxic effects of agents that produce O^6-MeG in DNA in both *E. coli* [426] and in human cells [427, 428] and a role for mismatch repair in modulating O^6-MeG mutagenesis in *E. coli* has been previously demonstrated [429-431]. MMR-mediated cytotoxicity in wild-type *E.coli* might result from both the formation of DSBs and through interference of their repair. Two separate mechanisms have been proposed for the formation of DSBs in *E.coli* [432]. Replication-independent DSBs arise from overlapping base excision and MMR repair tracts on complementary strands and constitute the majority of detectable DSBs in *dam⁻* mutants (*E.coli* strains that lack dam methylase). DSBs formed this way could be repaired efficiently by homologous recombination with a sister chromosome. In contrast, replication-dependent DSBs result from replication fork collapse at O^6-MeG containing base pairs undergoing MMR futile cycling and are more likely to contribute to cytotoxicity. This general model of "futile repair" in *E.coli* was later adopted to explain MMR-dependent DNA repair in mammalian cells [433].

Cytotoxicity induced by O^6-methylating agents in rodent fibroblasts, human lymphocytes and lymphoblastoid cells deficient in MGMT is due to apoptosis, which indicates that O^6-MeG is a critical pro-apoptotic DNA lesion [434-437]. Apoptosis induced by O^6-MeG requires cell proliferation [438] and competence for MMR [434, 437, 439]. How MMR is causally connected to apoptotic cell death and recombination has not been established, although mouse models indicate that the apoptotic response is mediated through p53 [440]. Two separate mechanisms have been proposed for the formation of DSBs in eukaryotes [425]. One possibility is that DSBs are produced when O^6-MeG residues are present on opposite DNA strands and the distance between the residues is short enough to allow the ss regions of DNA, resulting from excision of the newly replicated strand by MMR, to overlap. An alternative possibility is that some of the MMR ss intermediates are replicated (in the absence of cell cycle arrest) and that this results in a DSB in one of the daughter strands. It is well established from the effects of ionising radiation that double-strand breaks are lethal, only 40 being required to kill a cell [441]. Double-strand breaks are processed by the recombination repair system, of which there are two pathways (described in Chapter I); homologous recombination (HR) and non-homologous end joining (NHEJ). The molecular interactions occurring during these repair processes are highly complex and much has yet to be resolved.

3.1.2 Adduct site-specific mutagenesis

Mutation spectra yield information about the sites at which mutations can occur when DNA is damaged by a specific agent and about the molecular

nature of the resulting mutations. When this information is coupled with information concerning the chemical nature of the lesions introduced and distribution of these lesions in DNA, hypotheses can be formulated as to which lesions are *premutagenic* and as to the type of mutations they cause [442]. However, genotoxic agents such as chemical carcinogens and radiation commonly generate more than one and usually many, activated species. Additionally, DNA, the target of these species, is a chemical composite made up of many nucleophilic atoms and, thus, the range of the resulting adduct structures can be vast. For example, simple alkylating agents and radiation each produce >12 specific modified bases in nucleic acids [443, 444]. The complexity of adduct mutagenesis makes it a challenging task to (i) assess the impact of any individual adduct on local DNA structure and chemistry, (ii) determine the contribution of each adduct to the mutation spectrum generated by chemical or radiation treatment, (iii) assess the effect of sequence context on the pattern of mutations and (iv) ascertain the identities of the DNA repair proteins responsible from protecting cells for specific forms of damage [445].

One potential means to assess the consequences of individual DNA adducts involves their study in defined DNA sequences in biological systems. The ultimate goal of an adduct site-specific study is to investigate whether a particular adduct is responsible for a particular mutation obtained in an *in vivo* mutagenesis (or carcinogenesis) study and to understand the mechanisms governing lesion induced mutagenicity. Correlations between the qualitative and quantitative pattern of mutations induced by an adduct and that observed *in vivo* following treatment with the mutagen/carcinogen itself in the tissue of interest can provide information on the relative importance of each type of

DNA adduct and therefore on mechanisms of carcinogenesis [446]. The advent of strategies for the chemical synthesis of oligonucleotides containing single DNA adducts has revolutionized the study of how DNA lesions are processed by cellular systems. When DNA synthesizers became widely available in the mid 1980s, it was possible to design oligonucleotides of virtually any sequence containing normal or modified bases for which stable phosphoramidites could be made [447].

In brief, the field of adduct site-specific mutagenesis involves the synthesis of an oligonucleotide that contains a DNA adduct of known structure at a defined position, and the use of recombinant DNA techniques to incorporate this oligonucleotide into an autonomously replicating plasmid, or viral-based vector. This vector can then be studied *in vitro* or transfected into cells, where biological processing occurs, the nature of which can be deduced from analysis of the progeny [446]. The resulting mutations can be detected with a variety of methods, such as differential DNA hybridization [448-450], primer extension assays [451, 452], REAP analysis [453-455], or DNA sequencing [416]. Selection of mutant progeny for DNA sequencing can be achieved using an α-complementation (blue/white) screening assay in bacteria [456], or by antibiotic resistance [457].

In an adduct site-specific study, the choice of vector is significant. Both single stranded (ss) and double stranded (ds) vectors, as well as ds vectors where the lesion is in a ss gap, have been used. The advantage of studying a lesion in a ds vector is that the system comes as close as possible to resembling genomic replication and repair. However, a major drawback with a system

where one DNA strand contains a bulky lesion(s), while the complementary strand does not, is that progeny plasmids are preferentially derived from the unadducted strand. This phenomenon, referred to as "strand bias", may also occur with small DNA adducts and is due to damage avoidance mechanisms. In order to minimize the generation of plasmids from the unmodified strand, uracil bases can be incorporated opposite the region containing the adduct, or UV-induced damage can be introduced [446]. Repair of deoxyuridines by uracil DNA glycosylase and AP endonucleases creates a gap in the strand complementary to the adduct [431, 458]. The resulting constructs promote TLS past the modified base and eliminate possible preferential DNA synthesis over a normal complementary strand containing thymine residues [456]. The study of lesions in ss DNA has proven attractive, because it eliminates the problem of strand bias (only TLS can be performed in ssDNA). However, a potential problem with ss vectors is that there is no guarantee that the conformation of any given lesion is the same in ss DNA as it is in ds DNA, which may have an influence on the repair of the lesion and the pathway of mutagenesis when the DNA polymerase forms the ss/ds replication junction [446]. Finally, it is worth mentioning that although an adduct is replicated autonomously in an extrachromosomal genome in most adduct site-specific studies, there are strategies whereby an adduct can be placed in a vector that ultimately integrates into host cell genome, where it is replicated [459]. From such site-specific sequence context studies, empirical data are being generated that help understand the important variables in lesion processing as a function of its sequence environment, which may ultimately explain in

molecular terms the observed mutation spectra from exogenous and/or endogenous DNA damaging agents.

3.1.3 Adduct site-specific mutagenesis assays

While it falls outside the scope of this thesis to discuss in detail the intricacies of all methods for adduct site specific mutagenesis, various assays have been developed where specific adducts were placed in known locations within vectors that could subsequently be replicated in cellular systems (some earlier systems are reviewed in refs [446, 447, 460]). One such assay, developed by Delaney and Essigmann [453] uses an M13 derivative bacteriophage genome containing certain base sequences incorporating the adduct of choice which are transfected into *E.coli*. Restriction digestion and radioactive one-dimensional thin layer chromatography (TLC) can be used to identify which mutations are preferentially induced by DNA lesions in the site-specific sequences used. This assay has been used to study O^6-MeG [453, 454], 8-oxodG [455] and N^6-(2-deoxy-α,β-D-erythropentofuranosyl)-2,6-diamino-4-hydroxy-5-formamidopyrimidine (Fapy.dG) [461]. Although methods like this increase our understanding of the site selectivity and mutagenicity of DNA lesions in *E.coli* in ss DNA, they do not pertain to mammalian systems and in particular to human cells.

In recent years several assays have been developed for use in mammalian cells. Various reporter genes have been employed in these assays, such as the *HPRT* [462], *URA3* [463], *lacZ'* [456] or *supF* [464-466] genes. One such assay has been developed by McLuckie, *et al.* [466] and involves insertion of a short synthetic deoxyoligonucleotide containing a single O^6-methylguanine

adduct into the *supF* gene of the pSP189 plasmid using blunt ended ligation. The adduct containing plasmid is transiently transfected and replicated in human cells and subsequent screening for mutations is performed by MALDI-ToF mass spectrometry (MALDI-MS). A great advantage of this approach is that it can be applied to the study of any adduct in any short sequence of a biologically relevant gene (e.g. *p53*) in human cells. A separate site-specific approach utilizing the *supF* reporter gene has been developed by Upton, *et al.* [464, 465] and it involves placing the adduct on a position located in the acceptor stem of the transcribed tRNA. The adducted pLSX plasmids are transfected into Ad293 cells and screening of the recovered mutant progeny is achieved by the blue/white assay.

A major limitation of the existing adduct site-specific assays is that positioning of the adduct of interest in a gene is restricted to a small defined region of the gene. For example, it is often impossible to place a DNA adduct site-specifically at a position that manifested as a hotspot in a random mutagenesis study performed using the same system. This is largely because of the difficulty of devising assay systems that are tolerant of sequence alteration around a hotspot, since many nucleotide changes can be expected to inactivate the marker and preclude an experiment. This is particularly true of tRNAs which are quite sensitive to mutagenesis [467]. As a result, investigations of sequence context effects in mutagenesis have been limited to the cataloguing of bases adjacent to hotspots (based on random mutagenesis studies) and changing the local sequence environment surrounding the adducted base in site-specific *in vitro* studies. This strategy has been useful in elucidating the effects of neighbouring bases on the

mutagenesis of a DNA adduct. However, this type of approach does not take into account the effects of distal sequence context and secondary DNA structure of the reporter gene in the mutagenic processing of an adduct and consequently cannot replicate faithfully a random mutagenesis study. Another important constraint of existing assays is their inability to assess the mutagenicity of multiple DNA adducts and the influence of one DNA adduct on the mutagenic processing of another. Such methodological issues have been a large impediment to understanding the role of sequence context in adduct mutagenesis and can be potentially overcome with the development of the novel assay described in this Chapter. The main difference with this new approach is that instead of inserting a small adducted oligonucleotide (a foreign sequence) at a defined, restricted position on the reporter gene, the whole reporter gene (*supF* in this case) is recreated from ligation of smaller synthetic oligos. This strategy allows for insertion of a DNA adduct at any conceivable position in the *supF* gene without changing the plasmid sequence. It also permits insertion of multiple DNA adducts at desirable positions. Therefore, this novel assay offers a great potential to efforts aimed at delineating the complex nature of adduct mutagenesis.

3.1.3 Aim

The aim of the work described in this Chapter was to develop and validate an assay for determining the mutagenic potential and types of mutations caused by individual adducts situated in ds or gapped plasmids, in two different sequence contexts, in *E.coli*. The assay would be a site-specific variation of the *supF* assay (described in Chapter 1). The main difference compared to

the original *supF* assay is that this new approach would assess the mutagenicity of individual adducts, rather than compounds. The assay involves the use of DNA recombination techniques to excise a sequence of ds DNA from the original pSP189 plasmid and replace it with a chemically synthesized ds DNA oligomer containing the *supF* gene. The assay was validated using an oligomer that would completely inactivate *supF* function and with oligomers containing O^6-MeG adducts as a model lesion in two different DNA sequence contexts in double-stranded or gapped plasmids. Furthermore, the methods could be adapted to the study of any DNA adduct, as long as the particular DNA-adduct phosphoramidite is available for incorporation into the synthetic deoxyoligonucleotide insert.

3.2 PART I: EARLY METHODS

3.2.1 Experimental Design

Using DNA recombination techniques, a sequence of ds DNA was excised from the original pSP189 plasmid and replaced with a chemically synthesized ds DNA oligomer. The synthetic oligo was designed in such a way that it does not interrupt the activity of the *supF* gene and as a consequence the unadducted plasmid should produce blue colonies when transformed *E.coli* are grown on X-gal selection plates. A 114 basepair-long sequence (which includes the wild-type *supF* gene) was excised from the pSP189 plasmid using the XhoI and BamHI restriction endonucleases and replaced by a modified ds114mer. This modified ds oligo was constructed from the ligation of three smaller chemically synthesized ds oligos (Figure 3.2). The modified

114mer insert is slightly different from the wild-type sequence in that it has two

base pairs (at positions 99 and 179) changed. These two base changes

should not inactivate *supF* function and were introduced to serve as a marker

to prove that there was no contamination with the original or singly-digested

TOP 55MER

5′-AATTCGAGAGCCCTGCTCGAGCTGTAGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCG-3′
3′-GCTCTCGGGACGAGCTCGACATCACCCCAAGGGCTCGCCGGTTTCCCTCGTCTGAGATTTAGACGGCAGTA-5′

BOTTOM 55MER

TOP 11MER

5′-TCATCGACTTC-3′
3′-GCTGAAGCTTC-5′

BOTTOM 11MER

TOP 66MER

5′-GAAGGTTCGAATCCTTCCCCCACTACCACGGCCGAAATTCGTTACCCG-3′
3′-CAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

BOTTOM 66MER

religated plasmid [467].

Figure 3.2 Oligos used to construct the modified 114mer insert. The *supF*

sequence is shown in pink. Bases circled in red correspond to alterations in

the wild-type *supF* sequence and demonstrate incorporation of the synthetic

114mer oligo. The original bases are guanine in the place of adenine and

cytosine in the place of thymine.

This strategy takes advantage of the secondary structure of tRNA which

allows some alterations to be made in certain regions of the gene without loss

147

of function (Figure 3.3). During the early stages of assay development, only

unadducted oligos were used. The key aim of this early method development

was to successfully construct the recombinant vector and to ensure that the

activity of the *supF* gene is preserved when the unadducted modified 114mer

insert is present.
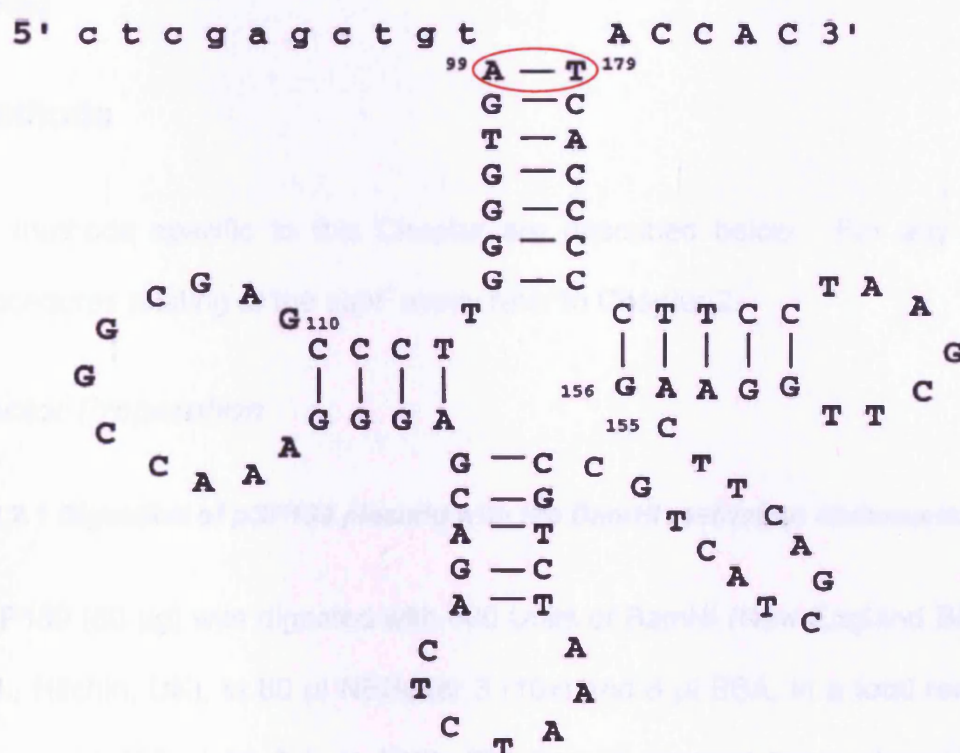
Figure 3.3 Sequence of the *supF* gene. The familiar cloverleaf form is

depicted as a DNA sequence. Bases circled in red correspond to alterations in

the wild-type *supF* sequence and demonstrate incorporation of the synthetic

114mer oligo. The original bases are guanine in the place of adenine and

cytosine in the place of thymine. These base changes do not distort the

secondary structure of the tRNA molecule and as a result do not inactivate the

supF function. Some of the pre tRNA bases are indicated in small letters.

3.2.2 Materials and methods

Materials

All chemicals were from Sigma-Aldrich, Poole, Dorset, UK unless otherwise

stated.

Methods

All methods specific to this Chapter are described below. For any other

procedures relating to the *supF* assay refer to Chapter 2.

Vector Preparation

3.2.2.1 Digestion of pSP189 plasmid with the BamHI restriction endonuclease

pSP189 (80 μg) was digested with 800 Units of BamHI (New England Biolabs

Ltd., Hitchin, UK), in 80 μl NEBuffer 3 (10x) and 8 μl BSA, in a total reaction

volume of 800 μl for 3 h at 37°C. The BamHI-digested linear plasmid was

immediately loaded on to a 1% agarose gel. BamHI has the following

recognition site:

$$5'...G^{\blacktriangledown}GATCC...3'$$

$$3'...CCATG_{\blacktriangle}G...5'$$
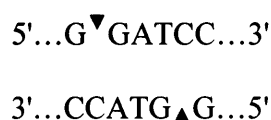
3.2.2.2 Agarose purification of BamHI-digested pSP189

The BamHI-digested plasmid was loaded on a 1% Seakem LE (Lonza Wokingham Ltd., Berkshire, UK) agarose gel for 24h at 40V. The gel was photographed using a Syngene Chemigenius 2 bio imaging system and genesnap image analysis software. The bands corresponding to linear digested psP189 were visualised under low intensity UV light using a GLLS-4 UV wand (Syngene, Cambridge, UK) and excised from the gel using a sterile scalpel. The DNA was purified from the agarose gel using the QIAquick Gel Extraction Kit according to the manufacturer's protocol (Qiagen, West Sussex, UK) and eluted in EB buffer (10mM Tris-Cl, pH 8.5).

3.2.2.3 Digestion of pSP189 plasmid with the XhoI restriction endonuclease

BamHI-digested linear plasmid (80 µg) was digested with 800 Units of XhoI (NEB), in 80 µl NEBuffer 2 (10x) and 8 µl BSA in a total reaction volume of 800 µl for 3 h at 37°C. The reaction was terminated by heat inactivating the enzyme for 20 min at 65°C. XhoI has the following recognition site:

$$5'...C^{\blacktriangledown}TCGAG...3'$$
$$3'...GAGCT_{\blacktriangle}C...5'$$

3.2.2.4 Agarose purification of BamHI+XhoI-digested pSP189

Doubly-digested vector was purified again by agarose gel electrophoresis as described in section 3.2.2.2. The eluted DNA (in EB buffer) was then ready to use for downstream applications.

3.2.2.5 Dephophorylation of BamHI+XhoI-digested pSP189 with Antarctic phosphatase

Doubly-digested vector (10 μg, in EB buffer) was dephosphorylated with 300 Units of antarctic phosphatase (NEB) in 60 μl antarctic Phosphatase Buffer (10x) (NEB) in a 600 μl reaction for 1 h at 37°C.

114mer modified insert preparation

The sequences of the deoxynucleotide oligos used for the construction of the 114 basepair insert are shown below:

Top 55mer (5'-3'): TCG AGC TGT AGT GGG GTT CCC G AG CGG CCA AAG GGA GCA GAC TCT AAA TCT GCC G

Bottom 55mer (5'-3'): TGA CGG CAG ATT TAG AGT CTG CTC CCC TTT GGC CGC TCG GGA ACC CCA CTA CAG C

Top 11mer (5'-3'): TCA TCG ACT TC

Bottom 11mer (5'-3'): CTT CGA AGT CG

Top 48mer (5'-3'): GAA GGT TCG AAT CCT TCC CCC ACT ACC ACG GCC GAA ATT CGG TAC CCG

Bottom 48mer (5'-3'): GAT CCG GGT ACC GAA TTT CGG CCG TGG TAG TGG GGG AAG GAT TCG AAC

Oligonucleotide synthesis was carried out by Biomers.net (Ulm, Germany) at a typical scale of 1μmol. The desalted oligonucleotides were dissolved in

sterile water to a final concentration of 1µg/µl and purified by denaturing PAGE (20%).

3.2.2.6 Polyacrylamide gel preparation

Denaturing polyacrylamide stock solution (20%) was prepared by dissolving 420.4 g urea in 200 ml TBE buffer (5x) and adding 498.8 ml 30% Acryl/Bis 37.5:1 (Biorad, Hertfordshire, UK). The solution was heated/stirred until completely dissolved and the volume was made up to 1 L by addition of distilled H_2O. Non-denaturing (native) polyacrylamide stock solution (20%) was prepared as above but the urea was excluded.

For each gel, 100 ml of stock solution was mixed with 500 µl 10% ammonium persulfate and 80 µl TEMED (Biorad, Hertfordshire, UK). The solution was stirred and poured straight away into the gel set up, where it was left to polymerise for 30 min. The oligomers were then loaded on to the gel (in 2x volume of formamide) together with loading dye as a marker at each side of the gel. The gel was left to run overnight in 0.5 x TBE buffer at 300V at 4°C. The slowest migrating bands were visualised under 254 nm UV light by placing the gels on a silica TLC GF25 fluorescent plate. The bands were excised and mixed with 300 µl water and left to elute for a minimum of 4 h. The solution containing DNA was isolated from the gel using a homemade device described by Delaney and Essigmann [453] by inverting into an eppendorf (with a small hole made in the bottom using a sterile needle) containing a small amount of sand at the bottom. This eppendorf was then placed into another eppendorf that was centrifuged for 3 min at 4000 rpm in a Sanyo Hawk 15/05 microcentrifuge (MSE, Watford, Herts, UK). The eluate

was desalted through an Illustra NAP-5 G-25 DNA grade sephadex column (GE Healthcare Ltd., Buckinghamshire, UK) and the concentration was calculated from the UV absorbance, then dried down under vacuum on a SC 210A speed vac (Savant Instruments, Inc. Holbrook, NY) and reconstituted to give a concentration of 100 ng/µl in TE buffer (pH 7.1).

3.2.2.7 Oligomer phosphorylation with T4 polynucleotide kinase (T4 PNK)

PAGE-purified oligomers (300 pmol) were 5'-phosphorylated with 100 units of T4 PNK (NEB) in 60 µl T4 DNA ligase buffer (10x) (NEB) in a total volume of 600 µl for 1 h at 37°C. The reaction was terminated by heat inactivating T4 PNK at 70°C for 20 min.

3.2.2.8 Annealing of ss oligos

The modified ds114mer insert was constructed from ligation of three smaller 5'-phosphorylated double-stranded (ds) oligos; a ds55mer, a ds11mer and a ds48mer (Figure 1). The top and bottom strands of each ds oligo were annealed together by heating at 95°C for 5 min then allowing to cool slowly to room temperature over a period of 1 h. Subsequently the three ds oligos were ligated together in two sequential ligation reactions.

3.2.2.9 Construction of the ds66mer

Equimolar amounts of ds55mer and ds11mer were ligated together using 2000 cohesive end units of T4 DNA ligase (NEB) in 100 µl T4 DNA ligase buffer (10x) (NEB) in a total volume of 1 ml for 1 h at room temperature. The ligation reaction was terminated by heat inactivation (65°C, 10min). The

resulting 66mer was loaded on to a 4% agarose gel (run at 40V, overnight) to separate it from the unligated 55mer and 11mer. The bands representing the ds66mer were visualized under 254 nm UV light and excised using a sterile scalpel. Excised DNA was subsequently purified using the QIAquick Gel Extraction Kit following the manufacturer's protocol.

3.2.2.9 Construction of the ds114mer

Equimolar amounts of the purified ds66mer and the ds48mer (~23 pmol) were then ligated together using the same protocol as above. The resulting 114mer was separated again from the 66mer, the 48mer and the concatemers formed in the reaction, using a 4% agarose gel (40V, overnight) and then purified with the QIAquick Gel Extraction kit.

3.2.2.10 Ligation of the ds114mer insert to the doubly-digested dephosphorylated vector

Doubly-digested dephosphorylated vector (1 µg) and ds114mer insert were ligated together in a 1:10 (vector:insert) molar ratio ligation reaction using 400 cohesive end units of T4 DNA ligase (NEB) in 2 µl T4 DNA ligase buffer (10x) (NEB) in a total volume of 20 µl for 3 h at room temperature.

3.2.2.11 Effect of UV light on mutagenic background

The ds114mer insert was exposed to low intensity 254 nm UV light for 1 and 3 min. These oligos were then ligated into the doubly-digested dephosphorylated vector and transformed into E.*coli* as usual, along with control samples prepared using the standard protocol, which were not exposed to any additional UV light.

3.2.3 Results of Part I

3.2.3.1 Digestion of pSP189 plasmid with the BamHI restriction endonuclease

The digestion product arising from incubation of pSP189 with BamHI is shown in Figure 3.4. The gel photo indicates that digestion with BamHI was highly efficient, resulting in almost complete conversion of circular plasmid to the linear form.
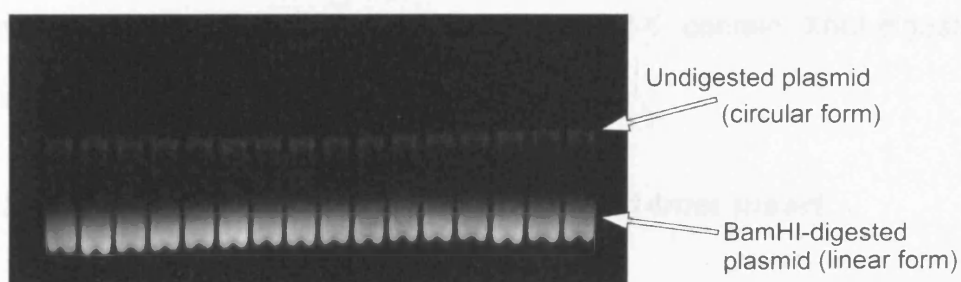
Figure 3.4 BamHI digestion of pSP189. Photograph of a typical 1% preparative agarose gel showing the BamHI-digested product (1µg of vector per lane).

3.2.3.2 Digestion of pSP189 plasmid with the XhoI restriction endonuclease

The results of digestion of pSP189 plasmid with XhoI are shown in Figure 3.5. The gel photo indicates that a high proportion of the circular plasmid is converted into the linear form upon digestion with XhoI under the conditions employed.
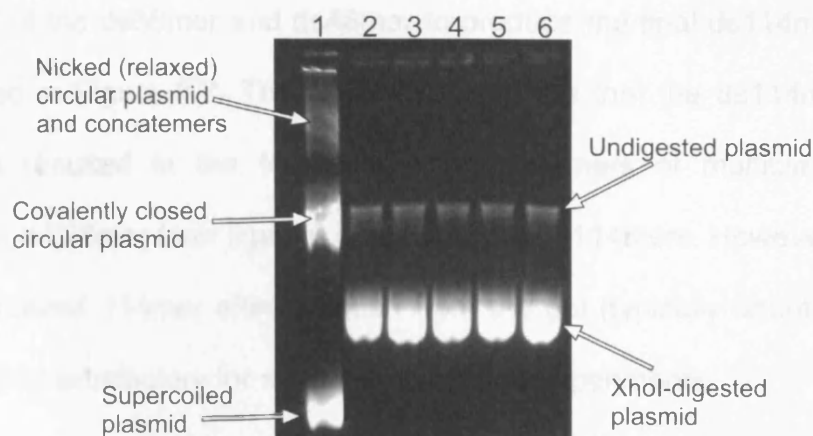
Figure 3.5 XhoI digestion of pSP189. Photograph of a typical 1% preparative agarose gel showing the XhoI-digested product. Lane 1 represents undigested control pSP189 (1 µg) and lanes 2-6 contain XhoI-digested plasmid (1µg of vector per lane).

3.2.3.3 Construction of the ds66mer and ds114mer insert

As shown in Figure 3.6, once the ligation reaction was optimized according to the protocol described in the methods section, the ligation reaction of the ds 55mer with the 11mer was highly efficient, giving the desired 66mer as the major product.
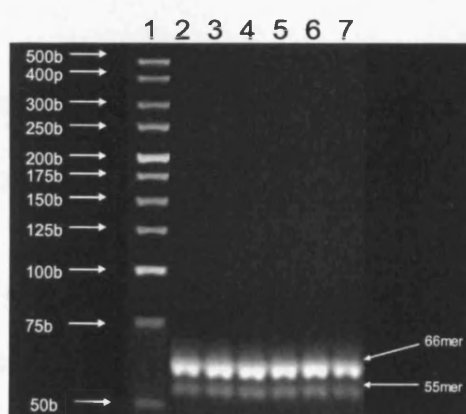
Figure 3.6 Construction of the ds66mer. Photograph of a typical 4% preparative agarose gel showing the ds66mer ligated product. Lane 1 contains a 500bp ladder (hyperladder V, Bioline, Bath, UK) and lanes 2-7 contain the ds66mer ligation reaction product (500 ng per lane).

Ligation of the ds66mer and ds48mer to produce the final ds114mer insert is illustrated in Figure 3.7. The gel photo suggests that the ds114mer ligation reaction resulted in the formation of concatemers of multiple sizes, for example a 228mer from ligation of two full length 114mers. However, the yield of the desired 114mer after isolation from the gel (typically about 2 µg) was considered satisfactory for subsequent ligation experiments.

Figure 3.7 Construction of ds114mer. Photograph of a typical 4% preparative agarose gel showing the ds114mer ligated product. Lane 1 contains a 500bp ladder and Lanes 2-5 contain the ds114mer ligation reaction product (1 µg per lane).

3.2.3.5 *Transformation of electrocompetent E.coli with the recombinant vector*

Once formed (see Section 3.2.2.10), the recombinant plasmid was used to directly transform *E.coli*. Transformation of electrocompetent *E.coli* cells with the recombinant plasmid DNA resulted in about 8.7 ± 3% white colonies (Table 3.1). When the bacteria were transformed with recombinant doubly cut dephosphorylated vector ligated in the absence of the modified insert (negative control sample) this produced few or more typically zero colonies.

This indicated that the vector was not religating to itself and also that there was not significant contamination with the parent undigested pSP189 which may have been present in small quantities in the gel purified plasmid.

Experiment	Sample	114mer insert	Doubly-digested Dephos vector	Colonies counted			
				Total	Blue	White	%White
1	1	+	+	1130	1022	108	9.6
	(-ve) control	-	+	2	2	0	
2	1	+	+	288	250	38	13.2
	2	+	+	406	355	51	12.6
	3	+	+	243	209	34	14.0
	(-ve) control	-	+	52	52	0	
3	1	+	+	762	702	60	7.9
	2	+	+	398	370	28	7.0
	(-ve) control	-	+	0	0	0	
4	1	+	+	762	718	44	5.8
	(-ve) control	-	+	0	0	0	
5	1	+	+	913	845	68	7.4
	(-ve) control	-	+	0	0	0	
6	1	+	+	830	760	70	8.4
Sum				5732	5231	501	
Average ± SD							9.4 ± 3.0

Table 3.1 Number of mutant and wild type colonies resulting from the transformation of electrocompetent *E.coli* cells with the recombinant vector. Each experiment was performed on a separate occasion, using freshly digested pSP189 plasmid and recently constructed 114mer insert. For most experiments only one ligation reaction of insert to plasmid (designated sample 1) was performed. In situations where multiple samples are listed (expt 2+3) samples 1, 2 and 3 were identical and represent separate vector to insert ligation reactions. For each sample a single transformation of *E.coli* was carried out. The average proportion of white mutant colonies across all experiments and the standard deviation are highlighted in yellow. The negative control represents plasmid ligated in the absence of insert. The high number of colonies observed in the (-ve) control sample in expt 2 may have resulted from a batch of plasmid that was inadequately purified.

3.2.3.6 DNA sequencing

In order to acquire information on the nature of the mutant white colonies and also to confirm that the blue colonies contained the synthetic 114mer oligo selected samples were sequenced. Three blue and three white colonies were randomly picked and sent for sequencing to PNACL.
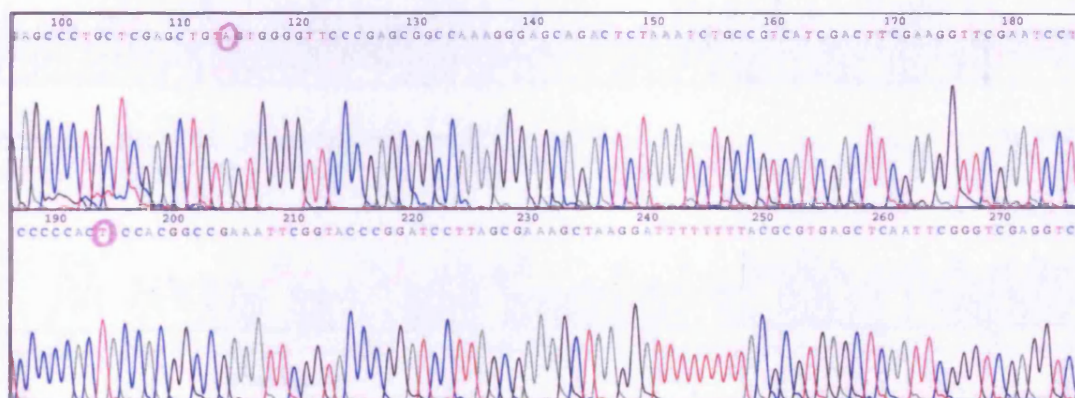
Figure 3.8 Electropherogram illustrating a sequence within a pSP189 recombinant plasmid derived from a blue colony. The bases circled in red correspond to alterations in the wild-type *supF* sequence and demonstrate incorporation of the synthetic 114mer oligo. The original bases are guanine in the place of adenine and cytosine in the place of thymine.

Importantly, three sequences derived from blue colonies (e.g. Figure 3.8) proved that the modified 114mer oligo was incorporated successfully into the vector in place of the original sequence and that the base changes made did not constitute inactivating mutations, as predicted. This is evident from the presence of the altered bases at positions 99 and 179 (see Figure 3.8 legend). The sequences of plasmid extracted from the white mutant colonies (Figure 3.9) confirmed that the modified 114mer insert was also successfully incorporated as indicated by the presence of the double base changes, and

that random mutations (e.g. base substitution, addition and deletion mutations) were responsible for inactivation of the *supF* gene.
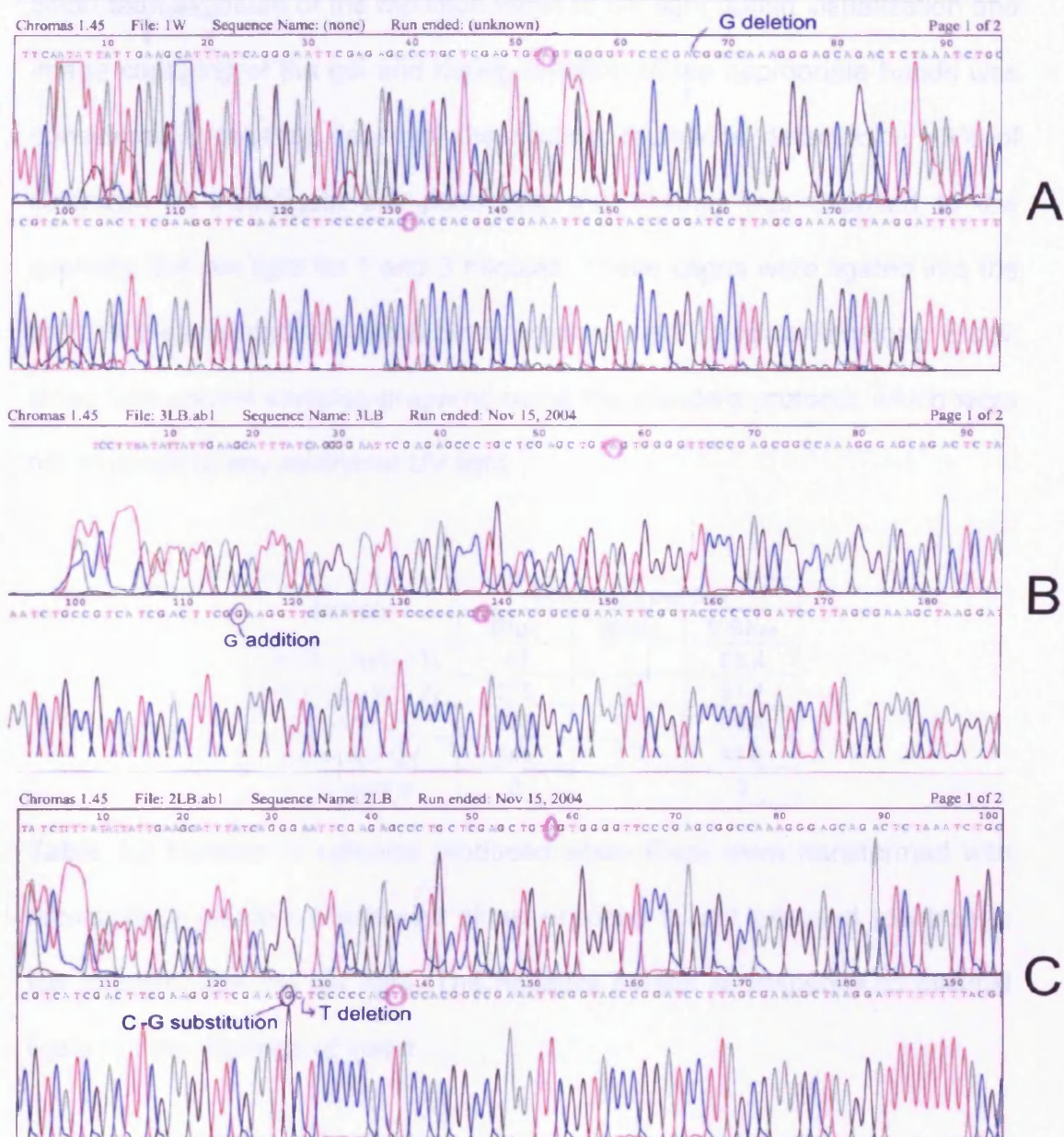
Figure 3.9 Electropherogram illustrating sequences of pSP189 recombinant plasmid derived from three separate white colonies (A, B and C). The bases circled in red correspond to the expected alterations in the wild-type *supF* sequence and demonstrate incorporation of the synthetic 114mer insert. The original bases are guanine in the place of adenine and cytosine in the place of thymine. Also observed are inactivating mutations in the form of a single G deletion, G addition and a C→G substitution coupled with a T deletion in samples A, B and C respectively.

3.2.3.7 Effect of short term exposure to UV light

Short term exposure of the modified insert to UV light during visualization and image capturing of the gel and during excision of the appropriate bands was considered a possible cause of the random mutations detected in ~9% of colonies. To investigate this possibility, the 114mer was exposed to low intensity 254 nm light for 1 and 3 minutes. These oligos were ligated into the psP189 plasmid and the recombinant vectors were used to transform *E.coli*, along with control samples prepared using the standard protocol, which were not exposed to any additional UV light.

Sample	Colonies counted		
	Blue	White	%Blue
no UV (control 1)	40	5	88.8
no UV (control 2)	575	56	91.1
1 minute UV	353	16	95.6
3 minutes UV	146	17	89.5
(-ve) control	0	0	0

Table 3.2 Number of colonies produced when *E.coli* were transformed with recombinant plasmid which was either exposed or not exposed (control) to low intensity 254 nm UV light. The negative control corresponds to plasmid ligated in the absence of insert.

As shown in Table 3.2, additional exposure to UV light did not induce an increase in the proportion of mutant colonies compared to those plasmid samples that were prepared using the standard protocol, indicating that UV exposure was not the cause of the typical mutation background observed (~9%).

3.3 PART II: METHODOLOGICAL IMPROVEMENTS

3.3.1 Experimental Design

Having established the basic protocol and ascertained that the synthetic oligo could be successfully incorporated in place of the original sequence without inactivating the *supF* gene several modifications were introduced into the methodology. First, the number of ss oligos used for the construction of the 114mer insert was increased from 6 to 14. This increase was carried out with the intention of reducing the mutation background and also to facilitate the introduction of short adducted or deoxyuridine-containing oligodeoxynucleotides. Second, additional 114mer constructs (Figure 3.10) were prepared containing either deoxyuridines in the complementary strand (114-B) or a single O^6-MeG DNA adduct at sequence context 1 (position 100, 114-C) or sequence context 2 (position 129, 114-D). Third, the ds66mer, ds48mer and ds114mer constructs were purified by non-denaturing PAGE (20%), instead of agarose purification, prior to ligating to the doubly-digested dephosphorylated vector. Finally, the recombinant vector was purified by ethanol precipitation using yeast tRNA as a co-precipitant, prior to transformation of *E.coli*.

SupF gene shown in red

114-A

5′-TCGAGCTGTAGTGGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACTTACCACGGCCGAAATTCGTTACCCG-3′
3′-CGACATCACCCCAAGGGCTCGCCGGTTTCCCTCGTCTGAGATTTAGACGGCAGTAGCTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

114-B

5′-TCGAGCTGTAGTGGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACTTACCACGGCCGAAATTCGTTACCCG-3′
3′-CGACAUCACCCCAAGGGCUCGCCGGTTTCCCTCGTCTGAGATTTAGACGGCAGTAGCTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

114-C

*

5′-TCGAGCTGTAGTGGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACTTACCACGGCCGAAATTCGTTACCCG-3′
3′-CGACATCACCCCAAGGGCTCGCCGGTTTCCCTCGTCTGAGATTTAGACGGCAGTAGCTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

114-D

*

5′-TCGAGCTGTAGTGGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACTTACCACGGCCGAAATTCGTTACCCG-3′
3′-CGACATCACCCCAAGGGCTCGCCGGTTTCCCTCGTCTGAGATTTAGACGGCAGTAGCTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

Figure 3.10 Sequence of the ds 114mer inserts. Bases circled in red correspond to alterations in the wild-type *supF* sequence and demonstrate incorporation of the synthetic 114mer oligo. Bases in green correspond to thymines replaced by uracil. Asterisks denote the presence of a single O^6-MeG adduct positioned at sequence context 1 (114-C) or at sequence context 2 (114-D).

3.3.2 Materials and methods

Materials

All chemicals were from Sigma-Aldrich, Poole, Dorset, UK unless otherwise stated.

Methods

All methods specific to part II are described below. For any other procedures refer to part I (Chapter 3) and Chapter 2.

114mer modified insert preparation

The sequences of the deoxyoligonucleotides used for the construction of the 114 basepair inserts are given below (Figure 3.11).

	TOP STRAND (5'-3'):	
T1:	TCG AGC TGT AGT GGG GTT	18MER
T1-O6-MeG:	TCG AGC TGT AXT GGG GTT	18MER
T2:	CCC GAG CGG CCA AAG GGA GC	20MER
T3:	AGA CTC TAA ATC TGC CGT C	19MER
T3-O6-MeG:	AXA CTC TAA ATC TGC CGT C	19MER
T4:	ATC GAC TTC	9MER
T5:	GAA GGT TCG AAT CCT T	16MER
T6:	CCC CCA CTA CCA CGG CC	17MER
T7:	GAA ATT CGG TAC CCG	15MER
	BOTTOM STRAND (5'-3'):	
B1:	CTC GGG AAC CCC ACT ACA GC	20MER
B1U:	CUC GGG AAC CCC ACU ACA GC	20MER
B2:	GAG TCT GCT CCC TTT GGC CG	20MER
B3:	TCG ATG ACG GCA GAT TTA	18MER
B4:	ACC TTC GAA G	10MER
B5:	TGG GGG AAG GAT TC G A	16MER
B6:	AAT TTC GGC CGT GGT AG	17MER
B7:	GAT CCG GGT ACC G	13MER

Figure 3.11 Sequences of the deoxyoligonucleotides used for the construction of ds 114mer inserts. X represents an O^6-MeG lesion and U represents a uracil base.

As before, oligonucleotide synthesis was carried out by Biomers.net (Ulm, Germany). O^6-MeG containing oligos T1-O6-MeG and T3-O6-MeG were synthesised by Trilink (TriLink BioTechnologies, San Diego, CA, USA). MALDI-ToF MS analysis of T1-O6-MeG and T3-O6-MeG oligodeoxynucleotides gave protonated molecular ions with a mass within 0.008% and 0.0017% of the expected theoretical mass respectively. Adducted and unadducted oligonucleotides were dissolved in sterile water to give a final concentration of 1μg/μl and purified by denaturing PAGE (20%), as described in section 3.2.2.6.

3.3.2.1 Oligonucleotide phosphorylation with T4 polynucleotide kinase (T4 PNK)

PAGE purified oligos T2, B2, T3, B3, T4, B4, T6, B6, T3-O6-MeG and B2U were phosphorylated as in 3.2.2.7.

3.3.2.2 Construction of the ds66mers

The ligation strategy followed was different from that adopted in part I. In particular, the annealing and ligation steps were combined in a single reaction. Four different ds66mers were constructed using the following ss oligos from Figure 3.11:

66-A: T1+T2+T3+T4+B1+B2+B3

66-B: T1+T2+T3+T4+B1U+B2+B3

66-C: T1-O6-MeG+T2+T3+T4+B1+B2+B3

66-D: T1+T2+T3-O6-MeG+T4+B1+B2+B3

Each ds66mer was constructed from 7 smaller ss oligos. Equimolar amounts of each ss oligo (~10 µg) were mixed with 100 µl T4 DNA ligase buffer (10x) (NEB) and the volume of the solution was made up to 1 ml with water. The 7 oligos were annealed together by heating at 95°C for 5 min then allowing to cool slowly to room temperature over a period of 1 h. Subsequently, 4000 cohesive end units of T4 DNA ligase (NEB) were added and the ligation reaction was carried out for 35 min at room temperature. The ligation reaction was terminated by heat inactivation of T4 DNA ligase (65°C, 10 min). The resulting ds66mers were purified by 20% non-denaturing PAGE.

3.3.2.3 Construction of the ds48mer

The ds48mer was constructed using the following oligos from Figure 3.11:

48MER: T5+T6+T7+B4+B5+B6+B7

The ligation strategy employed was the same as that described above in 3.3.2.2. The resulting ds48mer was purified by 20% non-denaturing PAGE.

3.3.2.4 Construction of the ds114mer inserts

Four different ds114mer inserts (Figure 3.10) were constructed using the following PAGE purified ds66mers and the ds48mer:

114-A: 66-A+48MER

114-B: 66-B+48MER

114-C: 66-C+48MER

114-D: 66-D+48MER

Equimolar amounts of each ds66mer and ds48mer (~23 pmol) were mixed with 100 µl T4 DNA ligase buffer (10x) (NEB) and the volume adjusted to 1 ml with water. The oligos were annealed together by heating at 95°C for 5 min then allowing to cool slowly to room temperature over a period of 1 h. Subsequently, 4000 cohesive end units of T4 DNA ligase (NEB) were added and the ligation reaction was carried out for 1 h at room temperature. The ligation reaction was terminated by heat inactivation of T4 DNA ligase (65°C, 10min). The resulting ds114mers were purified by 20% non-denaturing PAGE.

3.3.2.5 114mer insert phosphorylation with T4 polynucleotide kinase (T4 PNK)

PAGE purified ds114mers (200 ng) were 5'-phosphorylated with 20 units of T4 PNK (NEB) in 10 µl T4 DNA ligase buffer (10x) (NEB) in a total reaction volume of 100 µl for 1 h at 37°C. The reaction was terminated by heat inactivating T4 PNK at 70°C for 20 min.

3.3.2.6 Ligation of the phosphorylated ds114mer insert to doubly-digested dephosphorylated vector

Doubly-digested dephosphorylated vector (1 µg) and 5'-ds114mer insert (200 ng) were ligated together in a 1:10 (vector:insert) molar ratio ligation reaction using 800 cohesive end units of T4 DNA ligase (NEB) in 15 µl T4 DNA ligase buffer (10x) (NEB) in a 150 µl total volume for 24 h at room temperature. The ligation reaction was terminated by heat inactivation (65°C, 10 min). The

ligation reaction was subsequently purified by ethanol precipitation using yeast tRNA as a coprecipitant.

3.3.2.7 Ethanol precipitation of recombinant vector with yeast tRNA as a co-precipitant

Recombinant vector (150 µl) was mixed with 1500 µl 100% ethanol, 30 µl yeast tRNA (1 µg/µl, Invitrogen, Paisley, UK), and 420 µl water, then incubated at -20°C for 20 min. Yeast tRNA is suitable for use as a carrier in nucleic acid purification and precipitation procedures. Subsequently, plasmid DNA was precipitated by centrifuging at 14,000 g for 30 min in a microcentrifuge. The supernatant was removed and the pellet was washed once with 70% ethanol. The pellet was reconstituted in 10 µl water, 5 µl of which was used per transformation experiment.

3.3.3 Results of Part II

3.3.3.1 Construction of the ds66mers

PAGE-purified ds66mers were analysed by 4% agarose electrophoresis (Figure 3.12). The gel photo suggests that PAGE purification of the ligation reaction resulted in ds66mer fragments of high purity.
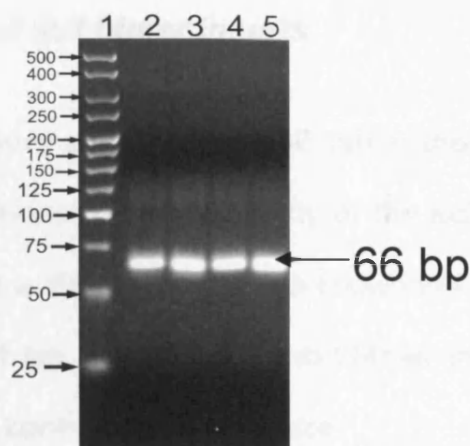
Figure 3.12 Construction of ds66mers. Photograph of a typical 4% agarose gel showing the ds66mer ligated products. Lane 1 contains a 500bp ladder and lanes 2-5 contain constructs 66-A, 66-B, 66-C and 66-D respectively.

3.3.3.2 Construction of the ds48mer

PAGE-purified ds48mer was analysed by 4% agarose electrophoresis and consistently found to be of high purity as shown by the representative gel photo (Figure 3.13).
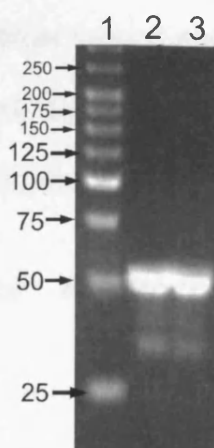
Figure 3.13 Construction of ds48mer. Photograph of a typical 4% agarose gel showing the ds48mer ligated product. Lane 1 contains a 500bp ladder and lanes 2+3 contain the same ds48mer construct.

3.3.3.3 Construction of ds114mer inserts

Ds 114mer constructs were purified by PAGE rather than the agarose gels
that were employed previously, then the purity of the isolated products was
assessed by analysis on a 4% agarose gel as shown in Figure 3.14. The gel
photo demonstrates that the PAGE-purified ds114mer inserts were of high
purity, revealing a single band of the desired size
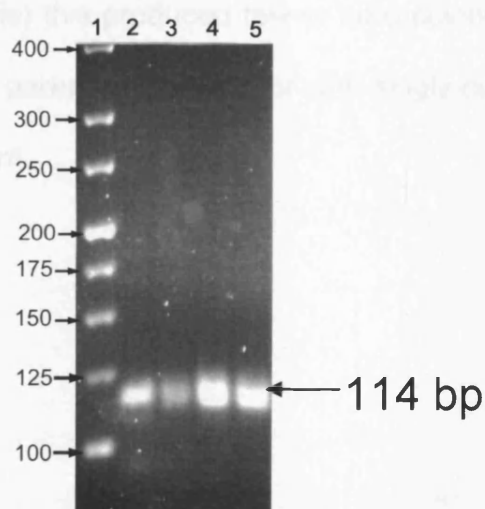
Figure 3.14 Construction of ds114mer inserts. Photograph of a typical 4%
agarose gel showing the ds114mer ligated and PAGE purified products. Lane
1 contains a 500bp ladder and lanes 2-5 contain constructs 114-A, 114-B,
114-C and 114-D (~200 ng each) respectively.

3.3.3.4 Transformation of electrocompetent E.coli with the recombinant vector

The results of *E.coli* transformation with the recombinant constructs produced
by ligation of the various ds 114mers with the digested vector are summarized
in Table 3.3. Transformation of electrocompetent *E.coli* cells with the 114-A

recombinant plasmid DNA resulted in about 3.2 ± 0.2% white colonies. The equivalent uracil containing construct (114-B) resulted in a 2.2-fold increased mutation frequency compared to the unadducted control (114-A). Likewise a similar increase in the mutant fraction of constructs containing a single O^6-MeG adduct in either sequence context 1 (114-C) or context 2 (114-D) was observed. When the bacteria were transformed with the recombinant doubly cut dephosphorylated vector ligated in the absence of the modified insert (negative control sample) this produced few or zero colonies, suggesting that contamination with the parent undigested or with singly-cut religated pSP189 plasmid was insignificant.

Experiment	Type of construct			Colonies counted			
		Sample	Total	Blue	White	%White	STDEV
1	114-A (T1)	1	1353	1311	42	3.1	
		2	1307	1264	43	3.3	
		3	1171	1124	47	4.0	
		4	437	424	13	3.0	
		5	656	636	20	3.0	
		SUM	4924	4759	165	3.4	0.4
	114-B (B1U)	1	996	922	74	7.4	
		2	500	466	34	6.8	
		SUM	1496	1388	108	7.2	0.4
	114-C (T1-O6)	1	1460	1354	106	7.3	
		2	1557	1454	103	6.6	
		SUM	3017	2808	209	6.9	0.5
	114-D (T3-O6)	1	498	435	63	12.7	
2	114-A (T1)	1	854	830	24	2.8	
		2	7307	7059	248	3.4	
		3	5633	5486	147	2.6	
		SUM	13794	13375	419	3.0	0.4
	114-D (T3-O6)	1	633	588	45	7.1	
		2	5213	4930	283	5.4	
		SUM	5846	5518	328	5.6	1.2
	(-ve) control	1	0	0	0		
		2	4	0	4		
1+2	114-A (T1)	Exp1+Exp2 SUM	18718	18134	584	3.2	0.2
	114-D (T3-O6)	Exp1+Exp2 SUM	6344	5953	391	9.1	5.0

Table 3.3 Number of mutant and wild type colonies resulting from the transformation of *E.coli* cells with recombinant constructs 114-A, 114-B, 114-C and 114-D. Samples numbered 1 through to 5 were identical but represent separate vector to insert ligation reactions. For each sample a single transformation of *E.coli* was carried out. Constructs 114-A and 114-D were assessed in two experiments, each one performed on separate ocasions, while constructs 114-B and 114-C were assessed in one. Average mutation frequencies and standard deviations are highlighted in yellow. The average percentages of white colonies and standard deviations from both experiments for constructs 114-A and D are highlighted in light green.

3.3.3.5 DNA sequencing

Three blue colonies from sample 114-A were picked and sent for sequencing by PNACL. All three sequences derived from these blue colonies confirmed the presence of the modified 114-A insert (Figure 3.15) as evidenced by the altered bases (circled), proving that it was successfully ligated into the vector and that there were no inactivating mutations in it, reinforcing the previous results.
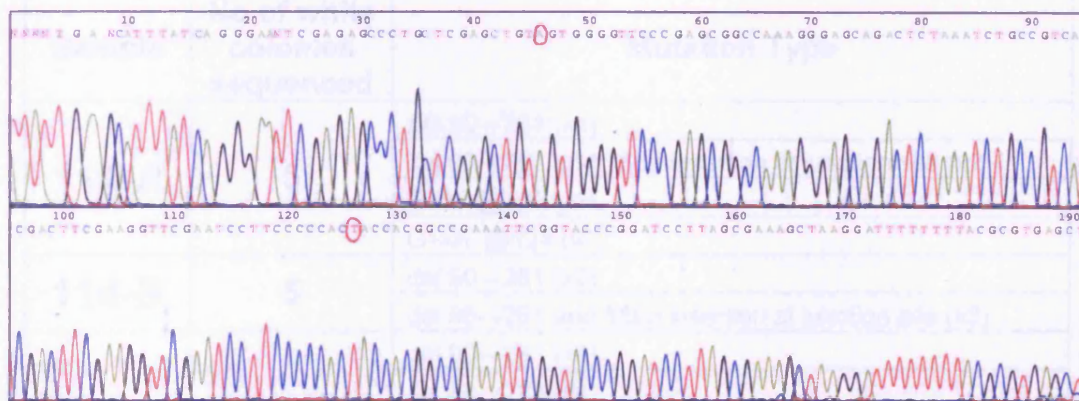
Figure 3.15 Electropherogram illustrating a sequence from a 114-A recombinant plasmid derived from a blue (non-mutant) colony. The bases circled in red correspond to alterations in the wild-type *supF* sequence and demonstrate incorporation of the desired synthetic 114mer oligo. The original bases are guanine in the place of adenine and cytosine in the place of thymine. As expected, no other changes were detected.

To characterize the types of alterations causing inactivation of the *supF* gene in the white colonies, a number of samples derived from recombinant vectors 114-A, 114-B, 114-C and 114-D were randomly picked and sent for sequencing to PNACL. The types of mutations observed in sequences of plasmid extracted from the white colonies are summarized in Table 3.4. DNA

sequencing of plasmid extracted from white colonies of unadducted control sample 114-A revealed that a variety of mutations were responsible for the mutagenic background (3.2 ± 0.2%) observed. The majority of mutations (3/5 sequences) were large deletions of the whole *supF* gene. All three deletions occurred between the XhoI restriction site and position 261, which is placed immediately after the 8 bp signature. The remaining inactivating mutations included a G deletion at position 102 and a G→A transition at position 124.

Sample	No of white colonies sequenced	Mutation Type
114-A	**5**	del 90→261 (x1)
		del 90→261 and 18bp insertion at junction site (x2)
		G del @102 (x1)
		G→A @124 (x1)
114-B	**5**	del 90→261 (x2)
		del 90→261 and 18bp insertion at junction site (x3)
114-C	**10**	del 90→261 (x5)
		del 88→203 (x1)
		del 129→203 (x1)
		G→A @100 (T1-O6-MeG-targeted) (x2)
		G→A @129
114-D	**10**	del 90→261 (x6)
		del 90→261 and 18bp insertion at junction site (x1)
		del 90→263 and 8bp insertion at junction site (x1)
		del 141→210 (x1)
		C→T @155 (x1)

Table 3.4 Types of mutations observed in sequences of plasmid extracted from white colonies.

Plasmid containing uracil residues in the complementary strand (114-B) was replicated in *E.coli* and DNA sequencing of the replicated plasmid revealed that all white colonies had resulted from the same large deletions seen in sample 114-A. Likewise, analysis of the mutation types observed in O^6-MeG

adducted plasmid at sequence context 1 (sample 114-C) suggested that the majority of inactivating mutations were large deletions between the XhoI site and position 261 (5/10) or between the XhoI and BamHI sites (1/10) or between position 129 and BamHI site (1/10). Placement of a single O^6-MeG adduct at position 100 (context 1) produced G→A targeted mutations in 2 out of 10 plasmids along with a G→A transition at position 129. Finally, sample 114-D produced no O^6-MeG targeted mutations at position 129 and the majority (8/10) of inactivating mutations occurred as deletions between the XhoI site and distant bases 261 or 263. The remaining mutation events observed in this sample were a large deletion of bases 141→210 and a random C→T transition at position 155. These results indicate that when the O^6-MeG adduct is situated in ds DNA the majority of mutations detected can be attributed to the background frequency, also observed with unadducted plasmid. The presence of the adduct only causes a small number (if any) of the expected targeted mutations.

3.4 PART III: FINAL METHODS

3.4.1 Experimental Design

Experimental results so far suggested that the majority of background mutations in unadducted plasmids were caused by religation of plasmid at XhoI sites (see discussion for a more in-depth analysis). To further minimize the mutation background, the recombination strategy was modified so that the EcoRI enzyme was used instead of XhoI to cut the vector. The pSP189 plasmid contains an EcoRI site 16 basepairs upstream of XhoI. As a result,

the modified insert would also have to be extended by 16 basepairs in order to create EcoRI-compatible overhangs. Furthermore, several additional 130mer modified inserts were constructed (Figures 3.16 and 3.17). These include inserts containing a single O^6-MeG DNA adduct at position 100 (sequence context 1) in the presence or absence of deoxyuridines in the complementary strand (130-H and 130-E respectively) and inserts containing a single O^6-MeG DNA adduct at position 129 (sequence context 2) in the presence (130-I) or absence (130-F) of deoxyuridines. By positioning uracil residues in the complementary strand, cellular uracil-DNA glycosylase and AP endonucleases create a gap in the complementary strand forcing translesion synthesis past the adducted base [430].

To study the mutagenicity of multiple DNA adducts, inserts containing two O^6-MeG adducts at both sequence contexts (positions 100 and 129) in the presence (130-J) or absence (130-G) of bottom strand uracil bases were also constructed. Moreover, control unadducted inserts containing bottom strand uracil residues alone (130-B, 130-C and 130-D) were prepared to assess the mutagenic effects of incorporating uracil in the plasmid. Finally, an insert with a GC→CG base alteration at position 150 was also prepared for validation purposes, as this point mutation should completely inactivate *supF* function, producing white colonies.

SupF gene shown in red

130-A

5′-AATTCGAGAGCCCTGCTCGAGCTGTAGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACTACCACGGCCGAAATTCGTTACCCG-3′
3′-GCTCTCGGGACGAGCTCGACATCACCCCAAGGGCTCGCCGGTTTCCCTCGTCTGAGATTTAGACGGCAGTAGCTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

130-B

5′-AATTCGAGAGCCCTGCTCGAGCTGTAGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACTACCACGGCCGAAATTCGTTACCCG-3′
3′-GCTCTCGGGACGAGCTCGACAUCACCCCAAGGGCUCGCCGGTTTCCCTCGTCTGAGATTTAGACGGCAGTAGCTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

130-C

5′-AATTCGAGAGCCCTGCTCGAGCTGTAGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACTACCACGGCCGAAATTCGTTACCCG-3′
3′-GCTCTCGGGACGAGCTCGACATCACCCCAAGGGCTCGCCGGTTTCCCUCGUCUGAGATTTAGACGGCAGTAGCTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

130- D

5′-AATTCGAGAGCCCTGCTCGAGCTGTAGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACTACCACGGCCGAAATTCGTTACCCG-3′
3′-GCTCTCGGGACGAGCTCGACAUCACCCCAAGGGCUCGCCGGTTTCCCUCGUCUGAGATTTAGACGGCAGTAGCTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

130-E

                          *
5′-AATTCGAGAGCCCTGCTCGAGCTGTAGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACTACCACGGCCGAAATTCGTTACCCG-3′
3′-GCTCTCGGGACGAGCTCGACATCACCCCAAGGGCTCGCCGGTTTCCCTCGTCTGAGATTTAGACGGCAGTAGCTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

Figure 3.16 Sequence of ds130mer inserts. Bases circled in red correspond to alterations in the wild-type *supF* sequence and act as a marker to demonstrate incorporation of the synthetic 130mer oligo insert. Bases in green (130-C and 130-D) correspond to uracil. Asterisks represent the site of O^6-MeG adduction.

SupF gene shown in red

130-F

5′-AATTCGAGAGCCCTGCTCGAGCTGTAGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACTACCACGGCCGAAATTCGTTACCCG-3′
3′-GCTCTCGGGACGAGCTCGACATCACCCCAAGGGCTCGCCGGTTTCCCTCGTCTGAGATTTAGACGGCAGTAGCTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

130-G

5′-AATTCGAGAGCCCTGCTCGAGCTGTAGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACTACCACGGCCGAAATTCGTTACCCG-3′
3′-GCTCTCGGGACGAGCTCGACATCACCCCAAGGGCTCGCCGGTTTCCCTCGTCTGAGATTTAGACGGCAGTAGCTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

130-H

5′-AATTCGAGAGCCCTGCTCGAGCTGTAGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACTACCACGGCCGAAATTCGTTACCCG-3′
3′-GCTCTCGGGACGAGCTCGACAUCACCCCAAGGGCUCGCCGGTTTCCCTCGTCTGAGATTTAGACGGCAGTAGCTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

130-I

5′-AATTCGAGAGCCCTGCTCGAGCTGTAGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACTACCACGGCCGAAATTCGTTACCCG-3′
3′-GCTCTCGGGACGAGCTCGACATCACCCCAAGGGCTCGCCGGTTTCCCUCGUCUGAGATTTAGACGGCAGTAGCTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

130- J

5′-AATTCGAGAGCCCTGCTCGAGCTGTAGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGAAGGTTCGAATCCTTCCCCCACTACCACGGCCGAAATTCGTTACCCG-3′
3′-GCTCTCGGGACGAGCTCGACAUCACCCCAAGGGCUCGCCGGTTTCCCUCGUCUGAGATTTAGACGGCAGTAGCTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

130-K

5′-AATTCGAGAGCCCTGCTCGAGCTGTAGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCCACTTCGAAGGTTCGAATCCTTCCCCCACTACCACGGCCGAAATTCGTTACCCG-3′
3′-GCTCTCGGGACGAGCTCGACATCACCCCAAGGGCTCGCCGGTTTCCCTCGTCTGAGATTTAGACGGCAGTAGGTGAAGCTTCCAAGCTTAGGAAGGGGGTGATGGTGCCGGCTTTAAGCCATGGGCCTAG-5′

Figure 3.17 Sequence of ds130mer inserts. Bases circled in red correspond to alterations in the wild-type *supF* sequence and act as a marker to demonstrate incorporation of the synthetic 130mer oligo insert. Bases in green (130-H, 130-I and 130-J) correspond to uracil. Bases in blue (130-K) represent a GC→CG inactivating alteration. The asterisks indicate the site of O^6-MeG adducts.

3.4.2 Materials and methods

Materials

All chemicals were from Sigma-Aldrich, Poole, Dorset, UK unless otherwise stated.

Methods

All methods specific to part III are described below. For any other procedures refer to parts I and II (Chapter 3) and Chapter 2.

Vector Preparation

3.4.2.1 Digestion of pSP189 plasmid with the EcoRI restriction endonuclease

pSP189 (80 µg) was digested with 800 Units of EcoRI (New England Biolabs Ltd., Hitchin, UK) in 80 µl EcoRI Buffer (10x) (NEB), in a total reaction volume of 800 µl for 3 h at 37°C. EcoRI has the following recognition site:

$$5'...G^\blacktriangledown AATTC...3'$$
$$3'...CTTAA_\blacktriangle G...5'$$

130mer modified insert preparation

The sequences of the deoxyoligonucleotides used for construction of the 130 basepair inserts are given below (Figure 3.18).

```
                            TOP STRAND (5'-3'):
T0:          AAT TCG AGA GCC CTG C                    16MER
T1:          TCG AGC TGT AGT GGG GTT                  18MER
T1-O6-MeG:   TCG AGC TGT AXT GGG GTT                  18MER
T2:          CCC GAG CGG CCA AAG GGA GC               20MER
T3:          AGA CTC TAA ATC TGC CGT C                19MER
T3-O6-MeG:   AXA CTC TAA ATC TGC CGT C                19MER
T4:          ATC GAC TTC                               9MER
T4C:         ATC CAC TTC                               9MER
T5:          GAA GGT TCG AAT CCT T                    16MER
T6:          CCC CCA CTA CCA CGG CC                   17MER
T7:          GAA ATT CGG TAC CCG                      15MER


                          BOTTOM STRAND (5'-3'):
B0:          TCG AGC AGG GCT CTC G                    16MER
B1:          CTC GGG AAC CCC ACT ACA GC               20MER
B1U:         CUC GGG AAC CCC ACU ACA GC               20MER
B2:          GAG TCT GCT CCC TTT GGC CG               20MER
B2U:         GAG UCU GCU CCC TTT GGC CG               20MER
B3:          TCG ATG ACG GCA GAT TTA                  18MER
B3G:         TGG ATG ACG GCA GAT TTA                  18MER
B4:          ACC TTC GAA G                            10MER
B5:          TGG GGG AAG GAT TC G A                   16MER
B6:          AAT TTC GGC CGT GGT AG                   17MER
B7:          GAT CCG GGT ACC G                        13MER
```

Figure 3.18 Sequences of the deoxyoligonucleotides used for the construction of the ds130mer inserts. X represents an O^6-MeG adduct and U represents a uracil base.

3.4.2.2 Oligomer phosphorylation with T4 polynucleotide kinase (T4 PNK)

PAGE purified oligos B0, T2, B2, T3, B3, T4, B4, T6, B6, T3-O6-MeG and B2U were phosphorylated as in section 3.2.2.7.

3.4.2.3 Construction of the ds66mers

The ligation strategy employed was the same as that outlined in part II (section 3.3.2.2). Eleven different ds66mers were constructed from the following ss oligos (Figure 3.18):

66-A: T1+T2+T3+T4+B1+B2+B3

66-B: T1+T2+T3+T4+B1U+B2+B3

66-C: T1+T2+T3+T4+B1+B2U+B3

66-D: T1+T2+T3+T4+B1U+B2U+B3

66-E: T1-O6-MeG+T2+T3+T4+B1+B2+B3

66-F: T1+T2+T3-O6-MeG+T4+B1+B2+B3

66-G: T1-O6-MeG+T2+T3-O6-MeG+T4+B1+B2+B3

66-H: T1-O6-MeG+T2+T3+T4+B1U+B2+B3

66-I: T1+T2+T3-O6-MeG+T4+B1+B2U+B3

66-J: T1-O6-MeG+T2+T3-O6-MeG+T4+B1U+B2U+B3

66-K: T1+T2+T3+T4C+B1+B2+B3G

3.4.2.4 Construction of the ds130mer inserts

Eleven different ds130mer inserts (Figures 3.16 and 3.17) were constructed
using the following PAGE purified ds66mers, ds48mer and ss16mers B0 and
T0:

130-A: 66-A+48MER+B0+T0

130-B: 66-B+48MER+B0+T0

130-C: 66-C+48MER+B0+T0

130-D: 66-D+48MER+B0+T0

130-E: 66-E+48MER+B0+T0

130-F: 66-F+48MER+B0+T0

130-G: 66-G+48MER+B0+T0

130-H: 66-H+48MER+B0+T0

130-I: 66-I+48MER+B0+T0

130-J: 66-J+48MER+B0+T0

130-K: 66-K+48MER+B0+T0

Equimolar amounts of each ds66mer, ds48mer, ssB0 and ssT0 (~23 pmol) were mixed with 100 µl T4 DNA ligase buffer (10x) (NEB) and the volume of the solution was adjusted to 1 ml with water. The oligos were annealed together by heating at 95°C for 5 min then allowing to cool slowly to room temperature over a period of 1 h. Subsequently, 4000 cohesive end units of T4 DNA ligase (NEB) were added and the ligation reaction was carried out for 1h and 30 min at room temperature. The ligation reaction was terminated by heat inactivation of T4 DNA ligase (65°C, 10 min). The resulting ds130mers were purified by 20% non-denaturing PAGE prior to ligating to the vector following the protocol detailed in Section 3.2.2.6.

3.4.3 Results of Part III

3.4.3.1 Digestion of pSP189 with the EcoRI restriction endonuclease

The products arising from digestion of pSP189 plasmid with EcoRI are illustrated in Figure 3.19. The gel photo indicates that the digestion reaction to linear plasmid is nearly complete.
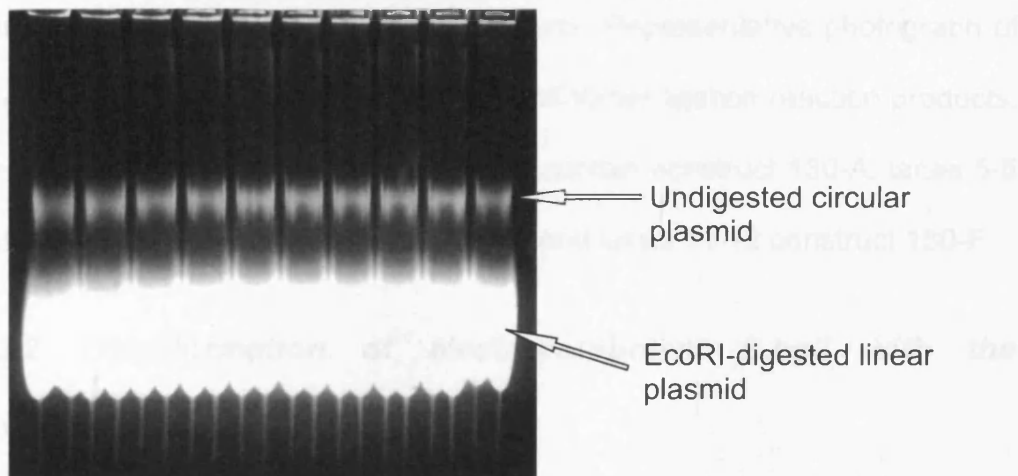
Figure 3.19 Digestion of pSP189 with EcoRI. Photograph of a typical 1% preparative agarose gel showing EcoRI-digested linear product.

3.4.3.2 Construction of the ds130mer inserts

The products of the ds130mer ligation reaction (3.4.2.4) are shown in Figure 3.20. The reaction resulted in bands representing the usual ds114mer fragments and in additional bands corresponding to fragments of the expected size (130 bp).
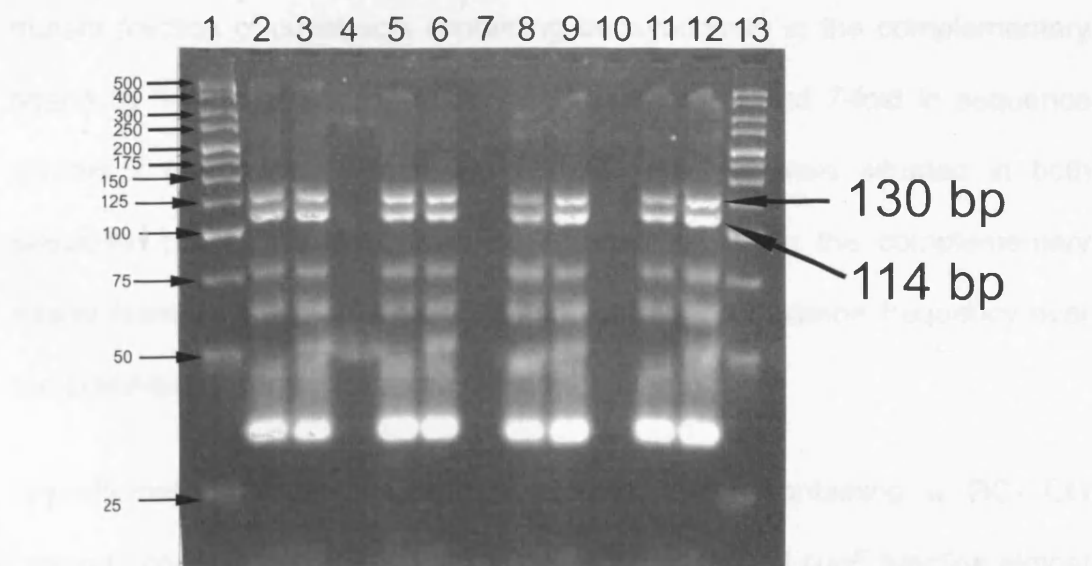
Figure 3.20 Construction of ds130mer inserts. Representative photograph of a 4% analytical agarose gel showing the ds130mer ligation reaction products. Lane 1 contains a 500bp ladder, lanes 2-3 contain construct 130-A, lanes 5-6 construct 130-B, lanes 8-9 construct 130-E and lanes 11-12 construct 130-F.

3.4.3.2 Transformation of electrocompetent E.coli with the recombinant vector

The results of *E.coli* transformation with the recombinant constructs produced from ligation of the 130mers with the digested vector are summarized in Table 3.5 and Figure 3.21. Transformation of *E.coli* with the unadducted recombinant plasmid (sample 1) resulted in a background mutation frequency of 2 ± 1.4%. The mutant fraction was increased slightly (2.3-3.6%) in uracil-containing constructs (samples 2, 3 and 4) and in O^6-MeG adducted plasmids (3.4-4.8%)in the absence of bottom strand uracil bases (samples 5, 6 and 7). However, the placement of deoxyuridines on the complementary strand at 5'- and 3'-positions flanking the O^6-MeG adducts resulted in significant enhancements of the mutation frequency. In particular, O^6-MeG increased the

mutant fraction of constructs containing deoxyuridines in the complementary strand by 14-fold in sequence context 1 (sample 8) and 7-fold in sequence context 2 (sample 9). When two O^6-MeG adducts were situated in both sequence contexts in the presence of uracil bases in the complementary strand (sample 10) there was a 17-fold increase in mutation frequency over the uracil-alone control sample (sample 4).

Transformation of *E.coli* with recombinant vector containing a GC→CG basepair change at position 150 (sample 11) inactivated *supF* function almost completely as is evident from the high mutation frequency observed (99.3 ± 0.2%). Validation samples 12 and 13 were prepared by mixing the 130-A and 130-K constructs in different proportions in the ligation reaction (75:25 and 50:50 respectively). When these samples were used to transform *E.coli*, they produced mutant fractions of 27.9 ± 5.2 and 46.1 ± 0.6 respectively. Finally, when the bacteria were transformed with the recombinant doubly cut dephosphorylated vector ligated in the absence of the modified insert (negative control sample) this produced very few colonies, confirming that contamination with the parent undigested pSP189 was insignificant.

Sample	Type of 130mer insert	Exp1 total	Exp1 whites	Exp1 % whites	Exp2 total	Exp2 whites	Exp2 % whites	Exp3 total	Exp3 whites	Exp3 % whites	Average M.F. from all exps	STDEV
1	130-A	2478	46	1.9	27832	981	3.5	727	5	0.7	**2.0**	1.4
2	130-B	5075	73	1.4	16315	650	4.0	856	12	1.4	**2.3**	1.5
3	130-C	1750	35	2.0	16135	948	5.9	421	4	1.0	**2.9**	2.6
4	130-D	4809	117	2.4	16110	1048	6.5	401	8	2.0	**3.6**	2.5
5	130-E	9890	377	3.8	17124	1047	6.1	786	19	2.4	**4.1**	1.9
6	130-F	13909	423	3.0	27629	1466	5.3	907	18	2.0	**3.4**	1.7
7	130-G	8948	463	5.2	7485	472	6.3	1003	30	3.0	**4.8**	1.7
8	130-H	3223	887	27.5	308	120	39.0				**33.2**	
9	130-I	1727	412	23.9	7450	2100	28.2	398	51	12.8	**21.6**	7.9
10	130-J	372	278	74.7	90	44	48.9				**61.8**	
11	130-K	16411	16274	99.2	6611	6578	99.5				**99.3**	
12	75% 130-A +25% 130-K				8517	2687	31.5	844	204	24.2	**27.9**	
13	50% 130-A +50% 130-K				3162	1471	46.5	844	386	45.7	**46.1**	
14	No insert	6	2		29	4		3	2			

Table 3.5 Number of mutant and total colonies resulting from the transformation of electrocompetent *E.coli* cells with recombinant constructs. Experiments 1 to 3 represent separate vector to insert ligation reactions. Average mutation frequencies and standard deviations are shown in blue font. Samples are colour-coded according to the type of modified construct they contain; red = unadducted, lavender = unadducted (+) uracil, pale blue = adducted (−) uracil, yellow = adducted (+) uracil, light green = validation oligos, grey = (−ve) control.
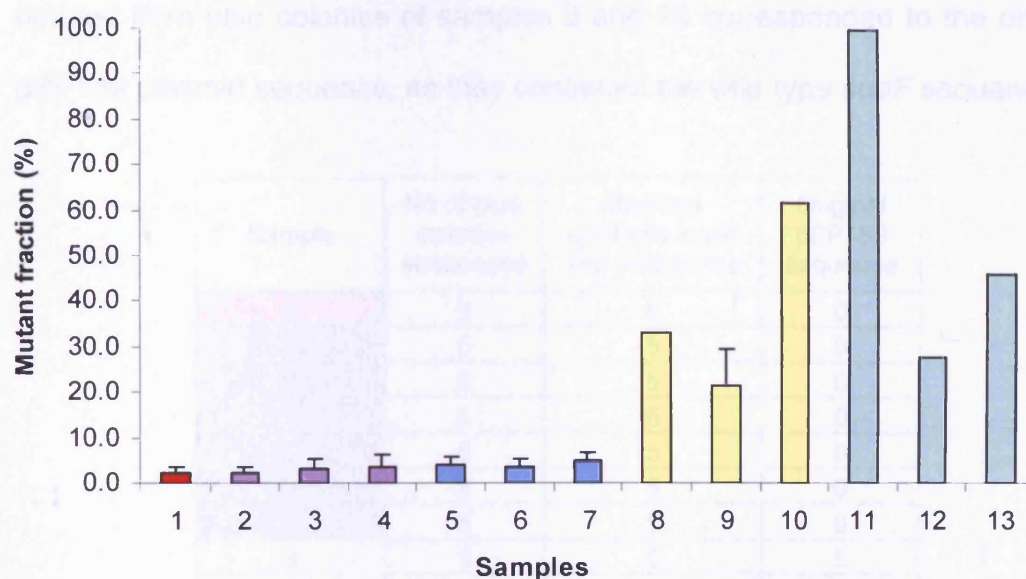
Figure 3.21 Graphical representation of *E.coli* transformation results. Average mutation frequencies from two (samples 8, 10, 11, 12 and 13) or three (samples 1, 2, 3, 4, 5, 6, 7 and 9) separate experiments. Error bars represent standard deviation, where three expts were performed. Samples are colour-coded according to the type of modified construct they contain; red = unadducted, purple = unadducted (+) uracil, light blue = adducted (−) uracil, yellow = adducted (+) uracil, light green = validation oligos.

3.4.3.3 DNA sequencing

A small number of blue colonies from all recombinant samples were picked randomly and sent for sequencing to PNACL. Sequences derived from blue colonies of all samples, except for samples 8 and 10, confirmed the presence of the modified 130mer insert illustrated by the presence of altered bases at positions 99 and 179, proving that it was successfully ligated into the vector

and that there were no inactivating mutations in it. In contrast, all sequences

derived from blue colonies of samples 8 and 10 corresponded to the original

pSP189 plasmid sequence, as they contained the wild type *supF* sequence.

Sample	No of blue colonies sequenced	Modified synthetic insert (no mutations)	Original pSP189 sequence
1	5	5	0
2	5	5	0
3	5	5	0
4	5	5	0
5	5	5	0
6	5	5	0
7	5	5	0
8	5	0	5
9	2	2	0
10	1	0	1
11	0	0	0
12	4	4	0
13	4	4	0

Table 3.6 Nature of the sequences derived from blue colonies of recombinant

samples.

A small number of white colonies from all samples were randomly picked and

sent for sequencing to PNACL. The types of mutations seen in sequences of

plasmid extracted from white colonies are summarized in Tables 7, 8, 9, 10

and 11.

Sample	No of white colonies sequenced	Mutation type
1	15	T DEL @153 (x7)
		DEL 79→263 + 10bp INS @junction
		DEL 76→158
		C DEL @110
		G DEL @160
		A DEL @177
		C DEL @110
		T-C @138
		A-C @158

Table 3.7 Mutations observed in DNA sequences derived from white colonies of sample 1 (unadducted control).

DNA sequencing of plasmid extracted from white colonies of unadducted sample 1 revealed that a variety of mutations were responsible for the mutagenic background (2 ± 1.4%) observed (Table 3.7). The major mutation observed (7/15) was a point deletion at position 153, followed by two large deletions (2/15), one of which occurred at the vicinity of the EcoRI restriction site. The remaining inactivating mutations were random point deletions (4/15) and random substitutions (2/15).

DNA sequencing results from control unadducted samples containing deoxyuridines on the complementary strand (samples 2, 3 and 4) are summarized in Table 3.8. Inclusion of deoxyuridines at sequence context 1 in the absence of adduct (sample 2) induced similar types of background mutations to those found in sample 1. The predominant mutation events were point deletions at positions 153 (6/15) and 152 (2/15). A large deletion of the whole *supF* occurred between the EcoRI restriction site and position 189 (1/15) while the remaining mutations were random point deletions (5/15) and one substitution (1/15).

Sample	No of white colonies sequenced	Mutation type
2	15	T DEL @153 (x6)
		C DEL @152(x2)
		DEL 73→189
		A DEL @130
		C DEL @178
		G DEL @111 and G DEL @115
		G DEL @113 and G DEL @116
		G DEL @126
		G-C @105
3	14	T DEL @153 (x3)
		DEL 74→177
		DEL 79-263 + 11bp INS @junction
		DEL 74→190
		G DEL @160
		G DEL @102
		C DEL@110 and C DEL@127
		G DEL @129
		C INSERTION @131
		A-C @120
		G-A @104
		T-A @140
4	14	DEL 74→189 (x2)
		T DEL @153 (x2)
		DEL 73→189
		G DEL @99 and DEL 171→210
		C DEL @152
		G DEL @102
		A DEL @130
		A DEL @177
		C DEL @127
		C-T @168
		G-A @102
		A-C @120

Table 3.8 Mutations observed in DNA sequences derived from white colonies

of uracil-containing control samples 2, 3 and 4

The presence of uracil bases at sequence context 2 (sample 3) again induced

a similar mutation spectrum, with the point deletion at T153 being the

predominant mutation (3/14), followed by large deletions between the EcoRI

190

site and distant positions 177, 190 and 263 (3/14). The remaining mutations were random point deletions (4/14), substitutions (3/14) and an insertion (1/14). Likewise, when uracils were present at both sequence contexts (sample 4) the prevailing mutation was again a deletion of thymine at position 153 (2/14), followed by large deletions (4/14) random point deletions (5/14) and substitutions (3/14). When the mutation spectra from the uracil-containing samples (samples 2, 3 and 4) are combined, the main mutation type observed are random point deletions (27/43 sequences), followed by large deletions (8/43), point substitutions (7/43) and a point insertion (1/43).

DNA sequences of plasmid extracted from 15 white colonies of samples 5, 6 and 7 were analyzed for mutations and the results are summarized in Table 3.9. When a single O^6-MeG adduct was placed site-specifically at sequence context 1 and the recombinant plasmid (sample 5) was replicated in *E.coli*, a targeted G→A transition was observed at position 100 in one third of the mutants sequenced (5/15). The remaining inactivating mutations were of the same type seen in previous control samples 1 to 4, namely a T deletion at 153 (2/15), large deletions between the EcoRI site and position 189 (2/15), random point deletions (4/15) and substitutions (2/15), indicating they are not related to the adduct. Positioning of the O^6-MeG adduct at sequence context 2 (sample 6) resulted in targeted G→A transitions at position 129 in more than half (8 out of 15) of the mutants sequenced. A point deletion at the same position was also observed. The remaining inactivating mutations were random point deletions (3/15) and substitutions (3/15). Incorporation of two O^6-MeG adducts on a single construct (130-G-sample 7) resulted in targeted G→A transitions in either sequence context 1 (1/15) or sequence context 2

(4/15). However, no sequences with both G→A transitions present on the same *supF* gene were observed. The rest of the mutation spectrum was comprised of large deletions between the EcoRI site and positions 189 or 190 (6/15), a large deletion of part of the *supF* gene (1/15), two point deletions (2/15, one of which was at position 153) and a random substitution.

Sample	No of white colonies sequenced	Mutation type
5	**15**	G-A @100 (T1-06-MeG targeted) (x5)
		T DEL @153 (x2)
		DEL 73→189 (x2)
		G DEL @113 and G DEL @116
		C DEL @155
		A DEL @118
		G DEL @122
		T-A @140
		C-T @182
6	**15**	G-A @129 (T3-O6-MeG targeted) (x8)
		G DEL @129 (T3-O6-MeG targeted)
		G DEL @159
		T DEL @153
		G DEL @122
		G-A @104
		G-A @124
		G-A @144
7	**15**	G-A @129 (T3-O6-MeG targeted) (x4)
		DEL 73→189 (x3)
		DEL 74→189 (x2)
		G-A @100 (T1-06-MeG targeted)
		DEL 74→190
		DEL 171→210
		T DEL @153
		C DEL @133
		C-T @139

Table 3.9 Mutations observed in DNA sequences derived from white colonies of samples 5, 6 and 7 which contain one or two O^6-MeG adducts.

Inclusion of deoxyuridines on the strand complementary to the adducts at 5' and 3' flanking positions (samples 8, 9 and 10) resulted in variable mutation patterns, depending on the sequence context on which the adduct(s) was placed. The mutation types observed in the DNA sequences from these samples are summarized in Table 3.10.

Sample	No of white colonies sequenced	Mutation type
8	20	DEL 73→189 (x9)
		DEL 74→189 (x4)
		T DEL @153 (x3)
		DEL 74→190
		DEL 74→157
		G-A @100 (T1-O6-MeG targeted) and T DEL @153
		G DEL @102
9	10	G-A @129 (T3-O6-MEG targeted) (x8)
		A DEL @128
		C DEL @131
10	11	DEL 72→207 (x3)
		DEL 73→207 (x2)
		DEL 73→203 (x2)
		DEL 71→208
		DEL 73→208
		DEL 74→204
		DEL 74→263 plus 17bp INS

Table 3.10 Mutations observed in DNA sequences derived from white colonies of samples 8, 9 and 10, which contain a single or two O^6-MeG adducts in gapped plasmids.

Positioning of the O^6-MeG adduct at sequence context 1 (position 100) in the presence of deoxyuridines (sample 8) resulted in large deletions between the EcoRI site and distant bases (at positions 189, 190 and 157) in the majority of mutants (15/20). The remaining mutations were comprised of the T153 point deletion (3/15), a double mutation consisting of a G→A targeted transition at

position 100 and a point deletion at 153 (1/15) and a random point deletion (1/15).In contrast, placement of the O^6-MeG adduct at sequence context 2 (position 129) in the presence of deoxyuridines resulted predominantly in G→A targeted transitions (8/10), followed by random point deletions (2/10). When O^6-MeG adducts and deoxyuridines were present in both sequence contexts (sample 10), the prevailing mutation obtained was a large deletion of bases between the EcoRI site and the BamHI site (10/11) or the EcoRI site and position 263 (1/11).

Mutation analysis of DNA sequences from plasmids containing a GC→CG inactivating basepair alteration at position 150 (samples 11, 12 and 13) is summarized in Table 3.11. DNA sequences from these validation samples confirmed the presence of the GC→CG inactivating mutation in 100% of the mutants sequenced.

Sample	No of white colonies sequenced	Mutation type
11	8	G-C @150 (Targeted) (x8)
12	8	G-C @150 (Targeted) (x8)
13	8	G-C @150 (Targeted) (x8)

Table 3.11 Mutations observed in DNA sequences derived from white colonies of validation samples 11, 12 and 13. Sample 11 was prepared using an inactivating 130mer construct (130-K), while samples 12 and 13 were prepared by mixing the 130-A and 130-K constructs in different proportions in the ligation reaction (75:25 and 50:50 respectively).

3.5 DISCUSSION

3.5.1 Discussion of Part I

The main aims of this early method development were to ensure (i) that the *supF* activity is preserved by the unadducted synthetic insert and (ii) that all colonies would result from the recombinant vector and not from contamination with the parent undigested vector or with singly-digested religated vector. Digestion of pSP189 with XhoI and BamHI restriction endonucleases was performed in two sequential reactions, each followed by agarose gel purification in order to maximize the purity of the linear digested product. Ligation of doubly-digested dephosphorylated plasmid in the absence of insert resulted in very few or more typically zero colonies, suggesting that dephosphorylation with phosphatase completely prevented religation of singly-cut vector and that contamination with undigested circular pSP189 was insignificant. Dephosphorylation of the doubly-cut plasmid was considered a necessary step as early experiments using phosphorylated plasmid (data not shown) produced high numbers of blue colonies resulting from religation of contaminant singly-digested linear plasmid, reducing thus the sensitivity of the assay.

The modified 114mer insert was constructed from 6 smaller PAGE-purified ss oligodeoxynucleotides. All ss oligos were phosphorylated at their 5' ends prior to ligation using T4 Polynucleotide kinase. This resulted in the formation of higher molecular weight multimers (concatemers) and consequently the ds114mer insert was constructed in two sequential ligation reactions with

agarose purification of intermediates in order to increase the purity and yield of the final construct. Selective phosphorylation of constituent oligonucleotides in later experiments (parts II and III) eliminated formation of concatemers.

Transformation of *E.coli* with the recombinant vector resulted in thousands of colonies, the majority of which were blue, non-mutants (90.6%). Plasmids from three separate blue colonies were sequenced and all three sequences confirmed the presence of the modified insert, indicating that blue colonies had resulted from replication of the recombinant plasmid. However, the background mutation frequency was still relatively high (9.4 ± 3%) and DNA sequencing analysis of plasmid derived from white colonies revealed that random point mutations were responsible. The source of these background inactivating mutations was unknown. One potential source considered was the short exposure of the modified insert to low intensity 254 nm UV light during band excision from the agarose gel or during photography of the gel. To investigate this possibility, the insert was exposed to low intensity UV light for different periods of time (one or three minutes) to mimic typical exposures encountered during oligo recovery prior to ligating to the vector and *E.coli* transformation and the resulting mutation frequency was compared to minimally exposed insert protected from UV light as much as possible. Additional exposure to UV light failed to increase the mutation frequency and consequently UV light exposure was excluded as a possible cause of the mutational background observed. A second potential source of random mutations could be errors introduced during the oligodeoxynucleotide synthesis process, as phosphoramidite chemistry is not 100% efficient. As oligo length increases, the number of incomplete synthesis products (n-1

species) increases and purity of oligos generally decreases with increasing length, which is Partly why the insert was constructed from smaller oligos rather than using a synthetic full length product. PAGE purification of crude oligos eliminates most of these by-products, but not all of them. This would result in point deletions in mutagenesis studies. Single base substitutions are more rare by-products and could be generated due to the purity of reagents used during the oligo synthesis.

3.5.2 Discussion of Part II

Construction of the unadducted modified insert (114-A) from smaller oligodeoxynucleotides and PAGE purification of the ds114mer inserts resulted in a reduced mutation frequency (~3%), suggesting that errors introduced during the chemical synthesis of longer oligomers were the main source of the point mutations observed in previous experiments (Part I).

DNA sequencing of mutant colonies from the unadducted plasmid (114-A) revealed that the predominant background mutation was a large deletion of bases 90→261. Inspection of the sequence at the junctions suggests that deletion of bases 90→261 resulted from perfect rejoining of DNA ends at the XhoI restriction site. Since there is only one XhoI restriction site in the pSP189 plasmid it would be virtually impossible for the ends to rejoin without any mutations introduced at the junction site. A potential mechanism for the generation of this large mutation involves the 8 bp signature of the pSP189 plasmid. pSP189 contains a randomly generated 8 bp signature 3' of the *supF* gene providing 4^8 (65,536) sequence possibilities [468]. Since the 8 bp signature is randomly generated, a certain proportion of plasmids (64/65536

copies) would have a signature ending in "ctc". Additionally, the sequence immediately 3' to the 8 bp signature is "gag". It is possible, therefore, that a second XhoI site was generated in a certain proportion of plasmids. As a result, digestion with XhoI would generate compatible sticky ends, *in vivo* phosphorylation and ligation of which would result in the 90→261 deletion observed. The remaining mutations seen in this sample (114-A) were random point deletions and substitutions, probably resulting from unavoidable errors introduced during the oligo synthesis.

Inclusion of deoxyuridines in the complementary strand (114-B) increased the background mutation frequency to ~7%. DNA sequencing of colonies arising from this sample showed that the major mutation was the same large deletion of bases 90→261, as in sample 114-A. A second mutation involving deletion of bases 90→261 and insertion of 18 bp at the junction was seen in 2 out of 5 colonies sequenced. Although the exact mechanism by which this mutation was generated is unclear, it is possibly similar to the one described above.

Plasmids adducted with a single O^6-MeG in sequence context 1 (sample 114-C) induced a mutation frequency of ~7%. The predominant mutation observed was again the large 90→261 deletion, followed by O^6-MeG targeted G→A substitutions. Two large deletions of bases 88→203 and 129→203 were also observed. Although the mechanism by which these latter two large deletions were generated is uncertain, they both involve the BamHI restriction site. Exonucleolytic hydrolysis from the GATC site (position 204) of the modified insert by MMR in response to the O^6-MeG adduct and subsequent recombination of the digested ends could possibly account for the large

deletions observed. Finally, a G→A transition was seen at position 129 and probably reflects a spontaneous mutation, as this mutation was very common in spontaneous *supF* spectra (Chapter II).

Plasmids containing a single O^6-MeG adduct at sequence context 2 (sample 114-D) induced about a 3-fold increase in the mutation frequency compared to unadducted sample 114-A. The small increase observed in the mutation frequency in response to the O^6-MeG adduct, in either sequence context, is consistent with previous investigations comparing adduct mutagenesis in ds vs gapped plasmids [430, 431, 465] that reported only small increases in mutation frequency in ds plasmids. The mutation types observed were similar to the ones detected in previous samples and primarily involved large deletions between the XhoI site and the 8 bp signature, followed by a large deletion of bases 141→210 and a random point substitution. Interestingly, the 3-fold increase in the mutation frequency was not accompanied by a similar increase in O^6-MeG targeted G→A transitions at position 129. A possible explanation for this increase could be an upregulation of DSB repair in response to the O^6-MeG adduct that would result in increased rejoining of compatible dephosphorylated XhoI overhangs *in vivo*.

3.5.3 Discussion of Part III

Experiments using the BamHI and XhoI endonucleases (part II) suggested that the majority of background mutations were large deletions resulting from *in vivo* recombination of XhoI compatible dephosphorylated ends and that the 8 bp signature was responsible for generating the second XhoI site. In order to avoid these nonspecific background mutations and increase the sensitivity

of the assay the XhoI endonuclease was replaced by EcoRI, as described in part III. This modified recombination strategy resulted in a further decrease of the mutation background (2 ± 1.4%) in the unadducted plasmid (sample 1) and successfully eliminated the 90→261 deletion from all samples, as it was absent from all white colonies sequences.

Mutation analysis of the unadducted control plasmid (sample 1) constructed using the new EcoRI protocol showed that the majority of background mutations were random point deletions and substitutions. These mutations were probably due to by-products of DNA synthesis that were not entirely removed by PAGE purification. Among the single base deletions, deletion of a thymine at position 153 was most frequent, suggesting poor efficiency of DNA synthesis and/or PAGE purification of the respective oligomer (T4). Inclusion of deoxyuridines in the complementary strand of unadducted plasmids (control samples 2, 3 and 4) resulted in small enhancements of the mutant fraction (1.1-fold, 1.4-fold and 1.8-fold respectively). Mutation spectra obtained from these samples were remarkably similar to each other and also to sample 1 and implicate DNA synthesis errors as the main source of background mutations.

Plasmids containing a single (samples 5 and 6) or two (sample 7) O^6-MeG adducts in the absence of deoxyuridines in the complementary strand, induced small increases in the mutation frequency compared to sample 1 (2-fold, 1.7-fold and 2.4-fold respectively). The large percentage of blue colonies produced by these double-stranded O^6-MeG-containing plasmids resulted from the lesion inducing preferential use of the unmodified complementary

strand as the DNA template (damage tolerance) instead of continuing synthesis past the adduct. The small enhancement of the mutant fraction observed in these samples was due to targeted O^6-MeG G→A transitions in a large proportion of the mutants, while the rest of the alterations were the same types of mutations seen in previous control samples, particularly single base deletions. More specifically, plasmids containing a single O^6-MeG adduct at sequence context 1 (sample 5) gave the expected targeted G→A transitions in one third of the mutants sequenced, while placing the O^6-MeG adduct at sequence context 2 (sample 6) resulted in targeted G→A transitions in more than half of the mutants sequenced. When two O^6-MeG adducts were placed on the same plasmid, polymerase replication induced G→A transitions in either sequence context, albeit 4 times more frequently in sequence context 2, suggesting that translesion synthesis occured more frequently in the second sequence context. Additionally, no plasmids harbouring both G→A transitions on the same *supF* gene were seen. This finding could be explained by preferential translesion synthesis of one O^6-MeG adduct by DNA polymerase and lesion bypass of the second. It is possible that if more samples were sequenced, however, a double substitution may have been detected.

Replacement of thymidines that flank the adducted positions on the complementary strand with deoxyuridines enhanced significantly the mutagenic potential of O^6-MeG in either (samples 8 and 9) or both (sample 10) sequence contexts, indicating that in the absence of an intact complementary strand translesion synthesis was continued out past the adducted templates. These findings suggest that the sequence context in which a single O^6-MeG adduct is placed can significantly influence its

mutagenicity in gapped plasmids and that replication past two O^6-MeG adducts (on the same *supF* gene) is more mutagenic in total than replication of a single O^6-MeG adduct. Moreover, mutation analysis of sequences derived from these three samples revealed that inactivation of *supF* was induced by different types of mutations, depending on the sequence context of the O^6-MeG adduct and the presence of one or two O^6-MeG adducts on the same *supF* gene.

Most of the mutations observed (14/20) in colonies derived from O^6-MeG adduction at sequence context 1 in gapped plasmids (sample 8) were large deletions of bases between the EcoRI site (positions 73 or 74) and distant positions 189 or 190. Point deletions reflecting errors introduced during chemical synthesis of the oligos were seen in four out of 20 mutants sequenced while only one O^6-MeG targeted G→A transition was observed. Positioning of the O^6-MeG adduct in sequence context 2 (sample 9) induced a strikingly different mutation spectrum compared to sample 8. The majority of mutations observed in sample 9 (8/10) were O^6-MeG targeted G→A transitions, followed by random point deletions (2/10) that presumably resulted from errors during the oligo synthesis. Interestingly, when two O^6-MeG adducts were placed on the same *supF* gene in gapped plasmids (sample 10) almost all of the resulting mutations observed (10/11) were large deletions between the EcoRI and BamHI sites.

The nature of the frequent large deletion between the EcoRI site and positions 189 or 190 in sample 8 indicates that recombination repair of adducted plasmid between regions of homology was carried out, as a 6 bp homologous

sequence (AATTCG) is present at both ends of the deleted sequence. The presence of homologous sequences from 6 to 20 bp long at the ends of deleted regions invokes mechanisms that cause deletion by annealing complementary regions from direct repeats on opposite ends of the deleted sequence [469]. This conclusion is further supported by the fact that all blue colonies sequenced from this sample had resulted from recombination of original singly-digested pSP189 contaminants and not from plasmid containing the modified insert. Although dephosphorylation of doubly-digested vector completely prevented religation of these contaminants *in vitro,* it would not prevent recombination in cells following activation of the homologous recombination repair system.

This large deletion was also detected in uracil control samples (samples 2, 3 and 4) and in samples containing a single O^6-MeG adduct at sequence context 1 (sample 5) or two O^6-MeG adducts (sample 7) in ds plasmids, albeit in much lower frequency. Interestingly, the deletion was absent in samples containing a single O^6-MeG adduct at sequence context 2 in both ds (sample 6) and gapped plasmids (sample 9). It is not entirely clear why this might have occurred. One possibility is that homologous recombination was more efficient when the adduct was placed at sequence context 2 and resulted in no mutations. A second option is that the presence of the O^6-MeG adduct at sequence context 2 did not induce a homologous recombination response at all. This scenario is favourable because of the fact that all blue colonies derived from these samples had resulted from plasmid containing the modified insert and not from recombined original pSP189. Nevertheless, it is not clear why recombination would not take place in these samples. It is possible that

the sequence context could affect the ability of the O^6-MeG adduct to generate DSBs.

It has been suggested that homologous recombination is essential to repair both DSBs and single-strand gaps in *E.coli* [432]. However, relatively little is known about the role of homologous recombination in response to methylation damage in *E.coli*. The RecA recombinase is the key player in homologous recombination as it alone can mediate homology recognition and exchange of DNA strands [470, 471]. Although there are numerous reports in the literature documenting the increased sensitivity of *recA* mutants to methylating agents (e.g. [472]), these results are difficult to interpret because of the multiple roles of RecA in cell physiology including regulation of the SOS response and translesion synthesis [473]. Unlike RecA, there are other genes that have well-defined roles in homologous recombination that are not part of SOS or TLS. For example, RecBCD and RecFOR are essential components of the two major recombinational repair pathways in *E.coli*, the former being primarily for the purpose of processing DSBs and the latter for processing single-strand gaps [474, 475].

In the present system, homologous recombination between the 6bp direct repeats presumably resulted in response to both single-strand gaps created by the action of AP endonucleases in uracil-containing control plasmids (samples 2, 3 and 4) and single or double strand breaks induced by O^6-MeG adduct(s) when placed at sequence context 1, but not at sequence context 2. DSBs could arise as a consequence of interference between MMR and BER. Overlapping BER and MMR repair tracts on complementary strands could

give rise to DSBs and DSBs have been detected in *E.coli dam recB ada ogt* cells exposed to MNNG in a dose- and MMR-dependent manner [432]. In the current recombinant system MMR should be activated in response to O^6-MeG:T mismatches, while BER could be induced by deoxyuridines in the complementary strand and by ss nicks in the EcoRI and BamHI sites that would result due to dephosphorylation of the doubly-cut plasmid prior to the ligation reaction. The presence of these lesions would induce repair by uracil DNA glycosylase and by AP endonucleases.

Deletion of bases between EcoRI and BamHI sites in plasmids containing both O^6-MeG adducts in gapped plasmids (sample 10) might have arisen as a result of a separate recombination mechanism, since this deletion was not observed in any of the sequences derived from other samples. Unlike in the case of the large deletion between the EcoRI site and position 189 or 190 (described above), there is no sequence homology between the EcoRI and BamHI sites of linear, doubly-cut plasmid molecules. In the absence of homology, *E.coli* repair systems will join noncomplementary ends by a mechanism that produces junctions identical to those produced by NHEJ in mammalian systems [476]. Since NHEJ is exceedingly rare or often absent in *E.coli* [477], the mechanism by which noncomplementary ends are joined has remained obscure. It has been proposed that DNA polymerases might facilitate illegitimate recombination by stabilizing transient constructs by primer extension [478]. DNA polymerases can join DNA ends of various structures in *vitro*. Short overlaps of one to four bases can be joined by an exonuclease-free Klenow fragment of *E.coli* DNA pol I [476]. Mismatched bases inhibited but did not prevent end-joining in these assays. In this assay system, it is

conceivable that DSBs generated by the two O^6-MeG adducts in gapped plasmids (sample 10) might have upregulated this unique recombination system and gave rise to the deletion observed.

Validation of the assay was achieved using a ds130mer modified construct containing a GC→CG inactivating mutation at position 150 (130-K). As expected, ligation of this construct to the doubly-cut dephosphorylated vector (sample 11) resulted in an almost complete inactivation of *supF* function. Furthermore, mixing of this insert with 130-A in two different proportions in the ligation reaction (75:25 and 50:50) produced mutant fractions close to the expected ones (~28 and 46% compared to the predicted 25 and 50% respectively). Importantly, mutation analysis of plasmid derived from white colonies of these three samples (samples 12, 13 and 14) confirmed that all 24 mutants sequenced had resulted from the targeted GC→CG substitution at position 150.

Mutagenesis of O^6-MeG has been previously studied both *in vitro* and *in vivo* site-specific systems, including ss [453, 454], ds [466] and gapped vectors [430, 431]. These studies have showed that the mutagenicity of O^6-MeG DNA adducts can be attributed to its potential to mispair with thymine, thereby introducing G→A transitions after two rounds of replication. However, most previous studies relied on differential oligonucleotide hybridization or mass spectrometry for detecting mutations, restricting their ability to detect substitution mutations and/or small deletions in short adducted oligonucleotides. Our current approach relies on DNA sequencing and can therefore detect larger deletions and rearrangements in response to DNA

damage and also assess the effect of DNA adducts on distal bases. Our results suggest that mutations induced by recombination repair in response to O^6-MeG adduct(s) in *E.coli* can vary significantly depending on the sequence context surrounding the adduct. Furthermore, the role of large deletions induced by homologous recombination might have been seriously overlooked in previous studies as they may constitute a significant proportion of the mutation spectrum induced by this lesion in *E.coli*.

3.5.4 Summary

The overall aim of this work was to develop and validate a novel adduct site-specific assay in *E.coli* that could then be adopted for use in human cells. Utilising DNA recombination techniques, the assay allows for insertion of a DNA adduct at any conceivable position in the *supF* gene. Therefore, it offers great potential in helping to unravel the role of both local and distal sequence context in the mutagenic processing of DNA adducts. The assay was successfully validated using an inactivating 130mer insert and was subsequently applied to the investigation of the mutagenic potential of a single or two O^6-MeG adducts at two different DNA sequence contexts in ds or gapped plasmids. Positioning of O^6-MeG adduct(s) at either site and subsequent replication in *E.coli* produced different types of mutations depending on the sequence context in which the O^6-MeG adduct(s) were placed. These results suggest that the mutagenic and cytotoxic potential of these adducts might have been underestimated and that deletions arising as a result of DNA recombination repair in response to O^6-MeG adducts might

play an important role in addition to the G→A transitions in initiating the

carcinogenic process.

CHAPTER 4

Development of a novel site-specific mutation

assay in Ad293 human kidney cells

4.1 AIMS

The aim of the work described in this Chapter was to develop and validate an assay for determining the mutagenic potential and types of mutations caused by individual adducts situated in double stranded or gapped plasmid DNA, in different sequence contexts, in human cells. The assay would be an adaptation of the *E.coli*-based site-specific assay (described in Chapter 3) for use in human cells. The main difference in the assay presented in this Chapter was the addition of a PCR step for selection of mutant plasmids after replication in Ad293 human kidney cells. The assay was validated as in Chapter 3 using an oligomer that would completely inactivate *supF* function and with oligomers containing O^6-MeG adducts in two different DNA sequence contexts in double-stranded or gapped plasmids. However, the methods could be adapted to the study of any DNA adduct, as long as a DNA-adduct phosphoramidite is available for insertion into the synthetic deoxyoligonucleotide insert.

4.2 PART I: EARLY METHODS

4.2.1 Experimental Design

The primary aim during the early stages of development was to assess the mutagenicity of unadducted recombinant plasmid after replication in human Ad293 kidney and GM00637 fibroblast cells. The recombinant vector was constructed as detailed previously from six smaller ss DNA oligos using the methods described in section 3.2.2 and the transfection procedure was similar

to the protocol described in section 2.2.10, although a smaller amount of plasmid DNA was transfected this time for the site-specific version (2 μg). Replicated plasmid DNA recovered from the human cells was digested with DpnI and subsequent screening of the replicated vector was performed in E.coli as described previously (sections 2.2.17-2.2.18).

4.2.2 Materials and methods

Materials

All chemicals were from Sigma-Aldrich, Poole, Dorset, UK unless otherwise stated.

4.2.2.1 Human cell lines

Human embryonic adenovirus-transformed kidney cells (Ad293) were cultured from cells provided by Dr. A. Dipple, National Cancer Institute, Frederick, MD, USA. Ad293 cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum (Life Technologies Ltd, Paisley, UK) at 37°C in 5% CO_2 in air. Human SV40 transformed apparently normal non-foetal cells (GM00637) were obtained from NIGMS Human Genetic Cell Repository, Camden, NJ, USA. Cells were grown in Dulbecco's modified Eagle's medium (with Earle's salts) supplemented with 15% fetal calf serum (Life Technologies Ltd, Paisley, UK) at 37°C in 5% CO_2 in air.

Methods

All methods specific to part I of this Chapter are described below. For any other procedures refer to Chapters 2 and 3.

4.2.2.2 Construction of the recombinant unadducted plasmid

Refer to Chapter 3, sections 3.2.2.1-3.2.2.10

4.2.2.3 Fugene-6 mediated transfection of Ad293 kidney and GM00637 fibroblast cells with recombinant unadducted plasmid

Ad293 cells (grown at 37°C, 5% CO_2, 90% confluent) were split 1 in 50 (approximately 4×10^5 cells) and plated in 21.5 cm^2 transfection plates in 2 ml Dulbecco's Minimum Eagle's medium (Life Technologies Ltd., Paisley, UK) (with glutamax 1 and pyridoxine, without glutamine and sodium pyruvate, 10% foetal calf serum) and grown for 68 h (37°C, 5% CO_2). Cells were 40-60% confluent prior to transfection. Three hours before transfection, used medium was aspirated and fresh medium added (2 ml). Fugene-6 (Roche Diagnostics Ltd, Lewes, East Sussex, UK) transfection reagent (6 μl) was slowly added to foetal calf serum free Dulbecco's Minimum Eagle's medium (100 μl). The tube was gently tapped to mix the contents. Plasmid DNA (2 μg) was added to the Fugene solution and incubated at room temperature for 45 min. The aliquot of DNA was carefully added dropwise to each transfection plate and swirled gently to mix. Plates were then returned to the incubator (37°C, 5% CO_2). After 24 h, used media was aspirated and replaced with fresh media (2 ml). The plasmid was reclaimed from the Ad293 cells between 45-48 h after

transfection as described in Chapter 2. The same protocol was followed for transfection of GM00637 fibroblst cells except the plasmid was reclaimed from the GM00637 cells 60 h after transfection, since GM00637 had a longer doubling time.

4.2.3 Results

4.2.3.1 Direct transformation of electrocompetent *E.coli* with recombinant vector recovered from Ad293 cells

Direct transformation of electrocompetent *E.coli* cells with recombinant plasmid recovered from the human kidney cells resulted in about 89% white colonies (Table 4.1). This high mutagenic background was unexpected, since direct transformation of *E.coli* with the same type of unadducted plasmid resulted in a much lower mutation frequency (9.3%).

Experiment	Sample	Colonies counted			
		Total	Blue	White	% White
1	1	868	74	794	91.5
	2	272	28	244	89.7
	3	167	22	145	86.8
	4	313	30	283	90.4
Average					**89.6**
2	1	278	39	239	86.0
	2	792	74	718	90.7
Average					**88.3**
Average of both experiments					**89.0**

Table 4.1 Number of mutant and wild type colonies resulting from transformation of *E.coli* cells with recombinant vector recovered from Ad293 cells. Each experiment was performed twice, using freshly digested pSP189 plasmid and recently constructed 114mer insert. In each experiment samples numbered 1-4 were identical and represent separate vector to insert ligation reactions. For each sample a single transformation of *E.coli* was carried out. The average percentage of white colonies from both experiments is highlighted in yellow.

4.2.3.2 Direct transformation of electrocompetent *E.coli* with the recombinant vector recovered from GM00637 fibroblast cells

To test whether the unexpectedly high mutagenic background observed in Ad293 cells transfected with the unadducted plasmid was specific to this cell line, GM00367 fibroblast were transfected using the same protocol. Direct transformation of electrocompetent *E.coli* cells with recombinant plasmid recovered from the human fibroblast cells resulted in about 77.6% white colonies (Table 4.2). Although this mutational background was lower compared to Ad293 cells it was still considered unacceptably high for the site-specific assay.

Experiment	Sample	Colonies counted			
		Total	Blue	White	% White
	1	74	18	56	75.7
	2	265	58	207	78.1
1	3	299	63	236	78.9
Average					77.6

Table 4.2 Number of mutant and wild type colonies resulting from the transformation of *E.coli* cells with recombinant vector recovered from fibroblast cells. Samples 1, 2 and 3 were identical and represent separate vector to insert ligation reactions. For each sample a single transformation of *E.coli* was carried out. The percentage of white colonies is highlighted in yellow.

4.2.3.3 DNA sequencing

One blue colony derived from the Ad293 cells was analysed by DNA sequencing (Figure 4.1) and this colony contained plasmid with the wild type *supF* sequence. This suggests that the colony resulted from contamination with either parent uncut or singly-cut recombined pSP189 plasmid.
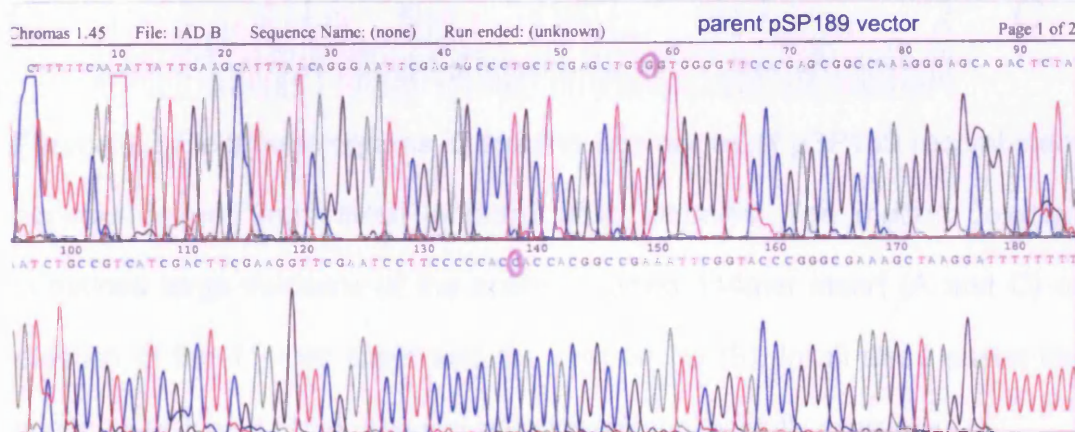
Figure 4.1 Electropherogram showing a sequence of Ad293 replicated pSP189 recombinant plasmid derived from a blue colony. The bases circled in red are unchanged from the wild type *supF* gene and show the plasmid does not contain the synthetic insert.

Mutational analysis of plasmid derived from three white colonies (Figure 4.2) produced after replication in Ad293 cells shows that these mutant sequences contained large deletions of bases between the XhoI and BamHI sites. In all three cases the original wild-type 114mer sequence had been successfully excised, suggesting that the mutations resulted from recombination of linear doubly-cut dephosphorylated vector contamination inside the human cells.
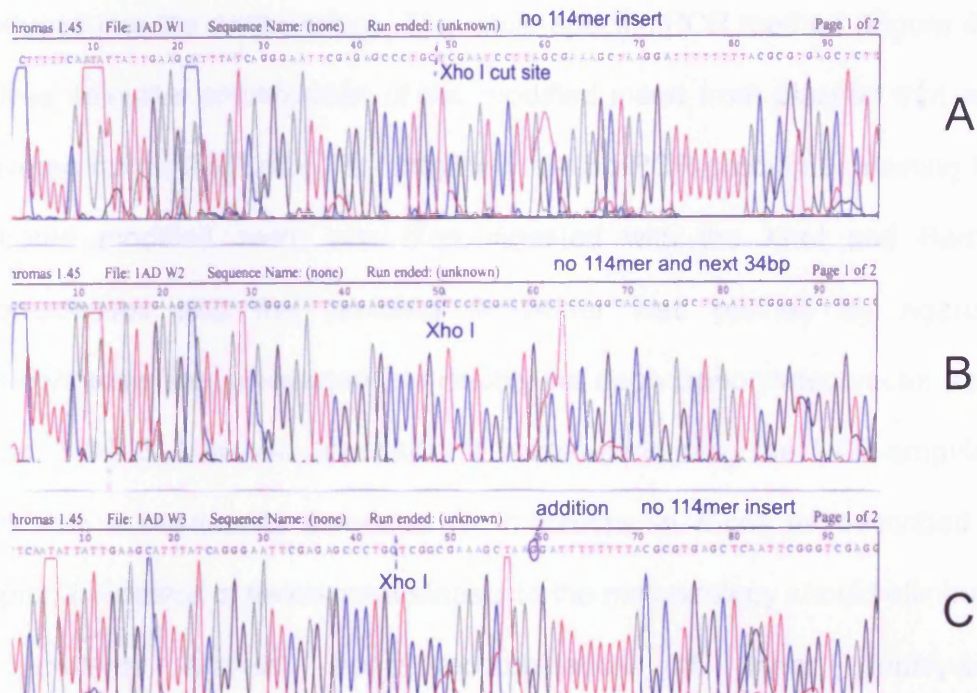
Figure 4.2 Electropherograms illustrating sequences of pSP189 recombinant plasmid derived from three different white colonies. The mutant colonies contained large deletions of the entire modified 114mer insert (A and C) or deletion of the 114mer insert and the next 34 bp (B). In all three cases the original wild-type 114mer sequence had been successfully excised.

4.3 PART II: METHODOLOGICAL IMPROVEMENTS

4.3.1 Experimental Design

Recombinant unadducted vector was constructed and transfected into Ad293 human kidney cells, then the replicated plasmid was recovered and digested with DpnI as described previously (4.2). At this stage, a PCR step was introduced into the methodology. The allele-specific PCR method (Figure 4.3) involves selective amplification of the modified insert from plasmid that was recovered from A293 cells. The amplified 134 bp PCR product containing the replicated modified insert was then digested with the XhoI and BamHI endonucleases and the resulting ds114mer was purified by agarose electrophoresis and re-inserted into doubly-cut dephosphorylated vector as in section 3.3.2.6. The new recombinant vector containing the PCR-amplified insert was subsequently screened for mutations in *E.coli* as described in Chapter 2. Addition of these extra steps into the methodology should eliminate any mutants derived from recombination of linear doubly-cut, dephosphorylated plasmid inside human cells, thereby reducing the high mutagenic background observed in previous experiments (see Sections 4.2.3.1-4.2.3.2).

XhoI   **g** in original pSP189   supF in green

5'- 1 gaattcgagagccctgctcgagctgtagtagggttcccgagcggccaaagggagcagact
        agccctgctcgagctgta
     Left Primer

   61 ctaaatctgccgtcatcgacttcgaaggttcgaatccttcccccaccacacggccgaaa

        BamH   **c** in original pSP189

  121 ttcggtacccggatccttagcgaaagctaagatttttttttacgcgtgagctcgactgact -3'
        atgggcctaggaatcgct
              Right Primer
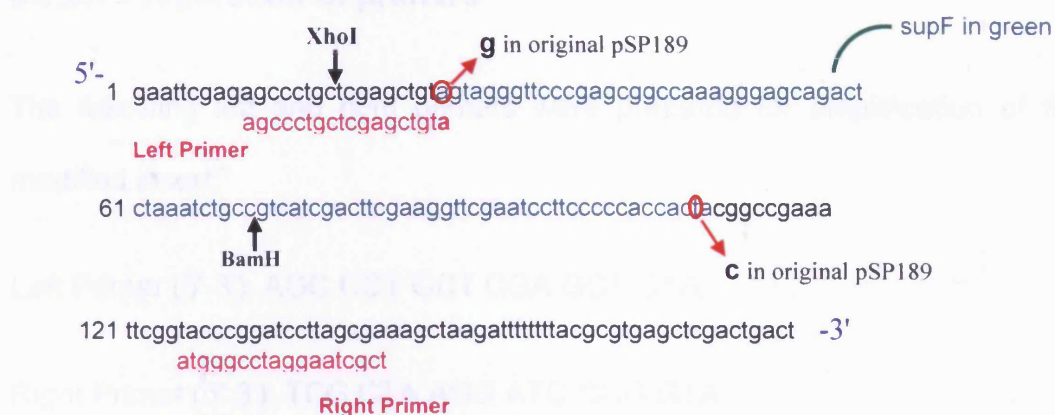
Figure 4.3 Allele-specific PCR amplification of the modified insert from unadducted plasmid replicated in Ad293 cells. Left and right primers flanking the modified insert are shown in pink and the *supF* gene sequence in green. The bases circled in red correspond to alterations in the wild-type *supF* sequence and demonstrate incorporation of the synthetic 114mer oligo. The original bases are guanine in the place of adenine and cytosine in the place of thymine.

4.3.2 Materials and methods

Materials

All chemicals were from Sigma-Aldrich, Poole, Dorset, UK unless otherwise stated.

Methods

All methods specific to part II of this Chapter are described below. For any other procedures refer to Chapters 2, 3 and 4 (part I).

4.3.2.1 Preparation of primers

The following left and right primers were prepared for amplification of the modified insert:

Left Primer (5'-3'): AGC CCT GCT CGA GCT GT**A**

Right Primer (5'-3'): TCG CTA AGG ATC CGG GTA

The oligonucleotides were prepared by Sigma-Proligo (Sigma-Proligo, the Woodlands, TX, USA) and supplied as HPLC-purified and desalted. The PCR primers were diluted to 50 pmol/µl with tissue culture grade water and stored at -20°C prior to use. The left primer contained a locked nucleic acid (LNA) at its 3' end (shown in red). LNA is a conformationally restricted nucleic acid analogue, in which the ribose ring is locked into a rigid C3'-*endo* (or Northern-type) conformation by a simple 2'-O, 4'-C methylene bridge. LNA may be used to enhance the performance of primers or probes used in allele-specific PCR by improving the stability and specificity of duplexes toward complementary DNA.

4.3.2.2 Polymerase chain reaction (PCR) amplification of unadducted plasmid replicated in Ad293 cells

Tissue culture grade water (78 µl) was mixed with 10x thermopol buffer (NEB) (10 µl), dNTPs (NEB) (4 µl), Left and Right primers (1 µl, 50 pmol each), *Taq* DNA polymerase (NEB) (5 µl, 1:10 diluted in water) and Ad293 recovered template plasmid DNA (1 µl, 1:5 diluted in water) in the order given. The PCR reaction (100 µl total volume) was performed in a 200 µl thin walled PCR tube

and subjected to the following program using a DNA engine DYAD Peltier thermal cycler (MJ Research, Inc., Waltham, MA, USA): 95°C for 5 min followed by 35 cycles of 95°C for 30 s, 59°C for 20 s, 72°C for 30 s, final extension at 72°C for 7 min, then cooled to 4°C (optimised method). Final concentrations in each tube were; 2.5 U *Taq* DNA polymerase, 0.4 μM Left and Right primers, 200 μM each of dATP, dCTP, dGTP and dTTP, 10 mM Tris-HCL (pH 8.3 @ 25°C), 50 mM KCl, 1.5 mM magnesium chloride. *Taq* DNA polymerase was removed after the reaction using Sureclean (Bioline, Bath, UK). Samples were then stored at -4°C until required.

4.3.2.3 Double digestion of the 134 bp PCR product with the BamHI and Xhol endonucleases

The purified 134 bp PCR product was digested with 200 Units of Xhol (NEB) in 25 μl NEBuffer 3 (10x) (NEB) with 2.5 μl BSA (NEB) and the reaction was made up to 250 μl using water. After incubation of the reaction at 37°C for 30 min, 200 units of BamHI (NEB) were added and the reaction was incubated at 37°C for 90 min.

4.3.2.4 Agarose purification of the digested PCR product

The reaction products were loaded on a 4% agarose gel and run for 6 h at 80V. The gel was photographed as outlined previously (Section 3.2.2.2). The bands corresponding to the 114 bp digested PCR product were visualised under low intensity UV light and excised from the gel using a sterile scalpel. The DNA was purified from the agarose gel using the QIAquick Gel Extraction Kit (Qiagen, West Sussex, UK).

4.3.2.5 Ligation of the 114 bp digested insert into the doubly-digested dephosphorylated vector

The 114 bp digested insert was ligated to doubly-cut dephosphorylated vector as in section 3.3.2.6. The recombinant vector containing the PCR-amplified insert was subsequently screened for mutations in *E.coli*, as described in Chapter 2.

4.3.3 Results

4.3.3.1 Direct transformation of electrocompetent *E.coli* with recombinant vector recovered from Ad293 cells

The experiment described in section 4.2.3.1 was repeated in order to generate fresh samples for the PCR part. Direct transformation of electrocompetent *E.coli* cells with recombinant plasmid recovered from human kidney cells resulted in about 86.4% white colonies (Table 4.3).

Experiment	Sample	Colonies counted			
		Total	Blue	White	% White
	1	68	10	58	85.3
	2	170	31	139	81.8
	3	458	85	373	81.4
	4	100	9	91	91.0
1	5	860	65	795	92.4
Average					86.4

Table 4.3 Number of mutant and wild type colonies resulting from the transformation of *E.coli* cells with recombinant vector recovered from Ad293 cells. Samples 1-5 were identical and represent separate vector to insert ligation reactions. For each sample a single transformation of *E.coli* was carried out. The average percentage of white colonies is highlighted in yellow.

4.3.3.2 PCR amplification of unadducted plasmid replicated in Ad293 cells

The 134 bp product of the PCR reaction performed with plasmid isolated from Ad293 cells was loaded on an analytical agarose gel (Figure 4.4) to confirm that the PCR amplification was successful. PCR amplification of the wild type vector was negligible in comparison (data not shown) confirming that the PCR method is specific for the synthetic insert.
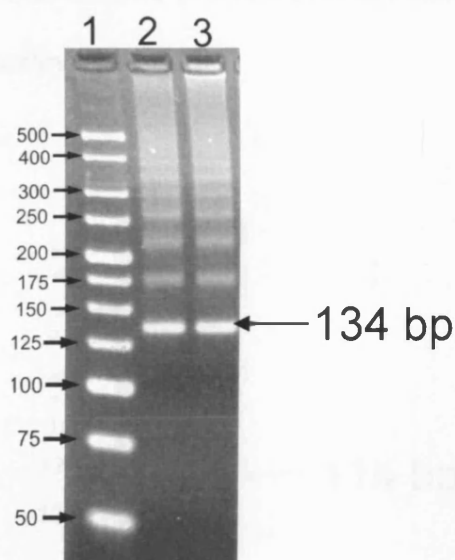
Figure 4.4 Allele-specific PCR amplification of the modified insert from plasmid replicated in Ad293 cells. Photograph of a 4% agarose gel showing the 134 bp PCR product. Lane 1 contains a 500bp ladder and lanes 2 and 3 contain the PCR amplified 134 bp product.

4.3.3.3 Double digestion of the 134 bp PCR product with the BamHI and XhoI endonucleases

The 134 bp PCR product was digested with BamHI and XhoI in a double digestion reaction, in order to produce a 114 bp insert with overhangs compatible to the doubly-cut dephosphorylated vector. The 114 bp product of the digestion reaction was purified on a 4% agarose gel (Figure 4.4) before ligation to the vector. Digestion was highly efficient as illustrated by a single band corresponding to the desired product and the absence of a visible band corresponding to the starting 134mer.
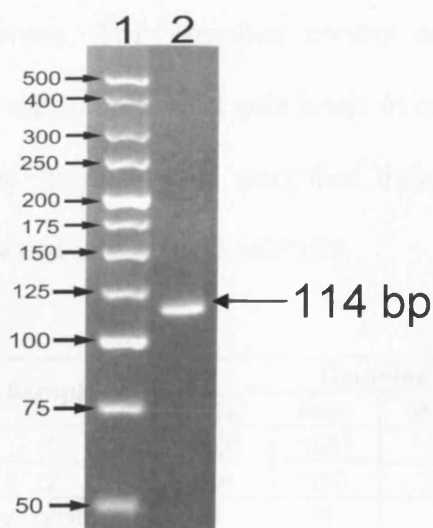
Figure 4.4 Digestion of the 134 bp PCR product with BamHI and XhoI to produce the 114 bp insert. Photograph of a 4% agarose gel showing the 114 bp digestion product. Lane 1 contains a 500bp ladder and lane 2 contains the 114 bp digested insert.

4.3.3.4 Transformation of electrocompetent *E.coli* with recombinant vector containing the PCR-amplified insert.

Transformation of electrocompetent *E.coli* cells with recombinant plasmid containing the PCR-amplified insert resulted in about 14.3% white colonies (Table 4.4). Although this was a significantly improved mutation background compared to directly transformed plasmid recovered from the Ad293 cells (86.4%) the number of white colonies was still too high for the assay to be used routinely in the investigation of DNA adduct mutagenesis. When the bacteria were transformed with recombinant doubly cut dephosphorylated vector ligated in the absence of the modified insert (negative control sample) this produced zero colonies. The negative control sample had not been previously replicated in Ad293 cells and was used in order to prove that the vector was not religating to itself and also that there was not significant contamination with the parent undigested pSP189.

Experiment	Sample	Colonies counted			
		Total	Blue	White	% White
	1	1208	1051	157	13
	2	616	520	96	15.6
1	-ve control	0	0	0	
Average					14.3

Table 4.4 Number of mutant and wild type colonies resulting from the transformation of electrocompetent *E.coli* cells with recombinant vector containing the PCR-amplified insert. Samples 1 and 2 were identical and represent separate *E.coli* transformations. The average percentage of white colonies from the two transformations is highlighted in yellow.

4.4 PART III: FINAL METHODS

4.4.1 Experimental Design

Experimental results so far suggested that the PCR method significantly reduced the high mutational background obtained after direct transformation of E.coli with recombinant plasmid replicated in Ad293 human cells from 86.4 to 14.3%. However, the mutation frequency obtained with this new PCR method was still too high. Subsequent experiments were focused on reducing this background further using 130mer PAGE purified constructs (as in Section 3.4) and also by replacing Taq polymerase with a high-fidelity polymerase (see discussion for further analysis). The assay was validated using the same types of constructs as in the E.coli experiments (Chapter 3, Part III), although construct 130-A was not included in the analysis because of methodological constrains.

4.4.2 Materials and methods

Materials

All chemicals were from Sigma-Aldrich, Poole, Dorset, UK unless otherwise stated.

Methods

All methods specific to Part III of this Chapter are described below. For any other procedures refer to Chapters 2, 3 and 4 (Parts I and II).

4.4.2.1 Transfection of Ad293 cells with recombinant plasmid samples

Ad293 cells were transfected with samples containing the same 130mer inserts described in Section 3.4.2.4 and also with a control sample ligated in the absence of insert (-ve control). However, construct 130-A was excluded from analysis due to insufficient quantities available. Additionally, constructs 130-H, 130-I, 130-J and 130-K were purified by 4% agarose electrophoresis instead of 20% PAGE because of time constraints.

4.4.2.2 Polymerase chain reaction (PCR) amplification of unadducted plasmid replicated in Ad293 cells

Tissue culture grade water (36.5 µl) was mixed with 5x Phusion HF Buffer (NEB) (10 µl), dNTPs (Bioline) (1 µl), Left and Right primers (1 µl), Phusion DNA Polymerase (NEB) (0.5 µl) and Ad293 recovered template plasmid DNA (1 µl) in the order given. The PCR reaction (50 µl) was conducted into a 200 µl thin walled PCR tube and subjected to the following PCR program using a DNA engine DYAD Peltier thermal cycler (MJ Research, Inc., Waltham, MA, USA): 98°C for 1 min followed by 40 cycles of 98°C for 10 s, 60°C for 20 s, 72°C for 5 s, final extension at 72°C for 5 min, then cooled to 4°C (optimised method). Final concentrations in each tube were; 1 U Phusion DNA polymerase, 0.5 µM Left and Right primers, 200 µM each of dATP, dCTP, dGTP and dTTP, 10 mM Tris-HCL (pH 8.3 @ 25°C), 50 mM KCl, 1.5 mM magnesium chloride. Phusion DNA polymerase was removed after the

reaction using Sureclean according to the manufacturers protocol (Bioline, Bath, UK). Samples were then stored at 4°C until required.

4.4.3 Results

4.4.3.1 Direct transformation of electrocompetent *E.coli* with recombinant vector recovered from Ad293 cells

The results of direct *E.coli* transformation with plasmid samples recovered from Ad293 cells are summarized in Table 4.5. Not all transfected samples were used to directly transform *E.coli*. Samples containing inserts 130-H, 130-I and 130-J were excluded due to the limited quantities recovered from the human cells.

Transformation of *E.coli* with uracil-containing constructs (samples Ad2, Ad3 and Ad4) resulted in similar mutation frequencies (46.1, 47.1 and 50.6% respectively). Plasmids containing single O^6-MeG adducts in two different DNA sequence contexts, in the absence of deoxyuridines in the complementary strand, (samples Ad5 and Ad6) produced mutation frequencies of 40.1 and 46.6% respectively. However, the presence of two O^6-MeG adducts on the same *supF* gene in ds plasmids increased the mutant fraction to 62.3%. Sample Ad12 was prepared by mixing constructs 130-B and 130-K in different proportions (75:25) in the ligation reaction. When this sample was recovered from the human cells and used to directly transform *E.coli*, it produced a mutation frequency of 50.7%. Finally, doubly cut dephosphorylated vector ligated in the absence of the modified insert (sample 14) was also replicated in Ad293 cells and direct transformation of *E.coli* with

the recovered plasmid resulted in a high mutation frequency (97.9%). Although the mutagenic background observed in unadducted recombinant plasmids prepared using PAGE-purified 130mer constructs were significantly lower than those obtained with agarose-purified 114mers (Parts I and II) in general, this background still interferes with the quantitative analysis of adducted plasmids when the PCR step is not included.

Sample	Type of 130mer insert	Colonies		
		Total	White	% White
Ad2	130-B	4082	1882	46.1
Ad3	130-C	922	434	47.1
Ad4	130-D	3176	1607	50.6
Ad5	130-E	5896	2366	40.1
Ad6	130-F	1535	716	46.6
Ad7	130-G	3417	2128	62.3
Ad12	75% 130-B + 25% 130-K	5236	2657	50.7
Ad14	No insert	12403	12147	97.9

Table 4.5 Number of mutant and total colonies resulting from direct transformation of electrocompetent *E.coli* cells with the recombinant constructs recovered from Ad293 cells. Mutation frequencies are highlighted in yellow. Samples are colour-coded according to the type of modified construct they contain; red = unadducted, lavender = unadducted (+) uracil, pale blue = adducted (–) uracil, yellow = adducted (+) uracil, light green = validation oligos, grey = (–ve) control.

4.4.3.2 DNA sequencing of mutants resulting from direct transformation of *E.coli* with plasmid samples recovered from Ad293 cells.

Blue colonies produced from samples Ad2 and Ad7 were picked randomly and sequenced. All DNA sequences derived from these blue colonies confirmed the presence of the modified 130mer insert, proving that it was successfully ligated into the vector and that there were no inactivating mutations in it. A selection of white colonies produced from samples Ad2 and Ad7 were also sequenced. The types of mutations seen in sequences of plasmid extracted from white colonies of these samples are summarized in Table 4.7

Sample	No of blue colonies sequenced	Modified insert (no mutations)	Original pSP189 sequence
Ad2	10	10	0
Ad7	10	10	0

Table 4.6 DNA sequences derived from blue colonies of samples Ad2 and Ad7.

The mutation spectrum observed in these two samples is comprised mainly of large deletions between the EcoRI and BamHI restriction sites, suggesting that these deletion mutations resulted due to recombination of linear doubly-cut dephosphorylated vector once in the Ad293 cells.

Sample	No of white colonies sequenced	Mutation type
Ad2	9	78→203 (x3)
		71→211
		78→207
		79→207
		78→211
		157→203
		CGAA insertion @155
Ad7	7	77→204
		77→203
		78→203
		78→205
		78→207
		156→203
		G-A @100 AND G-A@129 (both targeted)

Table 4.7 Mutations observed in DNA sequences derived from white colonies of samples Ad2 and Ad7.

4.4.3.3 PCR amplification of recombinant plasmid samples replicated in Ad293 cells

The plasmid samples isolated from Ad293 cells were used as templates for PCR to selectively amplify insert containing plasmid DNA. The 134 bp products (20 µl of each PCR reaction) were then loaded on an analytical agarose gel (Figure 4.5) to assess the efficiency of the PCR amplification. In order to confirm that all 134 bp PCR products resulted from specific amplification of the modified insert and not from amplification of the original pSP189 sequence, sample Ad14 (no insert control) was amplified in parallel to the recombinant samples. As the gel photo suggests, PCR amplification of this sample resulted in a detectable 134 bp band, however, the intensity of this band was much lower than that of bands resulting from recombinant samples indicating the method has specificity towards the synthetic insert.

Furthermore, samples Ad8, Ad9, Ad10 and Ad11 were not amplified as efficiently as the rest of the samples, as is evident from the low intensity of the respective 134 bp bands probably because they were constructed using agarose-purified 130mer constructs (instead of PAGE) that were less efficiently ligated into the vector.
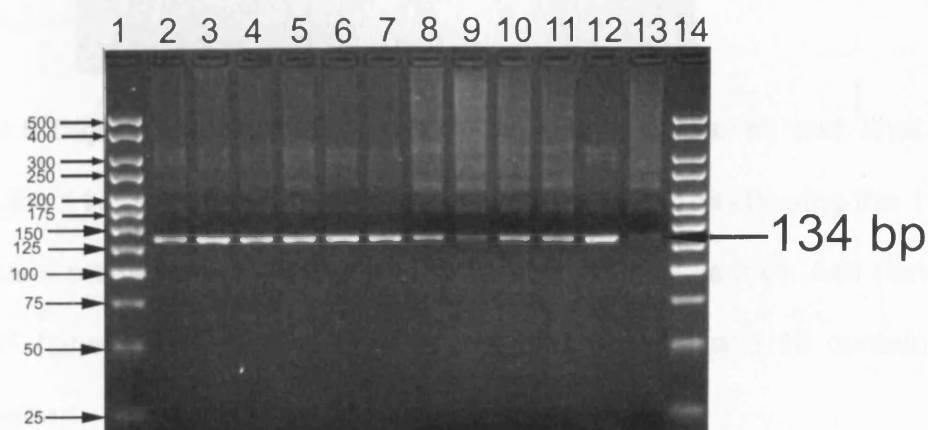
Figure 4.5 Allele-specific PCR amplification of the modified inserts from plasmid samples replicated in Ad293 cells. Photograph of a 4% agarose gel showing the 134 bp products of PCR reactions. Lanes 1 and 14 contain a 500bp ladder and lanes 2-13 contain the PCR 134 bp products amplified from samples Ad2 (lane 2), Ad3 (lane 3), Ad4 (lane 4), Ad5 (lane 5), Ad6 (lane 6), Ad7 (lane 7), Ad8 (lane 8), Ad9 (lane 9), Ad10 (lane 10), Ad11 (lane 11), Ad12 (lane 12) and Ad14 (lane 13).

4.4.3.4 Double digestion of the 134 bp PCR products with the BamHI and XhoI endonucleases

The 134 bp PCR products were digested with BamHI and XhoI in a double digestion reaction, in order to produce 114 bp inserts with overhangs compatible to the doubly-cut dephosphorylated vector. The 114 bp products of the digestion reactions were loaded on two 4% agarose gels (Figures 4.6 and 4.7).
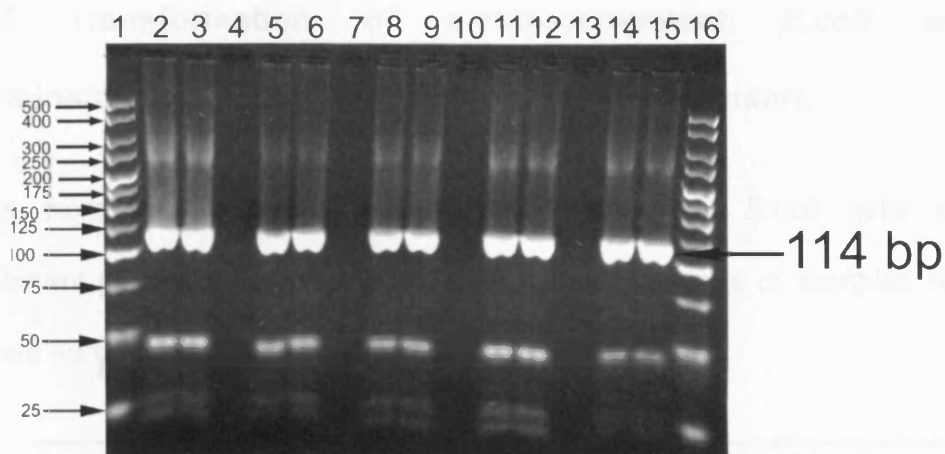
Figure 4.6 Digestion of the 134 bp PCR products with BamHI and XhoI to produce the 114 bp inserts. Photograph of a 4% agarose gel showing the 114 bp digestion products of sample Ad2 (lanes 2-3), Ad3 (lanes 5-6), Ad4 (lanes 8-9), Ad5 (lanes 11-12) and Ad6 (lanes 14-15). Lanes 1 and 16 contain a 500bp ladder.
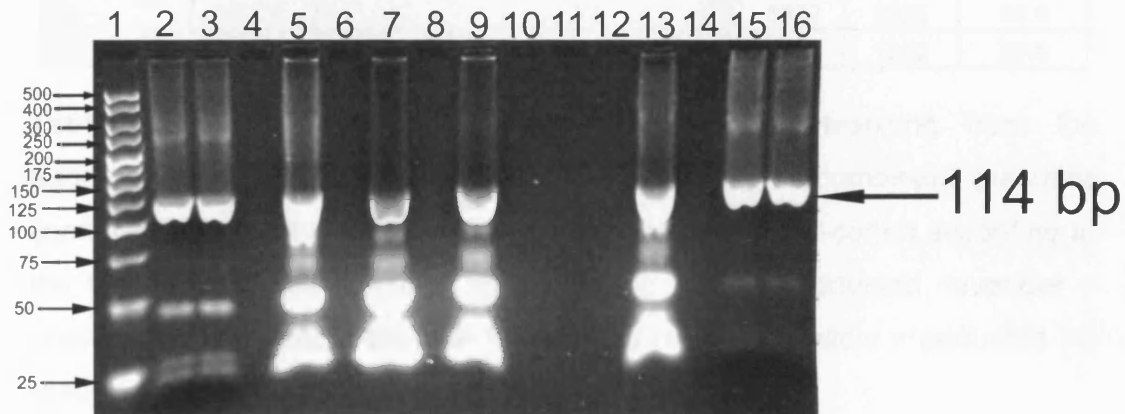
Figure 4.7 Digestion of the 134 bp PCR products with BamHI and XhoI to produce the 114 bp inserts. Photograph of a 4% agarose gel showing the 114 bp digestion products of sample Ad7 (lanes 2-3), Ad8 (lane 5), Ad9 (lane 7), Ad10 (lane 9), Ad11 (lane 13) and Ad12 (lanes 15-16). Lane 1 contains a 500bp ladder.

4.4.3.5 Transformation of electrocompetent *E.coli* with recombinant vector containing the PCR-amplified insert.

Results from the transformation of electrocompetent *E.coli* cells with recombinant plasmids containing the PCR-amplified inserts of samples Ad2-Ad12 are summarized in Table 4.8 and Figure 4.8.

Sample	Type of 130mer insert	Colonies		
		Total	White	% White
Ad2-PCR	130-B	2591	122	4.7
Ad3-PCR	130-C	6573	183	2.8
Ad4-PCR	130-D	4477	247	5.5
Ad5-PCR	130-E	1127	174	15.4
Ad6-PCR	130-F	3293	678	20.6
Ad7-PCR	130-G	3847	835	21.7
Ad8-PCR	130-H	3766	2576	68.4
Ad9-PCR	130-I	2606	412	15.8
Ad10-PCR	130-J	7846	7117	90.7
Ad11-PCR	130-K	5509	5286	96.0
Ad12-PCR	75% 130-B + 25% 130-K	3194	1232	38.6

Table 4.8 Number of mutant and total colonies resulting from the transformation of electrocompetent *E.coli* cells with recombinant plasmids containing the PCR-amplified inserts. Samples are colour-coded according to the type of modified construct they contain; red = unadducted, lavender = unadducted (+) uracil, pale blue = adducted (–) uracil, yellow = adducted (+) uracil, light green = validation oligos, grey = (–ve) control.

Transformation of *E.coli* with recombinant plasmids containing the PCR-amplified inserts resulted in significant reductions of the background mutation frequency in general, compared to directly transformed plasmids recovered from the Ad293 cells (Table 4.5). Transformation with control samples that contained deoxyuridines in the complementary strand prior to Ad293

transfection (Ad2-PCR, Ad3-PCR and Ad4-PCR) resulted in similar mutant

fractions (4.7, 2.8 and 5.5% respectively). Samples that contained single O^6-

MeG adducts in the absence of deoxyuridines in the complementary strand

prior to Ad293 transfection (samples Ad5-PCR, Ad6-PCR and Ad7-PCR)

produced mutation frequencies of 15.4, 20.6 and 21.7% respectively.

However, the placement of deoxyuridines on the complementary strand at 5'-

and 3'-positions flanking the O^6-MeG adducts (samples Ad8-PCR, Ad9-PCR

and Ad10-PCR) resulted in considerable enhancements of the mutation

frequency over the deoxyuridine-alone control samples (Ad2-PCR, Ad3-PCR

and Ad4-PCR respectively). In particular, O^6-MeG enhanced the mutant

fraction of constructs containing deoxyuridines in the complementary strand

14.5-fold in sequence context 1 (sample Ad8-PCR) and 5.6-fold in sequence

context 2 (sample Ad9-PCR). When O^6-MeG adducts were present in both

sequence contexts in the presence of uracil bases in the complementary

strand (sample Ad10-PCR) there was a 16.4-fold increase in the mutation

frequency over the uracil-alone control sample (Ad4-PCR). Positioning of the

adducts in gapped plasmids resulted in increased mutant fractions compared

to positioning of the same adducts in ds plasmids. Placement of a single O^6-

MeG adduct in sequence context 1 in a gapped plasmid (sample Ad8-PCR)

gave a 4.4-fold increase in the mutation frequency compared to placement of

the adduct at the same sequence context but in a ds plasmid (sample Ad5-

PCR). Similarly, when two O^6-MeG adducts were present in both sequence

contexts in a gapped vector (sample Ad10-PCR) they induced a mutation

frequency 4.2-fold higher than when present in a ds vector (sample Ad7-

PCR). In contrast, situating the adduct in sequence context 2 induced a 1.3-

fold higher mutation frequency in the ds plasmid (sample Ad6-PCR) compared to the gapped one (sample Ad9-PCR).
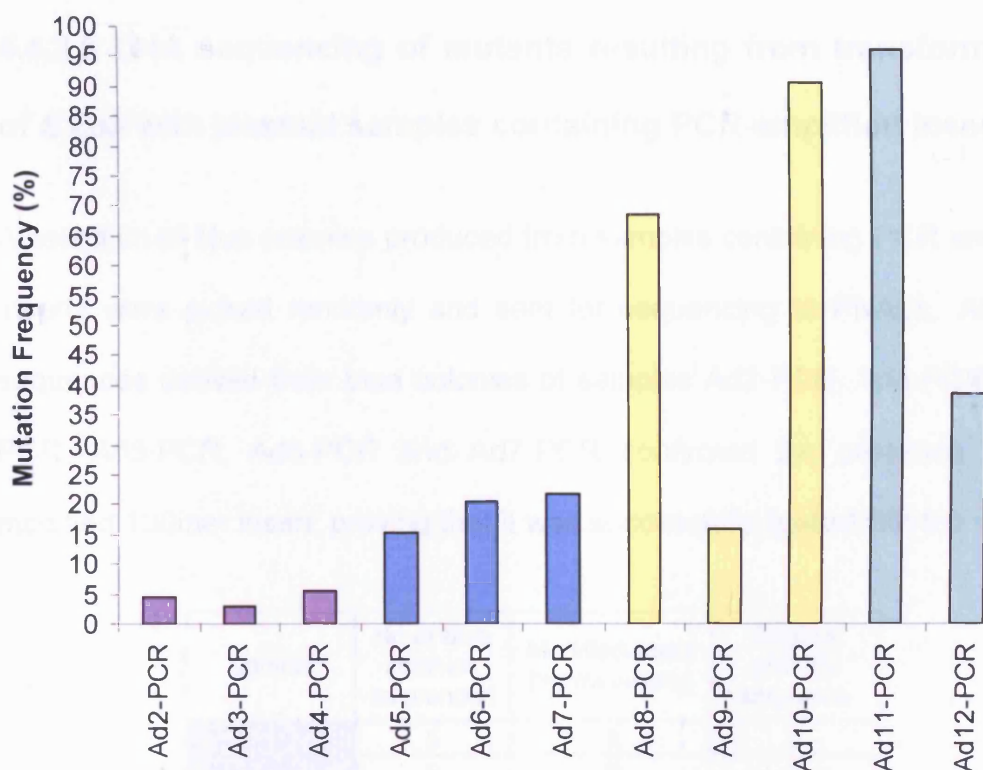
Figure 4.8 Graphical representation of *E.coli* transformation results using plasmid samples containing PCR-amplified inserts. Samples are colour-coded according to the type of modified construct they contain, as in Table 4.8.

Transformation of *E.coli* with a plasmid sample that contained a *supF* inactivating GC→CG basepair change at position 150 (sample Ad11-PCR) inactivated *supF* function almost completely as is evident from the high mutation frequency observed (96%). Sample Ad12-PCR was prepared by mixing by mixing the 130-B and 130-K constructs in different proportions (75:25) in the ligation reaction prior to transfection of Ad293 cells. Following PCR amplification, this sample produced a mutant fraction of 38.6%, which is

slightly higher than the expected 25%, possibly due to added background frequency.

4.4.3.6 DNA sequencing of mutants resulting from transformation of *E.coli* with plasmid samples containing PCR-amplified inserts.

A selection of blue colonies produced from samples containing PCR amplified inserts were picked randomly and sent for sequencing to PNACL. All DNA sequences derived from blue colonies of samples Ad2-PCR, Ad3-PCR, Ad4-PCR, Ad5-PCR, Ad6-PCR and Ad7-PCR confirmed the presence of the modified 130mer insert, proving that it was successfully ligated into the vector.

Sample	No of blue colonies sequenced	Modified insert (no mutations)	Original pSP189 sequence
Ad2-PCR	5	5	0
Ad3-PCR	5	5	0
Ad4-PCR	5	5	0
Ad5-PCR	5	5	0
Ad6-PCR	5	5	0
Ad7-PCR	5	5	0
Ad8-PCR	10	7	3
Ad9-PCR	10	5	5
Ad10-PCR	7	3	4
Ad11-PCR	10	0	10

Table 4.9 DNA sequences derived from blue colonies of plasmids containing PCR-amplified inserts.

In contrast, a certain proportion of DNA sequences derived from blue colonies of samples Ad8-PCR, Ad9-PCR, Ad10-PCR and Ad11-PCR corresponded to the original pSP189 plasmid sequence, as the base alterations at positions 99 and 179 were absent. These blue colonies clearly resulted from non-specific

PCR amplification of parent undigested or singly-cut linear pSP189 contaminants.

White colonies produced from samples containing PCR amplified inserts were also sequenced and the types of mutations detected in plasmid extracted from white colonies of these samples are summarized in Tables 4.10, 4.11 and 4.12.

DNA sequencing results from control unadducted samples containing deoxyuridines on the complementary strand (samples Ad2-PCR, Ad3-PCR and Ad4-PCR) are summarized in Table 4.10. Inclusion of deoxyuridines at sequence context 1 in the absence of adduct (sample Ad2-PCR) induced predominantly random substitutions (5/9) along with point deletions at positions 153 (3/9) and 152 (1/9).

The presence of deoxyuridines at sequence context 2 (sample Ad3-PCR) induced mainly random point deletions (4/8), substitutions (2/8) and large deletions (2/8). Likewise, when uracils were present at both sequence contexts (sample Ad4-PCR) the prevailing mutation type was again random point deletions (5/10), substitutions (4/10) and a large deletion (1/10). When the mutation spectra from the uracil-containing samples (samples 2, 3 and 4) are combined, the main mutation type observed are random point deletions (13/27 sequences), followed by random substitutions (11/27) and large deletions (3/27).

Sample	No of white colonies sequenced	Mutation type
Ad2-PCR	9	T DEL @153 (x3)
		C DEL @152
		G-A @156
		A-T @136
		C-T @108
		GG-AA @159
		T-A @170
Ad3-PCR	8	DEL 89→208
		DEL 89→207
		A DEL @158
		C DEL @110
		G DEL @124
		A DEL @121
		CC-TT @174
		G-C @105
Ad4-PCR	10	DEL 89→205
		A DEL @166
		C DEL @155
		C DEL @118
		C DEL @110
		C DEL @127
		C-T @155
		G-C @213
		T-A @138
		A-T @121

Table 4.10 Mutations observed in DNA sequences derived from white colonies of samples Ad2-PCR, Ad3-PCR and Ad4-PCR.

DNA sequences of plasmid extracted from white colonies of samples Ad5-PCR, Ad6-PCr and Ad7-PCR were analyzed for mutations and the results are summarized in Table 4.11. Placement of a single O^6-MeG adduct at sequence context 1 and subsequent PCR amplification of the replicated insert (sample Ad5-PCR) resulted in a targeted G→A transitions at position 100 in the majority of the mutants sequenced (10/14). The remaining mutations observed in this sample were point substitutions (3/14) and a point deletion (1/14). Positioning of the O^6-MeG adduct at sequence context 2 (Ad6-PCR)

resulted in targeted G→A transitions at position 129 in the majority of the mutants sequenced (9/11). A point deletion and a point substitution comprised the rest of the mutation spectrum of this sample.

Incorporation of two O^6-MeG adducts on a single construct (130-G-sample Ad7-PCR) induced primarily targeted G→A transitions in both sequence contexts at the same time (5/14 multiple mutations), followed by single targeted G→A transitions in either sequence context 1 (3/14) or context 2 (2/14), random substitutions (3/14) and a point deletion (1/14).

Sample	No of white colonies sequenced	Mutation type
Ad5-PCR	**14**	G-A @100 (T1-06-MeG Targeted) (x10)
		G-A @103
		A DEL @166
		CC-TT @174
		A -G @125
Ad6-PCR	**11**	G-A @129 (T3-06-MeG Targeted) (x9)
		C-A @174
		T DEL @153
Ad7-PCR	**14**	G-A @100 and G-A @129 (Both Targeted) (x5)
		G-A @129 (T3-06-MeG Targeted) (x3)
		G-A @100 (T1-06-MeG Targeted) (x2)
		T DEL @153
		C-A @173
		C-A @146
		G-A @156

Table 4.11 Mutations observed in DNA sequences derived from white colonies of samples Ad5-PCR, Ad6-PCR and Ad7-PCR.

Replication of O^6-MeG adducts in gapped plasmids, followed by PCR amplification of the replicated inserts and screening in *E.coli* (samples Ad8-PCR, Ad9-PCR and Ad10-PCR) resulted predominantly in targeted G→A

transitions. The mutation types observed in the DNA sequences from these samples are summarized in Table 4.12. When a single O^6-MeG adduct was placed in sequence context 1 (sample Ad8-PCR) it generated exclusively G→A targeted transitions at position 100. However, when the same adduct was positioned in sequence context 2 (Ad9-PCR) it resulted in both G→A targeted transitions at position 129 (16/24) and other mutation types, namely large deletions (3/24), random substitutions (3/24) and point deletions (2/24). Positioning of two O^6-MeG adducts in both sequence contexts in a gapped plasmid resulted in multiple targeted G→A transitions in the majority of mutants sequenced (31/39), while single targeted transitions were 3 times more frequent in sequence context 1 (6/39) than in sequence context 2 (2/39).

Mutation analysis of DNA sequences from plasmids containing a GC→CG inactivating basepair alteration at position 150 (sample Ad11-PCR) confirmed the presence of the GC→CG inactivating mutation in 100% of the mutants sequenced.

Sample	No of white colonies sequenced	Mutation type
Ad8-PCR	20	G-A @100 (T1-06-MeG Targeted) (x20)
		G-A @129 (T3-06-MeG Targeted) (x16)
		DEL 88→207
		DEL 89→207
		DEL 89→154
Ad9-PCR	24	T DEL @171
		C DEL @110
		T-A @145
		G-T @126
		T-A @167
		G-A @100 and G-A @129 (Both Targeted) (x31)
Ad10-PCR	39	G-A @100 (T1-06-MeG Targeted) (x5)
		G-A @129 (T3-06-MeG Targeted) (x2)
		G-T @100 (T1-06-MeG Targeted) (x1)
Ad11-PCR	9	C-G @150 (TARGETED) (x9)

Table 4.12 Mutations observed in DNA sequences derived from white colonies of samples Ad8-PCR, Ad9-PCR, Ad10-PCR and Ad11-PCR.

4.5 DISCUSSION

4.5.1 Discussion of Part I

The aim of work conducted during the early stage of development was to assess the mutagenicity of the unadducted recombinant plasmid in human cells. Replication of the recombinant vector in either Ad293 kidney or GM00637 fibroblast cells and subsequent screening in *E.coli* resulted in similarly high background mutation frequencies (89.0 and 77.6% respectively) which was unexpected. DNA sequencing of plasmid extracted from mutant colonies suggested that large deletions of bases between the XhoI and BamHI restriction sites were responsible for the high mutation frequencies

243

observed. These large deletions were presumably due to rejoining of the ends of linear doubly-cut dephosphorylated plasmid by non-homologous end-joining (NHEJ) in human cells. The ligation reaction of the modified insert to the doubly-cut dephosphorylated plasmid is never 100% efficient, in fact it is a low efficiency event (estimated to be less than 10% by agarose electrophoresis), a certain proportion of linear plasmid molecules will always be transfected along with the recombinant covalently closed relaxed plasmid. The ends of these linear plasmid molecules would be treated as DSBs inside the human cells and would be subject to DNA recombination repair. Assays involving the transfection of restriction enzyme-generated linear plasmids into host cells have been used to study joining of free DNA ends in intact cells [479-481]. Linear double-stranded deoxyribonucleotides, including transfected plasmids with DSBs are known inducers of DNA-PK activity [482]. DNA-PK is known to play an important role in NHEJ repair [482] and is a DNA end-binding protein, composed of a catalytic subunit (DNA-PKcs) and the Ku protein, a heterodimer of polypeptides of 70 and 80 kDa (Ku70 and Ku80 respectively). Ku is the DNA-targeting subunit of DNA-PK. In the current assay transfection of non-ligated recombinant plasmid may induce NHEJ repair, resulting in the large deletions observed.

4.5.2 Discussion of Part II

In order to reduce the high mutagenic background obtained with direct *E.coli* transformation of plasmid recovered from the human cells, an allele-specific PCR method was developed. The method takes advantage of the base changes introduced into the synthetic oligonucleotide and involves selective

amplification of the modified insert from plasmid that was recovered from the human cells, thereby eliminating the high mutagenic background associated with rejoined linear plasmids that do not contain the synthetic insert and improving the overall sensitivity of the assay. The amplified 134mer PCR product is then digested with BamHI and XhoI to produce a shorter 114mer insert with suitable overhangs for re-insertion into the vector and screening in *E.coli*. Therefore, any substitutions or small frameshift mutations fixed in the *supF* gene in response to replication of a DNA adduct would be detected by selective amplification of modified inserts. A potential disadvantage of this approach, however, is that any large deletions induced in response to DNA adducts, such as those observed in Chapter 3 (Part III) that encompass the primer binding sites, would escape detection, as these deletions would result in PCR products of shorter length or no PCR products at all.

At this stage of development only unadducted plasmid was used. Screening of vector containing the PCR-amplified insert resulted in a significant reduction of the mutagenic background (14.3%) compared to direct transformation of the recovered plasmid (89.0%). However, the mutagenic background was still higher than the background obtained from direct transformation of *E.coli* cells with the same type of recombinant vector (9.3%). A potential source of mutations could be errors introduced during PCR amplification of the modified insert by *Taq* Polymerase, and this was investigated subsequently.

4.5.3 Discussion of Part III

In order to further reduce the mutational background Phusion High Fidelity DNA polymerase was used instead of *Taq* DNA polymerase, since Phusion

reportedly has a 50-fold lower error rate than that of *Taq*, as determined using a modified *lacI*-based method [483]. In addition, the plasmid constructs were prepared using PAGE rather than agarose-purified 130mer inserts prior to transfection into Ad293 cells, as in Chapter 3 (part III).

Direct transformation of *E.coli* with gapped unadducted control plasmid samples recovered from Ad293 cells (samples Ad2, Ad3 and Ad4) resulted in a high mutagenic background (46-50%), although much lower compared to that observed in part II using unadducted ds control plasmid (89%). This finding suggests that construction of unadducted control recombinant plasmids using PAGE-purified 130mer inserts played a significant role in reducing the background mutation frequency. The mutagenic background arising from direct transformation of *E.coli* with O^6-MeG adduct-containing ds plasmid samples replicated in Ad293 (samples Ad5, Ad6 and Ad7) was similar to that obtained with the uracil controls, although there was a small increase in the sample containing both adducts (sample Ad7). DNA sequencing of mutant colonies from samples Ad2 and Ad7 revealed that the inactivating mutation type responsible for the high mutant fractions observed in both samples was large deletions between the EcoRI and BamHI restriction sites, similar to those seen in part II using unadducted control templates. In addition transfection of Ad293 cells with linear doubly-digested, dephosphorylated plasmid (sample Ad14, -ve control) and subsequent direct screening in *E.coli* resulted in the production of thousands of colonies the vast majority of which were white (97.9%). Since plasmid is required to be in a circular form in order to be replicated, it can be inferred that the linear vector was recircularized inside the human cells.

These findings reinforce the conclusion that these particular deletions resulted from rejoining of linear doubly-cut plasmid by NHEJ. Interestingly, this type of mutation was not observed in any of the samples from the experiments in *E.coli* (Chapter 3), except for sample 10 which contained two O^6-MeG adducts in a gapped plasmid, indicating that these large deletions occur only in response to heavy DNA damage in *E.coli*, whereas they are induced more readily in undamaged plasmid DNA replicated in Ad293 cells. There are several feasible non-exclusive reasons why this might happen. One possibility is that the longer replication time in human cells (48 h) as opposed to *E.coli* (45 min) might allow DNA recombination repair systems to join the ends of linear digested plasmids more efficiently. Another likely cause could be the low transformation efficiency of linear DNA molecules in *E.coli* compared to the transfection efficiency in human cells [484, 485], which would mean greater numbers of linear molecules are available for processing by the human DSB repair systems. A third reason could be the differences in the DNA recombination repair systems between *E.coli* and human cells, as it is known that *E.coli* are incapable of performing NHEJ .

Incorporation of the PCR step eliminated these non-specific large deletions and restored the mutagenic background to the levels seen in *E.coli* (see Chapter 3, part III) using gapped undamaged control plasmids (Ad2-PCR, Ad3-PCR and Ad4-PCR). DNA sequencing of mutants produced by these samples confirmed that the background mutation spectrum consisted of the same mutation types, such as point deletions (e.g. the T deletion at position 153) and substitutions, that were observed in spectra obtained from replication of the same constructs in *E.coli* (samples 2, 3 and 4 in Chapter 3),

while large deletions were rare. Additionally, the background mutation frequency obtained in part III experiments (Chapter 4) using unadducted control plasmids was ~3-fold lower than that of part II (Chapter 4). This reduction in mutagenic background could be due to both PAGE purification of the inserts and increased fidelity of the PCR reaction, although it is impossible to estimate the relative contribution of each factor.

The mutation frequency was increased moderately (3-4-fold) in samples that contained a single or two O^6-MeG adducts in ds context (Ad5-PCR, Ad6-PCR and Ad7-PCR) compared to the gapped unadducted controls when these samples were processed using the PCR approach. This increase in mutation frequency was accompanied by an increase in the proportion of O^6-MeG-targeted G→A transitions observed in DNA sequences generated from mutant colonies of these samples. The mutation fraction was enhanced dramatically in gapped plasmids when a single O^6-MeG adduct was positioned in sequence context 1 (sample Ad8-PCR) or when two O^6-MeG adducts were present in both contexts (sample Ad10-PCR), but not when a single adduct was placed at sequence context 2 (sample Ad9-PCR). This difference might be explained by the poor PCR amplification observed in this sample (Ad9-PCR) that could potentially cause an increase in the proportion of blue colonies by non-specific amplification of the wild type $supF$ gene. This hypothesis is further supported by the fact that half of the blue colonies sequenced from this sample contained the wild-type $supF$ sequence and not the modified insert. The wild-type $supF$ sequence was also observed in a certain proportion of blue colonies derived from plasmids Ad8-PCR, Ad10-PCR and Ad11-PCR indicating a low level of contamination with non-

specifically PCR-amplified products. This was probably due to the fact that PCR amplification of these samples was less efficient, resulting thus in a non-specific amplification of the wild type supF gene sequence.

Sequence analysis of the mutant white colonies generated by samples Ad8-PCR and Ad10-PCR proved that O^6-MeG targeted transitions were present in 100% of plasmids sequenced and that they were solely responsible for the enhancement of the mutant fraction over their respective unadducted gapped control samples (Ad2-PCR, Ad4-PCR) or ds adducted samples (Ad5 and Ad7 respectively). The increase in the mutation frequency observed with these samples is consistent with previous reports comparing the relative mutagenicity of adducts in ds intact vs gapped plasmids [486, 487] and is thought to reflect induction of TLS in gapped plasmids rather than damage avoidance. Interestingly, the majority of the mutants in sample Ad10-PCR had O^6-MeG-targeted transitions present in both sequence contexts, suggesting that the human polymerase(s) involved preferentially introduces a T opposite multiple O^6-MeG adducts, irrespective of sequence context, or number of adducts. O^6-MeG-targeted transitions were also prevalent in sample Ad9-PCR, although in a lower proportion (66.6%), with random point mutations and large deletions comprising the rest of the spectrum. It is unclear why this might have occurred, although it implies that the mutagenicity of the O^6-MeG adduct in sequence context 2 was not as high as in context 1. Alternatively, the lower occurrence of O^6-MeG-targeted transitions with this sample might have arisen due to an experimental error that involved the accidental incorporation of the thymine-containing oligo (B2) in the complementary strand in place of the uracil-containing one (B2U), essentially resulting in

duplication of sample Ad6-PCR. This could easily be investigated by simply repeating the experiment with newly prepared constructs. Finally, analysis of the DNA sequences derived from mutant colonies of the validation sample Ad11-PCR confirmed the presence of the inactivating insert in 100% of cases, providing further confidence in the sensitivity of the PCR method for detecting point mutations.

4.5.4 Summary

The overall aim of the work described in this Chapter was to develop and validate a novel adduct site-specific assay in Ad293 human kidney cells. The assay was successfully validated using an inactivating 130mer insert and was subsequently applied to the investigation of the mutagenic potential of a single or two O^6-MeG adducts at two different DNA sequence contexts in ds or gapped plasmids. Transfection of recombinant constructs and subsequent replication in Ad293 produced a high mutagenic background when vector recovered from the human cells was screened for mutations in E.coli. DNA sequencing of plasmid extracted from mutant colonies suggested that large deletions of bases between the XhoI and BamHI or EcoRI and BamHI restriction sites were responsible for the high mutation frequencies observed. These mutations presumably resulted from recombination of linear doubly-cut plasmid molecules that were co-transfected with the covalently closed circular plasmids by the NHEJ repair system. To overcome this impediment a PCR-based method was develop that involves selective amplification of the replicated modified insert and subsequent re-insertion into the vector and screening in E.coli. The method improved dramatically the sensitivity of the

assay, allowing detection of O^6-MeG-targeted G→A transitions in the majority of mutants derived from both ds and gapped O^6-MeG-adducted plasmids. However, a potential limitation of the assay is its inability to detect larger frameshift mutations (e.g. large deletions and insertions). This limitation could possibly be overcome by eliminating the linear plasmid molecules before the transfection reaction, e.g. by agarose electrophoresis purification of the covalently closed circular recombinant plasmid or by digestion with Exonuclease-It (Bayou Biolabs, Harahan, USA). This refinement could potentially improve the range of mutations detected and also simplify the assay by rendering the PCR step unnecessary.

CHAPTER 5

Conclusions and Summary

The aim of the work described in this thesis was to develop and utilise mutagenesis assays for investigating the mutagenicity of carcinogenic agents. The mechanisms through which tamoxifen causes endometrial cancer are not yet established but the potential contribution of tamoxifen-induced DNA damage is a subject of much interest. The first part of the thesis was aimed at investigating the mutagenicity of the tamoxifen-induced lesions dG-N^2-TAM and dG-N^2-4-OHTAM. The conventional version of the *supF* assay was employed to ascertain the mutations caused by these adducts in human target endometrial cells. A secondary aim was to employ the LwPy53 algorithm in order to enable *in silico* prediction of the G→T mutation distribution along the human *p53* gene, using α-acetoxytamoxifen and 4-OHtamQM-induced mutation data from the *supF* assay. Relative mutation frequencies increased proportionally with adduct level for α-acetoxytamoxifen, up to ~18 times the background frequency, indicating the dG-N^2-tam adduct, which is reportedly present in endometrial DNA of a proportion of tamoxifen treated women, poses a mutagenic risk. However, 4-OHtamQM induced damage did not cause a dose-related increase in mutation frequency above that of the untreated control, which is in contrast to analogous studies performed in different cell types and indicates the importance of cell specific factors in governing the biological effects of DNA adducts. Furthermore, only α-acetoxytamoxifen induced statistically different *supF* mutation spectra compared to the spontaneous pattern. Based on the *p53* G→T transversion predictions, treatment with α-acetoxytamoxifen resulted in two hotspots at positions 244 and 273. Further investigations to establish the *p53* mutation

spectrum in endometrial tumours arising in tamoxifen treated women are needed to determine whether dG-N^2-tam adducts play a causal role in cancer development.

The second part of this thesis was devoted to the development and validation of a novel site-specific assay in *E.coli*. The main difference compared to the original *supF* assay is that this new approach assesses the mutagenicity of individual adducts, rather than a profile of damage caused by a particular genotoxic compound. Once developed and validated, the site-specific assay was then used to investigate the mutagenic potential and types of mutations caused by individual O^6-MeG adducts situated in ds or gapped plasmids, in two different sequence contexts, in *E.coli*. Importantly, the assay facilitates investigations on the mutagenicity of multiple lesions (two O^6-MeG adducts) located on the same DNA strand, to ascertain whether the presence of additional damage influences the processing of each single adduct. This is a key feature given that humans are exposed to complex mixtures of genotoxic carcinogens that can potentially generate a wide variety of lesions in genomic DNA. The methods successfully established can be easily adapted to enable studies of any stable adduct that can be synthesized and incorporated into the plasmid. Positioning of O^6-MeG adduct(s) at either site and subsequent replication in *E.coli* produced different types of mutations depending on the sequence context in which the O^6-MeG adduct(s) were placed. These results suggest that the mutagenic and cytotoxic potential of these adducts might have been underestimated and that deletions arising as a result of DNA recombination repair in response to O^6-MeG adducts might play an important role, in addition to the G→A transitions, in initiating the carcinogenic process.

The final part of the thesis was dedicated to adapting the site-specific assay for use in human cells. The main modification here was the addition of a PCR step for selection of mutant plasmids after replication of adducted or control plasmids in Ad293 human kidney cells. The assay was validated using the same oligonucleotide constructs as in the *E.coli* version, and was subsequently used to examine the mutagenicity of O^6-MeG adducts as described previously. The human cell version of the assay exhibited high sensitivity in detecting O^6-MeG-targeted G→A transitions. A potential limitation of the assay, however, is its inability to detect longer frameshift mutations, which was not a problem evident in the *E.coli* version. Furthermore, small methodological refinements have been suggested to further improve the assay before routine use. The developed assay currently offers great potential to complement the traditional *supF* approach and provide further mechanistic information aimed at elucidating mutagenic processes in the target human cells most relevant to the particular carcinogen of interest.

CHAPTER 6

Appendices and References

Appendices

Appendix 1. Raw ^{32}P-postlabelling data from a-acetoxytamoxifen treatment of pSP189 plasmid
This data was used to produce graph A in Figure 2.7, page 89

	Value 1	Value 2	Value 3	Value 4	Average	S.D.	Standard Error
Solvent Control	0	0	0	0	0	0	0
10 mM	6295,85	20187,87	24484,95	18904,36	17468,26	7821,24	3910,62
25 mM	20689,64	26539,21	30748,18	41030,16	29751,80	8575,89	4287,94
50 mM	30279,13	37047,37	38062,35	45182,82	37642,92	6099,35	3049,67

Adducts per 10^8 nucleotides

Appendix 2. Raw ^{32}P-postlabelling data from 4-OHtamQM treatment of pSP189 plasmid
This data was used to produce graph B in Figure 2.7, page 89

	Value 1	Value 2	Value 3	Value 4	Average	S.D.	Standard Error
Solvent Control	0	0	0	0	0	0	0
50 mM	436,52	318,89	551,13		436	83,18	41,59
100 mM	1007,85	876,56	967,53	413,8	816	92,84	46,42
250 mM	4104,44	3205,34	2521,56	2564,12	3099	635,76	317,88

Adducts per 10^8 nucleotides

Appendix 3. Example of an Excel spreadsheet containing imars input data. The format of the data file is shown below where the first column describes the sequence where each base is preceded by a number. The remaining columns give the frequencies of the different types of point mutations observed for each base (where X is the current base and, for example, an A to G mutation is written A>G). Thus, in this example: for base 7G there were two mutations, both G A>.

	X-A	X-G	X-C	X-T	DEL	INS
1G	0	0	0	0	0	0
2G	0	0	0	0	0	0
3T	0	0	0	0	0	0
4G	0	0	0	0	0	0
5G	0	0	0	1	0	0
6G	0	0	0	1	0	0
7G	2	0	0	0	0	0
8T	1	0	0	0	0	0
9T	0	0	0	0	1	0
10C	0	1	0	1	0	0
11C	1	1	0	0	0	0
12C	0	0	0	0	0	0
13G	0	0	0	0	0	0
14A	0	0	0	0	0	0
15G	3	0	0	1	0	0
16C	0	0	0	0	0	0
17G	1	0	2	5	0	0
18G	1	0	0	0	0	0
19C	0	0	0	0	0	0
20C	2	1	0	0	0	0
21A	0	0	0	0	0	0
22A	0	0	1	0	0	0
23A	0	0	0	0	1	0
24G	0	0	0	7	0	0
25G	0	0	0	3	0	0
26G	0	0	0	1	7	1
27A	0	0	0	0	0	0
28G	2	0	0	0	0	0
29C	3	0	0	1	0	0
30A	0	0	0	0	0	0
31G	1	0	1	8	0	0
32A	0	0	0	0	0	0
33C	0	0	0	0	0	0
34T	0	0	0	0	0	0
35C	3	3	0	0	0	0

Appendix 4. HYPERG is a computer program that compares two mutation spectra. The input of the program is a text file containing the number and nature of mutations observed in the two spectra. The output of the program is a *P* value, which indicates the probability that the two spectra are drawn from the same population. A *P* value of <=0.05 indicates that the two spectra are most likely not drawn from the same population. The algorithm implemented here is described in [479]. This algorithm is currently viewed as one of the best for mutational spectra calculations. The types of substitution mutations present in the two spectra are entered in the form of a table, as shown in the example below;

bp position	wt base	specific mutation	spectra 1	spectra 2
44	A	A>C	0	1
		A>G	0	1
		A>T	0	0
45	C	C>A	0	0
		C>G	0	0
		C>T	0	0
46	G	G>A	3	1
		G>C	0	1
		G>T	1	1

Rows with no mutants (0 0) are discarded and the final two column table is set up, like this.

0	1
0	1
3	1
0	1
1	1

This table now contains the two input spectra. The file containing these two columns of numbers is then submitted to the program.

The program then generates a number of random tables that contains the same number of mutants in each row as the input spectra. The number of iterations that you request is the number of random tables to generate. 1700 random tables gives fairly good precision. The input table is compared to each random table. If the input spectra is "more unusual" than the random table, this is scored as a hit. The *P* value is "hits / (total number of random tables)" [479].

REFERENCES

1.   King, R.J.B., *Cancer biology, 2nd edition.* 2000.

2.   Lynch, H.T., R.M. Fusaro, and J. Lynch, *Hereditary cancer in adults.* Cancer Detect Prev, 1995. **19**(3): p. 219-33.

3.   Fearon, E.R., *Human cancer syndromes: clues to the origin and nature of cancer.* Science, 1997. **278**(5340): p. 1043-50.

4.   Cancer Research UK.[cited;Available from: http:\\www.cancerresearchuk.org.

5.   Doll, R. and R. Peto, *The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today.* J Natl Cancer Inst, 1981. **66**(6): p. 1191-308.

6.   Epstein, S.S. and J.B. Swartz, *Fallacies of lifestyle cancer theories.* Nature, 1981. **289**(5794): p. 127-30.

7.   Colditz, G.A., DeJong, D., Hunter, D.J., Trichopoulos, D. & Willett, W.C., *Causes of Human Cancer*, in *Harvard report on cancer prevention, vol.1.* 1996, Harvard University. p. 1-59.

8.   Barrett, J.C., *Mechanisms of multistep carcinogenesis and carcinogen risk assessment.* Environ Health Perspect, 1993. **100**: p. 9-20.

9.   Foulds, L., *Neoplastic Development.* 1975, New York: Academic Press.

10.   Nowell, P.C., *The clonal evolution of tumor cell populations.* Science, 1976. **194**(4260): p. 23-8.

11.   Alberts, B., *Molecular Biology of the Cell, 4th edition,.* 2002: Garland publishing.

12.   Jianling Yuan, P.M.G., *Mutagenesis induced by the tumor microenvironment.* Mutat Res. 400 (1998) 439-446, 1998.

13.   Bishop, J.M., *The molecular genetics of cancer.* Science, 1987. **235**(4786): p. 305-11.

14.   Loeb, L.A., *Mutator phenotype may be required for multistage carcinogenesis.* Cancer Res, 1991. **51**(12): p. 3075-9.

15.   Vogelstein, B., et al., *Genetic alterations during colorectal-tumor development.* N Engl J Med, 1988. **319**(9): p. 525-32.

16.   Vogelstein, B., et al., *Allelotype of colorectal carcinomas.* Science, 1989. **244**(4901): p. 207-11.

17.   Boyd, J.A. and J.C. Barrett, *Genetic and cellular basis of multistep carcinogenesis.* Pharmacol Ther, 1990. **46**(3): p. 469-86.

18.   Marks, F., G. Furstenberger, and K. Muller-Decker, *Tumor promotion as a target of cancer prevention.* Recent Results Cancer Res, 2007. **174**: p. 37-47.

19.   Holland et al., *Cancer Medicine, 6th edition.* 2003.

20.   Barrett, J.C., and Wiseman, R.W., *Cellular and molecular mechanisms of multistep carcinogenesis: relevance to carcinogen risk assessment.* Environ. Health Perspect., 1987. **76**: p. 65-70.

21.   *Multistage and multifactorial nature of carcinogenesis.* IARC Sci Publ, 1992(116): p. 9-54.

22.   Martinez, J.D., *Molecular biology of cancer; Chemotherapeutic Agents, 6th edition* Vol. 5. 2003.

23.   Jones, P.A., *DNA methylation and cancer.* Cancer Res, 1986. **46**(2): p. 461-6.

24.   Jones, P.A. and S.B. Baylin, *The fundamental role of epigenetic events in cancer.* Nat Rev Genet, 2002. **3**(6): p. 415-28.

25.   Counts, J.L. and J.I. Goodman, *Hypomethylation of DNA: a nongenotoxic mechanism involved in tumor promotion.* Toxicol Lett, 1995. **82-83**: p. 663-72.

26.   Pogribny, I.P., et al., *Breaks in genomic DNA and within the p53 gene are associated with hypomethylation in livers of folate/methyl-deficient rats.* Cancer Res, 1995. **55**(9): p. 1894-901.

27.   Surani, M.A., *Genomic imprinting: developmental significance and molecular mechanism.* Curr Opin Genet Dev, 1991. **1**(2): p. 241-6.

28.   Feinberg, A.P. and B. Tycko, *The history of cancer epigenetics.* Nat Rev Cancer, 2004. **4**(2): p. 143-53.

29.   Zhang, Y. and D. Reinberg, *Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails.* Genes Dev, 2001. **15**(18): p. 2343-60.

30.   Boveri, T., *Zur Frage der Entstehung Malignen Tumoren.* 1914, Jena, Germany: Gustave Fischer.

31.   Barrett, J.C., *Relationship between mutagenesis and carcinogenesis: In the origins of human cancer,* ed. T.C. J. Brugge, E. Harlow and F. McCormick. 1991, NY: cold spring harbor laboratory press.

32.   G. Peters, *Oncogenes and Tumour Suppressors,* ed. K.H. Vousden. 1997, New York: IRL Press, Oxford University Press.

33.   Srivastava, S., et al., *Germ-line transmission of a mutated p53 gene in a cancer-prone family with Li-Fraumeni syndrome.* Nature, 1990. **348**(6303): p. 747-9.

34.   Powell, S.M., et al., *APC mutations occur early during colorectal tumorigenesis.* Nature, 1992. **359**(6392): p. 235-7.

35.   Dunn, J.M., et al., *Identification of germline and somatic mutations affecting the retinoblastoma gene.* Science, 1988. **241**(4874): p. 1797-800.

36.   Kunkel, T.A., *DNA replication fidelity.* J Biol Chem, 2004. **279**(17): p. 16895-8.

37.   Errol C. Friedberg, G.C.W., Wolfram Siede, *DNA repair and mutagenesis.* 1995.

38.   Hoeijmakers, J.H.J., *Genome maintenance mechanisms for preventing cancer.* Nature, 411 (2001) 366-374, 2001.

39.   Jackson, A.L. and L.A. Loeb, *The contribution of endogenous sources of DNA damage to the multiple mutations in cancer.* Mutat Res, 2001. **477**(1-2): p. 7-21.

40.   Lindahl, T. and D.E. Barnes, *Repair of endogenous DNA damage.* Cold Spring Harb Symp Quant Biol, 2000. **65**: p. 127-33.

41.   Smith, M.B.M., J. , *Advanced Organic Chemistry 5th ed., pp 1218-1223,* ed. W. Interscience. 2001, New York.

42.   Anthony J.F. Griffiths, e.a., *An introduction to genetic analysis.* 7th Bk&Cdr edition ed. 2000: W. H. Freeman;.

43.   Lindahl, T., *Instability and decay of the primary structure of DNA.* Nature, 1993. **362**(6422): p. 709-15.

44.   Lindahl, T., *An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues.* Proc Natl Acad Sci U S A, 1974. **71**(9): p. 3649-53.

45.   Greer, S. and S. Zamenhof, *Studies on depurination of DNA by heat.* J Mol Biol, 1962. **4**: p. 123-41.

46.   Lindahl, T., *DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and base excision-repair.* Prog Nucleic Acid Res Mol Biol, 1979. **22**: p. 135-92.

47.   Loeb, L.A. and B.D. Preston, *Mutagenesis by apurinic/apyrimidinic sites.* Annu Rev Genet, 1986. **20**: p. 201-30.

48.   Lindahl, T. and B. Nyberg, *Rate of depurination of native deoxyribonucleic acid.* Biochemistry, 1972. **11**(19): p. 3610-8.

49.   Zoltewicz, J.A., et al., *Kinetics and mechanism of the acid-catalyzed hydrolysis of some purine nucleosides.* J Am Chem Soc, 1970. **92**(6): p. 1741-9.

50.   Ullman, J.S. and B.J. McCarthy, *Alkali deamination of cytosine residues in DNA.* Biochim Biophys Acta, 1973. **294**(1): p. 396-404.

51.   Lindahl, T. and O. Karlstrom, *Heat-induced depyrimidination of deoxyribonucleic acid in neutral solution.* Biochemistry, 1973. **12**(25): p. 5151-4.

52.   Boveris, A. and B. Chance, *The mitochondrial generation of hydrogen peroxide. General properties and effect of hyperbaric oxygen.* Biochem J, 1973. **134**(3): p. 707-16.

53.   Lehninger, A., *Principles of Biochemistry.* 1982, New York: Worth Publishers.

54.   Wang, D., D.A. Kreutzer, and J.M. Essigmann, *Mutagenicity and repair of oxidative DNA damage: insights from studies using defined lesions.* Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis, 1998. **400**(1-2): p. 99-115.

55.   Bartsch, H. and J. Nair, *Ultrasensitive and specific detection methods for exocylic DNA adducts: markers for lipid peroxidation and oxidative stress.* Toxicology, 2000. **153**(1-3): p. 105-14.

56.   Marnett, L.J., *Oxyradicals and DNA damage.* Carcinogenesis, 2000. **21**(3): p. 361-70.

57.   Esterbauer, H., P. Eckl, and A. Ortner, *Possible mutagens derived from lipids and lipid precursors.* Mutat Res, 1990. **238**(3): p. 223-33.

58.   Halliwell, B. and J.M. Gutteridge, *Lipid peroxidation, oxygen radicals, cell damage, and antioxidant therapy.* Lancet, 1984. **1**(8391): p. 1396-7.

59.   Imlay, J.A. and S. Linn, *DNA damage and oxygen radical toxicity.* Science, 1988. **240**(4857): p. 1302-9.

60.   Fenton, H.J.H., *Oxidation of tartaric acid in the presence of iron.* J. Chem. Soc. Trans, 1894. **65**: p. 899-905.

61.   Cohen, G., *The Fenton Reaction, p. 55-64.* R.A. Greenwald ed. 1985, Boca Raton, Fla: GRC Press, Inc.

62.   Dizdaroglu, M., *Oxidative damage to DNA in mammalian chromatin.* Mutat Res, 1992. **275**(3-6): p. 331-42.

63.   Breen, A.P. and J.A. Murphy, *Reactions of oxyl radicals with DNA.* Free Radic Biol Med, 1995. **18**(6): p. 1033-77.

64.   Henle, E.S. and S. Linn, *Formation, prevention, and repair of DNA damage by iron/hydrogen peroxide.* J Biol Chem, 1997. **272**(31): p. 19095-8.

65.   Bjelland, S. and E. Seeberg, *Mutagenicity, toxicity and repair of DNA base damage induced by oxidation.* Mutat Res, 2003. **531**(1-2): p. 37-80.

66.   Cheng, K.C., et al., *8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G----T and A----C substitutions.* J Biol Chem, 1992. **267**(1): p. 166-72.

67.   Kuchino, Y., et al., *Misreading of DNA templates containing 8-hydroxydeoxyguanosine at the modified base and at adjacent residues.* Nature, 1987. **327**(6117): p. 77-9.

68.   Shibutani, S., M. Takeshita, and A.P. Grollman, *Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG.* Nature, 1991. **349**(6308): p. 431-4.

69.   Floyd, R.A., et al., *Hydroxyl free radical adduct of deoxyguanosine: sensitive detection and mechanisms of formation.* Free Radic Res Commun, 1986. **1**(3): p. 163-72.

70.   Collins, A., et al., *Problems in the measurement of 8-oxoguanine in human DNA. Report of a workshop, DNA oxidation, held in Aberdeen, UK, 19-21 January, 1997.* Carcinogenesis, 1997. **18**(9): p. 1833-6.

71.   Gedik, C.M. and A. Collins, *Establishing the background level of base oxidation in human lymphocyte DNA: results of an interlaboratory validation study.* Faseb J, 2005. **19**(1): p. 82-4.

72.   Friedberg, E.C., Walker. G.C., Siede, W., *DNA repair mutagenesis.* 1995, Washington, DC: ASM Press.

73.   Trichopoulos, D., and Petridou, E. , *Epidemiologic studies and cancer etiology in humans.* Med. Exerc. Nutr. Health, 1994. **3**: p. 206-225.

74.   Lichtenstein, P., et al., *Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland.* N Engl J Med, 2000. **343**(2): p. 78-85.

75.   Czene, K., P. Lichtenstein, and K. Hemminki, *Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish Family-Cancer Database.* Int J Cancer, 2002. **99**(2): p. 260-6.

76.   Weisburger, J.H. and G.M. Williams, *Carcinogen testing: current problems and new approaches.* Science, 1981. **214**(4519): p. 401-7.

77.   Decordier, I., et al., *Elimination of micronucleated cells by apoptosis after treatment with inhibitors of microtubules.* Mutagenesis, 2002. **17**(4): p. 337-44.

78.   Wang, M.Y. and J.G. Liehr, *Induction by estrogens of lipid peroxidation and lipid peroxide-derived malonaldehyde-DNA adducts in male Syrian hamsters: role of lipid peroxidation in estrogen-induced kidney carcinogenesis.* Carcinogenesis, 1995. **16**(8): p. 1941-5.

79.   Liehr, J.G., *Role of DNA adducts in hormonal carcinogenesis.* Regul Toxicol Pharmacol, 2000. **32**(3): p. 276-82.

80.   Wanless, I.R. and A. Medline, *Role of estrogens as promoters of hepatic neoplasia.* Lab Invest, 1982. **46**(3): p. 313-20.

81.   Yager, J.D., J. Zurlo, and N. Ni, *Sex hormones and tumor promotion in liver.* Proc Soc Exp Biol Med, 1991. **198**(2): p. 667-74.

82.   Toman, Z., C. Dambly, and M. Radman, *Induction of a stable, heritable epigenetic change by mutagenic carcinogens: a new test system.* IARC Sci Publ, 1980(27): p. 243-55.

83.   Farmer, P.B., *Carcinogen Adducts - Use in Diagnosis and Risk Assessment.* Clinical Chemistry, 1994. **40**(7B): p. 1438-1443.

84.   Afshari, C.A. and J.C. Barrett, *Cell cycle controls: potential targets for chemical carcinogens?* Environ Health Perspect, 1993. **101 Suppl 5**: p. 9-14.

85.   IGHRC, *Assessment of chemical carcinogens: Background to general principles of a weight of evidence approach.* 2002, Institute for Environment and Health.

86.   Barrett, J.C., Lee T-C, *Mechanisms of arsenic-induced gene amplification. In: Gene amplification in Mammalian cells: a Comprehensive Guide, pp.441-446*, ed. K. RE. 1992, New York.

87.   Marsman, D.S. and J.C. Barrett, *Apoptosis and chemical carcinogenesis.* Risk Anal, 1994. **14**(3): p. 321-6.

88.   Phillips, D.H., *Environmental Mutagenesis.* 1995, Oxford: BIOS Scientific Publishers Ltd.

89.   Kirsch-Volders, M., et al., *Indirect mechanisms of genotoxicity.* Toxicol Lett, 2003. **140-141**: p. 63-74.

90.   Lock, E.A., A.M. Mitchell, and C.R. Elcombe, *Biochemical mechanisms of induction of hepatic peroxisome proliferation.* Annu Rev Pharmacol Toxicol, 1989. **29**: p. 145-63.

91.   Reddy, J.K., D.L. Azarnoff, and C.E. Hignite, *Hypolipidaemic hepatic peroxisome proliferators form a novel class of chemical carcinogens.* Nature, 1980. **283**(5745): p. 397-8.

92.   Ashby, J., et al., *Mechanistically-based human hazard assessment of peroxisome proliferator-induced hepatocarcinogenesis.* Hum Exp Toxicol, 1994. **13 Suppl 2**: p. S1-117.

93.   Hwang, J.J., M.T. Hsia, and R.L. Jirtle, *Induction of sister chromatid exchange and micronuclei in primary cultures of rat and human hepatocytes by the peroxisome proliferator, Wy-14,643.* Mutat Res, 1993. **286**(2): p. 123-33.

94.   Reisenbichler, H. and P.M. Eckl, *Genotoxic effects of selected peroxisome proliferators.* Mutat Res, 1993. **286**(2): p. 135-44.

95.   Tsutsui, T., E. Watanabe, and J.C. Barrett, *Ability of peroxisome proliferators to induce cell transformation, chromosome aberrations and peroxisome proliferation in cultured Syrian hamster embryo cells.* Carcinogenesis, 1993. **14**(4): p. 611-8.

96.   Williams, G.M., *Methods for evaluating chemical genotoxicity.* Annu Rev Pharmacol Toxicol, 1989. **29**: p. 189-211.

97.   Alvaro, P., and Wallace, K.B, *Molecular biology of the toxic response.* 1999, New York: Taylor & Francis.

98.   Dalton, T.P., H.G. Shertzer, and A. Puga, *Regulation of gene expression by reactive oxygen.* Annu Rev Pharmacol Toxicol, 1999. **39**: p. 67-101.

99.   Speit, G., et al., *Genotoxicity of hyperbaric oxygen.* Mutat Res, 2002. **512**(2-3): p. 111-9.

100.   Huseby, R.A., *Demonstration of a direct carcinogenic effect of estradiol on Leydig cells of the mouse.* Cancer Res, 1980. **40**(4): p. 1006-13.

101.   Dao, T.L. and P.C. Chan, *Hormones and dietary fat as promoters in mammary carcinogenesis.* Environ Health Perspect, 1983. **50**: p. 219-25.

102.   Yager, J.D. and J.G. Liehr, *Molecular mechanisms of estrogen carcinogenesis.* Annu Rev Pharmacol Toxicol, 1996. **36**: p. 203-32.

103.   Mitsumori, K., et al., *Modifying effects of ethinylestradiol but not methoxychlor on N-ethyl-N-nitrosourea-induced uterine carcinogenesis in heterozygous p53-deficient CBA mice.* Toxicol Sci, 2000. **58**(1): p. 43-9.

104.   Henderson, B.E. and H.S. Feigelson, *Hormonal carcinogenesis.* Carcinogenesis, 2000. **21**(3): p. 427-33.

105.   Schoeny, R., *Use of genetic toxicology data in U.S. EPA risk assessment: the mercury study report as an example.* Environ Health Perspect, 1996. **104 Suppl 3**: p. 663-73.

106.   Duesberg, P. and R. Li, *Multistep carcinogenesis: a chain reaction of aneuploidizations.* Cell Cycle, 2003. **2**(3): p. 202-10.

107.   Parry, J.M., et al., *In vitro and in vivo extrapolations of genotoxin exposures: consideration of factors which influence dose-response thresholds.* Mutat Res, 2000. **464**(1): p. 53-63.

108.   Rajagopalan, H. and C. Lengauer, *Aneuploidy and cancer.* Nature, 2004. **432**(7015): p. 338-41.

109.   Duesberg, P., et al., *The chromosomal basis of cancer.* Cell Oncol, 2005. **27**(5-6): p. 293-318.

110.   Foth, H., G.H. Degen, and H.M. Bolt, *New aspects in the classification of carcinogens.* Arh Hig Rada Toksikol, 2005. **56**(2): p. 167-75.

111.   Nora, J.J., and Fraser, F.C., *Medical genetics: Principles and Practice* 1989, Philadelphia.

112.   Langard, S., *One hundred years of chromium and cancer: a review of epidemiological evidence and selected case reports.* Am J Ind Med, 1990. **17**(2): p. 189-215.

113.   Hueper, W.C., *Occupational and environmental cancers of the respiratory system.* Recent Results in Cancer Research. 1966, New York: Springer-Verlag.

114.   Williams, G.M., *DNA reactive and epigenetic carcinogens.* Exp Toxicol Pathol, 1992. **44**(8): p. 457-63.

115.   Garner, R.C., *The role of DNA adducts in chemical carcinogenesis.* Mutat Res. 402 (1998) 67-75, 1998.

116.   Bradfield, *Mechanisms of Liver Carcinogenesis UNIT IV: Consequences and Targets of Alkylating Agents.*

117.   Dipple, A., *DNA adducts of chemical carcinogens.* Carcinogenesis, 1995. **16**(3): p. 437-41.

118.   Taverna, P. and B. Sedgwick, *Generation of an endogenous DNA-methylating agent by nitrosation in Escherichia coli.* J Bacteriol, 1996. **178**(17): p. 5105-11.

119.   Sedgwick, B., *Nitrosated peptides and polyamines as endogenous mutagens in O6-alkylguanine-DNA alkyltransferase deficient cells.* Carcinogenesis, 1997. **18**(8): p. 1561-7.

120.   Ferguson, L.R. and A.E. Pearson, *The clinical use of mutagenic anticancer drugs.* Mutat Res, 1996. **355**(1-2): p. 1-12.

121.   Lawley, P.D., *Carcinogenesis by alkylating agents, pp 325-484,* ed. C.E. In Searle. 1984, Washington DC: American Chemical Society.

122.   Pegg, A.E., *Methylation of the O6 position of guanine in DNA is the most likely initiating event in carcinogenesis by methylating agents.* Cancer Invest, 1984. **2**(3): p. 223-31.

123.   Singer B., G.D., *Molecular biology of mutagens and carcinogens.* 1983, NY.

124.   Beranek, D.T., *Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents.* Mutat Res, 1990. **231**(1): p. 11-30.

125.   Hecht, S.S., *DNA adduct formation from tobacco-specific N-nitrosamines.* Mutat Res, 1999. **424**(1-2): p. 127-42.

126.   Groopman, J.D. and T.W. Kensler, *Role of metabolism and viruses in aflatoxin-induced liver cancer.* Toxicol Appl Pharmacol, 2005. **206**(2): p. 131-7.

127.   Wogan, G.N., *Aflatoxins as risk factors for hepatocellular carcinoma in humans.* Cancer Res, 1992. **52**(7 Suppl): p. 2114s-2118s.

128.   Garner, R.C., et al., *Comparison of aflatoxin B1 and aflatoxin G1 binding to cellular macromolecules in vitro, in vivo and after peracid oxidation; characterisation of the major nucleic acid adducts.* Chem Biol Interact, 1979. **26**(1): p. 57-73.

129.   Greenblatt, M.S., et al., *Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis.* Cancer Res, 1994. **54**(18): p. 4855-78.

130.   Harvey, R.G., *Polycyclic Aromatic Hydrocarbons: Chemistry and Carcinogenicity.* 1991, Cambridge, UK: Cambridge University Press.

131.   Wynder, E.L. and D. Hoffmann, *A study of tobacco carcinogenesis. VII. The role of higher polycyclic hydrocarbons.* Cancer, 1959. **12**: p. 1079-86.

132.   Van Duuren, B.L., *Identification of some polynuclear aromatic hydrocarbons in cigarette-smoke condensate.* J Natl Cancer Inst, 1958. **21**(1): p. 1-16.

133.   Denissenko, M.F., et al., *Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53.* Science, 1996. **274**(5286): p. 430-2.

134.   Sabbioni, G. and C.R. Jones, *Biomonitoring of arylamines and nitroarenes.* Biomarkers, 2002. **7**(5): p. 347-421.

135.   Debnath, A.K., et al., *A QSAR investigation of the role of hydrophobicity in regulating mutagenicity in the Ames test: 1. Mutagenicity of aromatic and heteroaromatic amines in Salmonella typhimurium TA98 and TA100.* Environ Mol Mutagen, 1992. **19**(1): p. 37-52.

136.   Vineis, P. and R. Pirastu, *Aromatic amines and cancer.* Cancer Causes Control, 1997. **8**(3): p. 346-55.

137.   Delclos, K.B., and Kadlubar, F. F. , *Carcinogenic aromatic amines and amides.* Comprehensive Toxicology 1997. **12**: p. 141-170.

138.   Upton, A.C., *Historical perspectives on radiation carcinogenesis*, ed. U. A.C. 1986, New York: Elsevier.

139.   Fry, R.J.M.a.S., J.B., *External radiation carcinogenesis.* Advances in Radiation Biology, vol.13, ed. J.T. Lett. 1987, New York: Academic Press.

140.   Chadwick, K.H.e.a., *Cell Transformation and Radiation-Induced Cancer*, ed. A. Hilger. 1989, New York: Adam Hilger.

141.   Little, J.B., *The relevance of cell transformation to carcinogenesis in vivo.* Low Dose Radiation-Biological Basis of Risk Assessment, ed. K.F.a.S. Baverstock, J.W. 1989, London: Taylor and Francis.

142.   Ward, J.F., *DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability.* Prog Nucleic Acid Res Mol Biol, 1988. **35**: p. 95-125.

143.   Riley, P.A., *Free radicals in biology: oxidative stress and the effects of ionizing radiation.* Int J Radiat Biol, 1994. **65**(1): p. 27-33.

144.   de Gruijl, F.R., *Photocarcinogenesis: UVA vs UVB.* Methods Enzymol, 2000. **319**: p. 359-66.

145.   Matsumura, Y. and H.N. Ananthaswamy, *Toxic effects of ultraviolet radiation on the skin.* Toxicol Appl Pharmacol, 2004. **195**(3): p. 298-308.

146.   Gallagher, R.P., et al., *Trends in basal cell carcinoma, squamous cell carcinoma, and melanoma of the skin from 1973 through 1987.* J Am Acad Dermatol, 1990. **23**(3 Pt 1): p. 413-21.

147.   Setlow, R.B. and A.D. Woodhead, *Temporal changes in the incidence of malignant melanoma: explanation from action spectra.* Mutat Res, 1994. **307**(1): p. 365-74.

148.   Griffiths, H.R., et al., *Molecular and cellular effects of ultraviolet light-induced genotoxicity*. Crit Rev Clin Lab Sci, 1998. **35**(3): p. 189-237.

149.   Wei, H., et al., *DNA structural integrity and base composition affect ultraviolet light-induced oxidative DNA damage*. Biochemistry, 1998. **37**(18): p. 6485-90.

150.   Burnouf, D.Y., R. Miturski, and R.P.P. Fuchs, *Sequence context modulation of translesion synthesis at a single N-2-acetylaminofluorene adduct located within a mutation hot spot*. Chemical Research in Toxicology, 1999. **12**(2): p. 144-150.

151.   Matunis, M.J., *On the road to repair: PCNA encounters SUMO and ubiquitin modifications*. Mol Cell, 2002. **10**(3): p. 441-2.

152.   Augusto-Pinto, L., et al., *Escherichia coli as a model system to study DNA repair genes of eukaryotic organisms*. Genet Mol Res, 2003. **2**(1): p. 77-91.

153.   Carell, T., et al., *The mechanism of action of DNA photolyases*. Curr Opin Chem Biol, 2001. **5**(5): p. 491-8.

154.   Pegg, A.E., *Repair of O(6)-alkylguanine by alkyltransferases*. Mutat Res, 2000. **462**(2-3): p. 83-100.

155.   Hanawalt, S.D.C.a.P.C., *Who's on first in the cellular response to DNA damage?* Nature Reviews Molecular Cell Biology, 4 (2003) 361-372, 2003.

156.   Errol C. Friedberg, G.C.W., Wolfram Siede, *DNA repair and mutagenesis*. Second ed. 2006.

157.   Takata, M., et al., *Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells*. Embo J, 1998. **17**(18): p. 5497-508.

158.   Krejci, L., et al., *Mending the break: two DNA double-strand break repair machines in eukaryotes*. Prog Nucleic Acid Res Mol Biol, 2003. **74**: p. 159-201.

159.   Lieber, M.R., et al., *Mechanism and regulation of human non-homologous DNA end-joining*. Nat Rev Mol Cell Biol, 2003. **4**(9): p. 712-20.

160.   Fang, W.H. and P. Modrich, *Human strand-specific mismatch repair occurs by a bidirectional mechanism similar to that of the bacterial reaction*. J Biol Chem, 1993. **268**(16): p. 11838-44.

161.   Radman, M. and R. Wagner, *Mismatch repair in Escherichia coli*. Annu Rev Genet, 1986. **20**: p. 523-38.

162.   Modrich, P., *DNA mismatch correction*. Annu Rev Biochem, 1987. **56**: p. 435-66.

163.   Jascur, T. and C.R. Boland, *Structure and function of the components of the human DNA mismatch repair system*. Int J Cancer, 2006. **119**(9): p. 2030-5.

164.   Wood, R.D., *DNA repair in eukaryotes*. Annu Rev Biochem, 1996. **65**: p. 135-67.

165.   Ochs, K., et al., *Cells deficient in DNA polymerase beta are hypersensitive to alkylating agent-induced apoptosis and chromosomal breakage*. Cancer Res, 1999. **59**(7): p. 1544-51.

166.   Lindahl, T., *New class of enzymes acting on damaged DNA*. Nature, 1976. **259**(5538): p. 64-6.

167.   Siede, W., *DNA Damage Recognition*. 2006, New York: Taylor & Francis Group.

168.   Baynton, K. and R.P.P. Fuchs, *Lesions in DNA: hurdles for polymerases*. Trends in Biochemical Sciences, 2000. **25**(2): p. 74-79.

169.   Napolitano, R.L., I.B. Lambert, and R.P.P. Fuchs, *SOS factors involved in translesion synthesis*. Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(11): p. 5733-5738.

170.   Becherel, O.J. and R.P.P. Fuchs, *SOS mutagenesis results from up-regulation of translesion synthesis*. Journal of Molecular Biology, 1999. **294**(2): p. 299-306.

171.   Wagner, J., et al., *Genetics of mutagenesis in E-coli: various combinations of translesion polymerases (Pol II, IV and V) deal with lesion/sequence context diversity*. DNA Repair, 2002. **1**(2): p. 159-167.

172.   Rattray, A.J. and J.N. Strathern, *Error-prone DNA polymerases: when making a mistake is the only way to get ahead*. Annu Rev Genet, 2003. **37**: p. 31-66.

173.   Strathern, A.J.R.a.J.N., *Error-prone DNA polymerases: When making a mistake is the only way to get ahead*. Annu. Rev. Genet. 37 (2003) 31-66, 2003.

174.   Friedberg, E.C., R. Wagner, and M. Radman, *Specialized DNA polymerases, cellular survival, and the genesis of mutations*. Science, 2002. **296**(5573): p. 1627-30.

175.   Goodman, M.F., *Coping with replication 'train wrecks' in Escherichia coli using Pol V, Pol II and RecA proteins*. Trends Biochem Sci, 2000. **25**(4): p. 189-95.

176.   Livneh, Z., *DNA damage control by novel DNA polymerases: translesion replication and mutagenesis*. J Biol Chem, 2001. **276**(28): p. 25639-42.

177.   Kunz, B.A., A.F. Straffon, and E.J. Vonarx, *DNA damage-induced mutation: tolerance via translesion synthesis*. Mutat Res, 2000. **451**(1-2): p. 169-85.

178.   Bertram, J.S., *The molecular biology of cancer*. Molecular Aspects of Medicine 21 (2001) 167-223, 2001.

179.   Rogozin, I., F. Kondrashov, and G. Glazko, *Use of mutation spectra analysis software*. Hum Mutat, 2001. **17**(2): p. 83-102.

180.   Benzer, S., *On the Topography of the Genetic Fine Structure*. Proc Natl Acad Sci U S A, 1961. **47**(3): p. 403-15.

181.   Coulondre, C., et al., *Molecular basis of base substitution hotspots in Escherichia coli*. Nature, 1978. **274**(5673): p. 775-80.

182.   Betz, A.G., et al., *Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: clustering, polarity, and specific hot spots*. Proc Natl Acad Sci U S A, 1993. **90**(6): p. 2385-8.

183.   Hainaut, P. and G.P. Pfeifer, *Patterns of p53 G-->T transversions in lung cancers reflect the primary mutagenic signature of DNA-damage by tobacco smoke*. Carcinogenesis, 2001. **22**(3): p. 367-74.

184.   Perera, F.P. and I.B. Weinstein, *Molecular epidemiology: recent advances and future directions*. Carcinogenesis, 2000. **21**(3): p. 517-24.

185.   Seo, K.Y., S.A. Jelinsky, and E.L. Loechler, *Factors that influence the mutagenic patterns of DNA adducts from chemical carcinogens*. Mutation Research-Reviews in Mutation Research, 2000. **463**(3): p. 215-246.

186.   Moriya, M., et al., *Fidelity of translesional synthesis past benzo[a]pyrene diol epoxide-2'-deoxyguanosine DNA adducts: marked effects of host cell, sequence context, and chirality*. Biochemistry, 1996. **35**(51): p. 16646-51.

187.   Page, J.E., et al., *Sequence context profoundly influences the mutagenic potency of trans-opened benzo[a]pyrene 7,8-diol 9,10-epoxide-purine nucleoside adducts in site-specific mutation studies*. Biochemistry, 1998. **37**(25): p. 9127-37.

188. Latham, G.J., et al., *The replication fate of R- and S-styrene oxide adducts on adenine N6 is dependent on both the chirality of the lesion and the local sequence context.* J Biol Chem, 1993. **268**(31): p. 23427-34.

189. Jelinsky, S.A., et al., *The Major, N-2-Gua Adduct of the (+)-Anti-Benzo[a]Pyrene Diol Epoxide Is Capable of Inducing G-]a and G-]C, in Addition to G-]T, Mutations.* Biochemistry, 1995. **34**(41): p. 13545-13553.

190. Shukla, R., N.E. Geacintov, and E.L. Loechler, *The major, N2-dG adduct of (+)-anti-B[a]PDE induces G-->A mutations in a 5'-AGA-3' sequence context.* Carcinogenesis, 1999. **20**(2): p. 261-8.

191. Delaney, J.C. and J.M. Essigmann, *Context-dependent mutagenesis by DNA lesions.* Chem Biol, 1999. **6**(10): p. 743-53.

192. Boulikas, T., *Evolutionary consequences of nonrandom damage and repair of chromatin domains.* J Mol Evol, 1992. **35**(2): p. 156-80.

193. Gordenin, D.A. and M.A. Resnick, *Yeast ARMs (DNA at-risk motifs) can reveal sources of genome instability.* Mutat Res, 1998. **400**(1-2): p. 45-58.

194. Parris, C.N., et al., *Proximal and Distal Effects of Sequence Context on Ultraviolet Mutational Hotspots in a Shuttle Vector Replicated in Xeroderma Cells.* Journal of Molecular Biology, 1994. **236**(2): p. 491-502.

195. Horsfall, M.J., et al., *Mutational specificity of alkylating agents and the influence of DNA repair.* Environ Mol Mutagen, 1990. **15**(2): p. 107-22.

196. Krawczak, M., et al., *Somatic spectrum of cancer-associated single basepair substitutions in the TP53 gene is determined mainly by endogenous mechanisms of mutation and by selection.* Hum Mutat, 1995. **5**(1): p. 48-57.

197. Ponten, I., et al., *Sequence context effects on mutational properties of cis-opened benzo[c]phenanthrene diol epoxide-deoxyadenosine adducts in site-specific mutation studies.* Biochemistry, 1999. **38**(3): p. 1144-52.

198. Levy, D.D., A.D. Magee, and M.M. Seidman, *Single nucleotide positions have proximal and distal influence on UV mutation hotspots and coldspots.* Journal of Molecular Biology, 1996. **258**(2): p. 251-260.

199. Levy, D.D., et al., *The influence of single base changes on UV mutational activity at two translocated hotspots.* Journal of Molecular Biology, 1996. **255**(3): p. 435-445.

200. Kramata, P., et al., *A single site-specific trans-opened 7,8,9,10-tetrahydrobenzo[a]pyrene 7,8-diol 9,10-epoxide N2-deoxyguanosine adduct induces mutations at multiple sites in DNA.* J Biol Chem, 2003. **278**(17): p. 14940-8.

201. Hoffmann, G.R. and R.P. Fuchs, *Mechanisms of frameshift mutations: insight from aromatic amines.* Chem Res Toxicol, 1997. **10**(4): p. 347-59.

202. Burnouf, D., P. Koehl, and R.P.P. Fuchs, *Single Adduct Mutagenesis - Strong Effect of the Position of a Single Acetylaminofluorene Adduct within a Mutation Hot Spot.* Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(11): p. 4147-4151.

203. Lambert, I.B., R.L. Napolitano, and R.P.P. Fuchs, *Carcinogen-Induced Frameshift Mutagenesis in Repetitive Sequences.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(4): p. 1310-1314.

204. Broschard, T.H., N. Koffel-Schwartz, and R.P. Fuchs, *Sequence-dependent modulation of frameshift mutagenesis at NarI-derived mutation hot spots.* J Mol Biol, 1999. **288**(1): p. 191-9.

205. Bichara, M. and R.P. Fuchs, *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, 1985. **183**(3): p. 341-51.

206. Fuchs, R.P., N. Schwartz, and M.P. Daune, *Hot spots of frameshift mutations induced by the ultimate carcinogen N-acetoxy-N-2-acetylaminofluorene.* Nature, 1981. **294**(5842): p. 657-9.

207. Tebbs, R.S. and L.J. Romano, *Mutagenesis at a site-specifically modified NarI sequence by acetylated and deacetylated aminofluorene adducts.* Biochemistry, 1994. **33**(30): p. 8998-9006.

208. Patel, D.J., et al., *Nuclear magnetic resonance solution structures of covalent aromatic amine-DNA adducts and their mutagenic relevance.* Chem Res Toxicol, 1998. **11**(5): p. 391-407.

209. Shukla, R., et al., *How stereochemistry affects mutagenesis by N-2-deoxyguanosine adducts of 7,8-dihydroxy-9,10-epoxy-7,8,9,10-tetrahydrobenzo[alpha]pyrene: Configuration of the adduct bond is more important than those of the hydroxyl groups.* Biochemistry, 1997. **36**(43): p. 13263-13269.

210. Christner, D.F., et al., *Primer extension by various polymerases using oligonucleotide templates containing stereoisomeric benzo[a]pyrene-deoxyadenosine adducts.* Biochemistry, 1994. **33**(47): p. 14297-305.

211. Chary, P. and R.S. Lloyd, *In vitro replication by prokaryotic and eukaryotic polymerases on DNA templates containing site-specific and stereospecific benzo[a]pyrene-7,8-dihydrodiol-9,10-epoxide adducts.* Nucleic Acids Res, 1995. **23**(8): p. 1398-405.

212. Latham, G.J., et al., *The efficiency of translesion synthesis past single styrene oxide DNA adducts in vitro is polymerase-specific.* Chem Res Toxicol, 1995. **8**(3): p. 422-30.

213. Yang, W. and R. Woodgate, *What a difference a decade makes: Insights into translesion DNA synthesis.* Proc Natl Acad Sci U S A, 2007. **104**(40): p. 15591-8.

214. Hanawalt, P.C., *Preferential DNA repair in expressed genes.* Environ Health Perspect, 1987. **76**: p. 9-14.

215. Veaute, X. and R.P. Fuchs, *Greater susceptibility to mutations in lagging strand of DNA replication in Escherichia coli than in leading strand.* Science, 1993. **261**(5121): p. 598-600.

216. Fijalkowska, I.J., et al., *Unequal fidelity of leading strand and lagging strand DNA replication on the Escherichia coli chromosome.* Proc Natl Acad Sci U S A, 1998. **95**(17): p. 10020-5.

217. Radman, M., *DNA replication: one strand may be more equal.* Proc Natl Acad Sci U S A, 1998. **95**(17): p. 9718-9.

218. Milhe, C., et al., *NMR evidence of the stabilisation by the carcinogen N-2-acetylaminofluorene of a frameshift mutagenesis intermediate.* Nucleic Acids Res, 1994. **22**(22): p. 4646-52.

219. Garcia, A., I.B. Lambert, and R.P. Fuchs, *DNA adduct-induced stabilization of slipped frameshift intermediates within repetitive sequences: implications for mutagenesis.* Proc Natl Acad Sci U S A, 1993. **90**(13): p. 5989-93.

220. Loechler, E.L., *Adduct-Induced Base-Shifts - a Mechanism by Which the Adducts of Bulky Carcinogens Might Induce Mutations.* Biopolymers, 1989. **28**(5): p. 909-927.

221. Loechler, E.L., *The role of adduct site-specific mutagenesis in understanding how carcinogen-DNA adducts cause mutations: Perspective, prospects and problems.* Carcinogenesis, 1996. **17**(5): p. 895-902.

222. U.S. Department of Health and Human Services, t.R.o.C. 2001.

223. Igor B. Rogozin, Y.P., *Theoretical analysis of mutation hotspots and their DNA sequence context specificity.* Mutat Res. 544 (2003) 65-85, 2003.

224. Albertini, R.J., J.A. Nicklas, and J.P. O'Neill, *Somatic cell gene mutations in humans: biomarkers for genotoxicity.* Environ Health Perspect, 1993. **101 Suppl 3**: p. 193-201.

225. Seidman, M.M., et al., *A Shuttle Vector Plasmid for Studying Carcinogen-Induced Point Mutations in Mammalian-Cells.* Gene, 1985. **38**(1-3): p. 233-237.

226. Elion, G.B., *Symposium on immunosuppressive drugs. Biochemistry and pharmacology of purine analogues.* Fed Proc, 1967. **26**(3): p. 898-904.

227. DeMars, R., *Genetic studies of HG- PRT deficiency and the Lesch-Nyhan syndrome with cultured human cells.* Fed Proc, 1971. **30**(3): p. 944-55.

228. Elespuru, R.K., *Inducible responses to DNA damage in bacteria and mammalian cells.* Environ Mol Mutagen, 1987. **10**(1): p. 97-116.

229. Seetharam, S. and M.M. Seidman, *Modulation of Ultraviolet-Light Mutational Hotspots by Cellular Stress.* Journal of Molecular Biology, 1992. **228**(4): p. 1031-1036.

230. Boesen, J.J., et al., *Stress response induced by DNA damage leads to specific, delayed and untargeted mutations.* Mol Gen Genet, 1992. **234**(2): p. 217-27.

231. Patel, P.I., et al., *Fine structure of the human hypoxanthine phosphoribosyltransferase gene.* Mol Cell Biol, 1986. **6**(2): p. 393-403.

232. Jolly, D.J., et al., *Isolation and characterization of a full-length expressible cDNA for human hypoxanthine phosphoribosyl transferase.* Proc Natl Acad Sci U S A, 1983. **80**(2): p. 477-81.

233. Cariello, N.F. and T.R. Skopek, *In vivo mutation at the human HPRT locus.* Trends Genet, 1993. **9**(9): p. 322-6.

234. Lebkowski, J.S., J.H. Miller, and M.P. Calos, *Determination of DNA sequence changes induced by ethyl methanesulfonate in human cells, using a shuttle vector system.* Mol Cell Biol, 1986. **6**(5): p. 1838-42.

235. Levy, D.D., et al., *Sequence Specificity of Aflatoxin B(1)-Induced Mutations in a Plasmid Replicated in Xeroderma-Pigmentosum and DNA-Repair Proficient Human-Cells.* Cancer Research, 1992. **52**(20): p. 5668-5673.

236. Madzak, C., et al., *UV-induced mutations in a shuttle vector replicated in repair deficient trichothiodystrophy cells differ with those in genetically-related cancer prone xeroderma pigmentosum.* Carcinogenesis, 1993. **14**(7): p. 1255-60.

237. Seetharam, S., et al., *Ultraviolet Mutational Spectrum in a Shuttle Vector Propagated in Xeroderma-Pigmentosum Lymphoblastoid-Cells and Fibroblasts.* Mutation Research, 1991. **254**(1): p. 97-105.

238. Drinkwater, N.R. and D.K. Klinedinst, *Chemically induced mutagenesis in a shuttle vector with a low-background mutant frequency.* Proc Natl Acad Sci U S A, 1986. **83**(10): p. 3402-6.

239. DuBridge, R.B., et al., *Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system.* Mol Cell Biol, 1987. **7**(1): p. 379-87.

240. Eckert, K.A., et al., *Molecular analysis of mutations induced in human cells by N-ethyl-N-nitrosourea.* Mol Carcinog, 1988. **1**(1): p. 50-6.

241.   Ingle, C.A. and N.R. Drinkwater, *Mutational specificities of 1'-acetoxysafrole, N-benzoyloxy-N-methyl-4-aminoazobenzene, and ethyl methanesulfonate in human cells.* Mutat Res, 1989. **220**(2-3): p. 133-42.

242.   Provost, G.S., et al., *Transgenic systems for in vivo mutation analysis.* Mutat Res, 1993. **288**(1): p. 133-49.

243.   Leach, E.G., et al., *Tissue specificity of spontaneous point mutations in lambda supF transgenic mice.* Environ Mol Mutagen, 1996. **28**(4): p. 459-64.

244.   Kraemer, K.H. and M.M. Seidman, *Use of Supf, an Escherichia-Coli Tyrosine Suppressor Transfer- Rna Gene, as a Mutagenic Target in Shuttle-Vector Plasmids.* Mutation Research, 1989. **220**(2-3): p. 61-72.

245.   Seidman, M.M.a.D., K., *Analysis of genotoxic chemicals.* Molecular Biology of Toxic Response, ed. A. Puga. 1998, Philadelphia: Taylor & Francis.

246.   Munson, P.J., et al., *Test of models for the sequence specificity of UV-induced mutations in mammalian cells.* Mutat Res, 1987. **179**(1): p. 103-14.

247.   Bredberg, A., K.H. Kraemer, and M.M. Seidman, *Restricted Ultraviolet Mutational Spectrum in a Shuttle Vector Propagated in Xeroderma Pigmentosum-Cells.* Proceedings of the National Academy of Sciences of the United States of America, 1986. **83**(21): p. 8273-8277.

248.   Seetharam, S., et al., *Abnormal Ultraviolet Mutagenic Spectrum in Plasmid DNA Replicated in Cultured Fibroblasts from a Patient with the Skin-Cancer Prone Disease, Xeroderma-Pigmentosum.* Journal of Clinical Investigation, 1987. **80**(6): p. 1613-1617.

249.   McLuckie, K.I., et al., *Mutation spectra induced by alpha-acetoxytamoxifen-DNA adducts in human DNA repair proficient and deficient (xeroderma pigmentosum complementation group A) cells.* Biochemistry, 2005. **44**(22): p. 8198-205.

250.   P. D. Lewis, e.a., *Spontaneous mutation spectra in supF: comparative analysis of mammalian cell line base substitution spectra.* Mutagenesis 16, 6 (2001) 503-515, 2001.

251.   WHO, *Fact sheet No. 297: Cancer.* (February 2006).

252.   Cancer Research UK. [cited; Available from: http:\\www.cancerresearchuk.org.

253.   Peto, R., et al., *UK and USA breast cancer deaths down 25% in year 2000 at ages 20-69 years.* Lancet, 2000. **355**(9217): p. 1822.

254.   Putti, T.C., et al., *Estrogen receptor-negative breast carcinomas: a review of morphology and immunophenotypical analysis.* Mod Pathol, 2005. **18**(1): p. 26-35.

255.   Powles, T.J., *Efficacy of tamoxifen as treatment of breast cancer.* Semin Oncol, 1997. **24**(1 Suppl 1): p. S1-48-S1-54.

256.   Goss, P.E., et al., *Randomized trial of letrozole following tamoxifen as extended adjuvant therapy in receptor-positive breast cancer: updated findings from NCIC CTG MA.17.* J Natl Cancer Inst, 2005. **97**(17): p. 1262-71.

257.   *Tamoxifen for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group.* Lancet, 1998. **351**(9114): p. 1451-67.

258.   Poirier, D., et al., *Synthesis and antiestrogenic activity of diaryl thioether derivatives.* J Med Chem, 1994. **37**(8): p. 1115-25.

259.   Fisher, B., et al., *Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study.* J Natl Cancer Inst, 1998. **90**(18): p. 1371-88.

260.   Fisher, B., et al., *Endometrial cancer in tamoxifen-treated breast cancer patients: findings from the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-14.* J Natl Cancer Inst, 1994. **86**(7): p. 527-37.

261.   van Leeuwen, F.E., et al., *Risk of endometrial cancer after tamoxifen treatment of breast cancer.* Lancet, 1994. **343**(8895): p. 448-52.

262.   Fornander, T., et al., *Adjuvant tamoxifen in early breast cancer: occurrence of new primary cancers.* Lancet, 1989. **1**(8630): p. 117-20.

263.   Seoud, M.A., J. Johnson, and J.C. Weed, Jr., *Gynecologic tumors in tamoxifen-treated women with breast cancer.* Obstet Gynecol, 1993. **82**(2): p. 165-9.

264.   Bernstein, L., et al., *Tamoxifen therapy for breast cancer and endometrial cancer risk.* J Natl Cancer Inst, 1999. **91**(19): p. 1654-62.

265.   *IARC on tamoxifen.* Environ Health Perspect, 1996. **104**(7): p. 688.

266.   Lim, C.K., et al., *A comparative study of tamoxifen metabolism in female rat, mouse and human liver microsomes.* Carcinogenesis, 1994. **15**(4): p. 589-93.

267.   Boocock, D.J., et al., *Identification of human CYP forms involved in the activation of tamoxifen and irreversible binding to DNA.* Carcinogenesis, 2002. **23**(11): p. 1897-901.

268.   Jarman, M., et al., *The deuterium isotope effect for the alpha-hydroxylation of tamoxifen by rat liver microsomes accounts for the reduced genotoxicity of [D5-ethyl]tamoxifen.* Carcinogenesis, 1995. **16**(4): p. 683-8.

269.   Boocock, D.J., et al., *Alpha-hydroxytamoxifen, a genotoxic metabolite of tamoxifen in the rat: identification and quantification in vivo and in vitro.* Carcinogenesis, 1999. **20**(1): p. 153-60.

270.   Osborne, M.R., I.R. Hardcastle, and D.H. Phillips, *Minor products of reaction of DNA with alpha-acetoxytamoxifen.* Carcinogenesis, 1997. **18**(3): p. 539-43.

271.   Phillips, D.H., et al., *N-demethylation accompanies alpha-hydroxylation in the metabolic activation of tamoxifen in rat liver cells.* Carcinogenesis, 1999. **20**(10): p. 2003-9.

272.   Kim, S.Y., et al., *Alpha-hydroxylation of tamoxifen and toremifene by human and rat cytochrome P450 3A subfamily enzymes.* Chem Res Toxicol, 2003. **16**(9): p. 1138-44.

273.   Desta, Z., et al., *Comprehensive evaluation of tamoxifen sequential biotransformation by the human cytochrome P450 system in vitro: prominent roles for CYP3A and CYP2D6.* J Pharmacol Exp Ther, 2004. **310**(3): p. 1062-75.

274.   Sharma, M., et al., *Biotransformation of tamoxifen in a human endometrial explant culture model.* Chem Biol Interact, 2003. **146**(3): p. 237-49.

275.   Phillips, D.H., et al., *Organ specificity of DNA adduct formation by tamoxifen and alpha-hydroxytamoxifen in the rat: implications for understanding the mechanism(s) of tamoxifen carcinogenicity and for human risk assessment.* Mutagenesis, 2005. **20**(4): p. 297-303.

276.   Apak, T.I. and M.W. Duffel, *Interactions of the stereoisomers of alpha-hydroxytamoxifen with human hydroxysteroid sulfotransferase SULT2A1 and rat hydroxysteroid sulfotransferase STa.* Drug Metab Dispos, 2004. **32**(12): p. 1501-8.

277.   Osborne, M.R., et al., *Identification of the major tamoxifen-deoxyguanosine adduct formed in the liver DNA of rats treated with tamoxifen.* Cancer Res, 1996. **56**(1): p. 66-71.

278.   Kitagawa, M., et al., *Identification of tamoxifen-DNA adducts induced by alpha-acetoxy-N-desmethyltamoxifen.* Chem Res Toxicol, 2000. **13**(8): p. 761-9.

279.   Kim, S.Y., et al., *Formation of tamoxifen-DNA adducts via O-sulfonation, not O-acetylation, of alpha-hydroxytamoxifen in rat and human livers.* Drug Metab Dispos, 2005. **33**(11): p. 1673-8.

280.   Dasaradhi, L. and S. Shibutani, *Identification of tamoxifen-DNA adducts formed by alpha-sulfate tamoxifen and alpha-acetoxytamoxifen.* Chem Res Toxicol, 1997. **10**(2): p. 189-96.

281.   Martin, E.A., et al., *Evaluation of tamoxifen and alpha-hydroxytamoxifen 32P-post-labelled DNA adducts by the development of a novel automated on-line solid-phase extraction HPLC method.* Carcinogenesis, 1998. **19**(6): p. 1061-9.

282.   Poon, G.K., et al., *Analysis of phase I and phase II metabolites of tamoxifen in breast cancer patients.* Drug Metab Dispos, 1993. **21**(6): p. 1119-24.

283.   Poon, G.K., et al., *Identification of tamoxifen metabolites in human Hep G2 cell line, human liver homogenate, and patients on long-term therapy for breast cancer.* Drug Metab Dispos, 1995. **23**(3): p. 377-82.

284.   Ching, C.K., et al., *Tamoxifen-associated hepatocellular damage and agranulocytosis.* Lancet, 1992. **339**: p. 940.

285.   Rutqvist, L.E., et al., *Adjuvant tamoxifen therapy for early stage breast cancer and second primary malignancies. Stockholm Breast Cancer Study Group.* J Natl Cancer Inst, 1995. **87**(9): p. 645-51.

286.   Shibutani, S., et al., *Sulfation of alpha-hydroxytamoxifen catalyzed by human hydroxysteroid sulfotransferase results in tamoxifen-DNA adducts.* Carcinogenesis, 1998. **19**(11): p. 2007-11.

287.   Boocock, D.J., et al., *Major inter-species differences in the rates of O-sulphonation and O-glucuronylation of alpha-hydroxytamoxifen in vitro: a metabolic disparity protecting human liver from the formation of tamoxifen-DNA adducts.* Carcinogenesis, 2000. **21**(10): p. 1851-8.

288.   Williams, G.M., et al., *The triphenylethylene drug tamoxifen is a strong liver carcinogen in the rat.* Carcinogenesis, 1993. **14**(2): p. 315-7.

289.   Greaves, P., et al., *Two-year carcinogenicity study of tamoxifen in Alderley Park Wistar-derived rats.* Cancer Res, 1993. **53**(17): p. 3919-24.

290.   Hard, G.C., et al., *Major difference in the hepatocarcinogenicity and DNA adduct forming ability between toremifene and tamoxifen in female Crl:CD(BR) rats.* Cancer Res, 1993. **53**(19): p. 4534-41.

291.   Han, X.L., Liehr, J. G., *Induction of covalent DNA adducts in rodents by tamoxifen.* Cancer Res, 1992. **52**(5): p. 1360-3.

292.   Divi, R.L., et al., *Tamoxifen-DNA adduct formation in rat liver determined by immunoassay and 32P-postlabeling.* Cancer Res, 1999. **59**(19): p. 4829-33.

293.   Carthew, P., et al., *Tamoxifen induces short-term cumulative DNA damage and liver tumors in rats: promotion by phenobarbital.* Cancer Res, 1995. **55**(3): p. 544-7.

294.   Dragan, Y.P., et al., *Studies of tamoxifen as a promoter of hepatocarcinogenesis in female Fischer F344 rats.* Breast Cancer Res Treat, 1994. **31**(1): p. 11-25.

295.   Martin, E.A., et al., *Investigation of the formation and accumulation of liver DNA adducts in mice chronically exposed to tamoxifen.* Carcinogenesis, 1997. **18**(11): p. 2209-15.

296. Tucker, M.J., Adam, H.K. and Patterson, J.S. , *Tamoxifen*. Safety Testing of New Drugs, ed. D.R. Lawrence, McClean, A.E.M. and Wetherall, M. 1984, London: Academic Press. 125-161.

297. Newbold, R.R., et al., *Uterine carcinoma in mice treated neonatally with tamoxifen*. Carcinogenesis, 1997. **18**(12): p. 2293-8.

298. Diwan, B.A., Anderson, L. M., Ward, J. M., *Proliferative lesions of oviduct and uterus in CD-1 mice exposed prenatally to tamoxifen*. Carcinogenesis, 1997. **18**(10): p. 2009-14.

299. Bergman, L., et al., *Risk and prognosis of endometrial cancer after tamoxifen for breast cancer. Comprehensive Cancer Centres' ALERT Group. Assessment of Liver and Endometrial cancer Risk following Tamoxifen.* Lancet, 2000. **356**(9233): p. 881-7.

300. Mignotte, H., et al., *Iatrogenic risks of endometrial carcinoma after treatment for breast cancer in a large French case-control study. Federation Nationale des Centres de Lutte Contre le Cancer (FNCLCC)*. Int J Cancer, 1998. **76**(3): p. 325-30.

301. Magriples, U., et al., *High-grade endometrial carcinoma in tamoxifen-treated breast cancer patients*. J Clin Oncol, 1993. **11**(3): p. 485-90.

302. Dallenbach-Hellweg, G. and U. Hahn, *Mucinous and clear cell adenocarcinomas of the endometrium in patients receiving antiestrogens (tamoxifen) and gestagens*. Int J Gynecol Pathol, 1995. **14**(1): p. 7-15.

303. Kedar, R.P., et al., *Effects of tamoxifen on uterus and ovaries of postmenopausal women in a randomised breast cancer prevention trial.* Lancet, 1994. **343**(8909): p. 1318-21.

304. Umemoto, A., et al., *Identification and quantification of tamoxifen-DNA adducts in the liver of rats and mice*. Chem Res Toxicol, 2001. **14**(8): p. 1006-13.

305. Rajaniemi, H., et al., *Identification of the major tamoxifen-DNA adducts in rat liver by mass spectroscopy*. Carcinogenesis, 1999. **20**(2): p. 305-9.

306. Schild, L.J., et al., *Formation of tamoxifen-DNA adducts in multiple organs of adult female cynomolgus monkeys dosed with tamoxifen for 30 days*. Cancer Res, 2003. **63**(18): p. 5999-6003.

307. Shibutani, S., et al., *Identification of tamoxifen-DNA adducts in monkeys treated with tamoxifen*. Cancer Res, 2003. **63**(15): p. 4402-6.

308. Umemoto, A., et al., *Identification of hepatic tamoxifen-DNA adducts in mice: alpha-(N(2)-deoxyguanosinyl)tamoxifen and alpha-(N(2)-deoxyguanosinyl)tamoxifen N-oxide*. Carcinogenesis, 2000. **21**(9): p. 1737-44.

309. Mani, C., Kupfer, D., *Cytochrome P-450-mediated activation and irreversible binding of the antiestrogen tamoxifen to proteins in rat and human liver: possible involvement of flavin-containing monooxygenases in tamoxifen activation*. Cancer Res, 1991. **51**(22): p. 6052-8.

310. Moorthy, B., et al., *Tamoxifen metabolic activation: comparison of DNA adducts formed by microsomal and chemical activation of tamoxifen and 4-hydroxytamoxifen with DNA adducts formed in vivo*. Cancer Res, 1996. **56**(1): p. 53-7.

311. Marques, M.M., Beland, F. A., *Identification of tamoxifen-DNA adducts formed by 4-hydroxytamoxifen quinone methide*. Carcinogenesis, 1997. **18**(10): p. 1949-54.

312.   Beland, F.A., McDaniel, L. P., Marques, M. M., *Comparison of the DNA adducts formed by tamoxifen and 4-hydroxytamoxifen in vivo.* Carcinogenesis, 1999. **20**(3): p. 471-7.

313.   Li, D., et al., *Effects of chronic administration of tamoxifen and toremifene on DNA adducts in rat liver, kidney, and uterus.* Cancer Res, 1997. **57**(8): p. 1438-41.

314.   da Costa, G.G., et al., *DNA adduct formation and mutant induction in Sprague-Dawley rats treated with tamoxifen and its derivatives.* Carcinogenesis, 2001. **22**(8): p. 1307-15.

315.   Carthew, P., et al., *Tamoxifen induces endometrial and vaginal cancer in rats in the absence of endometrial hyperplasia.* Carcinogenesis, 2000. **21**(4): p. 793-7.

316.   White, I.N., et al., *Genotoxic potential of tamoxifen and analogues in female Fischer F344/n rats, DBA/2 and C57BL/6 mice and in human MCL-5 cells.* Carcinogenesis, 1992. **13**(12): p. 2197-203.

317.   Pathak, D.N., Pongracz, K., Bodell, W. J., *Activation of 4-hydroxytamoxifen and the tamoxifen derivative metabolite E by uterine peroxidase to form DNA adducts: comparison with DNA adducts formed in the uterus of Sprague-Dawley rats treated with tamoxifen.* Carcinogenesis, 1996. **17**(9): p. 1785-90.

318.   White, I.N., et al., *Comparisons of the binding of [14C]radiolabelled tamoxifen or toremifene to rat DNA using accelerator mass spectrometry.* Chem Biol Interact, 1997. **106**(2): p. 149-60.

319.   Hemminki, K., et al., *Tamoxifen-induced DNA adducts in endometrial samples from breast cancer patients.* Cancer Res, 1996. **56**(19): p. 4374-7.

320.   Shibutani, S., et al., *Tamoxifen-DNA adducts detected in the endometrium of women treated with tamoxifen.* Chem Res Toxicol, 1999. **12**(7): p. 646-53.

321.   Shibutani, S., et al., *Identification of tamoxifen-DNA adducts in the endometrium of women treated with tamoxifen.* Carcinogenesis, 2000. **21**(8): p. 1461-7.

322.   Martin, E.A., et al., *Tamoxifen DNA damage detected in human endometrium using accelerator mass spectrometry.* Cancer Res, 2003. **63**(23): p. 8461-5.

323.   Carmichael, P.L., et al., *Lack of genotoxicity of tamoxifen in human endometrium.* Cancer Res, 1996. **56**(7): p. 1475-9.

324.   Carmichael, P.L., et al., *Lack of evidence from HPLC 32P-post-labelling for tamoxifen-DNA adducts in the human endometrium.* Carcinogenesis, 1999. **20**(2): p. 339-42.

325.   Beland, F.A., et al., *Electrospray ionization-tandem mass spectrometry and 32P-postlabeling analyses of tamoxifen-DNA adducts in humans.* J Natl Cancer Inst, 2004. **96**(14): p. 1099-104.

326.   Phillips, D.H., et al., *Tamoxifen does not form detectable DNA adducts in white blood cells of breast cancer patients.* Carcinogenesis, 1996. **17**(5): p. 1149-52.

327.   Bartsch, H., et al., *Lack of evidence for tamoxifen- and toremifene-DNA adducts in lymphocytes of treated patients.* Carcinogenesis, 2000. **21**(4): p. 845-7.

328.   Umemoto, A., et al., *Determination of tamoxifen--DNA adducts in leukocytes from breast cancer patients treated with tamoxifen.* Chem Res Toxicol, 2004. **17**(12): p. 1577-83.

329.   Beland, F.A., et al., *Analysis of tamoxifen-DNA adducts in endometrial explants by MS and 32P-postlabeling.* Biochem Biophys Res Commun, 2004. **320**(2): p. 297-302.

330.   Sharma, M., et al., *Antioxidant inhibits tamoxifen-DNA adducts in endometrial explant culture.* Biochem Biophys Res Commun, 2003. **307**(1): p. 157-64.

331.   Brown, K., Carmichael, P. L., *Correspondence regarding M. Sharma et al., "Antioxidant inhibits tamoxifen-DNA adducts in endometrial explant culture".* Biochem Biophys Res Commun, 2003. **310**(4): p. 1039.

332.   Sharma, M., *Correspondence regarding October 31, 2003, Breakthroughs and Views by K. Brown and P.L. Carmichael.* Biochem Biophys Res Commun, 2004. **313**(1): p. 1-2.

333.   Yamamoto, Y., et al., *The tamoxifen-responsive estrogen receptor alpha mutant D351Y shows reduced tamoxifen-dependent interaction with corepressor complexes.* J Biol Chem, 2001. **276**(46): p. 42684-91.

334.   Hall, J.M., Couse, J. F., Korach, K. S., *The multifaceted mechanisms of estradiol and estrogen receptor signaling.* J Biol Chem, 2001. **276**(40): p. 36869-72.

335.   Shang, Y., Brown, M., *Molecular determinants for the tissue specificity of SERMs.* Science, 2002. **295**(5564): p. 2465-8.

336.   Saville, B., et al., *Ligand-, cell-, and estrogen receptor subtype (alpha/beta)-dependent activation at GC-rich (Sp1) promoter elements.* J Biol Chem, 2000. **275**(8): p. 5379-87.

337.   Castro-Rivera, E. and S. Safe, *17 beta-estradiol- and 4-hydroxytamoxifen-induced transactivation in breast, endometrial and liver cancer cells is dependent on ER-subtype, cell and promoter context.* J Steroid Biochem Mol Biol, 2003. **84**(1): p. 23-31.

338.   Hall, J.M. and D.P. McDonnell, *The estrogen receptor beta-isoform (ERbeta) of the human estrogen receptor modulates ERalpha transcriptional activity and is a key regulator of the cellular response to estrogens and antiestrogens.* Endocrinology, 1999. **140**(12): p. 5566-78.

339.   Monje, P. and R. Boland, *Characterization of membrane estrogen binding proteins from rabbit uterus.* Mol Cell Endocrinol, 1999. **147**(1-2): p. 75-84.

340.   Bombail, V., J.G. Moggs, and G. Orphanides, *Perturbation of epigenetic status by toxicants.* Toxicol Lett, 2004. **149**(1-3): p. 51-8.

341.   Kamiya, K., et al., *Evidence that carcinogenesis involves an imbalance between epigenetic high-frequency initiation and suppression of promotion.* Proc Natl Acad Sci U S A, 1995. **92**(5): p. 1332-6.

342.   Jaffe, L.F., *Epigenetic theories of cancer initiation.* Adv Cancer Res, 2003. **90**: p. 209-30.

343.   Feinberg, A.P., *The epigenetics of cancer etiology.* Semin Cancer Biol, 2004. **14**(6): p. 427-32.

344.   Karpinets, T.V. and B.D. Foy, *Tumorigenesis: the adaptation of mammalian cells to sustained stress environment by epigenetic alterations and succeeding matched mutations.* Carcinogenesis, 2005. **26**(8): p. 1323-34.

345.   Ushijima, T. and E. Okochi-Takada, *Aberrant methylations in cancer cells: where do they come from?* Cancer Sci, 2005. **96**(4): p. 206-11.

346.   Terashima, I., N. Suzuki, and S. Shibutani, *Mutagenic potential of alpha-(N2-deoxyguanosinyl)tamoxifen lesions, the major DNA adducts detected in endometrial tissues of patients treated with tamoxifen.* Cancer Res, 1999. **59**(9): p. 2091-5.

347.   Shibutani, S. and L. Dasaradhi, *Miscoding potential of tamoxifen-derived DNA adducts: alpha-(N2-deoxyguanosinyl)tamoxifen.* Biochemistry, 1997. **36**(42): p. 13010-7.

348.   Davies, R., et al., *Tamoxifen causes gene mutations in the livers of lambda/lacI transgenic rats.* Cancer Res, 1997. **57**(7): p. 1288-93.

349.   Davies, R., et al., *Tamoxifen induces G:C-->T:A mutations in the cII gene in the liver of lambda/lacI transgenic rats but not at 5'-CpG-3' dinucleotide sequences as found in the lacI transgene.* Carcinogenesis, 1999. **20**(7): p. 1351-6.

350.   McLuckie, K.I.E., et al., *DNA adducts formed from 4-hydroxytamoxifen are more mutagenic than those formed by alpha-acetoxytamoxifen in a shuttle vector target gene replicated in human Ad293 cells.* Biochemistry, 2002. **41**(28): p. 8899-8906.

351.   McLuckie, K.I., et al., *Mutation spectra induced by alpha-acetoxytamoxifen-DNA adducts in human DNA repair proficient and deficient (xeroderma pigmentosum complementation group A) cells.* Biochemistry, 2005. **44**(22): p. 8198-205.

352.   Yasui, M., et al., *Translesion synthesis past tamoxifen-derived DNA adducts by human DNA polymerases eta and kappa.* Biochemistry, 2006. **45**(39): p. 12167-74.

353.   Shibutani, S., et al., *Excision of tamoxifen-DNA adducts by the human nucleotide excision repair system.* Cancer Res, 2000. **60**(10): p. 2607-10.

354.   Kim, S.Y., et al., *Inefficient repair of tamoxifen-DNA adducts in rats and mice.* Drug Metab Dispos, 2006. **34**(2): p. 311-7.

355.   Lewis, P.D. and J.M. Parry, *In silico p53 mutation hotspots in lung cancer.* Carcinogenesis, 2004. **25**(7): p. 1099-107.

356.   Tornaletti, S., S. Bates, and G.P. Pfeifer, *A high-resolution analysis of chromatin structure along p53 sequences.* Mol Carcinog, 1996. **17**(4): p. 192-201.

357.   Denissenko, M.F., et al., *Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53.* Science, 1996. **274**(5286): p. 430-2.

358.   Denissenko, M.F., et al., *Cytosine methylation determines hot spots of DNA damage in the human P53 gene.* Proc Natl Acad Sci U S A, 1997. **94**(8): p. 3893-8.

359.   http://ecacc.org.uk/.

360.   Gravanis, A. and E. Gurpide, *Effects of estradiol on deoxyribonucleic acid polymerase alpha activity in the Ishikawa human endometrial adenocarcinoma cell line.* J Clin Endocrinol Metab, 1986. **63**(2): p. 356-9.

361.   Casadaban, M.J. and S.N. Cohen, *Analysis of gene control signals by DNA fusion and cloning in Escherichia coli.* J Mol Biol, 1980. **138**(2): p. 179-207.

362.   Foster, A.B., et al., *Hydroxy derivatives of tamoxifen.* J Med Chem, 1985. **28**(10): p. 1491-7.

363.   Parris, C.N. and M.M. Seidman, *A Signature Element Distinguishes Sibling and Independent Mutations in a Shuttle Vector Plasmid.* Gene, 1992. **117**(1): p. 1-5.

364.   Brown, K., et al., *Further characterization of the DNA adducts formed in rat liver after the administration of tamoxifen, N-desmethyltamoxifen or N, N-didesmethyltamoxifen.* Carcinogenesis, 1999. **20**(10): p. 2011-6.

365. Lowes, D.A., et al., *Site-specific tamoxifen-DNA adduct formation: lack of correlation with mutational ability in Escherichia coli.* Biochemistry, 1999. **38**(34): p. 10989-96.

366. Akman, S.A., F. Sander, and K. Garbutt, *In vivo mutagenesis of the reporter plasmid pSP189 induced by exposure of host Ad293 cells to activated polymorphonuclear leukocytes.* Carcinogenesis, 1996. **17**(10): p. 2137-41.

367. Niedernhofer, L.J., et al., *Malondialdehyde, a product of lipid peroxidation, is mutagenic in human cells.* J Biol Chem, 2003. **278**(33): p. 31426-33.

368. Cariello, N.F., et al., *Computer program for the analysis of mutational spectra: application to p53 mutations.* Carcinogenesis, 1994. **15**(10): p. 2281-5.

369. Courtemanche, C. and A. Anderson, *Multiple mutations in a shuttle vector modified by ultraviolet irradiation, (+/-)-7 beta,8 alpha-dihydroxy-9 alpha,10 alpha-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene, and aflatoxin B(1) have different properties than single mutations and may be generated during translesion synthesis.* Mutat Res, 1999. **430**(1): p. 23-36.

370. Randerath, K., et al., *Tamoxifen: evidence by 32P-postlabeling and use of metabolic inhibitors for two distinct pathways leading to mouse hepatic DNA adduct formation and identification of 4-hydroxytamoxifen as a proximate metabolite.* Carcinogenesis, 1994. **15**(10): p. 2087-94.

371. Bigger, C.A., et al., *Mutagenic specificities of four stereoisomeric benzo[c]phenanthrene dihydrodiol epoxides.* Proc Natl Acad Sci U S A, 1992. **89**(1): p. 368-72.

372. Boldt, J., et al., *Kinds of mutations found when a shuttle vector containing adducts of 1,6-dinitropyrene replicates in human cells.* Carcinogenesis, 1991. **12**(1): p. 119-26.

373. Yang, J.L., V.M. Maher, and J.J. McCormick, *Kinds of mutations formed when a shuttle vector containing adducts of (+/-)-7 beta, 8 alpha-dihydroxy-9 alpha, 10 alpha-epoxy-7,8,9, 10-tetrahydrobenzo[a]pyrene replicates in human cells.* Proc Natl Acad Sci U S A, 1987. **84**(11): p. 3787-91.

374. Routledge, M.N., et al., *Presence of benzo[a]pyrene diol epoxide adducts in target DNA leads to an increase in UV-induced DNA single strand breaks and supF gene mutations.* Carcinogenesis, 2001. **22**(8): p. 1231-1238.

375. Bigger, C.A., et al., *Mutational specificity of the anti 1,2-dihydrodiol 3,4-epoxide of 5-methylchrysene.* Carcinogenesis, 1990. **11**(12): p. 2263-5.

376. Juedes, M.J. and G.N. Wogan, *Peroxynitrite-induced mutation spectra of pSP189 following replication in bacteria and in human cells.* Mutat Res, 1996. **349**(1): p. 51-61.

377. Sikpi, M.O., et al., *Dependence of the mutation spectrum in a shuttle plasmid replicated in human lymphoblasts on dose of gamma radiation.* Int J Radiat Biol, 1991. **59**(5): p. 1115-26.

378. Keyse, S.M., F. Amaudruz, and R.M. Tyrrell, *Determination of the spectrum of mutations induced by defined-wavelength solar UVB (313-nm) radiation in mammalian cells by use of a shuttle vector.* Mol Cell Biol, 1988. **8**(12): p. 5425-31.

379. Strauss, B.S., *The "A" rule revisited: polymerases as determinants of mutational specificity.* DNA Repair (Amst), 2002. **1**(2): p. 125-35.

380. Kalnik, M.W., et al., *NMR studies of abasic sites in DNA duplexes: deoxyadenosine stacks into the helix opposite the cyclic analogue of 2-deoxyribose.* Biochemistry, 1988. **27**(3): p. 924-31.

381. Rachofsky, E.L., et al., *Conformation and dynamics of abasic sites in DNA investigated by time-resolved fluorescence of 2-aminopurine.* Biochemistry, 2001. **40**(4): p. 957-67.

382. Strauss, B., et al., *The role of DNA polymerase in base substitution mutagenesis on non-instructional templates.* Biochimie, 1982. **64**(8-9): p. 829-38.

383. Lewis, P.D., et al., *Spontaneous mutation spectra in supF: comparative analysis of mammalian cell line base substitution spectra.* Mutagenesis, 2001. **16**(6): p. 503-15.

384. Gordenin, D.A. and M.A. Resnick, *Yeast ARMs (DNA at-risk motifs) can reveal sources of genome instability.* Mutat Res, 1998. **400**(1-2): p. 45-58.

385. Albertini, A.M., et al., *On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions.* Cell, 1982. **29**(2): p. 319-28.

386. Shimizu, H., et al., *Short-homology-independent illegitimate recombination in Escherichia coli: distinct mechanism from short-homology-dependent illegitimate recombination.* J Mol Biol, 1997. **266**(2): p. 297-305.

387. Yamaguchi, H., et al., *A hotspot of spontaneous and UV-induced illegitimate recombination during formation of lambda bio transducing phage.* Mol Gen Genet, 1995. **248**(6): p. 637-43.

388. Ukita, T. and H. Ikeda, *Role of the recJ gene product in UV-induced illegitimate recombination at the hotspot.* J Bacteriol, 1996. **178**(8): p. 2362-7.

389. Huang, X., et al., *Conformational changes of a benzo[a]pyrene diol epoxide-N(2)-dG adduct induced by a 5'-flanking 5-methyl-substituted cytosine in a (Me)CG double-stranded oligonucleotide sequence context.* Chem Res Toxicol, 2002. **15**(3): p. 438-44.

390. Weisenberger, D.J. and L.J. Romano, *Cytosine methylation in a CpG sequence leads to enhanced reactivity with Benzo[a]pyrene diol epoxide that correlates with a conformational change.* J Biol Chem, 1999. **274**(34): p. 23948-55.

391. Seo, K.Y., S.A. Jelinsky, and E.L. Loechler, *Factors that influence the mutagenic patterns of DNA adducts from chemical carcinogens.* Mutation Research-Reviews in Mutation Research, 2000. **463**(3): p. 215-246.

392. Davies, A.M., et al., *Peroxidase activation of 4-hydroxytamoxifen to free radicals detected by EPR spectroscopy.* Free Radic Biol Med, 1997. **22**(3): p. 423-31.

393. Jones, R.M., Z.X. Yuan, and C.K. Lim, *Tamoxifen metabolism in rat liver microsomes: identification of a dimeric metabolite derived from free radical intermediates by liquid chromatography/mass spectrometry.* Rapid Commun Mass Spectrom, 1999. **13**(4): p. 211-5.

394. Dong, H., et al., *Mutagenic potential of benzo[a]pyrene-derived DNA adducts positioned in codon 273 of the human P53 gene.* Biochemistry, 2004. **43**(50): p. 15922-8.

395. Denissenko, M.F., et al., *Slow repair of bulky DNA adducts along the nontranscribed strand of the human p53 gene may explain the strand bias of transversion mutations in cancers.* Oncogene, 1998. **16**(10): p. 1241-7.

396. Vancutsem, P.M., P. Lazarus, and G.M. Williams, *Frequent and specific mutations of the rat p53 gene in hepatocarcinomas induced by tamoxifen.* Cancer Res, 1994. **54**(14): p. 3864-7.

397. IARC.

398. Druckrey, H., et al., *[Organotropic carcinogenic effects of 65 various N-nitroso- compounds on BD rats]*. Z Krebsforsch, 1967. **69**(2): p. 103-201.

399. Taverna, P. and B. Sedgwick, *Generation of an endogenous DNA-methylating agent by nitrosation in Escherichia coli*. J Bacteriol, 1996. **178**(17): p. 5105-11.

400. Shuker, D.E. and G.P. Margison, *Nitrosated glycine derivatives as a potential source of O6-methylguanine in DNA*. Cancer Res, 1997. **57**(3): p. 366-9.

401. Sedgwick, B., *Nitrosated peptides and polyamines as endogenous mutagens in O6-alkylguanine-DNA alkyltransferase deficient cells*. Carcinogenesis, 1997. **18**(8): p. 1561-7.

402. Loveless, A., *Possible relevance of O-6 alkylation of deoxyguanosine to the mutagenicity and carcinogenicity of nitrosamines and nitrosamides*. Nature, 1969. **223**(5202): p. 206-7.

403. Debiak, M., T. Nikolova, and B. Kaina, *Loss of ATM sensitizes against O6-methylguanine triggered apoptosis, SCEs and chromosomal aberrations*. DNA Repair (Amst), 2004. **3**(4): p. 359-68.

404. Beranek, D.T., *Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents*. Mutat Res, 1990. **231**(1): p. 11-30.

405. Kyrtopoulos, S.A., *DNA adducts in humans after exposure to methylating agents*. Mutat Res, 1998. **405**(2): p. 135-43.

406. Povey, A.C., *DNA adducts: endogenous and induced*. Toxicol Pathol, 2000. **28**(3): p. 405-14.

407. Szyfter, K., et al., *Tobacco smoke-associated N7-alkylguanine in DNA of larynx tissue and leucocytes*. Carcinogenesis, 1996. **17**(3): p. 501-6.

408. Foiles, P.G., et al., *Detection of O6-methyldeoxyguanosine in human placental DNA*. Cancer Res, 1988. **48**(15): p. 4184-8.

409. Umbenhauer, D., et al., *O(6)-methyldeoxyguanosine in oesophageal DNA among individuals at high risk of oesophageal cancer*. Int J Cancer, 1985. **36**(6): p. 661-5.

410. *O6-methylguanine in blood leucocyte DNA: an association with the geographic prevalence of gastric cancer and with low levels of serum pepsinogen A, a marker of severe chronic atrophic gastritis. The EUROGAST Study Group*. Carcinogenesis, 1994. **15**(9): p. 1815-20.

411. Wyatt, M.D. and D.L. Pittman, *Methylating agents and DNA repair responses: Methylated bases and sources of strand breaks*. Chem Res Toxicol, 2006. **19**(12): p. 1580-94.

412. Snow, E.T., R.S. Foote, and S. Mitra, *Base-pairing properties of O6-methylguanine in template DNA during in vitro DNA replication*. J Biol Chem, 1984. **259**(13): p. 8095-100.

413. Patel, D.J., et al., *Structural studies of the O6meG.C interaction in the d(C-G-C-G-A-A-T-T-C-O6meG-C-G) duplex*. Biochemistry, 1986. **25**(5): p. 1027-36.

414. Eadie, J.S., et al., *Mechanism of mutagenesis by O6-methylguanine*. Nature, 1984. **308**(5955): p. 201-3.

415. Abbott, P.J. and R. Saffhill, *DNA synthesis with methylated poly(dC-dG) templates. Evidence for a competitive nature to miscoding by O(6)-methylguanine*. Biochim Biophys Acta, 1979. **562**(1): p. 51-61.

416. Loechler, E.L., C.L. Green, and J.M. Essigmann, *Invivo Mutagenesis by O-6-Methylguanine Built into a Unique Site in a Viral Genome*. Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences, 1984. **81**(20): p. 6271-6275.

417.   Altshuler, K.B., C.S. Hodes, and J.M. Essigmann, *Intrachromosomal probes for mutagenesis by alkylated DNA bases replicated in mammalian cells: A comparison of the mutagenicities of O-4-methylthymine and O-6-methylguanine in cells with different DNA repair backgrounds.* Chemical Research in Toxicology, 1996. **9**(6): p. 980-987.

418.   Patel, D.J., et al., *Structural studies of the O6meG.T interaction in the d(C-G-T-G-A-A-T-T-C-O6meG-C-G) duplex.* Biochemistry, 1986. **25**(5): p. 1036-42.

419.   Singer, B., et al., *Effect of 3' flanking neighbors on kinetics of pairing of dCTP or dTTP opposite O6-methylguanine in a defined primed oligonucleotide when Escherichia coli DNA polymerase I is used.* Proc Natl Acad Sci U S A, 1989. **86**(21): p. 8271-4.

420.   Larson, K., et al., *Methylation-induced blocks to in vitro DNA replication.* Mutat Res, 1985. **150**(1-2): p. 77-84.

421.   Dosanjh, M.K., et al., *Kinetics of extension of O6-methylguanine paired with cytosine or thymine in defined oligonucleotide sequences.* Biochemistry, 1991. **30**(49): p. 11595-9.

422.   Pegg, A.E. and T.L. Byers, *Repair of DNA containing O6-alkylguanine.* Faseb J, 1992. **6**(6): p. 2302-10.

423.   Lindahl, T., et al., *Regulation and expression of the adaptive response to alkylating agents.* Annu Rev Biochem, 1988. **57**: p. 133-57.

424.   Samson, L., J. Thomale, and M.F. Rajewsky, *Alternative pathways for the in vivo repair of O6-alkylguanine and O4-alkylthymine in Escherichia coli: the adaptive response and nucleotide excision repair.* Embo J, 1988. **7**(7): p. 2261-7.

425.   Margison, G.P., M.F. Santibanez Koref, and A.C. Povey, *Mechanisms of carcinogenicity/chemotherapy by O6-methylguanine.* Mutagenesis, 2002. **17**(6): p. 483-7.

426.   Karran, P. and M.G. Marinus, *Mismatch correction at O6-methylguanine residues in E. coli DNA.* Nature, 1982. **296**(5860): p. 868-9.

427.   Branch, P., et al., *Defective mismatch binding and a mutator phenotype in cells tolerant to DNA damage.* Nature, 1993. **362**(6421): p. 652-4.

428.   Kat, A., et al., *An alkylation-tolerant, mutator human cell line is deficient in strand-specific mismatch repair.* Proc Natl Acad Sci U S A, 1993. **90**(14): p. 6424-8.

429.   Pauly, G.T., S.H. Hughes, and R.C. Moschel, *Response of Repair-Competent and Repair-Deficient Escherichia- Coli to 3 O-6-Substituted Guanines and Involvement of Methyl- Directed Mismatch Repair in the Processing of O-6-Methylguanine Residues.* Biochemistry, 1994. **33**(31): p. 9169-9177.

430.   Pauly, G.T., S.H. Hughes, and R.C. Moschel, *Mutagenesis in Escherichia-Coli by 3 O-6-Substituted Guanines in Double-Stranded or Gapped Plasmids.* Biochemistry, 1995. **34**(27): p. 8924-8930.

431.   Pauly, G.T., S.H. Hughes, and R.C. Moschel, *Comparison of mutagenesis by O-6-methyl- and O-6-ethylguanine and O-4-methylthymine in Escherichia coli using double-stranded and gapped plasmids.* Carcinogenesis, 1998. **19**(3): p. 457-461.

432.   Nowosielska, A. and M.G. Marinus, *DNA mismatch repair-induced double-strand breaks.* DNA Repair (Amst), 2007.

433.   Goldmacher, V.S., R.A. Cuzick, Jr., and W.G. Thilly, *Isolation and partial characterization of human cell mutants differing in sensitivity to killing and*

mutation by methylnitrosourea and N-methyl-N'-nitro-N-nitrosoguanidine. J Biol Chem, 1986. **261**(27): p. 12462-71.

434.   Kaina, B., et al., *Chromosomal instability, reproductive cell death and apoptosis induced by O6-methylguanine in Mex-, Mex+ and methylation-tolerant mismatch repair compromised cells: facts and models.* Mutat Res, 1997. **381**(2): p. 227-41.

435.   Tominaga, Y., et al., *Alkylation-induced apoptosis of embryonic stem cells in which the gene for DNA-repair, methyltransferase, had been disrupted by gene targeting.* Carcinogenesis, 1997. **18**(5): p. 889-96.

436.   Meikrantz, W., et al., *O6-alkylguanine DNA lesions trigger apoptosis.* Carcinogenesis, 1998. **19**(2): p. 369-72.

437.   Ochs, K. and B. Kaina, *Apoptosis induced by DNA damage O6-methylguanine is Bcl-2 and caspase-9/3 regulated and Fas/caspase-8 independent.* Cancer Res, 2000. **60**(20): p. 5815-24.

438.   Roos, W., M. Baumgartner, and B. Kaina, *Apoptosis triggered by DNA damage O6-methylguanine in human lymphocytes requires DNA replication and is mediated by p53 and Fas/CD95/Apo-1.* Oncogene, 2004. **23**(2): p. 359-67.

439.   Hickman, M.J. and L.D. Samson, *Role of DNA mismatch repair and p53 in signaling induction of apoptosis by alkylating agents.* Proc Natl Acad Sci U S A, 1999. **96**(19): p. 10764-9.

440.   Toft, N.J., et al., *Msh2 status modulates both apoptosis and mutation frequency in the murine small intestine.* Proc Natl Acad Sci U S A, 1999. **96**(7): p. 3911-5.

441.   Ward, J.F., *Some biochemical consequences of the spatial distribution of ionizing radiation-produced free radicals.* Radiat Res, 1981. **86**(2): p. 185-95.

442.   Errol C. Friedberg, G.C.W., Wolfram Siede, *DNA repair and mutagenesis.* 1995.

443.   Singer B., G.D., *Molecular biology of mutagens and carcinogens.* 1983, NY.

444.   Teebor, G.W., R.J. Boorstein, and J. Cadet, *The repairability of oxidative free radical mediated damage to DNA: a review.* Int J Radiat Biol, 1988. **54**(2): p. 131-50.

445.   Basu, A.K. and J.M. Essigmann, *Site-Specifically Modified Oligodeoxynucleotides as Probes for the Structural and Biological Effects of DNA-Damaging Agents.* Chemical Research in Toxicology, 1988. **1**(1): p. 1-18.

446.   Loechler, E.L., *The role of adduct site-specific mutagenesis in understanding how carcinogen-DNA adducts cause mutations: Perspective, prospects and problems.* Carcinogenesis, 1996. **17**(5): p. 895-902.

447.   Singer, B. and J.M. Essigmann, *Site-Specific Mutagenesis - Retrospective and Prospective.* Carcinogenesis, 1991. **12**(6): p. 949-955.

448.   Inouye S., I.M.N., *Synthesis and Applications of DNA and RNA, 181-206.* 1987, New York Academic Press

449.   Chary, P., et al., *In vivo and in vitro replication consequences of stereoisomeric benzo[a]pyrene-7,8-dihydrodiol 9,10-epoxide adducts on adenine N6 at the second position of N-ras codon 61.* J Biol Chem, 1995. **270**(10): p. 4990-5000.

450.   Fink, S.P., G.R. Reddy, and L.J. Marnett, *Mutagenicity in Escherichia coli of the major DNA adduct derived from the endogenous mutagen malondialdehyde.* Proc Natl Acad Sci U S A, 1997. **94**(16): p. 8652-7.

451.   Syvanen, A.C., *From gels to chips: "minisequencing" primer extension for analysis of point mutations and single nucleotide polymorphisms.* Hum Mutat, 1999. **13**(1): p. 1-10.

452.   Haff, L.A. and I.P. Smirnov, *Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry.* Genome Res, 1997. **7**(4): p. 378-88.

453.   Delaney, J.C. and J.M. Essigmann, *Context-dependent mutagenesis by DNA lesions.* Chemistry & Biology, 1999. **6**(10): p. 743-753.

454.   Delaney, J.C. and J.M. Essigmann, *Effect of sequence context on O-6-methylguanine repair and replication in vivo.* Biochemistry, 2001. **40**(49): p. 14968-14975.

455.   Henderson, P.T., et al., *Oxidation of 7,8-dihydro-8-oxoguanine affords lesions that are potent sources of replication errors in vivo.* Biochemistry, 2002. **41**(3): p. 914-921.

456.   Pauly, G.T., L.A. Peterson, and R.C. Moschel, *Mutagenesis by O-6-[4-oxo-4-(3-pyridyl)butyl]guanine in Escherichia coli and human cells.* Chemical Research in Toxicology, 2002. **15**(2): p. 165-169.

457.   Avkin, S., et al., *Quantitative measurement of translesion replication in human cells: evidence for bypass of abasic sites by a replicative DNA polymerase.* Proc Natl Acad Sci U S A, 2002. **99**(6): p. 3764-9.

458.   Prakash, S., R.E. Johnson, and L. Prakash, *Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function.* Annu Rev Biochem, 2005. **74**: p. 317-53.

459.   Ellison, K.S., et al., *Site-Specific Mutagenesis by O-6-Alkylguanines Located in the Chromosomes of Mammalian-Cells - Influence of the Mammalian O-6-Alkylguanine-DNA Alkyltransferase.* Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(22): p. 8620-8624.

460.   Seo, K.Y., S.A. Jelinsky, and E.L. Loechler, *Factors that influence the mutagenic patterns of DNA adducts from chemical carcinogens.* Mutation Research-Reviews in Mutation Research, 2000. **463**(3): p. 215-246.

461.   Patro, J.N., et al., *Studies on the replication of the ring opened formamidopyrimidine, Fapy.dG in Escherichia coli.* Biochemistry, 2007. **46**(35): p. 10202-12.

462.   Kramata, P., et al., *A single site-specific trans-opened 7,8,9,10-tetrahydrobenzo[a]pyrene 7,8-diol 9,10-epoxide N2-deoxyguanosine adduct induces mutations at multiple sites in DNA.* J Biol Chem, 2003. **278**(17): p. 14940-8.

463.   Zhao, B., et al., *Role of DNA polymerase eta in the bypass of abasic sites in yeast cells.* Nucleic Acids Res, 2004. **32**(13): p. 3984-94.

464.   Upton, D.C., et al., *Replication of N2-ethyldeoxyguanosine DNA adducts in the human embryonic kidney cell line 293.* Chem Res Toxicol, 2006. **19**(7): p. 960-7.

465.   Upton, D.C., et al., *Mutagenesis by exocyclic alkylamino purine adducts in Escherichia coli.* Mutat Res, 2006. **599**(1-2): p. 1-10.

466.   McLuckie, K.I., et al., *Development of a novel site-specific mutagenesis assay using MALDI-ToF MS (SSMA-MS).* Nucleic Acids Res, 2006. **34**(22): p. e150.

467.   Kraemer, K.H. and M.M. Seidman, *Use of Supf, an Escherichia-Coli Tyrosine Suppressor Transfer- Rna Gene, as a Mutagenic Target in Shuttle-Vector Plasmids.* Mutation Research, 1989. **220**(2-3): p. 61-72.

468.   Parris, C.N. and M.M. Seidman, *A Signature Element Distinguishes Sibling and Independent Mutations in a Shuttle Vector Plasmid.* Gene, 1992. **117**(1): p. 1-5.

469.   Kong, D. and W. Masker, *Deletion between direct repeats in T7 DNA stimulated by double-strand breaks.* J Bacteriol, 1994. **176**(19): p. 5904-11.

470.   Kowalczykowski, S.C., *Initiation of genetic recombination and recombination-dependent replication.* Trends Biochem Sci, 2000. **25**(4): p. 156-65.

471.   Lusetti, S.L. and M.M. Cox, *The bacterial RecA protein and the recombinational DNA repair of stalled replication forks.* Annu Rev Biochem, 2002. **71**: p. 71-100.

472.   Shanabruch, W.G., et al., *Mutagenesis, by methylating and ethylating agents, in mutH, mutL, mutS, and uvrD mutants of Salmonella typhimurium LT2.* J Bacteriol, 1983. **153**(1): p. 33-44.

473.   Sutton, M.D., et al., *The SOS response: recent insights into umuDC-dependent mutagenesis and DNA damage tolerance.* Annu Rev Genet, 2000. **34**: p. 479-497.

474.   Kuzminov, A., *Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda.* Microbiol Mol Biol Rev, 1999. **63**(4): p. 751-813, table of contents.

475.   Morimatsu, K. and S.C. Kowalczykowski, *RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: a universal step of recombinational repair.* Mol Cell, 2003. **11**(5): p. 1337-47.

476.   King, J.S., C.F. Fairley, and W.F. Morgan, *DNA end joining by the Klenow fragment of DNA polymerase I.* J Biol Chem, 1996. **271**(34): p. 20450-7.

477.   Bowater, R. and A.J. Doherty, *Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining.* PLoS Genet, 2006. **2**(2): p. e8.

478.   Stodolsky, M., M.E. Rae, and E. Mullenbach, *The addition of lac+ chromosome fragments to the E. coli proA-proB-lac deletion 13 chromosome.* Genetics, 1972. **70**(4): p. 495-510.

479.   Adams, W.T. and T.R. Skopek, *Statistical test for the comparison of samples from mutational spectra.* J Mol Biol, 1987. **194**(3): p. 391-6.