THE GENERATION AND CHARACTERISATION OF MAMMALIAN CELLS AND TRANSGENIC MICE CONTAINING THE *ESCHERICHIA COLI fpg* GENE

A

Thesis

submitted to the University of Manchester

for the Degree of Doctor of Philosophy

in the Faculty of Medicine

October, 1995

by

IAN MARTIN FRAYLING *M.A., M.B., B.Chir.*

This research was undertaken in:

The CRC Department of Carcinogenesis

Paterson Institute for Cancer Research

University Department of Oncology

Christie Hospital NHS Trust

Manchester, UK

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

ABSTRACT

Alkyl–formamidopyrimidine (alkFaPy) residues are formed in DNA subsequent to alkylation of the 7–position of purine bases by a variety of xenobiotics from simple methylating agents to complex heterocyclic compounds such as aflatoxins. There is evidence that alkFaPy lesions are cytotoxic, but it has not been defined to what extent they may be mutagenic or carcinogenic. Related lesions can also occur as part of the spectrum of oxidative DNA damage. Hydroxyl radicals and singlet oxygen attack the carbon at the 8–position of purine imidazole rings to form 8–oxopurines and simple, i.e. non–alkylated, FaPy residues. There is good evidence of a mutagenic rôle for 8–oxoguanine (8oxoG), and it is implicated as a carcinogenic lesion; 8oxoadenine appears to be neither toxic nor mutagenic.

The *E. coli fpg* gene was originally identified on the basis of it encoding formamidopyrimidine–DNA glycosylase (known as FPG protein), which is also an AP (apurinic/apyrimidinic) lyase. It can excise a wide range of alkFaPy lesions, including those formed from aflatoxin B_1. However, during the period when the work for this thesis was being carried out, it was found that the *mutM* mutator locus (characterised as conferring on *E. coli* a high rate of G:C→T:A transversions) was in fact *fpg*, leading to the identification of 8oxoG as perhaps the major mutagenic lesion removed by FPG protein *in vivo*.

In order to define the potential rôle of alkFaPy lesions as toxic, mutagenic and potentially carcinogenic, it was intended to introduce the *fpg* gene into mammalian cells *in vitro* and also generate *fpg*–transgenic mice. The FPG protein was first overexpressed 1800–fold in *E. coli* using a construct of the *fpg* gene in pKK223–3 and purified, enabling rabbit anti–FPG polyclonal antisera to be raised, some of which were potent liquid–phase inhibitors of FPG protein DNA–glycosylase activity and could also detect FPG protein in western analysis.

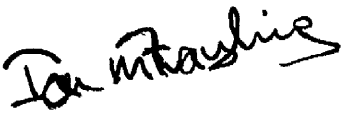
Constructs of the *fpg* gene were made in a retroviral–based vector (pLJ) and transfected into a V79–derived cell line (RJKO). Out of 42 clones isolated, 6 showed evidence of integration of the *fpg* gene, 4 of which gave a positive signal on northern analysis, 2 of these showing a significantly raised FaPy–DNA glycosylase level, but only by a factor of *ca.* 1.8–fold. None of these 6 clones was positive for FPG protein on western analysis. A construct of the *fpg* gene was also made in the vector pHMG (based on the mouse hydroxymethylglutaryl–Co–enzyme A reductase (HMGCR) gene) for use in the generation of transgenic mice. This was first transfected into RJKO cells (together with pLJ) and one clone, out of the six obtained, contained the *fpg* gene, with evidence of expression on northern but not western analysis. The clone did not have a significantly raised FaPy–DNA glycosylase level, and neither could it be stimulated to express FPG protein by treatment with HMGCR inhibitors.

Microinjection of C57BL/6J zygotes was carried out, and out of 20 potential founder mice born only one male was found to be transgenic, the transgene having integrated as a multicopy head–to–tail array. Subsequent Southern analysis showed this array to be highly methylated and there was no evidence of expression on northern or western analysis. There was also no consistent evidence of any increase in FaPy–DNA glycosylase levels. Treatment of the *fpg*–transgenic animals with either HMGCR inhibitors, or 5–azacytidine following partial hepatectomy, was unable to increase FaPy–DNA glycosylase levels.

Investigations were also carried out which showed that treatment of rodents with bromate or γ–irradiation did not up-regulate their endogenous FaPy-DNA glycosylase levels. In addition, repeated selection of RJKO cells by the photodynamic action of Methylene Blue did not produce sub-lines with either an increased clonal survival rate, or increased FaPy-DNA glycosylase levels.

DECLARATION

I declare that no portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.



Ian M. Frayling

COPYRIGHT

EDUCATION

May 1980	Graduated from the University of Cambridge with the Degree of B.A. (Hons.), Part I Tripos: Medical Sciences; Part II Tripos: Biochemistry.
December 1982	Graduated from the University of Cambridge School of Clinical Medicine with the Degrees of M.B. and B.Chir.
February 1983	House Officer, Medical Research Council Haematological Oncology Unit, Addenbrooke's Hospital, Cambridge.
February 1984	Senior House Officer in Pathology, North Manchester General Hospital.
March 1985	Locum Registrar, Department of Pathology, University Hospital of South Manchester (Withington).
May 1985	Senior House Officer in Pathology, Manchester Royal Infirmary.
January 1986	Registrar in Chemical Pathology (North Western Regional Health Authority): Bolton Royal Infirmary (to December 1986), University Hospital of South Manchester (to April 1988), Royal Manchester Children's Hospital (to September 1989).
June 1987	Passed Primary Examination for Membership of the Royal College of Pathologists (Chemical Pathology).
October 1989	Awarded Cancer Research Campaign Research Fellowship for Clinician, CRC Department of Carcinogenesis, Paterson Institute for Cancer Research, Christie Hospital, Manchester.

RESEARCH EXPERIENCE

1980	Biochemistry Honours Degree project: "The exogenous lipid requirement for rat peritoneal mast cell degranulation".
1981	Clinical elective in the Department of Clinical Immunology, Addenbrooke's Hospital: "The clearance from the circulation of rabbits of IgG–BSA immune complexes formed *in vivo*".
1985	Department of Pathology, University Hospital of South Manchester: "Assessment of human patellar cartilage degenerative changes, and *in vitro* culture of human cortical bone".
1987	Department of Chemical Pathology, University Hospital of South Manchester: "Development of assays for total protein in urine".
1988 - 1989	Department of Chemical Pathology, Royal Manchester Children's Hospital: "Methaemoglobinaemia induced by cutaneous application of local anaesthetic cream, and studies on the low-molecular weight proteinuria induced by chemotherapy".
October 1990	Admitted to study for the Degree of Doctor of Philosophy in the Faculty of Medicine, Department of Oncology, University of Manchester.

ACKNOWLEDGEMENTS

I would like to give especial thanks to Dr G.P. Margison for all his help, encouragement, support and excellent supervision during my work for this thesis. I am also grateful to Dr P.J. O'Connor for letting me work in his Department and sharing with me the benefit of his wide experience of the field of carcinogenesis. I would like to thank too Professor J. McClure and Dr G.M. Addison for their support and encouragement in my application for a Fellowship, and Professor D.G. Harnden for his advice and excellent direction, both in putting me in touch with Dr G.P. Margison, and suggesting that I later move into the area of cancer genetics. I would also like to thank Dr J.A. Rafferty, Dr D.P. Cooper and Dr R.H. Elder for their advice and many helpful discussions, Dr M. Santibanez-Koref for many stimulating discussions and the occasional translation from technical German, Dr C.Y. Fan for teaching me how to make transgenic mice, Dr L. Pearl (University College, London) for discussions on the structure of FPG protein and advice on purification techniques, and Mrs A.J. Watson for teaching me how to culture and clone mammalian cells. My thanks are also due to Miss J.L. Day, Mr M.A. Willington and Mrs L. Woolford for their much appreciated help in transgenic experiments. I am grateful to all those others, in the Department of Carcinogenesis and the Paterson Institute, which space does not allow me to mention by name, who helped in some way and made for such a good working environment. Thanks too to everyone for putting up with my sense of humour, which at times cannot have been easy. A special thank you is reserved for the Cancer Research Campaign, both for their "Research Fellowships for Clinicians", and awarding me one of them.

Finally, I would like to express my sincere thanks to my wife, Dr Ann Ager, for her unerring support and encouragement.

I dedicate this thesis to my parents.

LIST OF ABBREVIATIONS

A	Adenine
AFB_1	Aflatoxin B_1
AGE	Agarose gel electrophoresis
3alkA	3-alkyladenine
7alkG	7-alkylguanine
AP	Apurinic/apyrimidinic
ATase	Alkyltransferase
ATP	Adenosine triphosphate
A_λ	Spectrophotometric absorbance (wavelength (λ) in nm)
5azaC	5-azacytidine
B6	C57BL/6J
b.d.	Twice daily ("*bis diem*")
bp	Base pairs (in nucleic acid)
BRB	Bengal Rose B
BSA	Bovine serum albumin
b.w.	Body weight
C	Cytosine
CBG-250	Coomassie Blue G-250
CIAP	Calf intestinal alkaline phosphatase
cpm	Counts per minute
CV	Coefficient of Variation ((SD/mean) × 100%)
D2	DBA/2J
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
DEAE	Diethylaminoethyl-
D.F.	Degrees of freedom
dGTP	Deoxyguanosine triphosphate
DMS	Dimethyl-sulphate
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
dpm	Disintegrations per minute
dR	Deoxyribose
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
dTTP	Deoxythymidine triphosphate
EBSS	Earle's buffered salt solution
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
FaPy	Formamidopyrimidine
FCS	Foetal calf serum
FPG	*E. coli* FaPy-DNA glycosylase
fpg	The gene encoding the FPG enzyme

LIST OF ABBREVIATIONS (Cont.)

G	Guanine
hAT	Human ATase (*q.v.*)
hCG	Human chorionic gonadotrophin
HEPES	N-(2-Hydroxyethyl)piperazine-N′-(2-ethanesulphonic acid)
HMGCR	Hydroxy-methylglutaryl-Coenzyme A reductase
HRP	Horseradish peroxidase
i/p	Intra-peritoneal
IPTG	Isopropylthio-β-D-galactoside
kb	Kilo-base (in nucleic acid)
LB	Luria broth
LD50	Dose giving 50% lethality
LMP	Low melting point
LTR	Long terminal repeat
3mA	3-methyladenine
MAG	3-methyladenine-DNA glycosylase
Mb	Mega-base (in nucleic acid)
MB	Methylene Blue
5mC	5-methylcytosine
MCS	Multiple cloning site
7mG	7-methylguanine
MEM	Minimum essential medium
min	Minute (time)
MOPS	3-(N-Morpholino)propanesulphonic acid
MPR	Methylene Blue Photodynamic treatment-resistant
MMS	Methyl-methanesulphonate
M-MuLV	Moloney murine leukaemia virus
MNU	Methyl-nitrosourea
M	Molar (mol/L)
M_r	Relative molecular mass
mRNA	Messenger RNA (see below)
NDMA	N-nitrosodimethylamine
NMB	New Methylene Blue
NMU	N-nitrosomethylurea
NRS	Normal rat serum
O^6mG	O^6-methylguanine
OPA+	One-Phor-All+
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PICR	Paterson Institute for Cancer Research

LIST OF ABBREVIATIONS (Cont.)

PMS	pregnant mare's serum
PMSF	Phenylmethanesulphonyl fluoride
PolyGC	Poly(dG-dC)·poly(dG-dC)
Pu	Purine
PV	Pravastatin
Py	Pyrimidine
pZIP	pZIP-NeoSV(X)1
RBS	Ribosome binding site
RNA	Ribonucleic acid
RT	Room temperature
s/c	Sub-cutaneous
SCID	Severe combined immune deficiency
SD	Standard deviation
SDS	Sodium dodecylsulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SEM	Standard error of the mean ($SD/\sqrt{n}$)
ssDNA	Single-stranded DNA
ST	Sterigmatocystin
SV	Simvastatin (cyclic lactone form)
SV40	Simian virus 40
SVA	Simvastatin (hydroxy acid form)
T	Thymine
Taq	*Thermus aquaticus*
TBS	Tris-buffered saline
TdR	Thymidine
T_m	DNA melting (thermal denaturation) temperature
TP	Total protein
Tris	Tris-(hydroxymethyl)aminomethane
TS	Transcription start (site of)
UV	Ultraviolet

CHAPTER 1

INTRODUCTION

1.1 CANCER AND CARCINOGENESIS

Cancer is that collection of diseases characterised by the development of malignant neoplasia. A term derived from the Greek, neoplasm means "new growth", a definition which has been expanded by Willis (1952):

> *"A neoplasm is an abnormal mass of tissue, the growth of which exceeds and is uncoordinated with that of the normal tissues and persists in the same excessive manner after cessation of the stimuli which evoke the change."*

An important property of cancers is that they are malignant, i.e. they lead to the death of the host. Malignancy in this context is associated with rapid growth, invasion and destruction of contiguous structures, and ultimately dissemination throughout the body (i.e. the process of metastasis) leading to death. Benign neoplasms may occasionally be clinically malignant by virtue of their anatomical site, such as an intracranial meningioma, otherwise benign neoplasms are slow growing, respect tissue boundaries and are generally not life-threatening. The term "tumour" is commonly used in association with cancer, but by itself simply describes any gross morphological swelling of an organ or tissue; the oedema which follows physical trauma produces an inflammatory tumour, for example. Solid neoplasms cause tumours, though they may only be evident microscopically, and some forms of malignant neoplasia may cause death without the formation of a tumour, e.g. leukaemia. Malignant neoplasms might appear autonomous, but they are dependent on their host for a vascular system for the supply of nutrients, which may also take the form of endocrine support.

Carcinogenesis is the process by which malignant neoplasms arise and develop: there are many classes of carcinogenic agent, some apparently disparate. Infectious biological

agents, radioisotopes, ionizing radiation, as well as a wide range of chemicals, from heavy metals to complex organic compounds, have all been associated with carcinogenesis (Doll and Peto, 1981; Wynder and Gori, 1977). It might be said, with some justification, that the problem of studying cancer is not so much that the causes are unknown, rather that too many causes are known. While man is no different from other animals in having to exist in an environment containing symbionts, commensals, parasites and pathogens, in addition to a low level of background ionizing radiations, it is probably the chemical agents which contribute most to the occurrence of human cancer (Doll and Peto, 1981).

1.1.1 **A prehistorical perspective**

Living organisms have evolved on the earth in an environment containing a diverse range of hostile agents. Damage caused by some of these to the genetic material, DNA, can either have an immediate toxic effect and result in the cell's death, or have a more subtle effect which can be passed on to the next generation. In prokaryotes cell division is equated with production of a new generation, hence damage-induced alterations in the DNA resulting in gene mutations will cause a change in genotype, inevitably this will occasionally also involve a change in the phenotype. While a certain degree of variation in the phenotype is desirable for reasons of genetic diversity, too rapid or violent a change is harmful to the organism. To control the harmful effects of DNA damage, toxicity and mutagenicity, prokaryotes have therefore evolved systems for the repair of damaged DNA. Although conditions on the earth today almost certainly bear little resemblance to those pertaining when life arose, it would seem likely that DNA repair mechanisms were an early development in the evolution of DNA-based organisms. Among informational biological macromolecules DNA is unique in having the information stored in its primary structure in two complementary strands, thus damage to one strand can be repaired by reference to the other. Although open to debate, it may

be that life based on DNA became dominant over other systems because of its ability to repair its genetic material with greater efficiency.

The evolutionary steps involved when eukaryotes arose added a new complexity to the situation. The structure of a multicellular organism is determined and maintained by closely controlled and co-ordinated growth of the various groups of cells comprising it. Division of labour and functional diversity is achieved by the differentiation of cells within this framework. This organisation is only achieved under the direction of the information contained in the genetic material. Damage to this in a eukaryote is thus capable of disrupting the normal patterns of cell growth, differentiation and renewal. DNA repair acts to maintain the genes that are passed on from generation to generation, in both the germ-line and from stem-cells to their offspring. The necessity for such higher organisms to preserve their inheritance by repair of DNA is therefore of considerable importance and many repair mechanisms have now been identified.

1.1.2 **Historical background**

Cancer has affected humans throughout antiquity, evidence of osteosarcomas, osteochondromas, multiple myeloma, and metastases has been found in bones up to 5000 years old (Brothwell, 1967). While cancer has undoubtedly always been a major cause of death, its importance has increased in modern industrialised societies with the reduction in mortality, particularly amongst infants and children, due to infectious disease, and the consequent increase in longevity. Isolated reports exist of cancers associated with particular factors: Ramazzini described in 1700 a high incidence of breast cancer in nuns, though it was up to Lacassagne in 1932 to show that the influence of hormones was involved in the development of tumours, when he showed that mammary tumours could be induced in male mice with oestrone. Snuff was implicated in the development of (benign) nasal polyps by Hill in 1761. But it was Pott (1775) who made perhaps the first

public health recommendation based on an understanding of carcinogenesis, when he attributed the scrotal cancers of the boys who were chimney sweep's assistants to chronic exposure to soot and advised regular washing, which proved a most successful form of prevention. Lesions similar to those described by Pott were also found in workers handling arsenic-containing minerals in English factories in the early 19th century (Paris, 1820). Later in that century the association of various cancers with particular industries became apparent, Rehn observed in 1895 that malignant bladder tumours frequently occurred in aniline dye workers who were chronically exposed to aromatic amines, and skin cancers were linked to exposure to tars while working in the coal-gas industry (Daudel and Daudel, 1966).

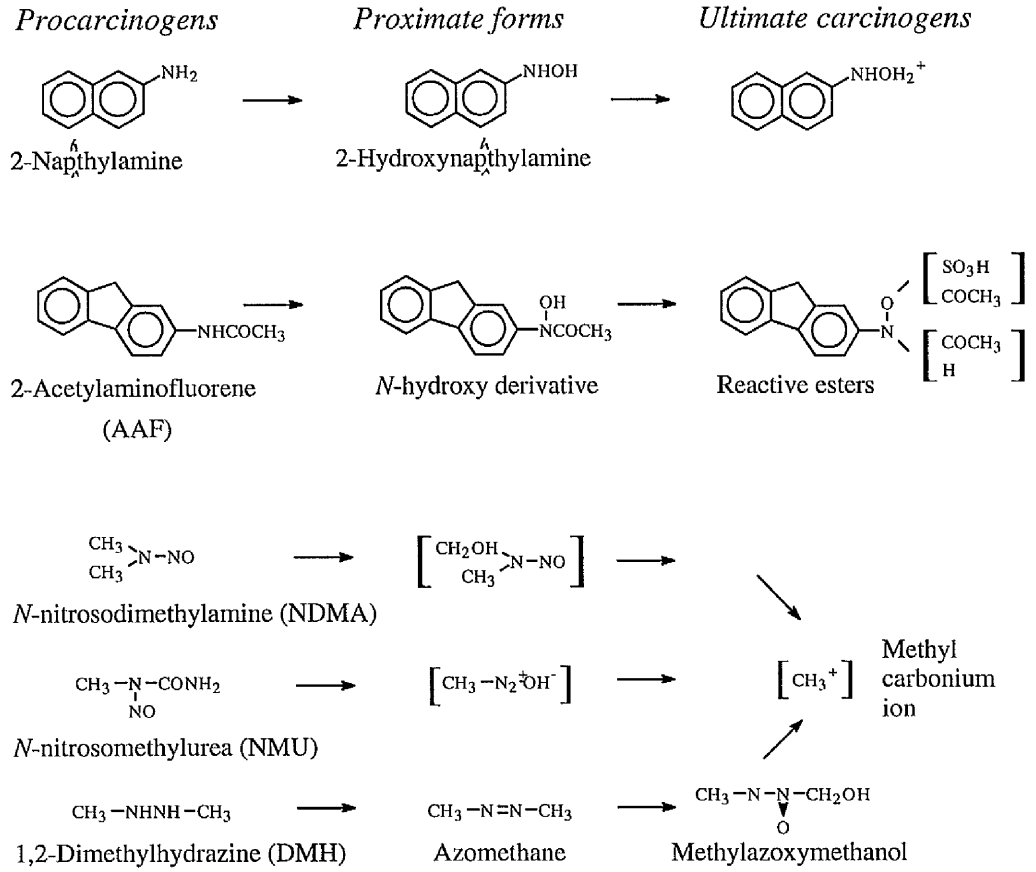
It was this increasing realisation that exposure to certain chemicals was perhaps the link between occupation and cancer that eventually led to Yamagawa and Ichikawa's demonstration (1915) that repeated application of coal tar to the ears of rabbits gave rise to skin sarcomas. Nearly two decades later 3,4-benzpyrene was isolated from coal tar and shown to be highly carcinogenic (Cook *et al.*, 1933), although the synthetic polycyclic hydrocarbon 1,2,5,6-dibenzanthracene had already been shown to be a carcinogen (Kennaway and Hieger, 1930). 2-Napthylamine, known for some time to be present in aniline dye preparations, was found to be a bladder carcinogen in dogs (Hueper *et al.*, 1938).

1.1.3 **Theories of carcinogenesis**

The work in the first half of the 20th century showing induction of cancers in animals by a wide range of chemical carcinogens gave rise to a number of hypotheses. Millers' hypothesis has proved to be the most germane. In 1969 they demonstrated that the carcinogen N-acetylaminofluorene underwent metabolism, first to an N–hydroxy derivative which subsequently became esterified, in which state it could interact with

nucleophilic sites on biological macromolecules, both proteins and nucleic acids (Miller and Miller, 1969). Later, they came to the conclusion that most chemical carcinogens underwent metabolism which generated electrophilic molecules capable of forming covalent bonds with DNA (Miller, 1970; Miller, 1978). This has led to the concept of procarcinogens, which are converted to proximate and then ultimate carcinogens (Fig. 1.1). Most carcinogens probably fall in to this class, but some can be direct-acting, e.g. MNU. Some procarcinogens may be formed endogenously, for example, nitrosamines can be formed from dietary nitrite and secondary amines reacting together in gastric acid (Hill *et al.*, 1973). Generally, procarcinogens exert maximum effect on the tissues capable of converting them to carcinogens; the liver, containing as it does so many pathways for metabolic conversion, such as detoxification, is a common site of cancer development in animal model systems.

Fig. 1.1
The activation of carcinogens



The structures of some procarcinogens are given (on the left of the diagram), with (in the centre) their proximate, and (on the right) their ultimate metabolic forms. The various steps may either occur spontaneously, or by the action of enzymes (adapted from Pitot, 1986).

1.1.4 **Tumour development**

Studies of cancer development have followed two approaches: experimental models in animals and observations of naturally occurring neoplasms.

1.1.4.1 **Experimental carcinogenesis: initiation, promotion and progression**

That experimental carcinogenesis proceeds in stages was first shown by Berenblum and Shubik (1947) who, by studying a mouse model of epidermal carcinogenesis, were able to identify two distinct stages which they termed initiation and promotion. Subsequently, a third stage has been described, i.e. progression (Potter, 1983; Hennings and Yuspa, 1985; O'Connell *et al.*, 1986). Initiation is the first stage identified in experimental

carcinogenesis and is due to a permanent, and hence irreversible, change(s) in the genome caused by covalent bonding of ultimate forms of carcinogens, which then become fixed by cell division (Hennings *et al.*, 1978; Ishikawa *et al.*, 1980; Ying *et al.*, 1982). Initiated cells are not considered capable of progressing to tumours until they have been stimulated by promoting agents. The interval between initiation and promotion can be considerable, showing that initiation is indeed heritable and permanent (Boutwell, 1964; Pitot *et al.*, 1978). Carcinogens are said to be "incomplete" if they are only capable of initiating carcinogenesis, and "complete" if they are able to produce cancer on their own. There are few truly incomplete carcinogens, the distinction is usually one of dose: most carcinogens in sufficient dose are complete carcinogens, at low doses only an initiating effect is seen.

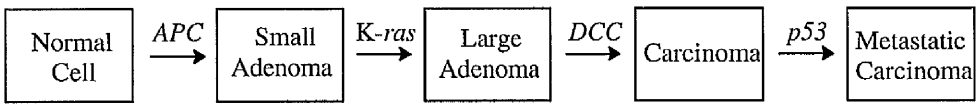
Promotion is the second stage of carcinogenesis and is characterised by being reversible, and hence not (directly) related to permanent covalent changes in DNA (DiPaulo *et al.*, 1981; Tatematsu *et al.*, 1983; Glauert *et al.*, 1986). It is considered to occur through the stimulation of growth, or clonal expansion, of initiated cells or clones of cells. The process is probably mediated through pre-existing mechanisms of cellular growth control, though the possibility exists that these may need to have been damaged or subverted by the prior process of initiation. It may also be the case that some mutagens are capable of acting as promoting agents. The final stage of carcinogenesis, progression, is distinguished by morphological changes to chromosomes (i.e. karyotypic abnormalities), smaller scale genetic changes (only measurable by recombinant DNA techniques), and alterations in enzyme activities (Pitot, 1986). These cells exhibit increased growth rate, invasiveness and metastatic potential, related to the appearance of gross neoplasms: benign and/or malignant (Schulte-Hermann, 1985). Thus experimental carcinogenesis shows two stages associated with permanent genetic changes, initiation and progression, together with an intermediate, reversible, process of promotion.

1.1.4.2 **The natural history of cancer**

Studies of a number of naturally occurring tumour types have indicated that malignant neoplasms do not arise *de novo*, but often from a preexisting benign neoplasm. For example, leukaemia is frequently preceded by one of the so-called myeloproliferative disorders. Polycythaemia (a condition of excessive production of red cells with a median survival (un-treated) of *ca.* 3 to 4 years) undergoes a transition to acute myeloid leukaemia in about 15% of patients. While it had been thought that this was due to treatment with [^{32}P] it is now accepted that this is part of the natural history of the illness and the high incidence is a result of their longer survival (median 10 to 16 years; Hoffbrand and Pettit, 1984). Chronic granulocytic leukaemia transforms in 70% of cases into acute myeloid leukaemia, but can also transform into acute lymphoblastic leukaemia (Hoffbrand and Pettit, 1984). Another, relatively easily observed and commonly occurring, tumour type which has received considerable attention is that of bowel cancer, specifically tumours of the colon and rectum. Benign adenomas of the colorectum show a correlation between size and dysplasia (the cellular atypia characteristic of malignant cells): large adenomas are more likely to show severe dysplasia, adenocarcinomas are frequently found to have arisen within adenomas, indeed the proportion so doing has been estimated to be *ca.* 90% (Morson, 1977), and patients with adenomas are more likely to develop carcinomas (Enterline, 1976). In its most severe form this tendency is inherited as an autosomal dominant trait known as Familial Adenomatous Polyposis (FAP), patients with this rare condition develop hundreds or thousands of adenomas, one or more of which, if untreated, inevitably develop into carcinomas by the age of 40 years (Jagelmann *et al.*, 1988). This is perhaps the first form of inherited predisposition to cancer to be described given the accounts of Cripps and Bickersteth over a century ago (Cripps, 1882; Bickersteth, 1890). Study of this disease led to the identification of the *APC* gene (Bodmer *et al.*, 1987; Leppert *et al.*, 1987), germ-line mutations of which are responsible

for FAP (Groden *et al.*, 1991). Mutations in *APC* are frequently (*ca.* 85%) observed in colorectal tumours, both benign and malignant (Miyaki, 1993). It was Fearon and Vogelstein (1990) who first proposed a scheme for the progression of adenomas to carcinomas as occurring in a stepwise fashion with the accumulation of defined genetic changes in a specific order (Fig. 1.2). Two classes of genes had by then been identified in connection with carcinogenesis: proto-oncogenes and tumour-suppressor genes.

Fig. 1.2
The genetic changes associated with the adenoma - carcinoma sequence in bowel tumours[a]



[a] after Fearon and Vogelstein (1990). This diagram shows the specific genetic changes associated with the progression of human large bowel mucosa cells from their normal state, first to form a small (benign) adenoma with loss of the tumour suppressor gene *APC*, then to a larger and often more dysplastic (but still benign) adenoma with activation of the K-*ras* oncogene. A malignant carcinoma may arise within the large adenoma (associated with loss of the *DCC* tumour suppressor gene), and eventually with the accumulation of further genetic changes (usually including loss of the tumour suppressor gene *p53*) progresses to a metastatic carcinoma. Not all possible genetic changes are shown, and not all tumours progress through all stages. Only a minority of adenomas progress to carcinomas, but most carcinomas arise from adenomas.

Proto-oncogenes are genes which when mutated, into oncogenes, act in a dominant fashion within a cell, i.e. the effect of the mutant protein is seen in the presence of normal protein - only one of the two alleles of such proto-oncogenes need be mutated to an oncogene. A considerable number of proto-oncogenes have now been identified, many being involved in control of cell growth and division (Bos *et al.*, 1987; Hsuan, 1993). Some proto-oncogenes, e.g. *src* are related to oncogenes found in viruses (Duesberg,

1987) and this is one way that viruses can be tumourigenic (in addition to acting *via* insertional mutagenesis). The family of *ras* genes (Ha-*ras*, K-*ras* and N-*ras*) have been extensively studied, Ha- and K-*ras* are of interest because the most commonly observed mutations occur in codons 11 and 12, facilitating molecular studies (Field and Spandidos, 1990; Boylan *et al.*, 1990).

Tumour-suppressor genes on the other hand act within cells in a recessive manner, i.e. the effect of the mutant protein is not seen in the presence of normal product - thus both copies of such genes need to be mutated (or one copy mutated and the other deleted or lost from the cell) for the cancer-promoting effect to be seen (Geiser and Stanbridge, 1989). Knudsen (1971) first proposed the "two-hit" hypothesis for tumour-suppressor genes after studying retinoblastoma (an eye cancer of infancy), which can occur in two forms: spontaneous (usually unilateral and around age 4 to 5 years) and familial (often bilateral and occurring at age 1 to 2 years). This important deduction has been borne out with the identification and isolation of the retinoblastoma gene, now known to be critically involved in the control of the cell cycle (Lee *et al.*, 1987; Wang *et al.*, 1994). The *APC* gene is another example of a tumour-suppressor gene. Mutations in such genes can be transmitted in the germ-line, giving rise to families with an apparent dominant predisposition to cancer (Hodgson and Maher, 1993). Germ-line mutations of the *p53* gene were identified in this way as causative of the Li-Fraumeni syndrome (Li and Fraumeni, 1969; Li *et al.*, 1991). Somatic mutations of the *p53* gene are frequently found in a wide variety of malignant tumours, and are most prevalent in late-stage, often metastatic, tumours from both humans and other animals (Thomas, 1994).

The stepwise, apparently sequential, accumulation of genetic changes in tumours is evidence of their evolution. Such a phenomenon implies variation between cells that can be acted upon by natural selection within the organism, the driving force to survive being the ever more successful growth of cells (Cohen and Ellwein, 1991). It can therefore be

considered that tumours evolve within organisms by a process analogous to the evolution of a species by Darwinian natural selection: heritable variations between cells that are generated by mutagens may give cells capable of greater growth or survival, which eventually become the dominant (most successful) cell type in the tumour. In this hypothesis promoting agents would act such as to provide a more favourable environment to those cells already carrying certain mutations (i.e. initiated), compared to normal cells. Thus, studies of naturally occurring tumours have shown that they evolve through a number of genetic steps in the progression from normality to benign neoplasia to, ultimately, malignant neoplasia.

1.1.4.3 **Experimentally-induced *v.* naturally occurring carcinogenesis**

There are points of both contrast and similarity between the experimentally-induced model of carcinogenesis and that observed to occur "naturally". Of course, how representative models in other species are of what occurs in humans will always be open to question, but it is to be hoped that more is gained by understanding such differences. In addition, the (intermittent) administration of (often synthetic) carcinogens in pharmacological doses may not accurately portray the natural processes of carcinogenesis. This is not to diminish the importance of experimental carcinogenesis, which has been instrumental in understanding the mode of action and consequences of carcinogens. Natural environmental exposure to carcinogens and promoting agents occurs all the time, even in laboratory animals. Exposure to high-levels of certain agents such as food contaminants, and nowadays ionizing radiation and therapeutic antitumour drugs, will occur from time to time, and in this context experimentally-induced carcinogenesis is of great relevance.

There are parallels between the stages of initiation, promotion and progression, and the steps observed in natural, apparently spontaneous, carcinogenesis. Permanent changes,

that is mutations, must occur, and in particular genes. These changes progress in a sequential fashion; the effect of promoting agents will vary depending on the mutations already accumulated. It is very unlikely that natural exposure to carcinogens and promoting agents follows the strict order suggested by the sequence initiation-promotion-progression, rather these agents are present all the time and evolution by selective advantage determines that the observed pattern of genetic changes occurs in a given order. Whatever the precise order of events, mutations are an absolute requirement for the process of carcinogenesis to occur. How some carcinogens may act to cause mutations, and the steps cells take to prevent that will be considered in the next sections.

1.2 DNA DAMAGE BY ALKYLATING AGENTS

The wide variety of agents capable of disrupting the normal covalent structure of DNA has already been mentioned; the common feature of mutagens is their ability to damage DNA in this way. Because of the great amount of work that has been carried out on DNA-alkylating agents, the mechanisms by which they damage DNA so as to cause mutations, and the processes by which that damage is repaired, are well characterised.

1.2.1 Alkylating agents

Alkylating agents are defined as chemicals able to transfer alkyl groups to nucleophilic sites in biological macromolecules, including DNA. A diverse range of compounds are so classified, ranging from simple agents such as halides, mustards, sulphates, sulphonates, and hydrazines, to more complex triazines, polycyclic aromatic compounds and heterocyclic agents, e.g. aflatoxins. The N-nitroso compounds (NOC), which include e.g. N-nitrosodimethylamine (NDMA) and N-nitrosomethylurea (NMU), are an important class of alkylating agents, given their environmental prevalence and frequent use experimentally. There are many environmental sources of NOC: dietary, industrial, therapeutic and recreational. The formation of NOC from secondary amines in gastric acid has already been mentioned, the Japanese diet contains many non-acid-pickled components, such as salt-fermented vegetables and fish, which have a high level of NOC (Kawabata *et al.*, 1983), and this has been correlated with the high incidence of gastric carcinoma amongst the Japanese (Hirayama, 1979). Exposure to NOC (NDMA and N-nitrosomorpholine) in the workplace has been shown to occur in the rubber and leather industries (Fajen *et al.*, 1979; Fine, 1980). Nitrite is used for its antioxidant properties in metal machining coolants, and in such a form has been shown to give rise to the carcinogen N-nitrosodiethanolamine (Fan *et al.*, 1977; Lijinsky and Kovatch, 1985). Therapeutic anti-tumour agents which are alkylating agents include bis-chloroethyl-N-

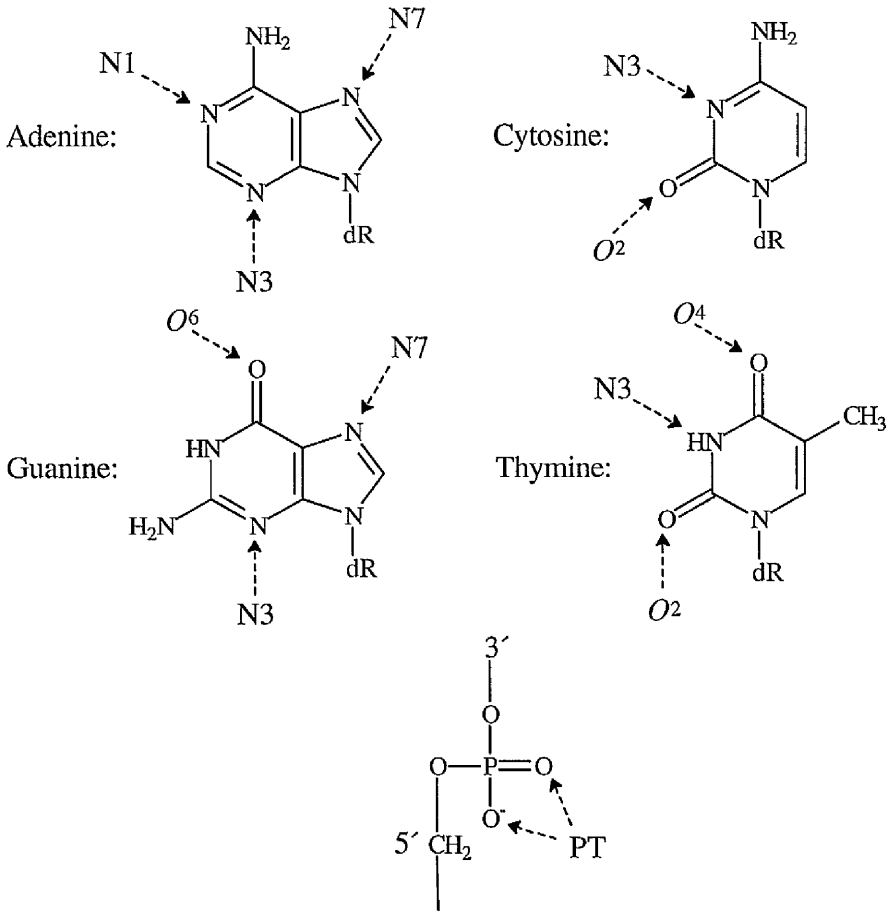
nitrosourea (BCNU), mitozolamide and procarbazine. But it is perhaps the recreational use of tobacco and its products which exposes humans to the largest quantities of NOC: of the eleven N–nitrosamines found in tobacco smoke, all but one (N-nitrosoproline) are carcinogenic in laboratory animals (US Govt. DHHS, 1982; Hecht *et al.*, 1983).

1.2.1.1 **The formation of alkylation damage**

Alkylating agents generally transfer alkyl groups to DNA by the production of intermediate highly-reactive electrophilic species, such as the methyl carbonium ion. The NOC can be divided into two groups: N-nitroso-N-alkylamides ("nitrosamides", e.g. NMU), which spontaneously decompose at physiological *p*H, and N–nitrosoalkylamines ("nitrosamines", e.g. NDMA) which require activation by microsomal mixed-function oxidases (Fig. 1.1). Most of the nucleophilic sites in DNA are liable to alkylation, and 13 sites in particular have been identified (Fig. 1.3; Lawley, 1976; Pegg, 1977; Margison and O'Connor, 1979; Singer and Kusmierek, 1982; Preussman and Stewart, 1984; Singer, 1984; Saffhill *et al.*, 1985).

Fig. 1.3

The sites of alkylation of DNA under physiological conditions



This shows the sites within DNA at which alkylation is observed to occur. Both ring nitrogens (e.g. N7 of purines) and exocyclic oxygen atoms (e.g. O^6 of guanine) of the bases are liable to alkylation, as well as the oxygen atoms of the backbone phosphate groups (forming phosphotriesters; PT).

Alkylphosphotriesters (alkPT) are produced when the phosphate groups of the DNA backbone are alkylated. Both ring nitrogen and exocyclic oxygen atoms are prone to alkylation and the relative amounts of product depend on the agent used (Table 1.1).

Table 1.1

The extent of reaction with different sites in DNA by selected alkylating agents[a]

Site of alkylation[b]	NMU	MMS[c]	NEU[d]	EMS[e]
N7-G	60 - 70[f]	81 - 85	11 - 14	58
O^6-G	6 - 8	0.3	8 - 10	2
N3-G	0.6 - 1.0	0.6 - 0.7	0.6	0.3
N3-A	8 - 9	9 - 11	3 - 5	4
N1-A	0.7 - 0.9	1.0 - 1.9	0.3	2
N7-A	0.8 - 2.0	0.9 - 1.8	0.4	2
O^2-T	0.1	nd[g]	7 - 8	nd
O^4-T	0.7	nd	1 - 4	nd
N3-T	nd	nd	nd	nd
O^2-C	0.1	nd	3 - 4	0.3
N3-C	0.5	nd	0.2	0.4
PT[h]	12	0.8	49 - 56	12

[a] adapted from Montesano (1981), [b] see Fig. 1.3, [c] methylmethanesulphate, [d] N-nitrosoethylurea, [e] ethylmethanesulphate, [f] % of total reaction products, [g] nd = not determined, [h] phosphotriesters.

It will be noted that the N7 position of guanine is particularly prone to alkylation, and this has been shown for a variety of other alkylating agents under *in vitro* conditions (Pegg, 1983). Although 7-methylguanine (7mG) was originally thought to be capable of causing mis-coding and hence mutations this has been since been disproved: 7mG has been shown to be neither toxic nor mutagenic, and persists in DNA through several rounds of replication *in vivo* (Abbott and Saffhill, 1979; Boiteux *et al.*, 1983). It later became apparent that the carcinogenic NOC were distinguished by their ability to alkylate at the O^6 position of guanine, leading to the hypothesis that this adducted base could mis-pair with thymine and hence give rise to mutations (Loveless, 1969). It would appear that O^6alkG miscodes because the O^6alkG:C base-pair has to adopt a wobble conformation, whereas O^6alkG:T retains the Watson-Crick alignment (Kalnik *et al.*, 1988;

Kalnik *et al.*, 1989). Thus the strength and extent of the hydrogen-bonding between bases may not be as important as first thought (Swann, 1990). In any event O^6alkG is a potent cause of G:C→A:T transitions, and in a similar fashion O^4alkT causes T:A→C:G transitions (Singer and Kusmierek, 1982; Saffhill *et al.*, 1985). While O^6alkG and O^4alkT are mutagenic, 3alkA is considered to be a toxic lesion because it causes premature termination of DNA synthesis (Loveless, 1969; Karran *et al.*, 1980; Sagher and Strauss, 1983). More recently Voigt and Topal (1995) have shown that O^6mG is capable of blocking the action of several DNA polymerases *in vitro*, implying that this lesion may be directly cytotoxic *in vivo*.

1.2.1.2 **Alkylation damage repair**

Three different mechanisms have been identified for the repair of alkylation damage in DNA. The first involves direct transfer of the alkyl group on DNA to a cysteine residue in a protein, thus restoring the structure *in situ*. The class of proteins involved (the DNA alkyltransferases; "ATase") promote the transfer of the alkyl group from DNA to protein and in the process they are autoinactivated in a stoichiometric reaction, hence they are not enzymes but rather large acceptor groups. The three lesions O^6alkG, O^4alkT and alkPT are all capable of repair by ATases, and ATase genes from a wide variety of species have now been cloned (Demple *et al.*, 1982; McCarthy *et al.*, 1983; Sedgwick, 1983; Margison *et al.*, 1985; Rydberg *et al.*, 1990; Xiao *et al.*, 1991; Santibanez-Koref *et al.*, 1992; Rafferty *et al.*, 1992), reviewed in Pegg and Byers (1992).

The second mechanism consists of enzyme catalysed removal of the alkylated bases by DNA glycosylases. This results in the formation of apurinic/apyrimidinic (AP) sites, which are acted upon by AP endonuclease activity to produce a gap in one strand of the DNA, subsequently repaired by DNA polymerase and DNA ligase. The process is termed base-excision repair, and many DNA glycosylases have now been identified (Brent, 1979;

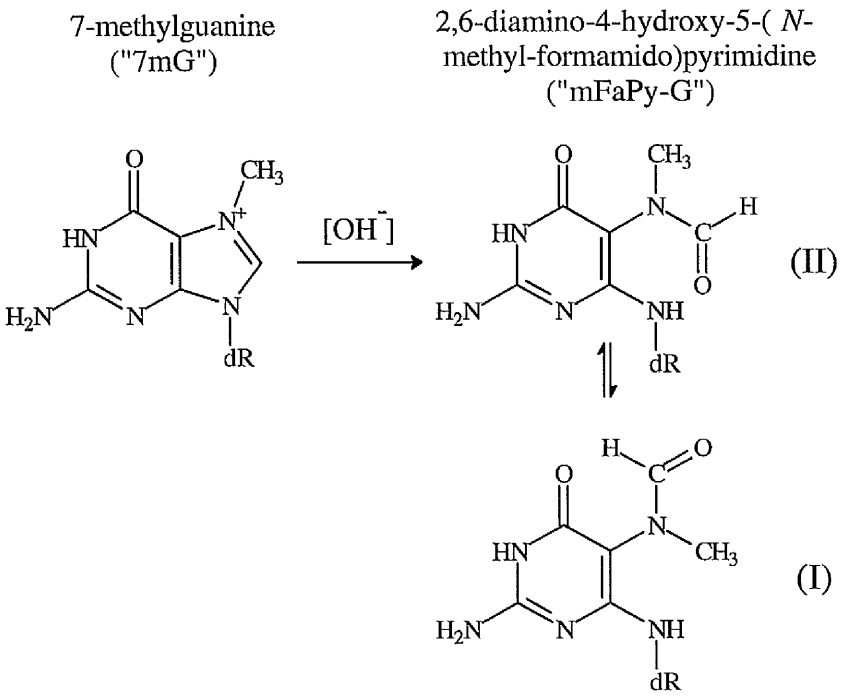
Breimer and Lindahl, 1980; Margison and Pegg, 1981; O'Connor and Laval, 1991; Eagleward *et al.*, 1993).

The third mechanism is that of nucleotide excision repair. In this complex process involving many components, specific enzymes recognise a wide variety of usually bulky or complicated DNA lesions and remove the whole nucleotide having a damaged base, plus a stretch of nucleotides either side of the original lesion. When this gap is filled in by DNA polymerase, and ultimately rejoined by DNA ligase, a large patch of DNA is resynthesised, in contrast to the simple gap in base-excision repair (Friedberg, 1987). Nucleotide excision repair works in concert with both ATases and the base excision repair pathway to eliminate DNA lesions, and there is a degree of overlap between them.

1.2.2 **The formation of alkyl-formamidopyrimidines in DNA**

It has been pointed out in the previous section how 7alkG is formed in relatively high yield by many alkylating agents, but that it is considered to be innocuous, being neither cytotoxic nor mutagenic and persisting in DNA *in vivo*. Methylation at the N7 position of purines, however, causes instability within the nucleotide: both the N9 glycosyl bond and the C8–N9 bond in the imidazole ring are weakened due to the quaternization of the N7 atom (Lawley and Brookes, 1961). Hydrolysis of the bonds is favoured producing either an AP site, or opening of the imidazole ring to produce 2,6-diamino-4-hydroxy-5-(N-methylformamido)pyrimidine (mFaPy-G), a form of secondary DNA damage. Of the four possible rotameric forms of mFaPy-G the same two are observed, whether the mFaPy-G is present in a polynucleotide, or free base form in solution (Boiteux *et al.*, 1984; Fig. 1.4).

Fig. 1.4
The imidazole ring-opening of 7-methylguanine in DNA



The imidazole ring of 7-methylguanine (7mG) residues in DNA is liable to opening, a process promoted by hydroxyl ions, so generating a methyl-formamidopyrimidine (mFaPy-G). Of the four possible rotamers, two are observed (I and II) with mFaPy-G in free base or polynucleotide form. Note that the bonds around N9 in 7mG are planar, but arranged tetrahedrally around the equivalent atom in mFaPy-G.

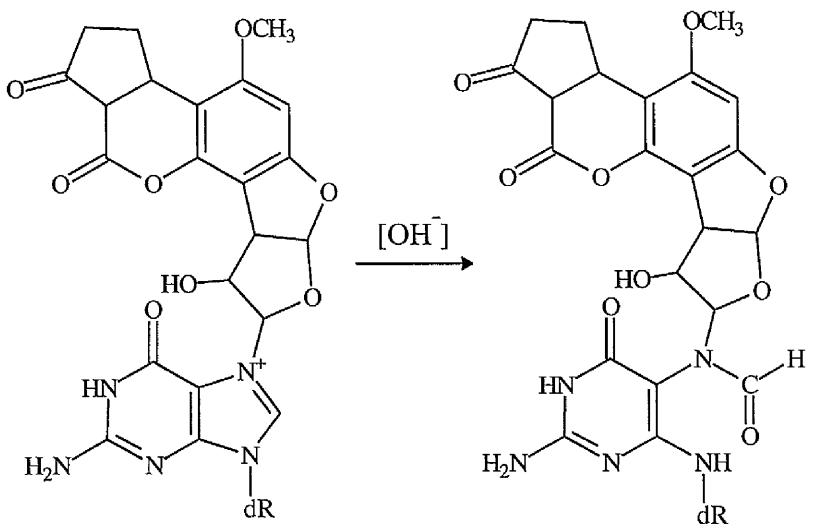
While the alkylated imidazole ring opens rapidly under alkaline conditions *in vitro* (Rainen and Stollar, 1978), it has also been found to occur at physiological *p*H, albeit at a slower rate (Beranek *et al.*, 1983). In view of the large amount of 7mG produced by methylating agents it is considered that ring-opening at even a slow rate *in vivo* is of significance (Boiteux *et al.*, 1983). A similar formamidopyrimidine (mFaPy-A; 4,6-diamino-5-(N-methylformamido)pyrimidine) can also be formed from 7alkA, but the amount of 7alkA formed by alkylating agents is much less than 7alkG (Table 1.1). More complex alkylating groups are also capable of giving rise to FaPy derivatives from guanine residues: phosphoramide mustard (PM) will generate PM-FaPy-G (Chetsanga *et al.*, 1982a), and the potent fungal hepatocarcinogens aflatoxin B_1 (AFB_1) and sterigmatocystin (ST) form AFB_1-FaPy-G and ST–FaPy-G respectively (Reddy *et al.*,

1985). While the N9 glycosyl bond is weakened in 7alkG, the equivalent bond in FaPy residues is much more stable and hence is less likely to generate AP sites (Chetsanga *et al.*, 1982a). Adduction at the *C*8 position of guanine is also possible, and this is the site at which AAF acts, giving rise to instability of the imidazole ring in the same way as N7 adduction (Kriek and Westra, 1980).

1.2.3 **The effects of alkyl-FaPy residues in DNA**

Methyl-FaPy-G residues in DNA have been shown to have significant deleterious effects, unlike the parent 7mG. Inhibition of DNA-directed DNA synthesis has been shown *in vitro*, suggesting, by analogy with 3mA, a rôle for FaPy lesions in cell killing (Boiteux *et al.*, 1983; O'Connor *et al.*, 1988; Tudek *et al.*, 1992). The formation and persistence of FaPy in DNA *in vivo* has been taken as evidence for its potential as a carcinogenic lesion (Beranek *et al.*, 1983). More convincing is the observation that rats treated with AFB_1 show increasing accumulation of AFB_1-FaPy-G in their liver DNA over the course of the following week, whereas the amount of AFB_1-G steadily diminishes (Croy and Wogan, 1981). The mutagenic effect of AFB_1 both *in vitro* and *in vivo* has been attributed to the production of AFB_1-G or its secondary derivative AFB_1-FaPy-G (Refolo *et al.*, 1985); AFB_1-G and AFB_1FaPyG have been found to be equally mutagenic in *E. coli* (Sambamurti *et al.*, 1988). Taken together, this would suggest an important carcinogenic rôle for the more persistent AFB_1-FaPy-G lesion. It is possible to see that bulky DNA lesions, such as those produced by AFB_1, should act as potent blocks to DNA replication (Fig. 1.5).

Fig. 1.5
Aflatoxin B_1 adducts in DNA



AFB_1-N7-G AFB_1-FaPy-G

Alkylation at the N7 position of guanine residues promotes imidazole ring-opening to form the equivalent formamidopyrimidine, this is so for methyl groups (7mG→mFaPy-G; Fig. 1.4) and also bulky lesions such as Aflatoxin B_1 (AFB_1-G→AFB_1-FaPy-G). The arrangements of the structures depicted above are conjectural, as the preferred rotameric forms of AFB_1 adducts in DNA have not been determined, but that for the FaPy derivative is based on that for mFaPy-G.

Why AFB_1 should be so mutagenic (Croy and Wogan, 1981) is not so evident. The arrangement of the bonds around the N9 atom in purine nucleotides when the imidazole ring is intact is planar, with opening of the ring the bonds will take up a tetrahedral arrangement, forcing rotation of the FaPy residue out of the plane of the base opposite as well as away from it (Dr L. Pearl, personal commununication). It is of note that AFB_1 reacts with guanine residues in DNA almost exclusively (>99%) in the sequence context 5′...GC...3′ (Refolo *et al.*, 1985), it is therefore probably the most selective of DNA damaging agents known. The amount of data available on the prevalence, formation and effects of FaPy lesions in DNA is not as extensive as that available on, for example, O^6alkG, O^4alkT and 3alkA. However, there is *in vitro* evidence that FaPy lesions are potentially cytotoxic, and there are suggestions that at least bulky FaPy lesions may be

carcinogenic *in vivo*.

1.2.4 **The repair of alkyl-formamidopyrimidines in DNA**

In 1979 Chetsanga and Lindahl described the isolation from *Escherichia coli* of an enzyme catalysing the release of mFaPy-G residues from DNA *in vitro*. The enzyme appeared to cleave the FaPy residues from DNA through the *N*9 glycosyl bond, needed no co-factors and had a molecular weight of the order of 30 000. Later work further characterised the enzyme: confirmed the molecular weight as 30 000; gave a Stokes radius of 2.64 nm; showed that the preferred substrate was dsDNA; that the activity was resistant to EDTA and maximal in the presence of 100 m*M* potassium chloride. The enzyme was termed FaPy-DNA glycosylase by Chetsanga *et al.* (1981). Subsequent reports have shown it to be also capable of removing PM-FaPy-G (Chetsanga *et al.*, 1982b), AFB_1-FaPy-G (Chetsanga *et al.*, 1983), the imidazole ring-opened form of AAF-*C*(8)-guanine (Boiteux *et al.*, 1989), and hydroxy-ethyl-FaPy-G (formed by chloroethylnitrosourea; Laval *et al.*, 1991). It is notable that *E. coli* FaPy-DNA glycosylase is more active at removing mFaPyG in G-rich regions of DNA, where N7-G methylation is also greatest (Graves *et al.*, 1992a).

Boiteux *et al.* reported in 1987 the cloning, nucleotide sequence and overexpression of the gene responsible for the production of the *E. coli* FaPy-DNA glycosylase, which they called *fpg* (Boiteux *et al.*, 1987). They showed that the *fpg* gene codes for a protein of 269 amino-acids with a predicted M_r of 30.2×10^3, and also that there were two regions of nucleotide and peptide sequence homology with the *denV* gene from bacteriophage T4, which codes for another DNA repair-associated protein (Valerie *et al.*, 1984). More recent work has since revealed the site of RNA transcription and that the N-terminal methionine residue is not present in the final *E. coli* product *in vivo* (Boiteux

et al., 1990). In addition, the enzyme appears to bind one zinc ion per protein molecule, which can be removed by EDTA when the enzyme is heat-denatured. The enzyme encoded by the *fpg* gene has been shown to be bi-functional in that it exhibits DNA-nicking activity at AP sites, such as produced by its own FaPy-DNA glycosylase activity or other DNA-glycosylases (O'Connor *et al.*, 1989; Boiteux *et al.*, 1990). The mechanism of nicking at AP sites is not by hydrolysis, but by β-elimination followed by δ-elimination of the AP deoxyribose, to leave the two phosphates 3′ and 5′ to the AP-site, i.e. an AP-lyase function (Bailly *et al.*, 1989). The two activities of the enzyme show similar broad *p*H optima, between *p*H 7.0 and 7.8, stimulation by 100 m*M* potassium chloride and 0.25 m*M* spermidine, inactivation by EDTA/heat denaturation, and inhibition by hydroxymercuryphenylsulphonic acid reversible by 2-mercaptoethanol (Boiteux *et al.*, 1990). This last result suggests that an accessible cysteine residue is involved in both enzymatic functions. In view of the bi-functional nature of the enzyme it has been more conveniently termed the FPG protein (Boiteux *et al.*, 1990), and will be referred to as such in this thesis.

The homologue of the *E. coli fpg* gene has been identified in and isolated from *Bacillus firmus* (Ivey, 1990), *Lactococcus lactis* (partial sequence from the 3′-region; Duwat *et al.*, 1992; Duwat *et al.*, 1995), and a cyanobacterium *Synechococcus sp.* (partial sequence from the 5′-region; Rhiel and Bryant, 1993). Alignment of the sequences of the (predicted) polypeptides from these genes reveals conservation at the amino-terminus (Fig. 1.6).

Fig. 1.6
Alignment of the amino-terminal regions of three prokaryotic FaPy-DNA glycosylases

```
        1         10         20         30         40         50
        .         .          .          .          .          .
EC:     MPELPEVETSRRGIEPHLVGATILHAVVRNGRLRWPVSEEIYRLSDQPVL
BF:     MPELPEVETVKRTLTELVIGKTIAGITVKWANIIKEPADVLEFETLLMNQ
SS:     LPELPEVETVRRGLTQQTLQRVCTGGEVLLSRTIATPTPELFLVALQ---
        :******** :*:::   :  :     *  ::     :  :
```

Alignment of the amino-terminal regions of three prokaryotic FaPy-DNA glycosylases reveals a region of identity between amino-acids 2 to 9 inclusive, as well as identical residues at positions 12 and 28 (shown in bold in the sequences, and by an asterisk (*) below the alignment). Amino-acid residues at a number of other positions show similarity (shown by a colon (:) below the alignment). The N-terminus methionine residue is not present in the mature protein from *E. coli* (Boiteux *et al.*, 1990). "EC" = *Escherichia coli*, "BF" = *Bacillus firmus*, and "SS" = *Synechococcus sp.* (the available DNA sequence data for this species only extends to the 47 amino-terminus codons), amino-acid residues shown by standard single-letter codes. The alignment was calculated by the CLUSTAL program in PC/GENE v6.8 (Intelligenetics Corp., USA) using the default computation parameters.

This conservation of primary structure at the amino-terminus of the various proteins suggests that the region is important for the correct functioning of the protein. It will be of interest to see whether the PELPEVET motif is conserved in the homologues of FPG from higher organisms; amongst the ATase proteins the acceptor cysteine residue (for methyl groups from O^6mG) is present in a highly conserved PCHRV motif (Harris *et al.*, 1992; Rafferty *et al.*, 1992; Santibanez-Koref *et al.*, 1992; Morgan *et al.*, 1993; Rafferty *et al.*, 1994). The zinc ion is held within the FPG protein by a zinc-finger motif of the C4 variety (O'Connor *et al.*, 1993; Tchou *et al.*, 1993): site-directed mutagenesis of each of these cysteine residues (at positions 244, 247, 264 and 267) has shown them all to be essential for glycosylase activity (Boiteux *et al.*, 1990). By analogy with similar structures in other DNA binding proteins, the zinc-finger structure in FPG may be necessary for DNA binding, alternatively the zinc ion may be necessary for catalytic activity (Dr L.Pearl, personal communication) Alignment of the three carboxyl-terminal sequences available from *B. firmus* and *L. lactis* with *E. coli* (Fig. 1.7) shows that they too have four identically-spaced cysteine residues, plus ten other identical and 14 similar

amino-acid residues. This conservation of structure (and presumably function) implies a critical rôle for this region in the functioning of FPG protein.

Fig. 1.7
Alignment of the carboxyl-terminus region of three prokaryotic FaPy-DNA glycosylases

```
      220        230        240        250        260        269
       .          .          .          .          .          .
EC:   FLQSDGKPGYFAQELQVYGRKGEPCRVCGTPIVATKHAQRATFYCRQCQ-K
BF:   YVNGQGEMGMFQQKLGVYGRKNEPCRQCGTDILKTVVGGRGTHFCPNCQ-L
LL:   -MQNE---------LQVYGKTGEKCSRCGAEIQKIKVAGRGTHFCPVCQQK
        ::::          * ***:::* *: **: *  :  : *:* :*: **
```

Alignment of the carboxyl-terminal regions of three prokaryotic FaPy-DNA glycosylases reveals a region of homology, including a number of identical residues (shown in bold in the sequences, and by an asterisk (*) below the alignment). Positions with similar amino-acid residues are shown by a colon (:) below the alignment. The numbering is for the residues in *E. coli* FPG protein (upper line). The four cysteine residues at positions 244, 247, 264 and 267 in FPG protein form a C4-type zinc-finger structure, necessary for glycosylase activity: they are conserved in the other two proteins. "EC" = *Escherichia coli*, "BF" = *Bacillus firmus*, and "LL" = *Lactococcus lactis* (DNA sequence data from this species is only available for the 41 C-terminal codons) , amino-acid residues shown by standard single-letter codes. The alignment was calculated by the CLUSTAL program in PC/GENE v6.8 (Intelligenetics Corp., USA) using the default computation parameters.

FaPy-DNA glycosylase activity has also been found in rodent liver extracts (Margison and Pegg, 1981), and purified from the yeast *Saccharomyces cerevisiae*, where it has a $M_r \approx 40\ 000$ (de Oliviera *et al.*, 1994). There is a single report of an enzyme activity in *E. coli* termed purine imidazole-ring cyclase, which is capable of reforming *in situ* the imidazole ring of purine residues in DNA which have been converted into FaPy residues (Chetsanga and Grigorian, 1985), however this has never been confirmed or described by other workers, in *E. coli* or other organisms, and so is open to doubt.

1.3 OXIDATIVE DNA DAMAGE

1.3.1 The origin of oxidative DNA damage: formamidopyrimidines and 8-oxopurines

During cellular respiration in both prokaryotic and eukaryotic cells, oxygen is metabolised by a series of one-electron reductions, resulting in reactive oxygen species (ROS) such as superoxide radical $(O_2^{\cdot-})$[1], hydrogen peroxide (H_2O_2) and singlet oxygen

($^1\Delta_g O_2$) (Halliwell and Gutteridge, 1989). These highly reactive intermediates may cause damage to cellular macromolecules, including DNA, and thus they may be mutagenic and carcinogenic (Ames, 1983; Cerutti, 1985; Weitzman *et al.*, 1985; Breimer, 1990). Under physiological conditions, however, neither superoxide nor hydrogen peroxide directly damages DNA (Lesko *et al.*, 1980; Rowley and Halliwell, 1983a; Aruoma *et al.*, 1989; Blakely *et al.*, 1990). The mechanism by which these species act is considered to be *via* metal ion-catalysed conversion into the extremely reactive hydroxyl radical (·OH) (Aruoma *et al.*, 1989; Halliwell and Gutteridge, 1986; Halliwell and Gutteridge, 1989) by a Fenton-type reaction:

$$M^{n+} + H_2O_2 \rightarrow M^{(n+1)+} + \cdot OH + OH^-$$

Hydrogen peroxide may be produced from superoxide by a redox reaction involving the same metal ions (Faux *et al.*, 1992a). Singlet oxygen is not a free radical, but it can be formed in radical reactions or together with radicals during oxidative stress (Sies and Mehlhorn, 1986; Halliwell and Gutteridge, 1989). Singlet oxygen is also produced by the process of photosensitisation, when certain molecules are irradiated with light of a specific wavelength (discussed below: 1.4.6), or by the spontaneous decomposition of particular compounds e.g. 3,3′-(1,4-napthylidene) dipropionate ($NDPO_2$; Di Mascio and Sies, 1989). Thus, the study of oxidative damage in DNA is complicated because a wide range of agents, including normal cellular metabolism, may all generate species capable of causing DNA damage. An additional complication is that a great variety of oxidative lesions has been observed in DNA, both to the bases and sugar moieties (Teoule and Cadet, 1978; von Sonntag, 1987), and also strand breaks and AP sites (Ward and Kuo, 1976; Schneider *et al.*, 1990). Experiments involving free nucleotides or DNA have shown more than 35 different base modifications (Hutchison, 1985; Fuciarelli *et al.*, 1989; Fraga *et al.*, 1990), many of which have been observed in DNA from oxidatively stressed cells (Hsie *et al.*, 1986). However, it has become apparent that a particular sub-class of

lesions may be critically involved in mediating the toxic, mutagenic and carcinogenic effects of oxidative DNA damage.

In 1984 Kasai and Nishimura first described the generation of a modified base in DNA treated with X-rays which on analysis proved to be 7,8-dihydro-8-oxoguanine (8oxoG; Kasai and Nishimura, 1984a). They also found that it could be generated by ascorbic acid plus hydrogen peroxide, or a number of other redox reagents (Kasai and Nishimura, 1984b) Also termed 8-hydroxyguanine, it is now known that the 6,8-diketo form is the energetically preferred enantiomer (Aida and Nishimura, 1987; Kasai and Nishimura, 1984; Kasai and Nishimura, 1987). After this, 8oxoG was found following exposure of DNA to a range of hydroxyl radical-generating sources both *in vivo* (Kasai *et al.*, 1986; Kasai and Nishimura, 1986; Kasai and Nishimura, 1987), or *in vitro* (Dizdaroglu, 1985; Floyd, 1986). The production of 8oxoG in DNA by iron and chromium has been directly attributed to such reactions (Faux, 1992a and 1992b; Aruoma *et al.*, 1989), and may explain how these and similar metals can act as carcinogens (Floyd, 1990). Other transition metal cations have been implicated, e.g. copper, nickel and cobalt (Rowley and Hall, 1983b; Halliwell and Gutteridge, 1986; Gutteridge and Halliwell, 1989; Aruoma *et al.*, 1989; Aruoma *et al.*, 1991; Nackerdian *et al.*, 1991; Kasprzak *et al.*, 1992). A particularly interesting recent report has shown evidence that [^{242}Pu] (which has a low specific activity as an α-particle emitter) is capable of generating 8oxoG in DNA *in vitro* with nearly five times the efficiency of iron, suggesting that the genotoxic effect of plutonium is not simply mediated *via* its radioactivity (Claycamp and Luo, 1994). Asbestos in the presence of hydrogen peroxide is capable of generating 8oxoG in DNA (Kasai and Nishimura, 1984b). The carcinogen 4-nitroquinoline 1-oxide has been shown to generate 8oxoG by the reaction of its proximate metabolite 4-hydroxy-aminoquinoline 1-oxide (Kohda *et al.*, 1986; Kohda *et al.*, 1987). In the rat bromate (BrO_3^-) is a renal carcinogen, and its administration specifically generates 8oxoG in kidney DNA (Kasai

et al., 1987), through the production of singlet oxygen by the kidney cells (Sai *et al.*, 1992). Radical-forming carcinogens in cigarette smoke tar can also form 8oxoG in DNA (Kasai and Nishimura, 1991). Kasai *et al.* (1986) found evidence of *in vivo* repair of 8oxoG, in that they discovered a time-dependent decrease in the level of 8oxoG in DNA from the livers of γ-irradiated mice. The free base 8oxoG is excreted in the urine of humans (Shigenaga *et al.*, 1989), and mice and rats (Fraga *et al.*, 1990), from which it has been estimated that 10^4 oxidative lesions occur per cell per day in humans and an order of magnitude more in rodents, the actual quantity of 8oxoG excreted being proportional to their specific oxygen consumption (i.e. $[O_2]$/g body weight; Fraga *et al.*, 1990). The endogenous level of 8oxoG in DNA, at 1 per 10^5 G, is also rather higher than that found for alkylation damage, i.e. 1 O^6mG per 10^7 to 10^8 G (Kasai and Nishimura, 1991). In a number of studies the equivalent oxidised form of adenine (8oxoA) has been found, in addition to the imidazole ring-opened forms of both guanine (FaPy-G) and adenine (FaPy-A), all in high yield compared to other oxidation products (Conley, 1963; Chetsanga and Grigorian, 1983; Bonicel *et al.*, 1980; Aruoma *et al.*, 1989). Singlet oxygen has been found to efficiently generate 8oxoG (and 8oxoA) in DNA (Floyd, 1990b; Schneider *et al.*, 1990). In summary, the oxidation of purines in DNA to their 8-oxo- and FaPy-derivatives occurs to a significant extent. The possible effects of this and the steps cells take to prevent and deal with such damage will be described in the following sections.

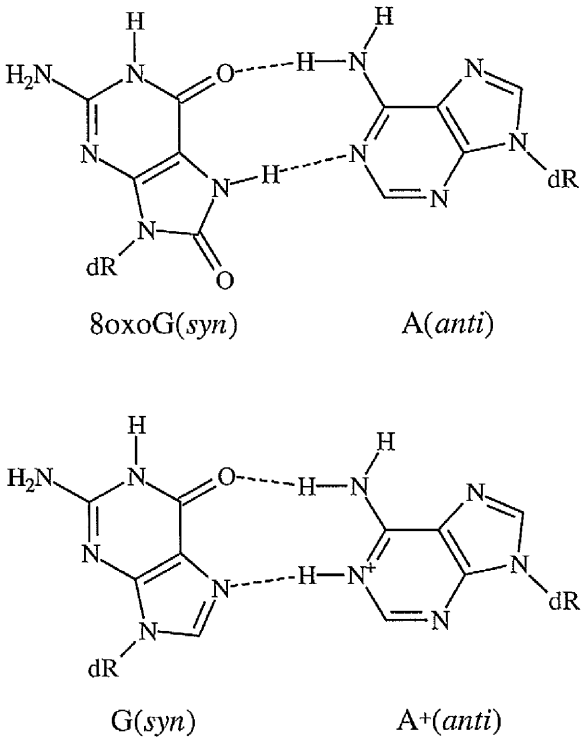
1.3.2 **The effects of 8-oxopurines**

Singlet oxygen-induced DNA lesions have been found to block DNA synthesis *in vitro* at the site of guanine residues (Ribeiro *et al.*, 1992), however this study did not differentiate between the formation of 8oxoG and FaPy-G. By analogy with the experience with mFaPy-G lesions it would be expected that FaPy-G is cytotoxic because

it prevents DNA synthesis. It has been shown that DNA synthesis proceeds past 8oxoA with the incorporation of thymine (Guschlbauer *et al.*, 1991), and thus it does not appear to be either a toxic or mutagenic lesion. There is a great deal of evidence, however, showing that 8oxoG is a mutagenic DNA lesion of some importance (Floyd, 1990). Misreading of DNA templates occurs at the site of 8oxoG residues, particularly when they are upstream of a cytosine (Kuchino *et al.*, 1987). Either adenine or cytosine is incorporated opposite 8oxoG lesions *in vitro* with approximately equal efficiency, the exact ratio depending on the polymerase tested (Shibutani *et al.*, 1991). By inserting a single 8oxoG residue into the *E. coli* M13 phage genome at a defined site, Wood *et al.* (1990) were able to show that point-mutations arose at high frequency *in vivo* (up to 1%) and that they were almost exclusively G→T transversions (96%). A similar system has been used to show that 8oxoG in DNA induces G:C→T:A transversions in mammalian cells (Moriya, 1993).

Structural studies have been carried out to investigate why 8oxoG is mutagenic. Nuclear magnetic resonance analysis of a double-stranded oligonucleotide containing 8oxoG:A mispairs has suggested that they take up an 8oxoG(*syn*):A(*anti*) conformation (Kouchakdjian *et al.*, 1991), in a manner analogous to G(*syn*):A^+(*anti*) mispairs (Brown *et al.*, 1990; Leonard *et al.*, 1990). In the G(*syn*):A^+(*anti*) mispair there is a requirement for the adenine residue to be protonated at the N1 position (which is unlikely at physiological *p*H), but this is not necessary in the 8oxoG(*syn*):A(*anti*) mispair, thus favouring its formation (Leonard *et al.*, 1992; Fig. 1.8).

Fig. 1.8
The structure of the 8oxoG:A and G:A$^+$ base pairs[a]



The structure (top) of the 8oxoG(*syn*):A(*anti*) base pair, and (bottom) of the analogous un-modified base pair. In order to form two G:A interbase hydrogen bonds protonation of the N1 of the adenine is necessary.
[a] Adapted from Leonard *et al.* (1992).

Oxidation of the guanine residue in dGTP can occur generating 8oxodGTP, and this can be misincorporated opposite adenine residues with almost equal efficiency as dTTP, thus providing a likely explanation for earlier data which showed that *in vitro* dGTP was misincorporated at a relatively high frequency opposite A (Sloane *et al.*, 1988; Maki and Sekiguchi, 1992). Thus, there are two ways in which 8oxoG may be mutagenic: 1) by its formation from G in DNA allowing the misincorporation of A (to cause G:C→T:A mutations), and 2) by its formation in dGTP allowing misincorporation opposite A (to cause A:T→C:G mutations). Additional evidence for the importance of 8oxoG as a mutagenic lesion comes from the mechanisms which cells have evolved to protect themselves from it, which will be described in the next two sections.

1.3.3 **Protection against oxidative DNA damage: antioxidants**

The first line of defense against ROS involves two different mechanisms. Cells contain many small molecules capable of reacting with ROS, which thus protect the more important macromolecules from attack. These include ascorbic acid (Vitamin C), glutathione (GSH), Vitamin E, β-carotene, and flavonoids (Michaels, 1992a). Dietary compounds from plants can also act in this way, e.g. curcumin and phenolic compounds (Shih and Linn, 1993). The other protective mechanism is a system of enzymes which rapidly degrade specific ROS: superoxide dismutase (SOD) converts $(O_2^-)^1$ to hydrogen peroxide, which (together with hydrogen peroxide from other sources) is then degraded by catalase; these enzymes are abundant, both in the cytosol and mitochondria (in eukaryotes) (Fridovich, 1974). In addition, the concentration of transition metal cations (e.g. Fe^{3+} and Cu^{2+}) is kept to a minimum by their tight binding to specific proteins (e.g. transferrin, ferritin and caeruloplasmin) which helps to minimise their activity in producing hydroxyl radicals from superoxide by Fenton-type reactions (Aruoma *et al.*, 1989).

1.3.4 **Repair of 8-oxopurines**

Hydroxyl radicals and singlet oxygen which manage to get past the first line of defense (1.3.3) may react with guanine residues, both in dGTP and DNA. A specific enzyme has evolved to degrade 8oxodGTP and prevent its misincorporation opposite A. In 1966 Yanofsky *et al.* determined the unusual specificity of a mutator locus in *E. coli*, originally described by Treffers *et al.* (1954) and later termed *mutT*: they found that the apparently spontaneous point mutations observed in this strain were A:T→C:G (Yanofsky *et al.*, 1966). The *mutT* gene was subsequently isolated and cloned (Akiyama *et al.*, 1987), and later overexpressed in *E. coli* (Bhatnagar *et al.*, 1988), where it was noted that the enzyme produced from the *mutT* gene (MutT) was active as a dGTPase *in vitro*.

However, it was Maki and Sekiguchi (1992) who were able to show that the MutT protein actually functioned as an 8oxodGTPase, thus preventing A:T→C:G mutations. The cDNA from the human homologue of the *mutT* gene (called *hMTH1*) has subsequently been isolated and overexpressed in *mutT* *E. coli*, where it complements the mutator defect (Sakumi, 1993); the *hMTH1* gene has four coding exons, spans approximately 9 kb, and is located on chromosome 7p22 (Furuichi *et al.*, 1994). A comparison of MutT with its human homologue shows a region of conserved structure in the central region of the protein (Sakumi *et al.*, 1993).

In 1988 Ylan Nghiem described a mutator locus in *E. coli* causing G:C→T:A transversions which was designated *mutY* and mapped near to 64 min on the genetic map, which it was proposed might act in a similar but reciprocal way to *mutT* which caused A:T→C:G transversions (Nghiem *et al.*, 1988). Later that year Milagros Cabrera, working in the same group, showed that there was a second locus in *E. coli* mapping between 80 and 82 min that also caused G:C→T:A transversions, and it was designated *mutM* (Cabrera *et al.*, 1988). In addition, it was proposed that *mutY* and *mutM* participated in the same pathway as a detailed examination of their mutational spectra showed them to be virtually identical and so efforts were directed at cloning the two genes (Cabrera *et al.*, 1988).

In the meantime an *fpg* mutant of *E. coli* was constructed (Boiteux and Huisman, 1989), but it was not unusually sensitive to the effects of either chemical methylation or γ-irradiation, implying that FaPy lesions were formed in low yield compared to major cytotoxic lesions such as 3mA (Boiteux *et al.*, 1984), and leading to the suggestion that the physiological substrate for FPG protein was yet to be discovered; it did, however, allow mapping of the *fpg* gene to 81.7 min on the genetic map of *E. coli*.

The following year it was shown that FPG protein was capable of removing singlet oxygen-damaged bases from DNA (Muller *et al.*, 1990), and although the precise form

of the damage was undefined it was proposed that singlet oxygen adduction at the *C*8 position of purines was involved, presumably generating FaPy lesions (Muller *et al.*, 1990). Simultaneously, an endonuclease activity was described in crude extracts of *E. coli* capable of specifically excising 8oxoG from DNA, but as it appeared to have negligible activity towards mFaPy-G lesions it could only be speculated that this 8oxoG endonuclease and FPG protein were the same enzyme (Chung *et al.*, 1991a). Coleman and Wild (1991) were able to place *fpg* adjacent to *rpmG* at 81.7 min by a comparison of their DNA sequences, giving the order: *radC-mutM-pcsA-fpg-rpmG,B-dfp-dut-pyrE-spoT*. The link was finally made when the *mutM* gene was cloned showing that it encoded FPG (Michaels *et al.*, 1991), and a close comparison of FPG protein with the 8oxoG endonuclease of Chung (1991a) showed the two enzymes to be identical (Tchou *et al.*, 1991). The *fpg* gene may have been identified by its product repairing a toxic lesion, but it was clear from the genetic (*mutM*) and later biochemical evidence (8oxoG glycosylase) that the major rôle of the FPG protein was to protect the cell from the mutagenic effects of 8oxoG in DNA. The similar apparent K_m and V_{max} values for FPG acting on mFaPy-G and 8oxoG (K_m 41 *v.* 8 n*M*, and V_{max} 0.11 *v.* 0.03 nmol/min) was explained by the similarity of the structures of FaPy-G and 8oxoG (Tchou *et al.*, 1991).

In addition to removing alkFaPy-G/A and imidazole ring-opened 8-AAF-G (1.2.4), plus 8oxoG/A lesions from DNA, it has been determined that FPG protein is also active against unsubstituted FaPy-A (generated by γ-irradiation; Breimer, 1984) and FaPy-G (Boiteux *et al.*, 1992), leading to the conclusion that FPG protein is able to act on all imidazole ring-opened and 8oxo-purines (Boiteux *et al.*, 1992), complementing the activity of endonuclease III on similarly damaged pyrimidines (Breimer and Lindahl, 1984; Asahara *et al.*, 1989). A report by Chung *et al.* (1991b) of an 8oxoG-DNA glycosylase in human neutrophils without FaPy-DNA glycosylase activity is open to

doubt, as the same group were unable to demonstrate FaPy-DNA glycosylase activity in what turned out to be FPG protein (Chung *et al.*, 1991a; Michaels *et al.*, 1991). It is likely that eukaryotic homologues of FPG have the dual ability to remove both FaPy- and 8oxo-purines (Boiteux *et al.*, 1992), though their relative rates of action on the two types of substrate may be different from FPG, which is effectively equally active on FaPy-G and 8oxoG (Tchou *et al.*, 1991; Dr S. Boiteux, personal commununication).

Further evidence for a functional connection between *fpg* (*mutM*) and *mutY* was provided by the observation that a *mutY* strain could be fully complemented by overexpression of FPG, and that a double mutant *mutY mutM* strain had a 25- to 75-fold greater mutation rate than either mutator alone (Michaels *et al.*, 1992b). With cloning of the *mutY* gene and its overexpression, it was possible to determine that MutY protein was a DNA-glycosylase (Michaels *et al.*, 1992a), subsequent work showing that it was able to remove A when this base was mispaired with G, 8oxoG, C or 8oxoA (Michaels *et al.*, 1992b). Thus, it was an example of a repair protein able to identify a normal base (A) as being misincorporated and remove it. It was also found that by binding to the mismatch MutY prevented FPG from removing the 8oxoG (or 8oxoA) which had caused it in the first place (Michaels, 1992b). MutY glycosylase is therefore able to prevent a potentially mutagenic effect of FPG protein, but this assumes and predicts that the amount of 8oxodGTP being misincorporated into DNA in normal circumstances is negligible. Repair mediated by MutY has since been shown to be mutagenic in *mutT*⁻ strains of *E. coli* showing that this is probably the case (Vidmar and Cupples, 1993), and indicating that imbalances in such a system might have deleterious effects. That a simple prokaryote like *E. coli* should have a system of protection against 8oxoG comprised of four lines of defence (antioxidants, MutT, FPG, and MutY) is powerful evidence of a strong selective advantage in minimising its presence in DNA. In other circumstances

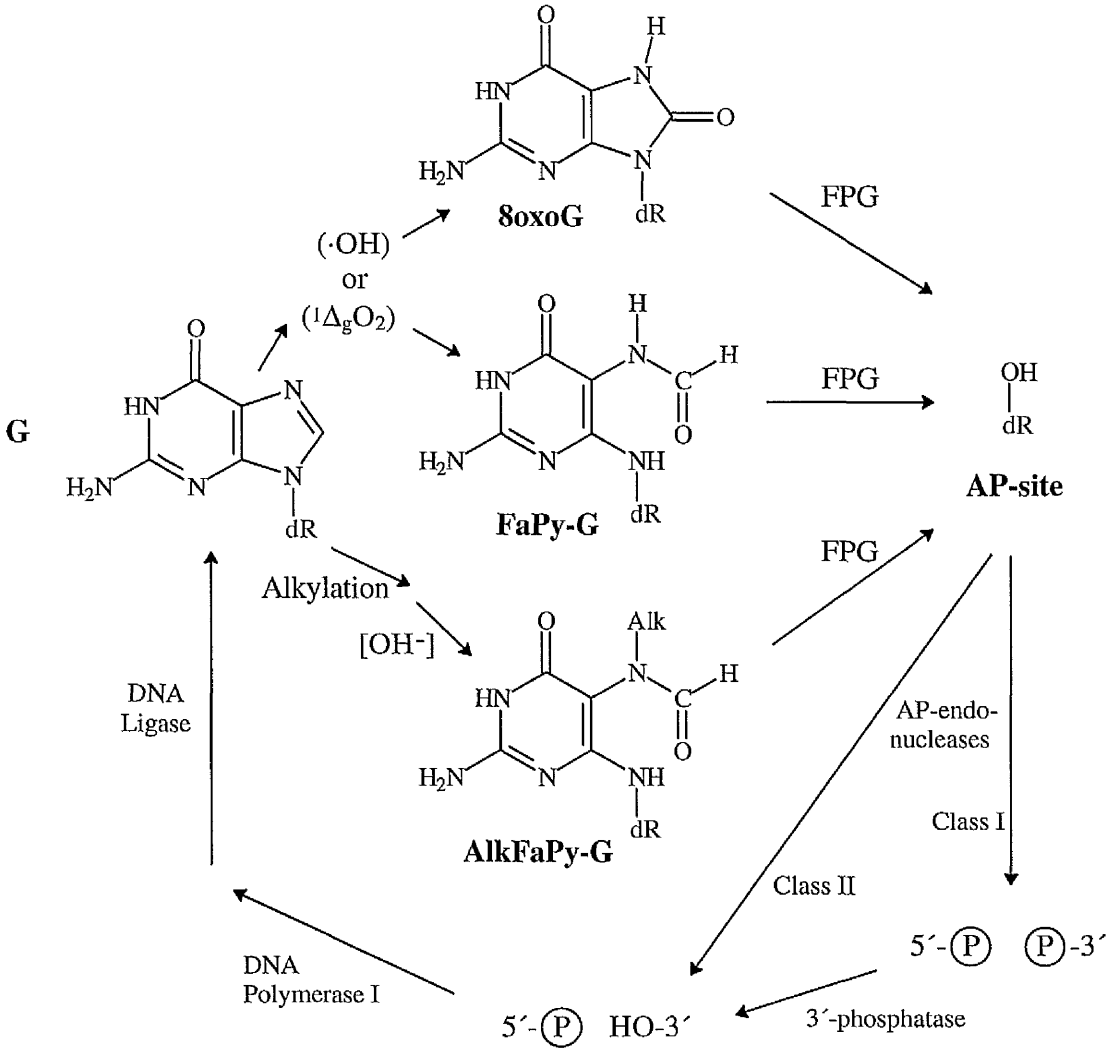
a similar system of defence in depth has been found to be extremely efficient (Middlebrook, 1971).

1.3.5 **Related repair mechanisms**

A number of other enzymes co-operate with FPG and MutY in order to effect complete repair of FaPy- and 8oxo-purine lesions. As mentioned above, FPG protein has intrinsic AP-lyase activity, as well as being able to excise 5′-terminal deoxyribose phosphate from nicked DNA (Graves *et al.*, 1992b), and so is capable of acting on the AP-sites that are produced by its glycosylase function. This leaves a gap bordered by both 3′ and 5′-phosphoryl groups, which is then presumably acted upon by a 3′-phosphatase to produce a terminus capable of extension by DNA polymerase: FPG is thus an example of a "Class I" AP endonuclease (Doetsch and Cunningham, 1990). Endonuclease III is another Class I DNA-glycosylase also active as an AP-lyase in *E. coli* (Doetsch and Cunningham, 1990). However, there are separate AP endonucleases that can carry this out directly, in both *E. coli* and higher organisms, so-called "Class II" AP endonucleases (Wallace, 1988; Doetsch and Cunningham, 1990; Robson and Hickson, 1991). AP-sites arise spontaneously at a significant rate (Lindahl and Nyberg, 1972; Lindahl and Karlstrom, 1973). After AP-sites have been processed to yield a terminus capable of extension by DNA polymerase I, a single nucleotide is added to the gap, and the repair is then completed by DNA ligase. This pathway is termed base-excision repair (Friedberg et al., 1978), and while best understood in *E. coli* has been conserved in mammalian cells (Price and Lindahl, 1991). In summary, the efficient repair of FaPy- and 8oxo-purine lesions in DNA by the base excision pathway is effected by the concerted action of a number of DNA repair and synthesis enzymes (Fig. 1.9).

Fig. 1.9

The base excision repair pathway of FaPy-G and 8oxoG in *E. coli*



Oxidation or alkylation of guanine residues in DNA results in 8oxoG or FaPy lesions. In *E. coli* the base excision repair pathway is initiated when FPG protein removes such damaged bases by its DNA-glycosylase activity. The resultant AP-sites are then acted upon by AP-endonucleases, either of Class I (e.g. FPG or endonuclease III: also known as AP-lyases), or Class II. The single base gap in the DNA is then filled in by DNA polymerase I followed by rejoining of the strand by DNA ligase. Two other enzymes (not shown) also act to prevent mutations by 8oxoG: MutT degrades 8oxodGTP preventing its possible misincorporation opposite A residues during DNA synthesis, and a DNA-glycosylase, MutY, removes A from A:8oxoG mispairs formed during DNA synthesis past 8oxoG residues. MutY also prevents FPG from removing the 8oxoG residue until a C residue has been correctly incorporated opposite the 8oxoG.

1.3.6 **Photosensitisation and oxidative DNA damage**

Raab was probably the first to describe the phenomenon of photosensitisation (also known as the photodynamic effect), whereby certain dyes in the presence of visible light are capable of inactivating living cells (Raab, 1900; Oster, 1959; Hill *et al.*, 1960). He observed that light killed protozoa placed in a solution of acridine dye, since then a large number of other dyes have also been found to have the property of inducing photosensitisation in the presence of light, in a wide variety of cells including protozoa, bacteria, yeasts, viruses, ova, sperm and mammalian cells in culture (Freeman *et al.*, 1970). Chick embryo cells were used to demonstrate the phototoxic effects of chemical carcinogens (Lewis, 1935; Lewis, 1945). Photosensitisation is of medical importance, being implicated in conditions such as porphyria and adverse drug reactions (Halliwell and Gutteridge, 1989). There is currently much interest in the possibilities of photodynamic therapy (PDT) whereby photosensitivity is used as an antitumour agent (Gill *et al.*, 1987; Yu *et al.*, 1990; Fowler *et al.*, 1990; Kleemann, 1990), given that a number of photodynamically active dyes, such as Methylene Blue (MB) and haematoporphyrin derivative (HpD), are preferentially taken up by tumour cells (Kessel, 1984; Konig *et al.*, 1987).

It has since become apparent that the photodynamic effect is mediated by the action of singlet oxygen on cellular macromolecules, including DNA (Ito, 1978; Joshi, 1985). Methylene Blue has become of particular interest. Mueller and Crothers (1975) showed that MB preferentially intercalates next to G:C base pairs in DNA, and it was later shown to generate base specific modification of DNA at guanine residues (Friedmann and Brown, 1978), and indeed as such was proposed as a G-specific DNA sequencing reagent. The nature of the lesions that MB plus light induced in DNA became evident when it was shown that 8oxoG was generated (Floyd *et al.*, 1990), and that this occurred preferentially over single strand nicks by a factor of at least 17-fold

(Schneider *et al.*, 1990). Shortly thereafter FPG protein was shown to be capable of recognising singlet oxygen-damaged purines in DNA and excising them, but whether these lesions were FaPy- or 8oxo-derivatives could not be established (Muller *et al.*, 1990). It was later demonstrated that the FPG protein did in fact excise FaPy-G, 8oxoG, FaPy-A and 8oxoA from γ-irradiated DNA, as well as FaPy-G and 8oxoG from DNA treated with MB plus light (Boiteux *et al.*, 1992). A total of thirteen products from damaged bases were detected by gas chromatography/mass spectrometry in the γ–irradiated DNA (the nine derived from pyrimidines were not excised by FPG protein), however, only FaPy-G and 8oxoG could be detected in the MB plus light-treated DNA, showing it to be a particularly "clean" form of DNA-damaging agent. Thus, increased understanding of the nature of the photodynamic effect, the origin of oxidative lesions in DNA and their repair has been made possible by a combination of approaches.

1.4 MODELS TO STUDY THE EFFECTS OF DNA DAMAGE

Cancer is a disease which affects animals, in particular mammals. Thus to determine the carcinogenicity of a compound or process inevitably requires experiments to be performed on animals, which is not only expensive but introduces a large number of variables. A number of *in vitro* model systems have been proposed to overcome this, but they can only be predictors of possible carcinogenicity, or other effects. *In vitro* systems, in both prokaryotic and eukaryotic cells, are valuable however, for the investigation of basic mechanisms, and are complementary to *in vivo* approaches. The conservation of DNA repair pathways, including many DNA repair enzymes, from prokaryotes up to the higher eukarotes is testimony to the merit of this approach.

Any given xenobiotic agent is capable of a wide range of effects, and some of these have already been described. The spectrum of possible effects includes toxicity,

clastogenicity (including sister-chromatid exchange; SCE), mutagenicity, teratogenicity, and carcinogenicity. The latter two effects can only be studied using animal models.

1.4.1 *In vitro* models: the use of cell lines

One of the major problems with attributing specific biological effects to specific lesions in DNA is that most xenobiotic agents which cause DNA damage produce a wide spectrum of lesions. The enhancement of DNA repair activity in mammalian cell lines, by the expression within them of genes encoding DNA repair enzymes, has been carried out as a way of elucidating the rôle of individual DNA lesions. If a particular biological effect is prevented, or diminished, by increasing the rate at which a specific lesion(s) is repaired then that is good evidence for a causative relationship between the two. Cell lines have been derived by transfection of ATase gene constructs, and in this way it has been possible to ascribe a direct mutagenic rôle to O^6mG and show its involvement in the induction of chromosome damage (Brennand and Margison, 1986a+b; White *et al.*, 1986; Fox *et al.*, 1987; Barrows *et al.*, 1987; Kaina *et al.*, 1987; Morten and Margison, 1988; Karran and Hall, 1989; Fan *et al.*, 1990; von Hofe *et al.*, 1992; Kaina *et al.*, 1993), as well as, indirectly, cytotoxicity through the activation of mismatch repair (Branch *et al.*, 1993). Similarly, the enhancement of MAG activity has allowed definition of the action of 3mA as a probable cytotoxic lesion (Ibeanu *et al.*, 1992; Klungland *et al.*, 1994).

1.4.2 *In vivo* models: the use of animals

The place of animals in experimental models of toxicity and carcinogenesis cannot be underestimated. It will only be by a full understanding of the interaction between genotype and environment that the significance of the many processes involved in cancer will become clear. A major problem with the use of animals is one of genetic variation between individuals. However, this has been largely overcome by the derivation of

inbred strains, particularly of mice, which allows an effectively clonal population of individuals to be used. The many such strains now available can differ widely in their susceptibility to different xenobiotic agents, but it is still very difficult to attribute a particular difference in response between two strains to specific genetic differences.

The introduction of technology allowing the generation of mice expressing specific genes (Brinster *et al.*, 1981; Palmiter *et al.*, 1982; Palmiter and Brinster, 1985; Hogan *et al.*, 1986; Gordon, 1993) has proved to be an experimental tool of the greatest utility. The term "transgenic" was coined by Gordon and Ruddle (1981) to describe those mice which had been so transformed, but the term has since been applied to any genes introduced into intact organisms (Gordon, 1993). The capability to produce animals precisely tailored to a particular purpose has long been a desire of breeders (Darwin, 1866), however the ability to now actually do it has allowed important advances to be made in the areas of genetics, developmental biology, gene regulation, immunology, biotechnology, as well as cancer (Gordon, 1993), which would not otherwise have been possible. The production of transgenic mice with specifically enhanced DNA repair functions has been achieved by a number of groups, allowing a start to be made in unravelling the complexities of how carcinogens interact within whole animals so as to produce cancer (Matsukuma *et al.*, 1989; Fan *et al.*, 1990; Walter *et al.*, 1993; Liu *et al.*, 1994).

1.6 **AIMS OF THE WORK IN THIS THESIS**

The genes for the *E. coli ada* and *ogt* ATases, as well as the human and rat ATase cDNAs had been independently isolated in the Department (Margison *et al.*, 1985; Potter *et al.*, 1987; Fan *et al.*, 1990). By their manipulation into a number of expression vectors, it had proved possible to derive cell lines with very high ATase levels (Brennand and Margison, 1986a and 1986b). These had enabled the establishment of O^6alkG as a

lesion in DNA which is toxic, mutagenic and capable of causing gross chromosome damage (Brennand and Margison, 1986a+b; Fox *et al.*, 1987; Morten and Margison, 1988; Fan *et al.*, 1990). This approach had been extended to the production of mice transgenic for the bacterial *ada* and *ogt* genes (Fan, 1991; Harris, 1990), and subsequently the human ATase (Fan *et al.*,1990; Fan, 1991), allowing studies of carcinogenicity and toxicity to be implemented in these genetically engineered animals. The involvement of FaPy, and other imidazole ring-opened purine, lesions in toxicity and carcinogenesis was by no means clear. However, with the recent isolation of the *fpg* gene (Boiteux *et al.*, 1987), it was realised that the successful approach to identifying the rôle of O^6alkG by the expression of ATase genes in cells and transgenic mice might also be applied to this question. At this stage the relevance of the FPG and associated proteins (MutT and MutY) to the repair of probably the most prevalent form of oxidative DNA damage (8oxoG) was unknown.

It was envisaged, therefore, to produce constructs of the *fpg* gene using appropriate mammalian cell expression vectors, of which there was already experience within the Department. These would then be introduced by transfection into mammalian cells *in vitro* to allow the derivation of cell lines stably expressing the FPG protein, thus allowing studies of the cytotoxic and mutagenic effects of selected agents. In addition, pronuclear microinjection of mouse zygotes would be carried out in order to generate *fpg*-transgenic mice. With such a model system available attempts could then be made in elucidating the place of FPG-repaired DNA lesions in toxicity and carcinogenesis, possibly including AAF and aflatoxins. Before the use of *fpg*-containing cell lines and transgenic mice in experiments such as these it would be necessary to show that expression of FPG was occurring to a significant level.

In the process of carrying this out the overexpression of FPG in *E. coli* would provide confirmation that the constructs made contained intact versions of *fpg*, and also a source

of protein for purification of FPG. This would enable the production of polyclonal antisera which would be of use in the analysis of cell lines and transgenic animals, as well as possibly providing a means of isolating and/or otherwise studying the homologous mammalian activity (Margison and Pegg, 1981). Purified FPG would also be useful in biochemical and perhaps also structural studies and assays of FPG-excised DNA lesions, though this was not a direct aim of this work.

CHAPTER 2

MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 Chemicals

All chemicals used were of either analytical reagent, molecular biology, chromatographic or cell culture grade, as appropriate. Water was deionised and double-distilled (ddH_2O). Solutions requiring sterilisation were either autoclaved (121°C, 1 Kg/cm^2, 15 minutes) or filtered through an Acrodisc 0.22 μm filter (Gelman, Ann Arbor, USA).

2.1.2 Enzymes

DNA modifying enzymes were obtained from BRL (Gaithersburg, MD, USA) or Promega (Southampton, UK), except for CIAP (Boehringer Mannheim, Lewes, UK).

2.1.3 Radio-isotopes

[^{35}S]-dATPαS and [^{3}H]-DMS were supplied by NEN (Dupont, Stevenage, UK), while [^{3}H]-MNU, [^{35}S]-L-Methionine, and [α^{32}P]-dCTP were supplied by Amersham International plc (Amersham, UK). Ecoscint A scintillation fluid was obtained from Mensura (Wigan, UK)

2.1.4 Buffers and solutions

For nucleic acid work buffers and solutions were generally as described in Sambrook *et al.* (1989). OnePhorAll-Plus (OPA+) buffer is produced by Pharmacia PL Ltd. The compositions given in Table 2.1 are for 1× strength (the factors in square brackets represent the stock solution concentration): the *p*H values quoted are for buffers at 20°C.

Optional components are given in ordinary parentheses, e.g. as for TBS(T). Simple solutions are not given here, but as necessary in individual methods.

Table 2.1
Buffers and solutions

Buffer/Solution	Composition
AGE loading buffer [10×]:	0.02% w/v Bromophenol blue 5 m*M* Disodium EDTA 8% w/v Sucrose
Buffer A [5×]: (FPG Purification)	50 m*M* Potassium-HEPES 2 m*M* Disodium EDTA 5 m*M* DTT 5% v/v Glycerol *p*H 7.6
Denhardt's solution [50×]:	10 g/L Polyvinylpyrollidone 10 g/L BSA 10 g/L Ficoll 400
ELISA substrate:	50 m*M* Sodium citrate (*p*H 5.0) 2 m*M* Hydrogen peroxide 0.42 m*M* Tetramethylbenzidine
FPG assay buffer [5×]:	70 m*M* Potassium-HEPES 100 m*M* Potassium chloride 2 m*M* Disodium EDTA 1 m*M* DTT 5% v/v Glycerol *p*H 7.6
FPG assay Stop solution:	2 *M* Sodium chloride 1 g/L BSA 0.5 g/L Salmon sperm DNA (sonicated)
FPG Lysis buffer:	300 m*M* Tris-chloride 10 m*M* Disodium EDTA *p*H 8.0
Ligation buffer [5×]:	50 m*M* Tris-chloride 10 m*M* Magnesium chloride 1 m*M* ATP 1 m*M* DTT 5% w/v PEG-8000 *p*H 7.6

Table 2.1 (cont.)

MOPS buffer [10×]:	20 m*M* Sodium-MOPS 5 m*M* Sodium acetate 1 m*M* Disodium EDTA *p*H 8.0
Mung Bean Nuclease buffer [10×]:	30 m*M* Sodium acetate 50 m*M* Sodium chloride 1 m*M* Zinc chloride 5% v/v Glycerol *p*H 5.0
OPA+ buffer [10×]:	50 m*M* Potassium acetate 10 m*M* Tris-acetate 10 m*M* Magnesium acetate *p*H 7.6
PBS(T):	10 m*M* Sodium phosphate 154 m*M* Sodium chloride *p*H 7.4 (0.05% v/v Tween-20)
PCR buffer [10×]:	50 m*M* Potassium chloride 10 m*M* Tris-chloride 1.5 m*M* Magnesium chloride 0.05% v/v Triton X-100 *p*H 8.4
PPI:	50 m*M* Glucose 25 m*M* Tris-chloride 10 m*M* Disodium EDTA *p*H 8.0
PPII:	200 m*M* Sodium hydroxide 1% w/v SDS
PPIII:	3 *M* Potassium acetate 2 *M* Acetic acid
RNA extraction buffer:	4 *M* Guanidine thiocyanate 12.5 m*M* Sodium citrate (*p*H 7.0) 0.5 g/L Sodium lauroylsarcosine 100 m*M* β-Mercaptoethanol
SDS-PAGE loading buffer [5×]:	60 m*M* Tris-chloride (*p*H 6.8) 2% w/v SDS 50 m*M* Dithiothreitol 4% v/v Glycerol 0.01% w/v Bromophenol blue

Table 2.1 (cont.)

SDS-PAGE resolving gel buffer [4×]:	375 m*M* Tris-chloride 0.1% w/v SDS *p*H 8.8
SDS-PAGE running buffer [10×]:	25 m*M* Tris 192 m*M* Glycine 0.1% w/v SDS
SDS-PAGE stacking gel buffer [4×]:	125 m*M* Tris-chloride 0.1% w/v SDS *p*H 6.8
T4 DNA polymerase buffer [10×]:	33 m*M* Tris-acetate 66 m*M* Potassium acetate 1 m*M* Magnesium acetate 0.5 m*M* DTT 0.2 g/L BSA
TAE [20×]:	40 m*M* Tris-acetate 1 m*M* Disodium EDTA
TBE [10×]:	90 m*M* Tris-borate 2 m*M* Disodium EDTA
TBS(T):	20 m*M* Tris-chloride 150 m*M* Sodium chloride *p*H 7.5 (0.05% Tween-20)
TE [10×]:	10 m*M* Tris-chloride 1 m*M* Disodium EDTA *p*H 7.4
TfbI:	30 m*M* Potassium acetate 50 m*M* Manganese (II) chloride 10 m*M* Calcium chloride 100 m*M* Potassium chloride 15% v/v Glycerol
TfbII:	10 m*M* Sodium-MOPS 75 m*M* Calcium chloride 10 m*M* Potassium chloride 15% v/v Glycerol *p*H 7.0

Table 2.1 (cont.)

Tissue digestion buffer:	100 m*M* Tris-chloride (*p*H 7.6) 10 m*M* Disodium EDTA 1% w/v SDS 2% w/v β-Mercaptoethanol
6 *M* Sodium iodide:	6 *M* Sodium iodide 40 m*M* Sodium sulphite
STBE:	100 m*M* Sodium chloride 90 m*M* Tris-borate 2 m*M* Disodium EDTA
STE:	100 m*M* Sodium chloride 10 m*M* Tris-chloride 1 m*M* Disodium EDTA *p*H 8.0
Western blocking buffer:	20 m*M* Tris-chloride (*p*H 7.5) 150 m*M* Sodium chloride 5% w/v Non-fat milk
Western transfer buffer:	25 m*M* Tris 192 m*M* Glycine 20% v/v Methanol

2.1.5 **Bacterial strains and plasmids**

Escherichia coli strains and genotypes:

DH5α: *supE*44 Δ*lacU*169(ø80 *lacZ*ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1

F26: *thyA his sulA*

HB101: *supE*44 *hsdS*20($r_B^-m_B^-$) *recA*13 *ara*-14 *proA*2 *lacY*1 *galK*2 *rpsL*20 *xyl*-5 *mtl*-1

JM105: *supE endA sbcB*15 *hsdR*4 *rpsL thi* Δ(*lac-proAB*)

F'[*traD*36 *proAB*$^+$ *lacI*^{q} *lacZ*ΔM15]

MC1061: *hsdR mcrB araD*139Δ(*araABC-leu*)7679 Δ*lacX*74 *galU galK rpsL thi*

Plasmids:

pUCD (Clark, 1989), pKK223-3 (Brosius, 1984), pZIP-NeoSV(X)1 (Cepko *et al.*, 1984), pLJ (Stair *et al.*, 1991), and pHMG (Gautier *et al.*, 1989).

All bacteria and plasmids used were taken from stocks held in the Department.

2.1.6 **Bacterial media**

LB broth, TB broth, TYM broth, were as given in Sambrook *et al.* (1989). Zinc chloride (100 m*M*) solution and 50% v/v glycerol used to supplement media were sterilised by autoclaving. Isopropylthio-β-D-galactoside (IPTG; BRL, Gaithersburg, MD, USA) was made up as a 100 m*M* stock in ddH_2O, filter sterilised and stored in aliquots at -20°C. Uridine (Aldrich, Gillingham, Dorset, UK) was dissolved at 100 g/L in ddH_2O, filter sterilised and used immediately. Ampicillin was used at a final concentration of 100 mg/L, chloramphenicol at 10 mg/L and kanamycin at 50 mg/L. Stock solutions of antibiotics were as follows; Ampicillin: 50 g/L in 50% v/v methanol (dissolved by neutralising with 10 *M* sodium hydroxide), *p*H 7.4; Chloramphenicol: 34 g/L in ethanol; Kanamycin: 25 g/L of the sulphate in ddH_2O; all were stored at -20°C. Oxygen (Medical grade) was from the British Oxygen Company Ltd. (UK).

2.1.7 **Mammalian cell lines**

The RJKO cell line used in mammalian cell culture experiments was from stocks of this line maintained in the Department, and is derived from V79 Chinese Hamster lung fibroblasts (A. Watson, personal communication).

2.1.8 **Mammalian cell culture media and solutions**

Minimum Essential Medium with Earle's salts (MEM) was made up as follows:

10× MEM (Gibco, Paisley, UK):	100 mL
7.5% w/v Sodium bicarbonate (1 *M*; Gibco):	27 mL
L-glutamine (200 m*M*)	10 mL
Foetal calf serum (Sigma, Poole, UK):	100 mL
ddH_2O	to 1000 mL

Filter sterilised and stored at 4°C.

L-glutamine (BRL, Gaithersburg, MD, USA) was made up as a 200 m*M* stock in ddH_2O, filter sterilised and stored in aliquots at -20°C. For any given set of experiments the same batch of foetal calf serum (FCS; Biological Industries, Ltd.) and 10× MEM were used throughout. Trypsin solution (0.5 g/L) was made up in PBS, filter sterilised and stored in aliquots at -20°C. The transfection reagent Lipofectin (BRL) was used to introduce DNA into cells. The antibiotic G–418 (Geneticin; BRL, Gaithersburg, MD, USA) was made up as a 100 g/L stock, filter sterilised and stored frozen at -20°C. The hydroxymethylglutaryl-CoA reductase inhibitors were kindly provided by The Squibb Institute for Medical Research (Pravastatin; PV) and Merck Sharp and Dohme Development Laboratories (Simvastatin; SV), they were made up as 100 m*M* stock solutions in DMSO and diluted accordingly before use. Normal rat serum (NRS) was obtained from male outbred Wistar rats (150 to 180 g b.w.; 2.1.9). For experiments with the photodynamic agents Earle's Buffered Salt Solution (EBSS; without phenol red) was used (Gibco, Paisley, UK). The dyes used in the photodynamic experiments: Methylene Blue (MB), New Methylene Blue N (NMB), and Bengal Rose B (BRB) were all supplied by Fluka, and were made up as 2 m*M* solutions in PBS, filter sterilised, and stored in the dark at 4°C, before dilution in EBSS. The MB was checked for purity by measurement of its $A_{665/610\ nm}$ ratio which equalled 2.08; an $A_{665/610\ nm}$ ratio exceeding 2.0 indicates absence of the toxic demethylated form of MB (Bergmann and O'Konski, 1963). Cells

were incubated at 37°C in water-jacketed humidified incubators in an atmosphere of 5% v/v CO_2 in air. Cell culture flasks and Petri dishes were from Falcon, and 6- and 24-well plates from Costar. Giemsa stain was prepared as a 10% v/v solution in *p*H 6.8 buffer as supplied by the manufacturer (BDH), and filtered through Whatman No.1 filter paper.

2.1.9 **Animals**

All animals used came from the stocks or colonies held in the Institute. For antiserum production outbred New Zealand Red rabbits were used. Mice were of strains C57BL/6J (B6), DBA/2J (D2), BDF_1 (B6 × D2), and BALB/c. Rats were of outbred Wistar (OBW) stock. All were maintained on standard laboratory diets, with water *ad libitum*.

2.1.10 **Materials used in animal work**

Incomplete Freund's adjuvant (Sigma) was used for the immunisation of rabbits with FPG protein. Anaesthesia was either by inhalation (Enflurane) or i/p injection (Avertin). Stock ("100%") Avertin was prepared from tribromoethanol (10 g; Fluka) and *tert*-amyl alcohol (10 mL; Sigma; Hogan *et al.*, 1986) and stored in the dark. For injection it was diluted to "2.5%" in PBS. Potassium bromate (BDH,) was made up as a 100 g/L solution in 0.9% w/v saline, filter sterilised and stored in aliquots at -20°C. Experiments involving γ-irradiation used the [^{60}Co] source in the Institute. Sutures were of 5/0 silk on cutting 16 mm curved needles (W500; Ethicon, UK), or stainless steel clips (9 mm Autoclips, #7631; Becton Dickinson, NJ, USA); ligatures were of 3/0 linen (W132; Ethicon, UK). Sterile scalpel blades for tail tip biopsies were from Swann Morton (Sheffield, UK).

2.1.11 **Materials for transgenic work**

Materials and reagents for transgenic work were as given in Hogan *et al.* (1986). The gonadotrophin preparations of Pregnant Mare's serum (PMS; "Folligon") and human chorionic gonadotrophin (hCG; "Chorulon") were supplied by Intervet, and made up at 500 IU/mL in the sterile PBS supplied by the manufacturers. They were stored in 0.1 mL aliquots at -20°C, and diluted 10-fold with saline immediately prior to use. Media M2 and M16 were prepared from concentrated stocks as in Hogan *et al.* (1986). Bovine serum albumin (Fraction V) and hyaluronidase (Type IV-S, from bovine testis) were both obtained from Sigma. For overlaying egg/embryo cultures Paraffin Oil BP was used. Pipettes for general handling of mouse eggs and embryos were made by drawing out glass pasteur pipettes over a Bunsen burner. Cannulas for egg holding pipettes and implantation of embryos into oviducts were made from glass capillaries (diameter *ca.* 1 mm, as used for melting point determinations; BDH) drawn out over a low bunsen flame. Microinjection needles were formed from internally-beaded capillaries (Type GC100TF-15; Clarke Electromedical Instruments, Reading, UK) using a moving-coil microelectrode puller (Model 753; Campden Instruments Ltd, UK) with the following settings: Pre-pull = 300, Initial pull force = 500, Main pull force = 500, Main pull delay = 200, Heater control = 360, and Optical switch = 140. Microinjection was carried out using an Olympus Inverted microscope with Nomarski optics (IMT-2), while egg handling and implantation were carried out with an Olympus VMZ dissecting microscope (Olympus Optical Co. Ltd., UK). Microinjection buffer (10 m*M* Tris-chloride, 0.2 m*M* disodium EDTA) was made using Aristar grade Tris and hydrochloric acid (BDH), in Cell culture grade sterile water (Sigma) according to Hogan *et al.* (1986). For i/p injection Pravastatin (20 m*M*) was made up in PBS, while SV was prepared as a suspension in Arachis oil BP (≡ 100 m*M*). 5-Azacytidine (BDH) was made up at 50 g/L in PBS and stored in aliquots at -20°C, before being diluted with saline immediately before use.

2.2 NUCLEIC ACID RELATED METHODS

2.2.1 General techniques

General methods of DNA and RNA handling such as precipitation from alcohol, extraction of contaminants with phenol/chloroform, quantitative and qualitative analysis by measurement of $A_{260/280\ nm}$ were carried out according to Sambrook *et al.* (1989), or according to the manufacturers' instructions, unless otherwise attributed.

2.2.2 Oligonucleotide synthesis and purification

Oligonucleotides were kindly synthesised by Dr M. Mackett (Dept. of Molecular Biology, PICR) using standard phosphoramidite chemistry on a Coder 300 DNA synthesizer (Dupont, Wilmington, DE, USA) and supplied with the 5′-trityl group left attached. The oligonucleotides were released from the resin bead substrate by incubation overnight at 60°C in concentrated ammonia solution. They were then purified by affinity chromatography on Nensorb™ columns (Dupont), which bind *via* the trityl group, according to the manufacturer's instructions. The fractions eluted from the columns with 35% v/v methanol were assayed by $A_{260\ nm}$ measurement: those containing nucleic acid were pooled and the methanol removed by evaporation. After dissolving in ddH_2O the $A_{260\ nm}$ was remeasured. The concentration of oligonucleotide was calculated by taking 1 $A_{260\ nm}$ unit = 33 mg/L of ssDNA. The stock oligonucleotides were adjusted to 200 μ*M* with ddH_20 and then frozen in aliquots at -20°C.

2.2.3 Bacterial DNA purification

Escheri chia coli, strain F26, were grown to static phase in 100 mL of LB broth, pelleted cells were then resuspended in 10 mL PPI, to which 2 g/L lysozyme was then added, and left at RT for 10 minutes. To this was mixed 17 mL TE, 1 mL 10% w/v SDS, 1 mL 0.5 *M* disodium EDTA and 1 mL 1 *M* Tris-chloride (*p*H 8.0), followed by

the addition of Proteinase K to 0.1 g/L, with incubation overnight at 37°C. The resultant digested lysate was then extracted 4 times with phenol/chloroform, twice with chloroform and sodium acetate added to 0.3 *M*. Two volumes of ethanol were added, the nucleic acid spooled out with a sterile plastic inoculating loop and rinsed twice in 70% ethanol. The nucleic acid was then allowed to redissolve in 20 mL TE overnight at 4°C. When dissolved, RNase A was added to 20 mg/L and the solution incubated at 37°C for 4 h, followed by two phenol/chloroform and two chloroform extractions. The DNA was precipitated as before from 0.3 *M* sodium acetate, washed with 70% ethanol and redissolved in 3 mL TE overnight at 4°C, before assaying by measurement of the $A_{260/280\ nm}$ ratio.

2.2.4 Mammalian cell preparation for PCR analysis

The method of Saiki (1990) was used. Briefly, 10^2 to 10^4 cells obtained at passage or cloning (2.4.1, 2.4.2) were washed with PBS, resuspended in 50 µL ddH_20 and the DNA released by heating to 95°C for 5 minutes. After centrifugation (25 000 × *g*/10 minutes/4°C) the supernatant was used immediately in PCR, or stored at -20°C.

2.2.5 PCR amplification of DNA

Amplification of DNA using the polymerase chain reaction (PCR) method of Saiki *et al.* (1985) was carried out using the heat-stable *Thermus aquaticus* DNA-directed DNA polymerase (*Taq*-polymerase). For amplification of *fpg* from *E. coli* genomic DNA for cloning purposes the PCR was set up as follows:

E. coli F-26 genomic DNA	1 μg
Oligonucleotide 184	0.50 μg (= 1.6 μ*M*)
Oligonucleotide 185	0.58 μg (= 1.6 μ*M*)
dNTP s (total)	0.4 m*M*
PCR buffer (10×)	5 μL
Taq-polymerase	5 U
ddH_20	to 50 μL

The mix was overlaid with 50 μL of mineral oil, briefly vortexed and centrifuged for 5 s in a microcentrifuge. Thermal cycling was then carried out in a Techne PHC-2 controlled heating block, with 15 cycles of [93°C × 90 s (denaturation), 45°C × 120 s (annealing), 72°C × 180 s (extension)]. After the thermal cycling the tubes were briefly centrifuged and then 5 μL of the aqueous phase (10% product) taken for analysis by agarose gel electrophoresis (AGE).

For detection of the *fpg* gene sequence in transfected mammalian cell lines the reaction mix was altered according to Innis and Gelfand (1990):

Cell extract (≈ 1μg genomic DNA)	25 μL
Oligonucleotide 184	78 ng (= 0.25 μ*M*)
Oligonucleotide 185	90 ng (= 0.25 μ*M*)
dNTP s (total)	0.4 m*M*
PCR buffer (10×)	5 μL
Taq-polymerase	2 U
ddH_20	to 50 μL

(overlaid with 50 μL mineral oil)

Thermal cycling was carried out in a Techne PHC-2 controlled heating block or Hybaid Omnigene with 10 minutes denaturation at 95°C, followed by 32 cycles of [93°C × 30 s, 58°C × 30 s, 72°C × 60 s], finishing with 72°C for 5 minutes. For analysis by AGE 5 μL was loaded per lane, as before. When analysing extracts from mice for the presence of the *fpg*-transgene, either *ca.* 100 to 200 ng of purified DNA (2.2.23) or 2 μL of crude tail tip extract (2.2.24) were used. Minor differences were seen in the yield of side-reaction products, attributable to the PCRs being performed in different thermal cyclers.

2.2.6 **Agarose gel electrophoresis**

Agarose gel electrophoresis (AGE) of DNA was carried out essentially as given in Sambrook *et al.* (1989), except that when TBE was used it was at 1× concentration, not 0.5× as suggested, and the ethidium bromide concentration was reduced to 0.32 mg/L to minimise background fluorescence. For general use 1% w/v agarose (10 g/L) was used, although for improved resolution of smaller DNAs (0.1 to 2 kb) the concentration was increased to 2% w/v. Commercially supplied DNA markers of defined size (*Hin*d III digested λ DNA (BRL), *Hae* III digested ϕX174 RF DNA (BRL), or pGEM™ (Promega)) were made up in OPA+ buffer so as to minimise differences in migration rate due to differences in salt concentrations. Samples were loaded onto gels in 2× AGE loading buffer.

The size of DNA fragments of unknown size, migrating in agarose gels, was estimated by comparing their migration distance with that of bands of known size. On those occasions when a more precise estimate of fragment size was required then a plot of migration distance *v.* the $\log_{10}$[DNA size in bp] was constructed. For a 1% gel these were typically linear over the range 0.3 to 5 kb. UV-transilluminated gels were recorded either by photography with Polaroid Type 667 film (ISO 3000) or by means of a video gel-recording system (UVP Products, UK), which also allowed densitometry.

2.2.7 **Isolation of DNA from agarose gels (phenol)**

Agarose, to which has been covalently attached hydroxy-ethyl groups, has a lowered melting point and is termed LMP agarose. DNA electrophoresed in LMP agarose can be isolated by removal from the gel as follows. After electrophoresis and documentation, DNA bands of the correct size were excised from the gel with a sterile scalpel, chopped into small pieces and placed in a 1.5 mL microcentrifuge tube. The gel volume was estimated by weighing the tube, sodium chloride was added to 0.1 *M* and the

tube then heated to 65°C for 10 minutes, with occasional vortexing, to melt the gel. This was then extracted twice with phenol equilibrated against STBE. The resultant aqueous phase was then reduced to 1/10th its original volume by *n*-butanol extraction, followed by the addition of 1/5th the original volume of ethanol. DNA was precipitated by centrifugation (25000 × *g*/20 minutes/4°C), washed with 70% ethanol and then resuspended in 0.1 mL of 0.3 *M* sodium acetate followed by reprecipitation with 0.2 mL ethanol, centrifugation as before and washing with 70% ethanol.

2.2.8 **Isolation of DNA from agarose gels (glasspowder)**

Double-stranded DNA binds reversibly to glass in the presence of a high salt concentration. Sodium iodide is a chaotropic salt which will dissolve agarose gel at room temperature. By combining these two phenomena DNA can be extracted from agarose gels (Hogan *et al.*, 1986). DNA was electrophoresed in 0.8% agarose in 1× TAE buffer (the borate in TBE interferes with the adsorbtion process), the band of interest was excised and the gel volume estimated. To this was added two volumes of 6 *M* sodium iodide followed by incubation at 37°C, with occasional vortexing, until the gel melted. For every 2 μg of DNA, 1 μL of glasspowder suspension (BIO 101, La Jolla, CA, USA) was added to the melted gel and chilled on ice for 1 h with occasional mixing. The glasspowder was then washed twice with 6 *M* sodium iodide and then four times with wash solution (0.1 *M* sodium chloride, 10 m*M* Tris-chloride and 1 m*M* disodium EDTA in 50% ethanol (*p*H 8.0)). As much as possible of the wash was removed and then TE or microinjection buffer (10 m*M* Tris-chloride, 0.2 m*M* disodium EDTA, *p*H 7.4) added to elute the DNA (37°C/20 minutes). The concentration of DNA was estimated by comparison of the band fluorescence after AGE, running aliquots of similar-sized DNA fragments of known concentration (*via* spectrophotometry) on the same gel. DNA

prepared in this way is pure enough for microinjecting into fertilised mouse eggs (Hogan *et al.*, 1986; Dr C.Y.Fan, personal communication).

2.2.9 **Small-scale plasmid preparation**

For small-scale preparation of plasmids the alkaline lysis method was used, *E. coli* cells were routinely washed once with ice-cold STE and the optional phenol/chloroform extraction step was included, otherwise the method was as given in Sambrook *et al.* (1989).

2.2.10 **Large-scale plasmid preparation**

Large-scale plasmid preparation was also carried out by the alkaline lysis method (Sambrook *et al.*, 1989). Bacteria were grown by inoculating 150 to 500 mL of TB broth with 1/100th of its volume of a fresh overnight 5 mL culture in LB broth, both containing the appropriate antibiotic. After incubation for 2 h in a rotary-shaker at 37°C ($A_{600\ nm} \approx 0.2$) the culture was supplemented with 1 g/L of uridine, and then allowed to grow for a further 4 to 5 h ($A_{600\ nm} \approx 1.4$) before the addition of 10 mg/L chloramphenicol. After overnight incubation of this, the bacteria were harvested and washed once in ice-cold STE.

Plasmid DNA was purified by equilibrium centrifugation in caesium chloride-ethidium bromide gradients; using a starting solution density of 1.60 g/cm^3, ethidium bromide concentration of 67 mg/L and centrifugation at 206 000 × *g* for 14 to 16 h at 20°C. The plasmid was precipitated out from the isolated buoyant band by the addition of six volumes of 50% isopropanol. Any ethidium bromide remaining in the DNA was removed by phenol/chloroform extraction followed by precipitation from ethanol/0.3 *M* sodium acetate and washing several times with 70% ethanol. The plasmid DNA was then

dissolved in TE, assayed by $A_{260/280\ nm}$ measurement and diluted to 1.0 g/L before storage at -20°C.

2.2.11 **Digestion of DNA with restriction endonucleases**

Digestion with restriction endonucleases was carried out by reference to the manufacturer's recommended conditions. Typically, OnePhorAll+ (OPA+) buffer (Pharmacia PL Biotechnology Ltd.) was used at a dilution of 1×, except for *Eco*R I and *Hha* I which required 2×. Incubation was for 1 h (with 1 to 2 U/μg DNA) when dealing with small-scale plasmid preparations, or 14 to 16 h for digestion of genomic DNA for Southern analysis (2.2.20). By using OPA+ no buffer change was needed in order to use many other DNA modifying enzymes, e.g. calf intestinal alkaline phosphatase.

2.2.12 **Polynucleotide Kinase treatment of DNA**

To enable PCR generated DNA to be ligated into phosphatase digested plasmid it was treated with T4 polynucleotide kinase to reestablish the 5′-phosphate groups (Sambrook *et al.*,1989). DNA was used at 0.1 μg/μL, with 1 m*M* dATP, 1× OPA+ and 40 U of T4 polynucleotide kinase in a total volume of 100 μL. After 30 minutes at 37°C the reaction was stopped by adding disodium EDTA to 25 m*M* followed by precipitation of the DNA with *iso*-propanol. After washing with 70% ethanol the DNA pellet was dissolved in 20 μL water for use in ligation reactions.

2.2.13 **5′-Phosphatase treatment of restricted plasmids**

As recommended by Pharmacia PL Ltd., after restriction enzyme digestion of between 1 and 10 μg plasmid in OPA+ buffer, 0.1 U of calf intestinal alkaline phosphatase (CIAP; Boehringer Mannheim, Lewes, UK) was added. Incubation was then continued for 30 minutes at 37°C, followed by 85°C for 30 minutes to inactivate the enzymes, before

allowing the reaction to cool to RT for 1 h. This reduced the background rate of religation of plasmid to much less than 1%, in fact typically less than 0.05%.

2.2.14 **Ligation of DNA**

Following the recommendations of Sambrook *et al.* (1989) T4 DNA ligase was used at the rate of 10 U/cohesive-end ligation or 15 U/blunt-end ligation, in a total reaction volume of 15 μL (of 1× Ligation buffer), with plasmid and insert DNAs at optimal concentrations whenever possible. The reactions were incubated either at 10°C overnight or 16°C for 6 h.

2.2.15 **Mung-bean nuclease digestion**

In order to remove the 5′-overhangs resulting from *Eco*R I digestion, the DNA was treated with mung-bean nuclease. Under the appropriate conditions this enzyme's 5′ to 3′ exonuclease activity will digest back single-stranded DNA and A:T pairs in dsDNA until the first G:C pair is reached (Hammond and D'Alessio, 1986), conveniently the first base-pair in dsDNA cleaved by *Eco*R I is G:C. Two μg of pKK223-3 cut with *Eco*R I were digested with 40 U mung-bean nuclease in 100 μL of MBN buffer, for 10 minutes at 30°C. The reaction was stopped by adding disodium EDTA to 5 m*M*, which inactivates the nuclease by chelating the Zn^{2+} ions necessary for its activity. The enzyme was then removed by extraction with phenol and the DNA precipitated from 0.5 *M* sodium chloride with ethanol.

2.2.16 **T4 DNA polymerase treatment**

The enzyme T4 DNA polymerase, in the absence of dNTPs, has a strong 3′ to 5′ directed exonuclease function. Thus it can be used to remove the 3′-overhangs generated by restriction enzymes such as *Pst* I, although it also cuts back into the dsDNA. When

dNTP s are added the exonuclease activity is replaced by DNA-directed DNA polymerase activity, therefore allowing replacement of the strand just removed by the exonuclease function (Sambrook *et al.*,1989). T4 DNA polymerase (5 U) was added to 2 μg of *Pst* I cut plasmid in 28 μL of T4 DNA polymerase buffer and well mixed. Immediately afterwards dNTPs (2 μL) were added to give a concentration of 25 μM each and the reaction incubated for 5 minutes at 37°C before being stopped by adding disodium EDTA to 25 m*M*. After phenol/chloroform extraction the DNA was precipitated from 0.5 *M* sodium chloride with ethanol.

2.2.17 **Preparation of competent *E. coli***

Competent *E. coli* cells were made by a modified method of Hanahan (Hanahan, 1983; Dr P.Potter, personal communication). The required strain of *E. coli* was grown by overnight culture in 5 mL of LB broth. This was then used to inoculate 500 mL TYM broth. Growth was allowed to proceed to mid-log phase ($A_{600\ nm} \approx 0.6$), typically 2 h, at 37°C in a rotary-shaker before rapid cooling of the flask and contents by swirling in a water/ice mixture. The cells were then pelleted (2000 × *g*/10 minutes/4°C), the supernatant discarded and the bacteria carefully resuspended in 100 mL of TfbI on ice before being repelleted (2000 × *g*/10 minutes/4°C). The bacteria were then gently resuspended in 20 mL TfbII on ice, followed by snap-freezing (with dry-ice/methanol) of 420 μL aliquots in sterile 0.5 mL microcentrifuge tubes. These were stored at -80°C until needed.

2.2.18 **Transformation of competent *E. coli***

Sufficient aliquots of competent cells were thawed on ice and then 200 μL pipetted into the required number of sterile 15 mL polypropylene transformation tubes (Sarstedt, Leicester, UK) kept on ice. An aliquot from a ligation reaction, diluted if necessary and

typically containing 2 ng vector DNA, was then added, mixed by gentle swirling and left on ice for 30 minutes. At the end of this time the tubes were warmed in a 37°C water-bath for 5 minutes before being put back on ice. LB broth (0.8 mL) was then added and the tubes incubated for a further 60 minutes at 37°C in a rotary-shaker, in order to allow expression of antibiotic resistance in transformed cells. Aliquots of the reactions were then spread on LB agar containing the appropriate antibiotic and incubated overnight at 37°C. Individual colonies were then picked and grown in 5 mL LB broth (plus antibiotic) for plasmid analysis; clones were stored at -80°C by mixing 80 µL of DMSO with 920 µL of such cultures.

2.2.19 **Nucleotide sequencing of DNA**

The dideoxy-termination method of Sanger (1977) was used in the form of a kit produced by Amersham International plc (Amersham,UK). This utilises microtitre plates in which all the necessary reagents (except labelled nucleotide) have been freeze-dried in the wells. The T7 DNA polymerase is used which gives high processivity and non-discrimination of nucleotide analogues. For sequencing of pUC based plasmids pUC/M13 forward and reverse sequencing primers (Boehringer Mannheim, Lewes, UK) were used as recommended by the manufacturers. The radiolabelled nucleotide used was [^{35}S]-dATPαS (10 µCi/µL; NEN Dupont, Stevenage, UK). Electrophoresis was carried out on 6% v/v polyacrylamide/8.3 *M* urea/1× TBE gels cast, using Sequagel reagents (Mensura, Wigan, UK), between the 40 cm long glass plates of a BRL S2 sequencing gel apparatus (BRL, Gaithersburg, MD, USA), using 1× TBE. Gels were pre-run at a constant power of 50 W for 30 minutes.

When the sequencing reactions were complete, as per the manufacturer's instructions, they were heated to 95°C for 5 minutes, to denature the DNA in the formamide based loading buffer provided, and then loaded without delay onto the polyacrylamide

sequencing gel. The gel was then run at a constant power of 70 W (typically 30 mA at ≈ 1 kV), though this was varied, depending on the ambient temperature, to maintain the gel at 55°C. The electrophoresis was terminated when the bromophenol blue marker dye had migrated to the bottom of the gel, usually about 4 h. After fixation, with 10% v/v methanol/10% v/v acetic acid for 30 minutes, gels were vacuum dried at 80°C onto Whatman 3MM paper and then put up against Kodak X-Omat AR film, for between 15 h and 2 days, to generate autoradiographs. The film was developed by inspection using Kodak D-19 high-contrast developer.

2.2.20 **Southern analysis**

This was performed essentially as described in Sambrook *et al.* (1989). DNAs were digested as in 2.2.11 and checked by AGE (2.2.6) for completeness before being used for Southern analysis. Marker DNAs were labelled with [^{35}S]-dATPαS (NEN) by nick-translation (see below). Typically 1% w/v agarose was used, although some experiments required higher strength gels for better resolution of small DNAs. Initial experiments used capillary transfer (with 5× SSC), but later on vacuum transfer was performed (Vacugene Vacuum Blotting Unit #2016; LKB Bromma, Sweden), which enabled transfers to be completed in one hour rather than overnight. With vacuum blotting, gels were first soaked for 30 minutes in 0.3 *M* hydrochloric acid to assist DNA transfer *via* depurination (as per the apparatus manufacturer's instructions). Hybond N nylon hybridisation membrane (Amersham) was used throughout, with UV-crosslinking to bind DNAs. Pre-hybridisation was carried out for 4 to 14 h with 5× SSC/5× Denhardt's/0.1 g/L sonicated salmon sperm DNA/65°C (in a Hybaid rotary oven). Probes were prepared by labelling the 0.85 kb *Bam*H I *fpg*4 DNA fragment (purified from LMP agarose gel by glasspowder; 2.2.8, 3.3.3) with [^{32}P]-αdCTP (Amersham; 3000 Ci/mmol) using a nick-translation kit (BRL). Unincorporated

radioisotope was removed by gel filtration chromatography with Sephadex G-25. Probes were denatured in the presence of 1 g/L salmon sperm DNA (in 0.5 mL of 1× SSC) by heating to 99°C for 3 to 4 minutes, with immediate addition to the pre-hybridisation solution. After hybridisation for at least 14 h the membranes were washed to a final stringency of 0.2× SSC/0.1% SDS/65°C, covered in Saran wrap (Dow, USA), and put up against X-ray film (Kodak X-Omat AR) in cassettes, containing calcium tungstate intensifying screens, kept at -80°C for between 4 h to 14 days, depending on the strength of the hybridisation signal and intensity of exposure required.

2.2.21 **RNA purification**

The method of Chomczynski and Sacchi (1987) was used to prepare total RNA from mammalian cells and tissues. Sterile disposable plasticware was used for handling samples and solutions, where this was not possible glass vessels were treated with diethylpyrocarbonate (DEPC), as was ddH_2O (Sambrook *et al.*, 1989). Cultured cells were briefly rinsed with ice-cold PBS before immediate lysis *in situ*. Tissue biopsies were collected straight into liquid nitrogen and stored at -70°C; they were crushed while frozen in liquid nitrogen and then lysed in a tissue homogeniser. Purified total RNAs were dissolved in DEPC-treated ddH_2O, and purity and concentration were assayed by measurement of the $A_{260/280\ nm}$ ratio. They were stored at -70°C after the addition of 5 U/mL RNAsin (Promega).

2.2.22 **Northern analysis**

This was performed as described in Sambrook *et al.* (1989). Electrophoresis in 1% w/v agarose gels was carried out under RNA-denaturing conditions in 1× MOPS/2.2 *M* formaldehyde. Samples (10 μg total RNA) were denatured by heating to 65°C after the addition of formamide (to 50% v/v), MOPS buffer (to 1×), ethidium

bromide (to 0.25 mg/L), and AGE loading buffer (to 1×), followed by immediate loading onto the gels. RNA markers of defined size (Promega) were included on the gels. Vacuum blotting was used to transfer the RNA onto Hybond N hybridisation membrane (Amersham), as per manufacturer's instructions, followed by UV-crosslinking. Pre-hybridisation, probing and autoradiography were carried out as for Southern analysis (2.2.20), except that the strips of hybridisation membrane onto which the markers had been transferred were removed, before pre-hybridisation of the rest of the blot, and stained with 0.5% w/v methylene blue.

2.2.23 **Purification of DNA from tissue biopsies and cultured cells**

The purification of DNA from large tail tip biopsies (1 to 2 cm) was essentially as described in Hogan *et al.* (1986), except that the tissue digestion buffer was modified by the addition of 2% v/v β-mercaptoethanol to assist in disruption of keratins. The 1 to 2 cm tail-tip biopsies were placed in 0.6 mL of digestion buffer containing Proteinase K (0.1 g/L), and incubated at 55°C for 1 h, followed by 14 to 16 h at 45°C. Phenol extraction was then carried out to purify the DNA. The method was found to work well with cultured cells (washed, and then resuspended in *ca.* 200 µL PBS, prior to the addition of digestion buffer), as well as other tissue samples, e.g. liver and kidney, which were first either crushed while frozen in liquid nitrogen or minced with a scalpel blade, adjusting the volumes of digestion buffer *pro rata.*

2.2.24 **Preparation of small tail tip biopsies for PCR**

Small tail tip biopsies (⩽ 0.5 mm) were digested in 100 µL of 1× PCR buffer containing 0.1 g/L Proteinase K by incubation at 55°C for 2 h. Tissue disruption was ensured by repeatedly drawing each sample up into a disposable pipette tip and

dispensing it back out. The Proteinase K was then denatured by heating the tubes to 99°C for 10 minutes, followed by centrifugation of the tubes in a microcentrifuge (25 000 × *g*/10 minutes/4°C). The supernatant (90 µL) was removed from the pellet of tissue debris and (1 µL) used either used immediately in PCR (2.2.5) or stored at -20°C until required.

2.2.25 **Preparation of blood samples for PCR**

The method described by Kawasaki (1990) was used. Approximately 50 µL of blood was collected into 0.5 mL of TE buffer, enabling red cell lysis. The white cells were then collected by centrifugation (12 000 × *g*/30 s), and resuspended in TE and repelleted. Resuspension and pelleting was repeated twice more, and the pellet then resuspended in 100 µL of 1× PCR buffer containing 0.1 g/L Proteinase K. This was incubated at 55°C for 1 h, before being denatured by heating to 95°C for 10 minutes; 10 µL of this was used per PCR (2.2.5).

2.3 PROTEIN RELATED METHODS

2.3.1 **Preparation of bacterial lysates for FaPy-DNA glycosylase assay**

Crude bacterial lysates were prepared essentially as described by Boiteux *et al.* (1987). In brief, the bacteria in 1.5 mL of culture were pelleted, washed in ice-cold STE, resuspended in 0.2 mL FPG Lysis buffer and then frozen at -80°C. After thawing, 10 µL of 20 g/L lysozyme solution was added with incubation for 15 minutes at 37°C, followed by refreezing at -80°C. After thawing, the lysates were centrifuged (25 000 × *g*/20 minutes/4°C) and the cell-wall debris removed with a sterile toothpick. The lysate was then stored on ice, if for immediate assay of FPG activity and total protein, otherwise they were frozen at -20°C. When estimating total protein, due

allowance was made for any lysozyme which had been added by including an equivalent amount of the lysozyme solution in the total protein assay.

2.3.2 DNA-based FaPy-DNA glycosylase assay

This assay for FaPy DNA-glycosylase was essentially as described by Boiteux *et al.* (1987) except for the substrate. This consisted of calf-thymus DNA which had been methylated with [^{3}H]-MNU (23.0 Ci/mmol; Dr G.P. Margison, personal communication) which was dissolved in 50 m*M* sodium phosphate buffer (*p*H 11.4) and left at RT for 24 h to open the imidazole rings of the 7mG residues. The DNA was precipitated, extensively washed and then dissolved at 2 g/L ($\approx 10^6$ cpm/mL) in 1× FPG assay buffer (without DTT) and stored at 4°C. For use in the assay this stock substrate was diluted 1 + 9 with 1× FPG buffer (containing DTT) to give approximately 10^5 cpm/mL. The standard assay mixture consisted of 1 to 20 µL of bacterial lysate or cell extract, 12.5 µL of 4× FPG buffer plus ddH_2O to 50 µL. The assay reaction was started by adding 50 µL of diluted substrate ($\approx 5 \times 10^3$ cpm) with brief vortexing, and allowed to proceed for 15 minutes at 37°C. At the end of this time 25 µL of FPG assay Stop solution was added with vortex mixing, followed immediately by 250 µL of ice-cold ethanol. After cooling for 20 minutes at -20°C the tubes were centrifuged (25 000 × *g*/15 minutes/4°C) and 300 µL of supernatant taken for scintillation counting (after mixing with 3 mL Ecoscint), using an LKB Rackbeta scintillation counter (counting efficiency = 30%). Ethanol soluble radioactivity released during the assay was related to total protein in the cell lysate and time of incubation. From the counting efficiency and specific radioactivity of the substrate it was calculated that 18.0×10^3 cpm were equivalent to the release of 1.0 nmol of mFaPy-G residues.

Units of FaPy-DNA glycosylase activity as stated in this thesis conform to the Recommendations of the Commission on Clinical Chemistry of the International Union

of Pure and Applied Chemistry (IUPAC) and of the International Federation of Clinical Chemistry (Dybkaer and Jorgensen, 1967), in relation to enzymes which act on substrates of undefined M_r , i.e. enzymes defined in terms of product rather than substrate: on this basis one Unit (U) of FaPy-DNA glycosylase activity is defined as producing one micromole of FaPy per minute, in 1× FPG assay buffer at 37°C. The FaPy-DNA glycosylase activity is usually given corrected for the total protein (TP; 2.3.5) concentration; the precise calculation of this is described in 2.3.4 (equation 1).

Initial experiments determined that the assay conditions described were substrate and not enzyme limiting (data not shown). Analysis of the substrate by prolonged incubation with pure FPG protein (kindly supplied by Dr S. Boiteux), cold perchloric acid precipitation and heat depurination showed that [^{3}H]-FaPy residues accounted for 75% of the radioactivity, the balance being mostly a mixture of alkyl-phosphotriesters, 3mA and O^6mG (data not shown), which was similar to the consensus values for MNU methylation given in Margison *et al.* (1987).

2.3.3 **Preparation of mammalian cell and tissue extracts for FaPy-DNA glycosylase assay**

For the assay of FPG activity in RJKO cell lines a 175 cm^2 flask of each line was grown to confluence, the cells harvested in PBS, pelleted (200 × *g*/5 minutes) and then resuspended in FPG assay buffer containing 1 m*M* PMSF. Cells were lysed by sonication on ice (10 s × 10 μm peak-peak, then 10 s × 12 μm), followed by centrifugation (25 000 × *g*/15 minutes/4°C), assay of TP and storage at -20°C. Tissue extracts were made by sonication of biopsies in 0.5 mL FPG assay buffer, containing 5% v/v glycerol and 1 m*M* PMSF, in the presence of an approximately equal amount of glass beads (*ca.* 0.2 mm, as used for gas-liquid chromatography; BDH). The sonicates were then

clarified by centrifugation (25 000 × *g*/15 minutes/4°C), assayed for TP and stored at -20°C.

2.3.4 **PolyGC-based FaPy-DNA glycosylase assay**

This was carried out as described by Boiteux *et al.* (1987). The substrate was synthesised from poly(dG-dC)·poly(dG-dC) (referred to hereafter as polyGC) as supplied by Pharmacia, the average size of the molecules was quoted as approximately 805 bp, which was confirmed by AGE in the presence of ethidium bromide, showing the spread in size to be *ca.* 0.7 to 1.0 kb (data not shown). In order to [^{3}H]-methylate this polymer the method of Uhlenhopp and Krasna (1971) was followed, as referred to by Boiteux *et al.* (1984). PolyGC (2.5 mg) was dissolved in 0.7 mL 50 m*M* sodium cacodylate/1 m*M* sodium EDTA buffer (*p*H 7.0) and placed in a 15 mm diameter round-bottom glass flask, containing a magnetic stirring bar made from a 4 mm length of steel paper clip sealed into a glass pasteur pipette tip. The polyGC solution was overlaid by 5 mCi of [^{3}H]-DMS in hexane (in *ca.* 1 mL, as supplied by the manufacturer; 5.0 Ci/mmol), and a 3 cm silica gel drying column contained by glass wool attached immediately. The flask was placed over a magnetic stirrer and the reaction allowed to procede at RT for 4 h, by which time all the hexane had evaporated.

The now [^{3}H]-labelled polyGC was precipitated from 0.5 *M* sodium chloride with ethanol and the pellet washed extensively with 70% v/v ethanol until, as assessed by scintillation counting, the [^{3}H] in the washes was minimal. The pellet was dissolved in 300 µL of 50 m*M* sodium phosphate (*p*H 11.4) and left at RT for 66 h to convert 7mG residues to FaPyG. A fraction of this corresponding to *ca.* 5000 cpm was digested (in 1× FPG buffer) with FPG protein (2.3.11) with the release of 40% of the radioactivity (data not shown), thus suggesting that purine imidazole ring-opening was incomplete. The substrate was therefore reprecipitated and dissolved in 0.2 *M* sodium hydroxide and

left to incubate at RT for 16 h, followed by precipitation from 0.5 *M* sodium chloride and washing, before being dissolved in 10 m*M* potassium-HEPES (*p*H 7.6). After this treatment 98.5% of the [^{3}H] could be released from the polymer by FPG protein (data not shown), indicating effectively complete ring-opening of [^{3}H]-7mG residues to generate [^{3}H]-methylFaPy-Gua-polyGC.

From the total radioactivity in the substrate (155×10^6 dpm; 70.1 μCi) the overall efficiency of the methylation reaction could be calculated as 1.4%. Assuming that none of the polyGC had been lost, it could also be calculated that 14.0 nmol of [^{3}H] was present in 2.5 mg of polyGC; from the specific activity of the [^{3}H]-DMS (5.0 Ci/mmol) this indicated that 5.6 nmol of [^{3}H] had reacted per mg of polyGC; 2.5 mg of polyGC contains 3.57 μmol of dG residues; thus, from the average size of the polyGC molecules (805 bp), it could be calculated that 1 in 255 dG residues had been methylated (i.e. 0.39%), corresponding to 3.1 FaPyG residues per double-stranded polyGC molecule. The [^{3}H]-FaPy-polyGC substrate was stored frozen in aliquots at –20°C.

In view of the absence of reported data on the performance of the FaPy-DNA glycosylase assay using this substrate, when used to measure such activity in mammalian cell extracts, a number of validation experiments were carried out. Extracts were made from male BDF_1 liver and kidney, and were used to assess the linearity of the assay with respect to time and TP (2.3.5) added to the reaction (Fig. 2.1). This showed that over the range of 1 to 4 h and 0 to 1.5 mg TP added the reaction was almost linear, with some evidence of diminishing response with increasing time or TP. It was therefore decided to limit the assay incubation time to no more than 2 h and the TP added to no more than 1 mg. In order to economise on substrate use, commensurate with using sufficient such that counting times were not excessive, it was decided to use 1 pmol [^{3}H] (≈ 3300 cpm) per 50 μL reaction. Scintillation counting of assay samples was carried out so that a constant 10^4 counts were accumulated, thus minimising the counting error to 1%. With

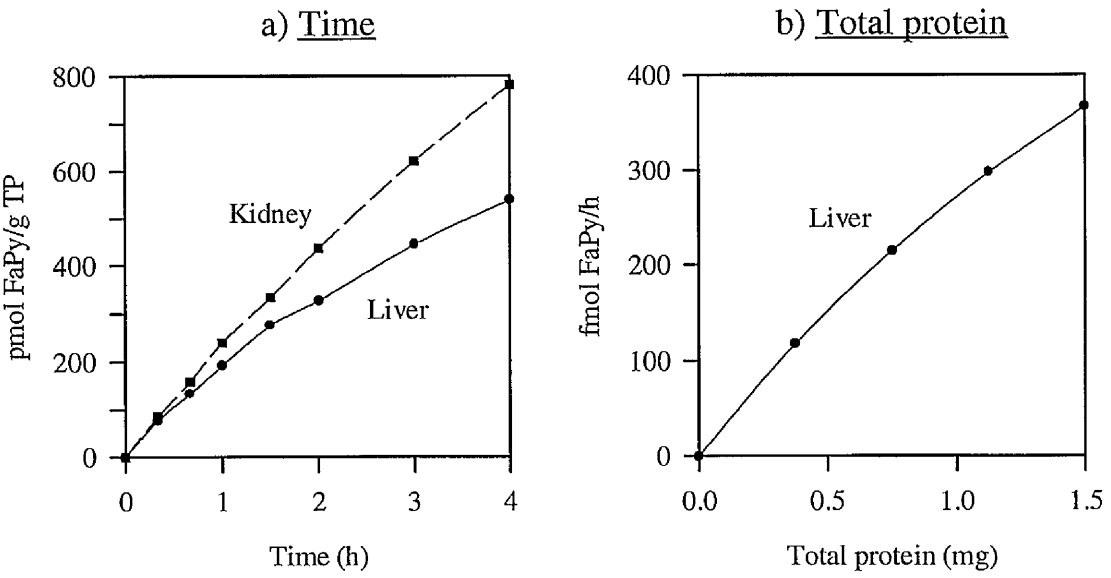
each assay four "zero" tubes were included to assess the counts released from the substrate in the absence of protein, the mean of which was deducted from the counts released from each test sample (this mean gradually increased from *ca.* 1.4 to 2.6% of total counts added over a period of about 18 months; data not shown). The limit of detection in any particular assay could also be assessed by calculating the upper 2.8× SD limit for the mean of these "zero" tubes. The total counts added was assessed in duplicate with each assay and any test samples causing release of more than 80% of the [^{3}H]–FaPy were deemed to have caused substrate depletion, and were repeated after dilution (± shorter incubation). The FaPy-DNA glycosylase activity was calculated from equation 1:

$$Activity\ (\mu U/g\ TP) = \frac{(cpm - zero) \times \frac{d}{k}}{V \times c \times t} \qquad (Eqn.\ 1)$$

Where *cpm* = radioactivity in counted supernatant in cpm, *zero* = mean counts in "zero" tubes, k = dilution factor e.g. when counting 300 µL out of a total supernatant of 375 µL k = 1.25, k = specific activity of substrate allowing for counting efficiency of β-counter (30%) i.e. 3330 cpm/pmol FaPy, V = volume of sample extract used (L), c = TP concentration in extract (g/L), and t = assay incubation time (minutes).

This was set up in the form of a computerised spreadsheet using Excel v3.0 (Microsoft, Redmond, WA, USA), which also automatically calculated the limit of detection, means of duplicate or triplicate samples (with SEM), and flagged samples suspected of causing substrate depletion. No correction was made for the change in specific activity with radioactive decay of the substrate in view of the relatively short time period involved (18 months) compared to the half-life of tritium (12.6 years).

Fig. 2.1
PolyGC-based FaPy-DNA glycosylase assay:
Linearity of response with time and total protein



Extracts from BDF_1 liver and kidney were made (2.3.3) and assayed for total protein (TP; 2.3.5). These were then used to test the linearity of response of the polyGC-based FaPy-DNA glycosylase assay (2.3.4), with a) varying time of incubation, and b) varying the TP added to the reaction.

2.3.5 **Total protein assay**

Total protein (TP) in column fractions, cell and tissue extracts was assayed by the method of Bradford (1976), which utilises the change in absorbance of Coomassie Brilliant Blue G-250 at 595 nm in acidic solution when bound to protein. Patel and Lott's modification (1985) of Bradford's original method was used. This was to add the surfactants SDS and Triton X-100, which has the effect of minimising inter-protein differences in response, improving calibration, and giving the assay a linear response over the range 0 to 4.0 g TP/L (with a sample to reagent volume ratio of 1:100), compared to 0.3 to 1.2 g/L for unmodified reagent (Patel and Lott, 1985). Calibration curves were established using 1.00 and 2.00 g/L BSA standards (plus others as necessary), made up in FPG Lysis buffer or FPG buffer as appropriate, from a 10.0 g/L standard solution of BSA in PBS (stored frozen in aliquots at -70°C). Stock CBG-250 dye reagent (5× final concentration) was prepared by dissolving Coomassie Brilliant Blue G-250 (Sigma or IBI) at 0.5 g/L in a solution of 50% v/v methanol/25% v/v orthophosphoric acid, with the addition of 75 mg/L each of SDS and Triton X-100, with storage at 4°C in the dark. Prior to use this was diluted to 1× CBG-250 with ddH_2O, leaving overnight at 4°C in the dark, and then filtering through Whatman No. 3 filter paper.

In view of the assay's linearity, the precise TP of samples with an $A_{595\ nm}$ between that of the zero and 2.00 g/L standards concentration was simply calculated from the $A_{595\ nm}$ of the 1.00 g/L standard. Total protein concentrations of samples greater than this were either reassayed after dilution or calculated from the $A_{595\ nm}$ of the 2.00 to 4.00 g/L standards, as interpolated from the graph of the $A_{595\ nm}$ *v.* TP standard concentration. To measure the TP in column fractions the assay was adapted to be carried out in 96-well plates using an ELISA plate reader set to determine $A_{600\ nm}$ (5 µL of column fraction + 200 µL of 1× CBG-250 reagent), with 0, 1.0, 2.0, 3.0 and 4.0 g/L BSA standards. The

TP concentrations of individual samples were calculated by the internal software of the instrument.

2.3.6 **Protein analysis of bacterial extracts by SDS-PAGE**

The proteins in bacterial extracts were analysed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE), according to Laemmli (1970). Bacteria, washed in ice-cold STE, were resuspended in ddH_2O and lysed by the addition of 1/4 volume of 5× SDS-PAGE loading buffer (Sambrook *et al.*, 1989). The total protein content was estimated by lysing an equivalent amount of bacteria as for the FPG assay (2.3.3) assaying the total protein and subtracting that due to the added lysozyme. A commercially available pre-stained marker mix was run with each gel (BRL), however it was later found that these did not run according to their M_r (manufacturer's data). Electrophoretic separation was achieved with a field strength of 8 to 10 V/cm. Gels were fixed and stained with Coomassie Brilliant Blue R-250 as described (Sambrook *et al.*, 1989), and analysed by densitometry using a Shimadzu CS930 dual-wavelength TLC scanner (Shimadzu Corporation, Kyoto, Japan) in transmission mode at 595 nm.

2.3.7 ***In vitro* transcription-translation of plasmid DNA**

A cell-free transcription-translation system, produced in the form of a kit (P&S Biochemicals, Liverpool, UK), was used to study the polypeptides produced from various plasmids. The kit is based on a modified Zubay system (Zubay, 1973). A crude cell-free extract from *E. coli*, containing all the enzymes and factors necessary for transcription and translation is provided, which when supplemented with amino-acids (without methionine), [^{35}S]-L-Methionine, an energy regenerating system and circular plasmid DNA will synthesize protein labelled with [^{35}S]. Each reaction was carried out as per the manufacturer's instructions, with 20 μCi of [^{35}S]-L-Methionine. An aliquot of each was

assessed for incorporation of radioactivity in protein by perchloric acid precipitation of polypeptides after hydrolysis of [^{35}S]-L-Met-tRNA with dilute alkali. In this way an equal amount of incorporated radioactivity was loaded into each lane of an SDS-polyacrylamide gel. An equivalent amount of [^{14}C]-labelled marker proteins (BRL, Gaithersburg, MD, USA) was also loaded. After electrophoresis the gel was fixed (10% v/v acetic acid, 30% v/v methanol, 3% v/v glycerol) and dried under vacuum at 80°C onto photocopier paper. An autoradiograph was made by exposure of Kodak X-Omat AR film for six days, followed by development by inspection in Kodak D-19 high-contrast developer.

2.3.8 **Induction of FPG protein expression in *E. coli* harbouring pKK223-3/*fpg* constructs**

The growth of *E. coli* JM105 clones containing pKK223-3/*fpg* constructs was monitored by means of the $A_{600\ nm}$. The medium was inoculated with 1/100th volume of stationary phase culture, when the $A_{600\ nm}$ reached ≈ 0.6 (i.e. mid-log phase, typically 2 to 3 h) the cultures were induced to overexpress protein by the addition of IPTG to the medium to a final concentration of 0.5 or 1 m*M*. After a given interval the bacteria were harvested by centrifugation (12 000 × *g*/15 minutes/4°C), followed by resuspension in ice-cold STE and recentrifugation. The washed pellets were then resuspended in ice-cold Buffer A containing 1 m*M* PMSF and sonicated (60 s × 10 μm peak-peak, then 300 s × 12 μm, in 30 s pulses with 10 s intervals) with the tube cooled in a water/ice bath. The sonicate was clarified by centrifugation (12 000 × *g*/15 minutes/4°C) before being processed further.

2.3.9 **Purification of FPG protein: DEAE-cellulose treatment**

DEAE-cellulose (Whatman, UK) was prepared according to the manufacturer's instructions and finally suspended in Buffer A (≈ 50% v/v) containing 50 m*M* or 250 m*M* sodium chloride. Sodium chloride was added to the sonicated bulk preparation of induced bacteria (2.3.8) to a final concentration of 50 m*M* or 250 m*M*, as required, and then 0.5 volumes of DEAE-cellulose slurry was added and the mixture left on ice for 10 minutes. The DEAE-cellulose was removed by centrifugation (12 000 × *g*/15 minutes/4°C), and the supernatant used immediately for cellulose-phosphate chromatography.

2.3.10 **Purification of FPG protein: Cellulose phosphate chromatography**

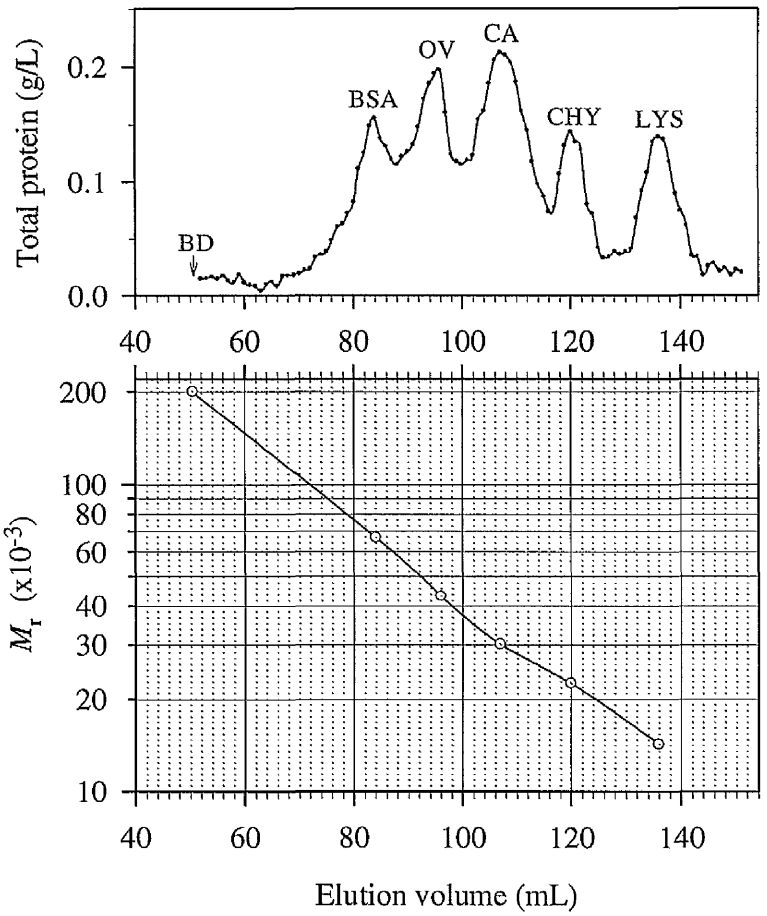
Cellulose phosphate (Sigma) was prepared according to the manufacturer's instructions, by first being suspended in 0.5 *M* sodium hydroxide (washing after with ddH_2O), then 0.5 *M* hydrochloric acid (washing after with ddH_2O), 0.4 *M* potassium phosphate (*p*H 7.6), 55 m*M* potassium phosphate (*p*H 7.6) (washing after with ddH_2O), and finally suspension in Buffer A containing 50 m*M* sodium chloride. In this form it was poured into a 4.5 cm diameter × 12 cm high column (volume 190 cm^3). This column was then washed for 16 h with Buffer A containing 50 m*M* sodium chloride, using a peristaltic pump giving a flow rate of 20 mL/h. After this the DEAE-cellulose-treated sonicate was loaded onto the column and the column developed with a gradient of sodium chloride concentration (in Buffer A) from 50 to 1000 m*M* or 200 to 600 m*M* as required (flow rate 40 mL/h), with the collection of 4 mL fractions. A new cellulose phosphate column was used in each experiment. Fractions were assayed for TP and FaPy-DNA glycosylase activity (2.3.2 and 2.3.5), and also by SDS-PAGE (2.3.6), those containing FPG protein were pooled and concentrated using a positive-pressure ultrafiltration device (Ultrafiltration cell Model 8200 with YM-10 filter (M_r cut-off at 10 000); Amicon, USA)

with compressed nitrogen gas to apply the pressure. This and the columns were kept and operated in a cold cabinet at 4°C.

2.3.11 **Purification of FPG protein: Gel filtration chromatography**

Sephacryl HR-200 (Pharmacia) was prepared according to the manufacturer's instructions in Buffer A + 50 m*M* sodium chloride and used to pour a 1.5 cm diameter × 82 cm high column (kept at 4°C). This was washed through with Buffer A + 50 m*M* sodium chloride for 24 h at 6 mL/h. A mixture of marker proteins (1 g/L each of lysozyme, α-chymotripsinogen, carbonic anhydrase, ovalbumin, BSA and Blue dextran (Sigma)) was then loaded onto the column and 1 mL fractions collected at the same nominal flow rate. From the manufacturer's data (Pharmacia) Blue dextran when subjected to gel filtration chromatography on Sephacryl migrates with an apparent M_r of 200×10^3, i.e. it elutes with the void volume; this took 8 h at a flow rate of 6.3 mL/h giving a void volume of 50.4 mL. The elution profile is shown in Fig. 2.2, which demonstrated that over the range of M_r of between 200 to 14×10^3 elution of proteins was linearly related to $\log_{10} M_r$.

Fig. 2.2
Calibration of gel filtration column:
Low salt: 50 m*M* Sodium chloride



After equilibration with Buffer A containing 50 m*M* sodium chloride, the Sephacryl HR-200 gel filtration column was calibrated with a mixture of marker proteins: BD = position of Blue Dextran elution (= "Void volume"; determined visually), BSA = bovine serum albumin, OV = ovalbumin, CA = carbonic anhydrase, CHY = bovine chymotrysinogen, and LYS = lysozyme; M_r values are those quoted by the manufacturer (Sigma). Protein concentrations determined by Coomassie Blue G-250 assay (2.3.5).

The pooled and concentrated fractions from cellulose phosphate chromatography (2.3.10) were dialysed against 2 L Buffer A + 50 m*M* sodium chloride for 24 h at 4°C, and then a portion was loaded onto the Sephacryl column. The elution profile of this suggested oligomerisation/aggregation of FPG protein (4.2.4) and so the column was equilibrated with Buffer A + 1 *M* sodium chloride for 72 h at 6 mL/h, and then re-calibrated (Fig. 4.5; 4.2.4), when it was found that the void volume had reduced slightly to 47 mL. This column was used for the final purification of FPG protein. Fractions

were assayed for TP and FaPy-DNA glycosylase, and those containing FPG protein were pooled and concentrated, as before (2.3.10). This purified preparation of FPG protein had glycerol added to a final concentration of 50% v/v and was stored at -20°C.

2.3.12 Enzyme-linked immunosorbent assay (ELISA) for anti-FPG titre

The ELISA assay of antisera was based on the method of Douillard and Hoffman (1983). Purified FPG protein (2.3.11) was diluted in PBS and used to coat the wells of 96-well ELISA plates by overnight incubation at RT. Wells were coated with 0, 1, 10, 25, 50 and 100 ng of FPG protein. Following this the excess solution was tipped out and the plates dried at 37°C, after which they were stored at -20°C wrapped in aluminium foil. Before use the wells were rinsed four times with PBST and then diluted serum (1:1000 to 1:128 000 in 50 µL PBST containing 2 g/L BSA) was pipetted into each well and left at RT for 90 minutes. The wells were then washed six times with PBST, before the addition of 50 µL of a 1:1000 dilution of horseradish peroxidase(HRP)-labelled swine anti-rabbit immunoglobulin (Sigma) in PBST, again left to incubate for 90 minutes. After six washes with PBST and one with ddH_2O 50 µL of freshly prepared ELISA substrate was added and left to incubate at RT for 15 minutes, following which the reaction was stopped by the addition of 50 µL of 1 *M* hydrochloric acid. The $A_{450\ nm}$ of the reaction mixtures was measured using an ELISA plate reader (Multiskan MCC/340 P).

2.3.13 Liquid-phase inhibition of FPG with rabbit antisera

To investigate the ability of the different antisera (2.5.2) to inhibit the FaPy-DNA glycosylase function of FPG protein, the polyGC-based assay (2.3.4) was carried out using a constant amount of purified FPG protein (70 nU; 0.2 µg TP) in a total reaction volume of 100 µL (1× FPG buffer), with and without preincubation with the different pre-immune and immune sera, in duplicate. The antisera were used in amounts varying from

0.1 μL ("1:1000") to 10 μL ("1:10"), the preimmune sera were used at the 1:10 dilution, and the total activity controls included 10 μL of 50 g/L BSA, with preincubation for 10 minutes on ice before the addition of substrate in 10 μL of 1× FPG buffer (prewarmed to 37°C). The reaction was allowed to proceed for 10 minutes before being stopped by the addition of 25 μL FPG Stop solution plus 270 μL ethanol (2.3.2). Activity in the tubes containing sera was expressed as a percentage of the mean activity in the total activity controls.

For the investigation of the ability of the antisera to inhibit mammalian FaPy-DNA glycosylase a constant amount of BALB/c liver (1.28 nU FaPy-DNA glycosylase) or human kidney (1.26 nU) extract was used, again with preincubation on ice for 10 minutes. The activities in this experiment were expressed as a percentage of the mean of the control tubes for each extract, which in this case did not contain added BSA.

2.3.14 **Western analysis**

For western analysis (Dr R.H.Elder, personal communication), mammalian cell and tissue extracts (made for FaPy-DNA glycosylase assay; 2.3.3) were subjected to SDS-PAGE using 12% w/v polyacrylamide gels (2.3.6). The marker protein mix used (Rainbow™ markers, low range; Amersham), however, was one in which the apparent M_r of the individual components was defined, unlike that previously employed (2.3.6). The proteins in the gels were transferred by electroblotting, using western transfer buffer (100 V, 1 h, 4°C), to nitrocellulose membranes (Hybond C, Amersham). After a brief rinse in TBS, the membranes were blocked using western blocking buffer (30 minutes, RT), with gentle continuous agitation (in this and all subsequent steps). The primary antisera were diluted 1:1000 in blocking buffer and incubated with the blocked membranes for 2 to 3 h at RT, following which the blots were washed with TBST (3 × 10 minutes). The membranes were then incubated with a 1:1000 dilution of the

secondary antibody (HRP-labelled goat anti-rabbit immunoglobulin; Sigma) in blocking buffer (1 h, RT), followed by washing with TBST (3 × 10 minutes) and then TBS (3 × 10 minutes). Excess TBS was blotted off (without drying the membrane) immediately prior to detection by enhanced chemiluminescence (ECL Kit; Amersham). Blots were exposed to Kodak X-Omat AR film for between 20 s and 10 minutes, as necessary, and developed as for Southern analysis (2.2.20).

2.4 CELL CULTURE METHODS

2.4.1 General techniques

The RJKO cell line (2.1.7) was grown as a monolayer. Cells were plated at an approximate density of 1×10^6 per 75 cm^2 flask in 10 mL of medium (2.1.8; *pro rata* for other sizes of flask: 25 cm^2 and 175 cm^2 flasks were also used), and the medium was changed twice weekly. Incubation was at 37°C in a humidified atmosphere of 95% air/5% CO_2. At or approaching confluence of the cells, they were subcultured by removing the medium, washing with PBS and then adding 1 mL 0.05% w/v trypsin in PBS to release the cells from the floor of the flask (typically 3 to 5 minutes at RT). The trypsin was neutralised by the addition of 10 mL of medium and the cells harvested by centrifugation (1500 × *g*/5 minutes/RT), following which the pellet was resuspended in 1 mL of medium. The cells were counted by means of a haemocytometer, if required, before being added to fresh medium in a new flask. Freezing, thawing and storage of cells were carried out according to Freshney (1987).

2.4.2 Mammalian cell transfection and clonal selection

Lipofectin is an equimolar mix of the cationic lipid N-(1-(2,3-dioleyloxy)propyl)-N,N,N-trimethylammonium chloride and dioleoylphosphatidylethanolamine, which forms lipid-DNA complexes facilitating the introduction of DNA into cells *in vitro* (Felgner

et al., 1987). Transfection of cells was carried out as follows: RJKO cells were passaged in MEM + 10% FCS and allowed to grow to approximately 50% confluence in 75 cm^2 flasks. Plasmid DNA (20 μg) in sterile water (50 μL) was diluted with 2.5 mL Optimem serum-free medium (BRL, Gaithersburg, MD, USA): 50 μL of Lipofectin reagent (BRL, Gaithersburg, MD, USA) was similarly diluted in 2.5 mL of Optimem, added to the 2.5 mL of diluted plasmid and left to stand at RT for 30 minutes. In the meantime, the RJKO cells were washed with Optimem to remove traces of serum which interfere with the lipofection process. The Optimem containing the plasmid/lipofectin complex was then added to the cells and the flask returned to the incubator for 24 h. After this time 5 mL of MEM + 20% FCS was added and the cells grown to near confluence. The medium was then replaced with MEM + 10% FCS containing 1 g/L G–418 to enable selection of resistant cells. After eight days, with medium changes every other day, visible colonies were picked. This was done by cutting open the flask with a hot wire, removing the bulk of the medium and placing sterile 5 mm cloning rings around individual colonies. Into these was placed 200 μL of 0.05% w/v trypsin in PBS, left to incubate at RT for 5 minutes. (For PCR analysis of clones at this stage 10 μL was taken of this suspension; 2.2.4.) The detached cells were transferred to 25 cm^2 flasks containing 5 mL of medium, before being subcultured into 75 cm^2 flasks (with medium containing 0.2 g/L G–418, as required).

2.4.3 **Treatment with HMGCR inhibitors**

To convert the HMGCR inhibitor Simvastatin (SV) into its active hydroxy-acid form (SVA) it was treated with normal rat serum (NRS; 2.1.8; Vickers *et al.*, 1990). Stock SV (100 m*M* in DMSO) was added to medium (MEM + 10% FCS + 20% NRS) to a final concentration of 0.1 m*M* (i.e. 1:1000), incubated at 37°C for 16 h, and used immediately. For the treatment of RJKO and clone 10H4 the cells were subcultured (2×10^6 per

175 cm^2 flask) and then Pravastatin or SV added to the medium (MEM + 10% FCS) 16 h later. Treatment with SVA was achieved by exchanging the MEM + 10% FCS medium for MEM + 10% FCS + 20% NRS + 0.1 m*M* SV (diluted as necessary with MEM + 10% FCS + 20% NRS). The cells were harvested (2.4.2) 48 h later, and extracts made (2.3.3) for TP and FaPy-DNA glycosylase assay in 0.5 mL 1× FPG assay buffer containing 1 m*M* PMSF (2.3.4, 2.3.5).

2.4.4 **Treatment with photodynamic agents**

RJKO cells were treated with photodynamic agents as follows: cells were plated at the required density (e.g. 200/6 cm × 1.5 cm dish in 3 mL MEM + 10% FCS) and allowed to adhere for 4 h. Conditions of minimal lighting for further handling of the cells were achieved by turning off the fluorescent lighting in the laminar flow cabinet, excluding daylight by covering the room windows with black plastic sheeting and restricting artificial lighting to indirect illumination by a 40 W tungsten bulb placed *ca.* 3 m away. After the cells had adhered the medium was removed, the cells washed once with 5 mL EBSS, then 3 mL EBSS containing the dye (MB, NMB, or BRB) added, and the dishes (with lids on) immediately placed either in a light-tight metal box ("Dark" conditions), or under the light source (an Ilford 400HL photographic enlarger with both colour filters removed but the diffuser/infra-red filter present, 16 cm above the dishes), at RT (22°C, maintained by air-conditioning). The surface on which the dishes were placed for light treatment was of white plastic laminate; six dishes were treated at any one time, in a 2 × 3 array. After treatment the dye solution was removed, the cells were briefly washed again with 5 mL EBSS, and then the MEM + 10% FCS medium (5 mL) was replaced and the dishes immediately placed back in the incubator. Surviving colonies were counted 5 to 7 days later (medium change every 2 to 3 days) after Giemsa staining (Freshney, 1987).

2.5 ANIMAL AND TRANSGENIC MOUSE METHODS

2.5.1 Immunisation of rabbits and antiserum production

Polyclonal FPG antisera were obtained by immunisation of rabbits, carried out according to the recommendations of Harlow and Lane (1988). Purified FPG protein (2 mg; 2.3.11) was dialysed for 24 h at 4°C with two changes of 1 L of sterile PBS containing 5% v/v glycerol, and then emulsified with an equal volume of incomplete Freund's adjuvant by rapid transfer between two 1 mL syringes connected at 90° by a three-way tap. This mixture (0.5 mL equivalent to 1 mg FPG/animal) was then injected sub-cutaneously at multiple sites on the flank and back limbs. Pre-immune sera were obtained from marginal ear veins 24 h prior to the first injection, and similarly immune sera were obtained 10 days after the second and third immunisations. Ten days after the fourth (final) immunisation the animals were exsanguinated by cardiac puncture while under general anaesthesia (administered intravenously). The blood samples were allowed to coagulate in glass vessels for 2 to 3 h at 37°C, the clots were separated from the walls with a pipette and then were left at 4°C overnight to maximise clot retraction. The serum was then drawn off and the red cells and other debris removed by centrifugation (12 000 × *g*/30 minutes/4°C), following which it was filtered under pressure, first through a 0.45 µm filter and then a 0.22 µm filter (Acrodisc, Gelman, USA). Sera were stored frozen in aliquots at -20°C.

2.5.2 General transgenic animal techniques

All techniques involved in transgenic mouse production were as described in Hogan *et al.* (1986). In summary, eggs were collected into M2 medium containing hyaluronidase, from superovulated females who had mated the night before with stud males of the same strain (2.5.3). After removal of the follicular cells the eggs were washed in M2 medium, placed into drops of M16 medium under oil, and incubated in

humidified 95% air/5% CO_2 at 37°C. Microinjection of DNA into the male pronuclei of the fertilised zygotes then took place, and those surviving were incubated overnight in M2 medium (2.5.4). The next morning pseudopregnant females were selected from those females housed with vasectomised males by examination for vaginal plugs. The surviving 2-cell embryos were then implanted into these females by oviduct transfer (2.5.5). Pups were generally weaned at 3 to 4 weeks at which time tail tip biopsies were taken (2.5.6), and the DNA extracted (2.2.23).

2.5.3 **Superovulation of mice**

Virgin female mice were stimulated to superovulate by sequential treatment with gonadotrophins: 5 IU Folligon i/p (acting as FSH) between 13:00 and 15:00, followed by 5 IU Chorulon i/p (acting as LH) 47 h later (on a 07:00 to 19:00 light/dark cycle), immediately after which they were each placed with a separate stud male (which were not used more than twice a week). The next morning the animals were killed by cervical dislocation and the oviducts removed into M2 medium containing 300 mg/L of hyaluronidase, whence the eggs were dissected out and the follicular cells surrounding them allowed to fall free (typically 3 to 5 minutes at RT). The eggs were then collected up and placed in drops of M16 medium under oil, and incubated (2.5.2).

2.5.4 **Microinjection of zygotes**

Fertilised zygotes were selected a few hours after collection (*ca.* 13:00 to 14:00 for BDF_2, and 14:30 to 15:30 for B6) when the pronuclei became visible, and microinjection of the male (larger) pronucleus carried out. Microinjection buffer (with or without DNA) was centrifuged (25 000 × *g*/10 minutes/4°C) to sediment debris, and only the upper part of the supernatant was used to back-fill microinjection needles. Microinjection was judged successful when the pronucleus was observed to increase in diameter by 1.5 to

2–fold; care was taken to avoid the pronucleoli; zygotes were injected in batches of approximately 24 to minimise time spent out of the incubator, and in M2 (HEPES-buffered) rather than M16 (bicarbonate-buffered) medium.

2.5.5 Implantation of embryos into pseudopregnant females

A supply of pseudopregnant BDF_1 females was obtained by housing 8 to 16 week old animals with vasectomised BDF_1 males: one male to four females. The females were randomly reassigned to the cages every 2 to 3 weeks to stimulate oestrous cycling. Pseudopregnant females were selected by examination for the presence of a vaginal plug; any males consistently failing to plug were replaced. The surviving zygotes (2.5.4) were transferred into the oviducts of the anaesthetised pseudopregnant females, after these had been brought out of the abdomen through a para-lumbar incision on each side, made *via* a single dorsal skin incision. Care was taken to avoid blood vessels when dissecting open the bursa membrane which covers the ovary and oviduct entrance. The para-lumbar muscle incisions were closed with 5/0 silk sutures, the dorsal skin incision with stainless steel clips (2 or 3). The females were allowed to recover from the anaesthetic and then rehoused (singly or in pairs) in cages on the highest shelf in the animal room. The ventilation there was from top to bottom, and this minimised the possibility of the females being disturbed by the scent of males (only housed on the lower shelves).

2.5.6 Tail tip biopsies

Large (1 to 2 cm) tail tip biopsies were taken under light general anaesthesia (Enflurane), by first ligature with 3/0 linen and then sectioning with a sterile scalpel (with

a new blade for each animal). Small (≯ 0.5 mm) tail-tip biopsies were taken under the same conditions except that no ligature was necessary.

2.5.7 **Partial hepatectomy**

Mice for partial hepatectomy (Higgins and Anderson, 1931; Roberts, 1978; Wright, 1984) were anaesthetised with i/p Avertin (Hogan *et al.*, 1986). The abdomen was opened by a transverse incision between the lower costal margins, care being taken to avoid the superior epigastric vessels. The anterior lobes (median and left lateral) of the liver were mobilised and brought out through the opening, when a 3/0 linen ligature was passed around them and tightened on their origin, which could now be sectioned and removed within 2 to 3 mm of the ligature, and the stump returned to the abdomen. The peritoneum and anterior abdominal musculature was closed with 5/0 silk sutures, and skin closure effected with stainless steel clips. Animals were allowed to recover under the warmth of a 60 W tungsten light bulb.

CHAPTER 3

ISOLATION AND OVEREXPRESSION OF *fpg* IN *Escherichia coli*

3.1 GENERAL STRATEGY

The following strategy was adopted in order to obtain the protein-coding region of the *fpg* gene in a form suitable for ligation into pro- and eukaryotic expression vectors. Initial isolation and amplification of the gene would be achieved by means of the polymerase chain reaction (Saiki *et al.*, 1985) using oligonucleotide primers and *E. coli* genomic DNA. The design of the primers would:

1) Allow amplification of the protein-coding region of the gene devoid of potentially interfering up- and downstream prokaryotic controlling sequences,

2) Remove an upstream and out-of-frame initiation codon, and

3) Introduce unique restriction endonuclease sites to facilitate ligation into plasmid vectors.

Changes in the nucleotide sequence of the gene enabling 2) and 3) would be generated by specific mismatches in the base-pairing of the oligonucleotides, when compared with the genomic sequence, i.e. a form of site-directed mutagenesis. The DNA fragment so obtained could then be digested, using the restriction endonucleases specific for the introduced sites, and ligated into a suitable intermediate vector. Choice of such a vector would permit clonal amplification of the gene and checking of the nucleotide sequence, the latter necessary in view of the low, but finite, error rate of the non-proof-reading *Taq*-polymerase used in PCR. Cloning into a suitable vector would also add a number of other useful restriction endonuclease sites; a number of eukaryotic expression vectors only allow insertion of the desired sequences into a unique *Bam*H I site, hence the introduction of *Bam*H I compatible restriction endonuclease sites at each end of the gene would be required.

Having checked the DNA sequence of the insert it would then be possible to engineer the gene into a variety of other vectors. A prokaryotic expression vector construct containing *fpg* would enable the production of milligram quantities of FPG protein, which after purification could be used for subsequent antibody generation and the study of physico-chemical properties of the enzyme. By means of suitable eukaryotic expression vectors the *fpg* gene could then be introduced into mammalian cells *in vitro*, as well as allowing the production of *fpg*-transgenic mice.

3.2 PCR SITE-DIRECTED MUTAGENESIS OF THE *fpg* GENE

3.2.1 Oligonucleotide primer design

The *E. coli fpg* gene sequence according to Boiteux *et al.* (1987), is shown in Fig. 3.1. This indicates the position of possible promoter sequences (Pribnow, 1975), ribosome binding site (RBS) (Shine-Dalgarno sequence; Shine and Dalgarno, 1975), potentially regulatory palindromic sequences and selected restriction endonuclease sites. The site of mRNA transcription initiation (Boiteux *et al.*, 1990) and the upstream out-of-frame initiation codon are also shown. The out-of-frame ATG overlaps with the RBS and no suitable alternative RBS is upstream of it: hence it does not act as a translation initiation codon in *E. coli*, however it might cause problems if it was cloned into an expression vector, next to a prokaryotic RBS, without the upstream controlling sequences. For expression in mammalian cells, therefore, it would be necessary to delete it as it could cause out-of-frame message translation in the eukaryotic environment.

Fig. 3.1 The DNA sequence of the *E. coli fpg* gene

The sequence is as given in Boiteux *et al.* (1987). Untranslated sequence is shown in lower case. Potential stem-and-loop structures are shown: --> <--. The underlined sequences labelled **-35** and **-10** are transcription regulation sites, **TS>** indicates the transcription start site and **RBS** is the ribosome binding site. Selected restriction endonuclease sites are shown underlined and labelled accordingly.

```
                                1                  20                  40
                                .                  .                   .
                               5'-gatctacaaagaagcgaaaatcaaataattctcgctttga

                  60                   80                  100
          ---------->    <----------    .                   .
tgtaacaaaaaaacctcgctccggcggggtttttgttatctgcttgcccccatattgactgcatctgttc
                                  -35                  -10  TS>
       120                  140                 160                 180
        .                    .                   .                   .
attcctggagatgctATGCCTGAATTACCCGAAGTTGAAACCAGCCGCCGCGGCATAGAACCGCATCTCG
      RBS
                  200                  220                 Bgl II/Xho II
        .                    .                   .                   .
TTGGTGCAACCATTCTTCATGCAGTGGTGCGCAACGGACGCTTGCGCTGGCCGGTTTCAGAAGAGATCTA

       260                  280                 300                 320
        .                    .                   .                   .
CCGTTTAAGCGACCAACCAGTGCTTAGCGTGCAGCGGCGGGCTAAATATCTGCTGCTGGAGCTGCCTGAG

                  340                  360                 380
        .                    .                   .                   .
GGCTGGATTATCATTCATTTAGGGATGTCTGGCAGCCTGCGCATCCTTCCAGAAGAACTTCCCCCTGAAA

       400                  420                 440                 460
        .                    .                   .                   .
AGCATGACCATGTGGATTTGGTGATGAGCAACGGCAAAGTGCTGCGCTACACCGATCCGCGCCGCTTTGG

                 BstX I              500                 520
        .                    .                   .                   .
TGCCTGGCTGTGGACCAAAGAGCTGGAAGGGCATAATGTGCTGACCCATCTTGGACCGGAGCCGCTTAGC

       540                  560                 580                 600
        .                    .                   .                   .
GACGATTTCAATGGTGAGTATCTGCATCAGAAGTGCGCGAAGAAAAAAACGGCGATTAAACCGTGGCTGA

                  620                  640                 660   BamH I
        .                    .                   .                   .
TGGATAACAAGCTGGTGGTAGGGGTAGGGAATATCTATGCCAGCGAATCACTGTTTGCGGCGGGGATCCA

       680                  700                 720                 740
        .                    .                   .                   .
TCCGGATCGGCTGGCGTCATCACTGTCGCTGGCAGAGTGTGAATTGTTAGCTCGGGTGATTAAAGCGGTG

                  760                  780                 800
        .                    .                   .                   .
TTGCTGCGTTCGATTGAGCAGGGTGGTACAACGCTGAAAGATTTTCTGCAAAGTGATGGTAAACCGGGCT

       820                  840                 860                 880
        .                    .                   .                   .
ATTTCGCTCAGGAATTGCAGGTTTACGGGCGAAAAGGTGAGCCGTGTCGGGTGTGCGGTACGCCGATTGT

                  900                  920                 940
        .                    .                   .                   .      ------
GGCGACTAAACATGCGCAGCGGGCAACGTTTTATTGTCGGCAGTGCCAGAAGtaattcatgcgcgccgga

       960                  980                1000
--->  <----------           .                   .                   .
tggcataccatccggcataaacgctacgctaacttcgccatcagcgcctgatggacattc-3'
```

Two oligonucleotides, designated 184 and 185 (Fig. 3.2), were therefore designed (and synthesized; 2.2.2) which spanned the protein-coding region but excluded controlling sequences. By means of base mismatches they would introduce unique restriction endonuclease sites (*Pst* I at the 5′, *Nsi* I at the 3′ end) to enable ligation into the intermediate vector pUCD (3.3.1). The 5′-oligonucleotide primer 184 would also eliminate the out-of-frame start codon. The T_m of the primers was estimated by allowing 4°C per G/C pair plus 2°C per A/T pair. Mismatched bases were taken not to affect the T_m and therefore were not included in the calculation.

Fig. 3.2
The Oligonucleotide Primers 184 and 185

a) Oligonucleotide 184 (19-mer; T_m = 54°C)

```
                              120                    140
                               .                      .
E. coli genomic DNA:  5′...ATTCCTGGAGatgCTATGCCTGAATTACCC...3′
                                 *  *     |-----> Translation
Oligonucleotide 184:        5′- CCTGCAGGTGCTATGCCTG -3′
                                Pst I
```

b) Oligonucleotide 185 (22-mer; T_m = 60°C)

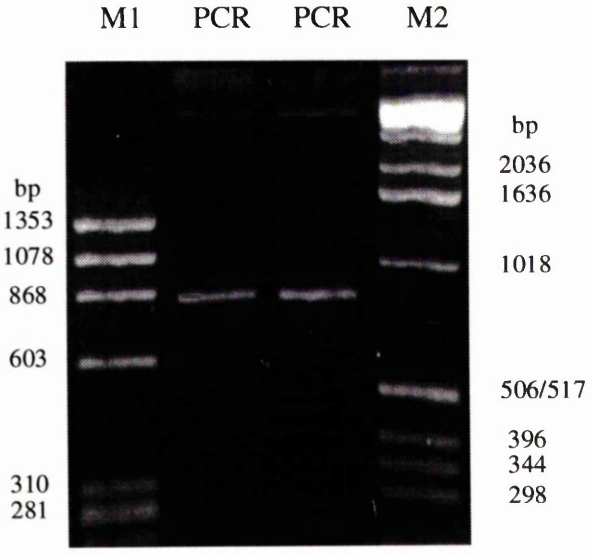
```
                          920                    940
                           .                      .
E. coli genomic DNA:  5′...CAGTGCCAGAAGTAATTCATGCGCGCCGGA...3′
                                                  **
Oligonucleotide 185:        3′- GGTCTTCATTAAGTACGTACGG -5′
                                           Nsi I
```

Note: 1) The *E. coli* genomic sequence numbering is taken from that in Fig. 3.1.
2) The genomic out-of-frame start codon is in lower case (atg).
3) Mismatched bases in the primers are shown in **bold**.
4) Introduced restriction endonuclease sites are shown underlined.

3.2.2 **Amplification of *fpg* using oligonucleotides 184 and 185**

Escherichia coli B strain F26 genomic DNA was isolated and purified (2.2.3) for use in PCR experiments. The annealing temperature (T_a) used in the initial PCR was 45°C, 9°C below the lower calculated T_m of the two primers (2.2.5). The maximum potential yield of *fpg* sequence, using 15 cycles of amplification, from 1 µg *E. coli* genomic DNA, was estimated at 2 to 3 µg; the size of the DNA fragment was predicted to be 834 bp (947 minus 113 bp; see Fig. 3.2). Analysis of the initial PCR product by agarose gel electrophoresis showed a sharp band which migrated at a rate corresponding to 830 (± 20) bp, when compared to markers of known size (Fig. 3.3; 2.2.6). From the intensity of the band on the gel the total yield of this fragment was estimated at approximately 1 µg.

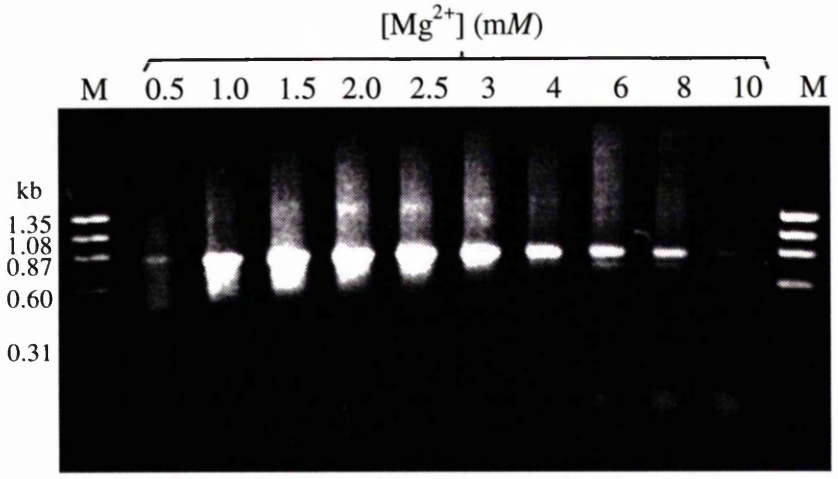
Fig. 3.3
The result of the PCR using *E. coli* DNA and oligonucleotide primers 184 and 185



E. coli F26 DNA was used as the template for a PCR using the oligonucleotide primers 184 and 185, in order to amplify the protein-coding region of the *fpg* gene and at the same time introduce base changes by site-directed mutagenesis (see text for details). Reaction conditions were: 1 µg *E. coli* F26 DNA, each [primer] = 0.25 µ*M*, [Mg^{2+}] = 1.5 m*M*, 50 µL total volume, T_a = 45°C, 15 cycles (see 2.2.5 for further details), 4 µL of product loaded per lane ("PCR"). Markers: "M1" = ϕX174/*Hae* III, "M2" = 1 kb ladder. The product obtained was of the predicted size for successful amplification from the *fpg* gene, i.e. 830 bp.

In view of this apparently successful amplification of DNA of the correct size, it was considered useful to optimise the reaction yield in relation to the magnesium concentration. It was considered that the greatest fidelity during amplification would be found in PCRs giving the highest yield of the desired product and containing as little as possible of side reaction products. In the initial experiment the total concentration of magnesium was 1.5 m*M*, however, as each dNTP binds an equimolar amount of magnesium the effective "unbound" concentration would have been 1.1 m*M*. The PCR was therefore repeated, but with 30 cycles of amplification and the total magnesium concentration varied from 0.5 m*M* to 10.0 m*M* (Fig. 3.4). It can be seen that the major product at all magnesium concentrations is the band at 0.83 kb, as found in the initial experiment. The yield of this band shows a broad maximum between 1.0 and 3.0 m*M*; minimal amounts are present at 0.5 m*M* ("unbound" $[Mg^{2+}]$ = 0.1 m*M*) and 10.0 m*M*, presumably due to the very low "unbound" magnesium level in the first case and some form of reaction inhibition in the other. In contrast, a minor side-reaction product of approximately 0.7 kb is evident at 6.0 and 8.0 m*M*. The 0.83 kb DNA fragment from the four reactions carried out with 1.0 to 2.5 m*M* magnesium was isolated by purification from LMP-agarose and stored for use in ligation experiments (2.2.7).

Fig. 3.4
The effect of magnesium concentration on the *fpg* PCR



The *fpg* PCR (2.2.5 and Fig. 3.3) was repeated, but with the [Mg^{2+}] concentration varied from 0.5 to 10 m*M* and 30 cycles instead of 15 being performed; 2.5 µL of product loaded per lane. Maximum yield of the 0.83 kb product was obtained with [Mg^{2+}] concentrations between 1 and 3 m*M*.

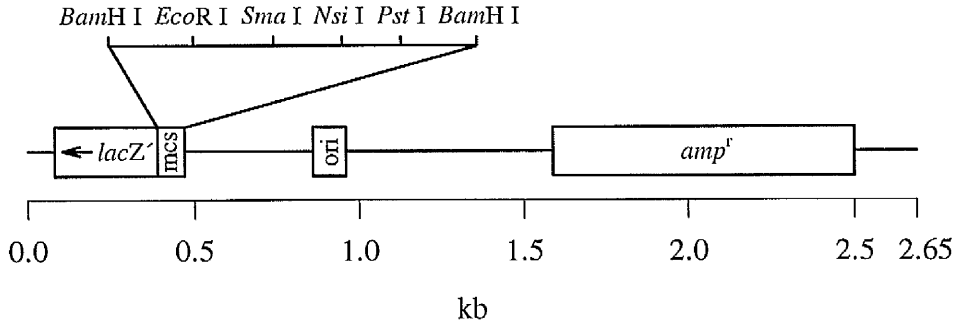
3.3 CLONING OF THE 0.83 kb PCR FRAGMENT IN pUCD

3.3.1 Choice of pUCD as the intermediate vector

The choice of intermediate vector in which to amplify the PCR product was dependent on two factors: first, it should introduce *Bam*H I (compatible) sites at both ends of the inserted DNA and, secondly, it should facilitate direct sequencing of the insert. The intermediate vector pUCD (Fig. 3.5) was produced in the Paterson Institute (Clark, 1989) from the small, high copy-number, *E. coli* plasmid cloning vehicle pUC19 (Yanisch-Perron *et al.*, 1985). By replacing the pUC19 multiple cloning site (from *Hin*d III to *Eco*R I) with one flanked by two *Bam*H I sites, Clark produced a vector which would enable the generation of recombinant DNA fragments with a *Bam*H I site at each end, facilitating their ligation into other plasmids containing a unique *Bam*H I (compatible) site. The other restriction endonuclease sites in the pUCD polylinker allow a variety of cloning strategies to be adopted when either inserting fragments or cutting them out for use in other plasmids. In addition, pUCD has the advantage of permitting sequencing of

inserts in the polylinker with commercially available pUC/M13 universal sequencing primers.

Fig. 3.5
The intermediate plasmid vector pUCD



Key: *lacZ′* = β-galactosidase, mcs = multiple cloning site: shown expanded with those restriction endonuclease sites used in cloning *fpg* (in this and all subsequent plasmid maps, only those sites used in this study are indicated, although all occurrences of a particular enzyme's site(s) are shown), ori = ColE1 origin of replication, *amp*^{r} = β-lactamase. The plasmid is shown as if linearised (its native form iscircular), the forward and reverse pUC/M13 primer sites flank the mcs (thus allowing sequencing of inserts ligated into restriction sites in the mcs, ← = direction of transcription (5′ to 3′). pUCD was derived by Clarke (1989) from pUC19 by replacement of the mcs.

3.3.2 **Ligation of the 0.83 kb PCR product into pUCD**

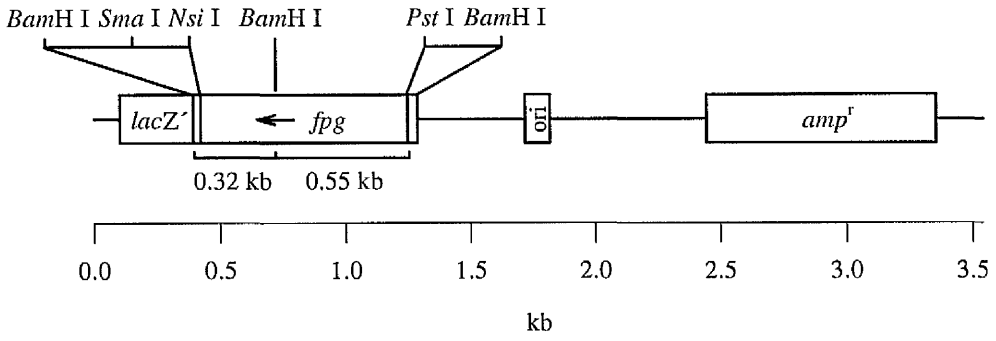
Two cloning strategies were adopted in order to clone the 0.83 kb PCR product in pUCD. It was decided to ligate the PCR product directly into the *Sma* I site of the pUCD polylinker by blunt-end ligation (*Sma* I cuts dsDNA to leave flush ends to the two strands, i.e. a "blunt" end), as well as using the *Pst* I and *Nsi* I sites as originally planned (ligation *via* overlapping or cohesive ends) (Fig. 3.6). This was in view of the proximity of the engineered restriction endonuclease sites, *Pst* I and *Nsi* I, to the ends of the 0.83 kb PCR product (particularly the *Pst* I site), which meant that there was a possibility that digestion with the restriction endonucleases might prove inefficient or even impossible.

Fig. 3.6

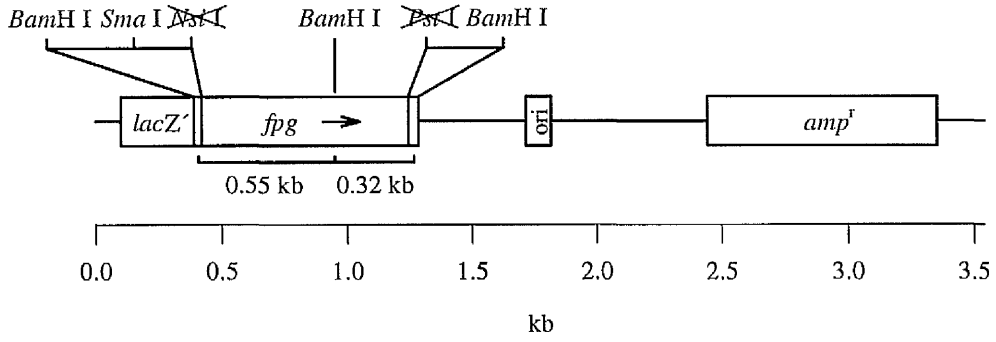
The predicted structures of the pUCD/*fpg* constructs

a) Cohesive-end ligation (pUIF):

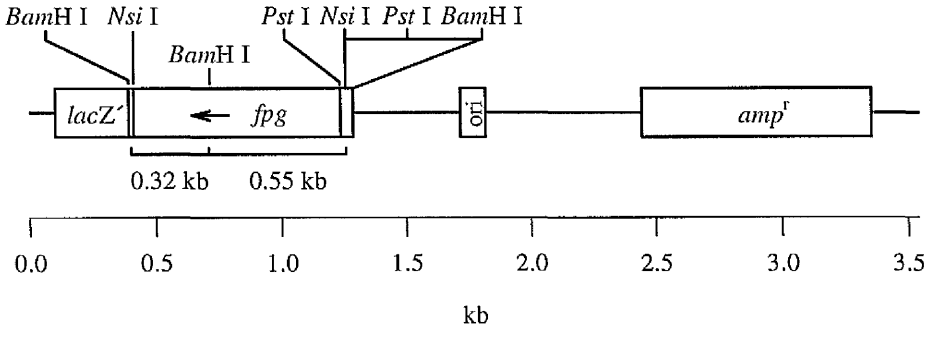
i) Correct orientation



ii) Incorrect orientation



b) Blunt-end ligation (pUIFb): correct orientation



Key: *lacZ′* = β-galactosidase, *fpg* = *E. coli* FaPy-DNA glycosylase gene obtained by PCR (see text for details), ori = ColE1 origin of replication, *amp*^{r} = β-lactamase. The plasmids are shown as if linearised, ← = direction of transcription (5′ to 3′). Two strategies were taken to ligate the *fpg* PCR product into pUCD: a) by cohesive-end ligation into the *Pst* I and *Nsi* I sites in the mcs, or b) by blunt-end ligation into the *Sma* I site.

The blunt-end ligation strategy necessitated digesting pUCD with *Sma* I and CIAP (2.2.11 and 2.2.13). As the primers used in the PCR were not 5′-phosphorylated, T4 polynucleotide kinase was used to 5′-phosphorylate the purified 0.83 kb PCR product so that it could be ligated into 5′-dephosphorylated vector (2.2.12). The cohesive-end ligation was carried out by digesting pUCD, and separately some of the 0.83 kb PCR product, with *Pst* I and *Nsi* I. The cut plasmid was also treated with CIAP. Three ligation reactions were set up with each kind of cut vector: 1) cut, but not CIAP treated, plasmid, 2) CIAP treated plasmid, and 3) CIAP treated plasmid plus appropriately treated PCR product (2.2.14). Competent *E. coli* MC1061 were transformed to ampicillin resistance (2.2.17 and 18) with the products of the ligation reactions (Table 3.1).

Table 3.1

Efficiency of *E. coli* MC1061 transformation by pUCD/PCR product ligations

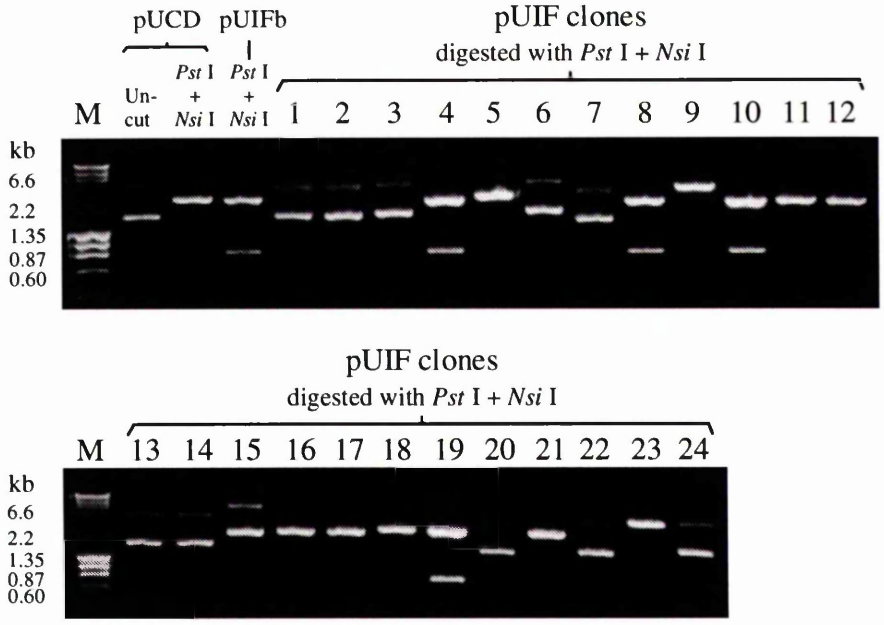
	Vector alone	**Vector(CIAP) alone**	**Vector(CIAP)+PCR product**
pUCD/*Sma* I	12×10^6 [a]	60×10^3	250×10^3
pUCD/*Pst* I/*Nsi* I	8×10^6	20×10^3	80×10^3
pUCD[b]	50×10^6	n/a	n/a
No vector	0	n/a	n/a

[a] values given are ampicillin resistant colonies per μg vector DNA, [b] undigested purified supercoiled plasmid, n/a = not applicable.

Small-scale plasmid preparations (2.2.9) were made from 24 colonies (designated pUIF1 to 24) from the cohesive-end ligation. A combination of *Pst* I and *Nsi* I was used to check the orientation of the insert: because *Pst* I and *Nsi* I produce the same cohesive end it was possible for the insert to have been ligated in the opposite orientation to that intended, but this would not reform the restriction endonuclease sites (Fig. 3.6a-ii). The agarose gel electrophoresis results of the digested pUIF plasmids is shown in Fig. 3.7. A recombinant plasmid from the blunt-end ligation (pUIFb) was included as a positive

control. It can be seen that four (17%) of the 24 clones (pUIF4, 8, 10 and 19) contained recombinant plasmid with the insert oriented as intended (as in Fig. 3.6a-i). Eleven clones were resistant to digestion (pUIF1-3, 6, 7, 13-15, 20 and 22), while nine (pUIF5, 9, 11, 12, 16-18, 21 and 23) gave single bands consistent with simple linearisation, with clones pUIF9 and 23 giving bands larger than linearised pUCD suggesting ligation of insert but with reformation of only one restriction endonuclease site. The four plasmids containing correctly ligated insert were designated pUIF4, pUIF8, pUIF10 and pUIF19 respectively. A large-scale preparation of pUIF4 (2.2.10) was made for further analysis.

Fig. 3.7
Restriction analysis of cohesive end-ligated plasmids



The structures of the plasmids was tested by making small-scale preparations of them from the various *E. coli* MC1061 clones, and then digesting the DNAs with a combination of *Pst* I and *Nsi* I. "pUIFb" = plasmid derived by blunt-end ligation, "pUIF" = plasmids derived by cohesive-end ligation.

3.3.3 **Initial analysis of the pUIF clones**

The predicted sequence of the *fpg* gene after site-directed mutagenesis by PCR with oligonucleotides 184 and 185, digestion with *Pst* I and *Nsi* I, ligation into pUCD and partial digestion with *Bam*H I, is shown in Fig. 3.8. The nucleotide numbering in this figure will be used in all subsequent discussion about the sequence of the inserts in the different pUIF clones. To facilitate sequencing (2.2.19) it was envisaged to take advantage of the "internal" *Bam*H I site (at position 554) for sub-cloning of the insert. Partial digestion of pUIF4 with *Bam*H I was predicted to give fragments of 0.85 kb, 0.55 kb and 0.30 kb, in addition to parent plasmid (2.65 kb). The 0.85 kb fragment would be used for cloning into other unique *Bam*H I site-containing vectors, while the 0.55 kb and 0.30 kb fragments would be sub-cloned in pUCD for sequencing.

Fig. 3.8

The predicted sequence of the 0.85 kb *fpg* gene DNA fragment

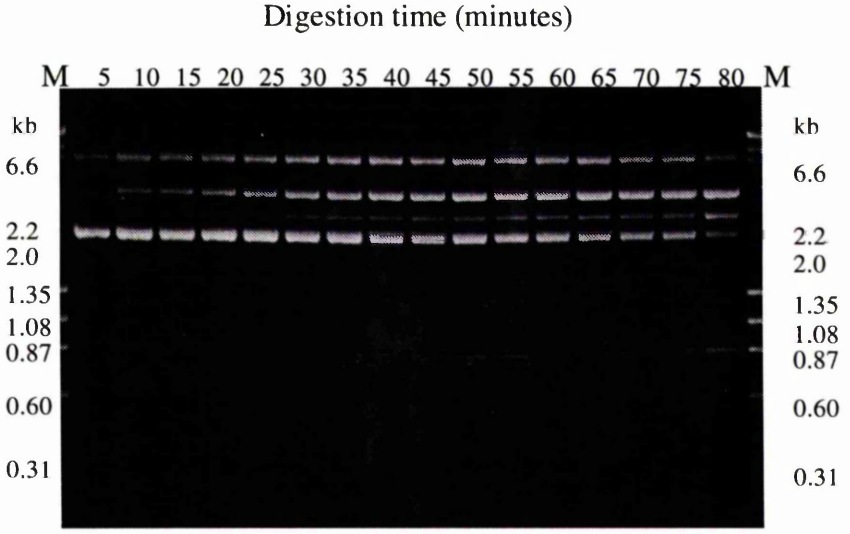
The DNA fragment obtained by site-directed mutagenesis of the *E. coli fpg* gene using PCR with oligonucleotides 184 and 185, digestion with *Pst* I and *Nsi* I, ligation into pUCD and partial digestion with *Bam*H I:

```
1     Pst I          20                  40                  60
.     .              .                   .                   .                   .
gatcctgcaggtgctATGCCTGAATTACCCGAAGTTGAAACCAGCCGCCGCGGCATAGAACCGCATCTCG

         80                  100                 120         Bgl II/Xho II
         .                   .                   .           .                   .
TTGGTGCAACCATTCTTCATGCAGTGGTGCGCAACGGACGCTTGCGCTGGCCGGTTTCAGAAGAGATCTA

                  160                 180                 200
                  .                   .                   .                   .
CCGTTTAAGCGACCAACCAGTGCTTAGCGTGCAGCGGCGGGCTAAATATCTGCTGCTGGAGCTGCCTGAG

        220                 240                 260                 280
        .                   .                   .                   .           .
GGCTGGATTATCATTCATTTAGGGATGTCTGGCAGCCTGCGCATCCTTCCAGAAGAACTTCCCCCTGAAA

                  300                 320                 340
                  .                   .                   .                   .
AGCATGACCATGTGGATTTGGTGATGAGCAACGGCAAAGTGCTGCGCTACACCGATCCGCGCCGCTTTGG

        360       Bst XI    380                 400                 420
        .                   .                   .                   .           .
TGCCTGGCTGTGGACCAAAGAGCTGGAAGGGCATAATGTGCTGACCCATCTTGGACCGGAGCCGCTTAGC

                  440                 460                 480
                  .                   .                   .                   .
GACGATTTCAATGGTGAGTATCTGCATCAGAAGTGCGCGAAGAAAAAAACGGCGATTAAACCGTGGCTGA

        500                 520                 540         BamH I
        .                   .                   .           .                   .
TGGATAACAAGCTGGTGGTAGGGGTAGGGAATATCTATGCCAGCGAATCACTGTTTGCGGCGGGGATCCA

                  580                 600                 620
                  .                   .                   .                   .
TCCGGATCGGCTGGCGTCATCACTGTCGCTGGCAGAGTGTGAATTGTTAGCTCGGGTGATTAAAGCGGTG

        640                 660                 680                 700
        .                   .                   .                   .           .
TTGCTGCGTTCGATTGAGCAGGGTGGTACAACGCTGAAAGATTTTCTGCAAAGTGATGGTAAACCGGGCT

                  720                 740                 760
                  .                   .                   .                   .
ATTTCGCTCAGGAATTGCAGGTTTACGGGCGAAAAGGTGAGCCGTGTCGGGTGTGCGGTACGCCGATTGT

        780                 800                 820         Nsi I
        .                   .                   .           .                   .
GGCGACTAAACATGCGCAGCGGGCAACGTTTTATTGTCGGCAGTGCCAGAAGtaattcatgcatcccggg

        EcoR I
        .
gagctcgaattcg
```

(Note: Non-protein coding sequence is shown in lower case and selected restriction endonuclease sites are shown underlined. Oligonucleotides 184 and 185 correspond, respectively, to nucleotides 4 to 22 and 816 to 837, inclusive.)

Pilot experiments were carried out to define the conditions necessary for optimal production of the three fragments. Initially, pUIF4 was digested with 0.5 U *Bam*H I/µg DNA and aliquots were taken every 5 minutes up to 80 minutes. When these aliquots were analysed by agarose gel electrophoresis the 0.85 kb band was only just apparent after digestion for 80 minutes, with no detectable 0.55 kb or 0.30 kb bands. The experiment was therefore repeated, this time increasing the *Bam*H I to 2 U/µg DNA, but still only the 0.85 kb fragment appeared (Fig. 3.9).

Fig. 3.9
Digestion of plasmid pUIF4 with *Bam*H I

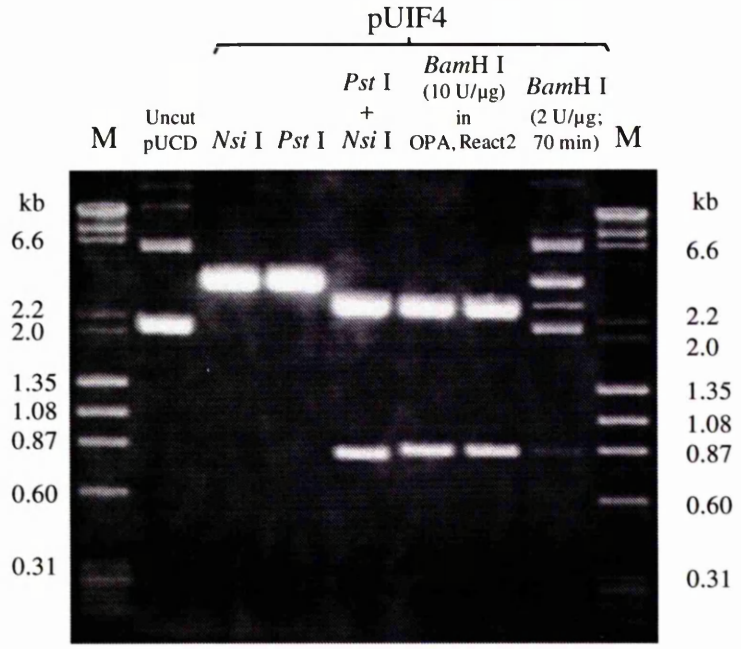


In order to define the conditions required for partial digestion, caesium chloride gradient-purified plasmid pUIF4 was digested with *Bam*H I at a ratio of 2 U/µg DNA. Aliquots containg 0.5 µg of plasmid were taken at 5 minute intervals, up to 80 minutes incubation time (at 37°C), and analysed by AGE. Although the 0.85 kb band was apparent, the predicted 0.55 and 0.30 kb bands were not.

As *Bam*H I may have been sensitive to impurities the pUIF4 DNA was repurified by phenol/chloroform extraction and precipitation from 0.3 *M* sodium acetate with ethanol, followed by extensive washing with 70% ethanol. This cleaned-up DNA was then digested with *Nsi* I, *Pst* I, *Nsi* I + *Pst* I, *Bam*H I in One-Phor-All buffer, and *Bam*H I in "React2" buffer (BRL, MD, USA), using 10 U of each enzyme per µg DNA for 1 hour. The digests were analysed by agarose gel electrophoresis, with some of the 70 minute digest from the 2 U/µg DNA experiment (Fig. 3.10). Digestion with 10 U/µg DNA

appeared complete in all cases (compared with the 2 U/μg DNA lane), however there were still no bands evident at 0.55 kb or 0.30 kb.

Fig. 3.10
Further restriction analysis of plasmid pUIF4

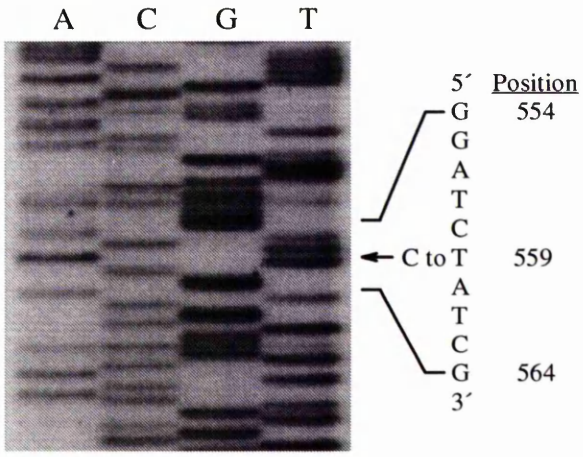


Caesium chloride gradient-purified pUIF4 (1 μg) was digested with the various restriction endonucleases at a ration of 10 U/μg DNA, and analysed by AGE; *Bam*H I was tested with two buffers: "OPA" (One-Phor-All+) and "React2". Uncut (i.e. supercoiled plus nicked circular forms of) pUCD and an aliquot of the 70 minute digestion (= 0.5 μg plasmid) from the previous experiment were included as controls.

It was concluded from this that the internal *Bam*H I site was either missing, corrupted or subject to marked site-preference by the restriction endonuclease (Thomas and Davis, 1975), and that by determining the nucleotide sequence of pUIF4 in this region the cause would be revealed. An aliquot was therefore alkali-denatured and sequenced (2.2.19). Reading the sequence from the resulting autoradiographs (Fig. 3.11) showed that the cytosine residue at the 3′ end of the *Bam*H I site (position 559 in Fig. 3.8) had been replaced by a thymidine residue, thus explaining the inability of *Bam*H I to restrict pUIF4 at this site. This mutation also predicted a change in the amino-acid at codon 182 from histidine (CAT) to tyrosine (TAT). Fortunately this mutation still allowed digestion by

Xho II (Pu/GATCPy) at position 554. The resulting cohesive end being the same as *Bam*H I (GATC) allowing sub-cloning via this site as intended.

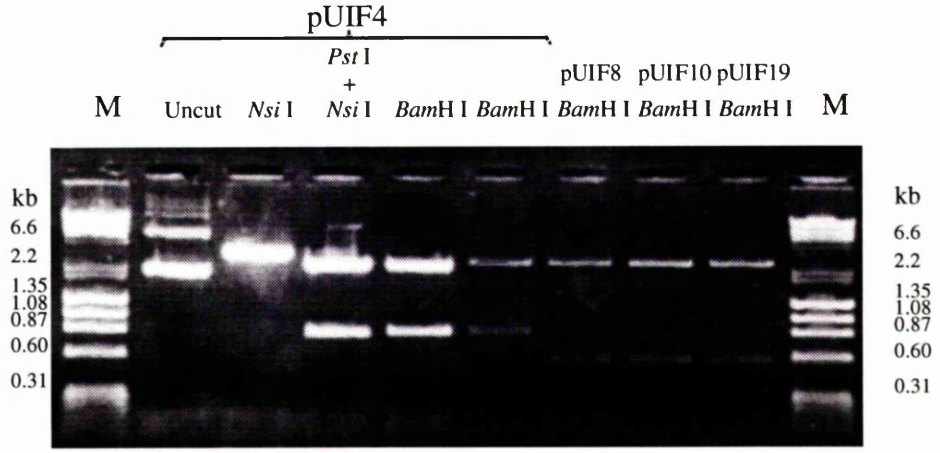
Fig. 3.11
Sequencing autoradiograph showing the mutation in the internal *Bam*H I site of the *fpg* insert in plasmid pUIF4



DNA sequencing (2.2.19) of the *fpg* insert in plasmid pUIF4 demonstrated that most 3′ cytosine residue in the internal *Bam*H I site (GGATCC) had been mutated to a thymidine residue, thus explaining the inability to achieve the predicted pattern of bands by digestion with this enzyme. "A" = ddATP, "C" = ddCTP, "G" = ddGTP, and "T" = ddTTP reaction (in this and all subsequent sequencing autoradiographs).

To investigate whether the other pUIF clones also had this mutation their plasmids were digested with *Bam*H I and analysed by agarose gel electrophoresis (Fig. 3.12).

Fig. 3.12
Restriction analysis of plasmids pUIF4/8/10 and 19

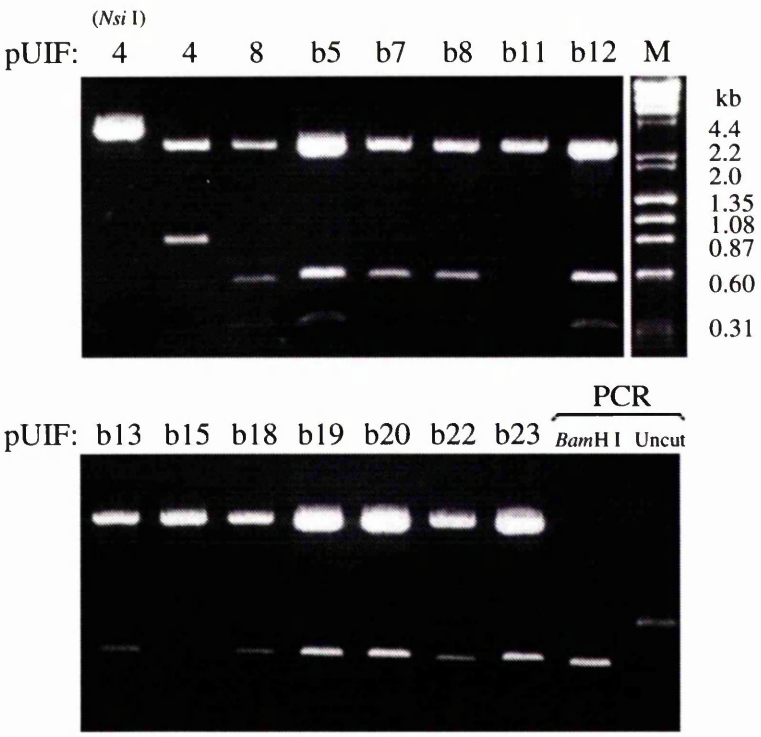


Caesium chloride gradient-purified plasmids were digested with the various restriction endonucleases at a ratio of 10 U/μg DNA, and analysed by AGE. The four plasmid-containing lanes on the right hand side were loaded with 0.5 μg DNA; those on the left were loaded with 2 μg. The plasmids pUIF8/10 and 19 all produced the predicted 0.55 and 0.30 kb fragments with *Bam*H I.

Only pUIF4 appeared to have a mutation affecting the internal *Bam*H I site. The analysis was repeated on ten of the blunt-end ligated plasmids (pUIFb) containing insert as well as two not containing insert. In addition, an aliquot of the gel-purified 0.83 kb PCR product used in the ligations was digested with *Bam*H I (10 U/μg DNA) for four hours, to try to determine what proportion of the PCR generated DNA did not contain the internal *Bam*H I site (Fig. 3.13). All of the ten pUIFb clones digested fully with *Bam*H I: thus only one (7%) of the fourteen blunt- and cohesive-end insert-positive clones carried a mutation affecting the internal *Bam*H I site. The digested PCR product showed a faint band, at the limit of visual detection, corresponding to the undigested 0.83 kb product. Given that the limit for visual detection for this system was approximately 20 ng DNA and approximately 1 μg of PCR product was loaded on the gel, this showed that at least 98% of the DNA could be *Bam*H I digested. The *Bam*H I resistant fragments therefore appeared to be present in a minority.

Fig. 3.13

*Bam*H I restriction analysis of blunt end-ligated plasmids



The blunt end-ligated plasmids (1 µg) together with pUIF4, pUIF8 and some of the 0.85 kb PCR product were digested with *Bam*H I (10 U). Some pUIF4 (1 µg) was linearised with *Nsi* I.

3.3.4 **Subcloning of pUIF4 and pUIF8**

In order to further analyse the plasmid pUIF8 a large-scale preparation of it was made. Using 10 U *Bam*H I/µg DNA it was found that digestion for 50 minutes produced roughly equal amounts of the 0.85, 0.55 and 0.30 kb fragments. A scaled-up digestion of 5 µg pUIF8 was therefore carried out, as well as a complete digestion of pUIF4. The 0.85, 0.55 and 0.30 kb bands were isolated by LMP-agarose gel electrophoresis (2.2.7). In order to sub-clone pUIF4 it was digested with *Xho* II to produce 0.55 and 0.30 kb fragments.

Ligation reactions were set up with *Bam*H I/CIAP-digested pUCD plus the 0.55 kb or 0.30 kb pUIF4/*Xho* II or pUIF8/*Bam*H I fragments. These were then used to transform *E. coli* HB101 to ampicillin resistance. Two pUIF4 *Xho* II sub-clones, one containing a 0.55 kb insert (designated pUIF4.X5), the other a 0.30 kb insert (designated pUIF4.X3)

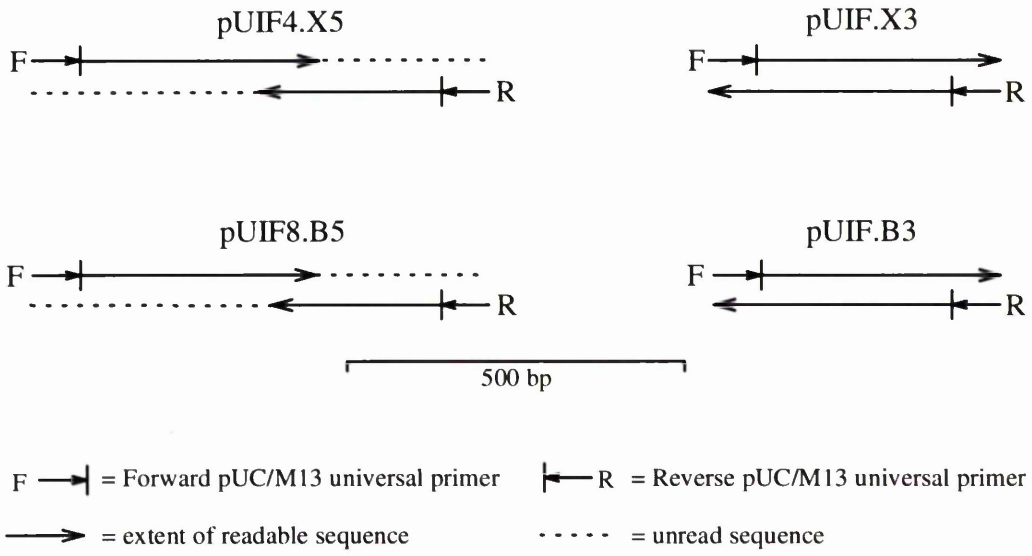
were selected for sequencing. Similarly, one of the pUIF8 sub-clones from each of the 0.55 and 0.30 kb ligations were chosen, designated pUIF8.B5 and pUIF8.B3 respectively.

3.3.5 DNA sequencing of the pUIF4 and pUIF8 sub-clones

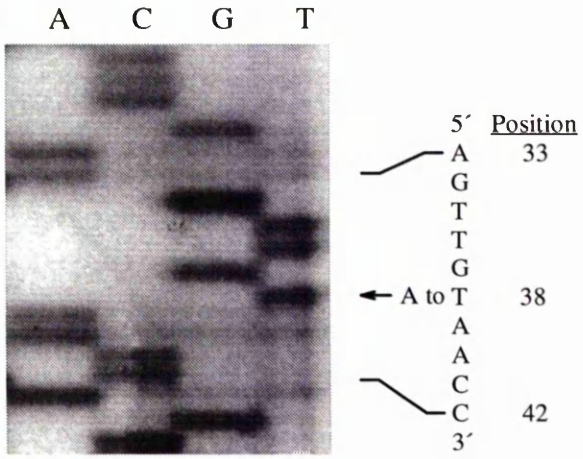
The four sub-clones of pUIF4 and pUIF8 were sequenced by the dideoxy-termination method (2.2.19). Using forward and reverse pUC/M13 universal primers it was possible to sequence the inserts, including the *Xho* II and *Bam*H I sites used for sub-cloning (Fig. 3.14a). It was thus possible to confirm that the original ligation involving *Pst* I/*Nsi* I cut 0.85 kb PCR product had occurred as planned, in that the restriction endonuclease sites had been reformed accurately on splicing into pUCD. It was also possible to confirm that the adenine residue in the upstream out-of-frame start codon (at position 121, Fig. 3.1) had been altered to a guanine residue as intended.

Fig. 3.14
Sequencing of pUIF4 and pUIF8 sub-clones

a) Extent of readable sequence in sub-clones



b) Sequencing autoradiograph showing the mutation at position 38 in plasmid pUIF4.X5



Sequencing showed that the plasmid pUIF4.X5 had an A to T mutation at nucleotide 38.

The sequencing of pUIF4.X5 revealed a total of three base changes, including the previously known mutation in the internal *Bam*H I site (Table 3.2 and Fig. 3.14b). No mutations were observed in pUIF4.X3.

Table 3.2
The Mutations Observed in pUIF4

Position[a]	Change	Codon[a]	Amino-acid change
38	A → T	8	Glu → Val
351	T → C	112	(None)
559	C → T	182	His → Tyr

[a] as in Fig. 3.8.

In contrast, the sequencing of the pUIF8 sub-clones did not reveal any base changes when compared to the expected sequence (Fig. 3.8). As the two amino-acid substitutions predicted by the mutations in pUIF4 would produce a novel polypeptide, whose properties might reveal something of the structure/activity relationships of the FPG enzyme, it was decided to use the mutant version of the *fpg* gene (designated *fpg*4) in parallel with that from pUIF8 (*fpg*8) in subsequent experiments.

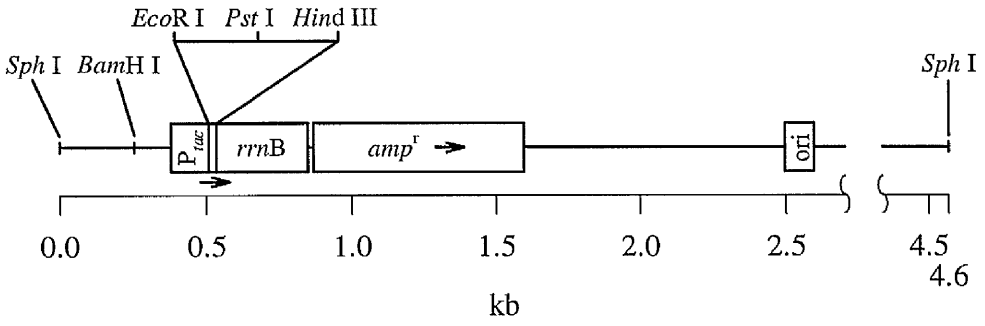
3.4 OVEREXPRESSION OF *fpg*4 AND *fpg*8 IN *E. coli*

3.4.1 The *E. coli* expression vector pKK223-3

The expression of intact native proteins to high levels in *E. coli* is dependent on the presence of a strong regulated promoter (Pribnow, 1975) coupled with an efficient RBS (Shine and Dalgarno, 1975). Thus, as the protein-coding regions in *fpg*4 and *fpg*8 were devoid of these it would be necessary to put them under the control of such elements within a suitable plasmid. Although the pUCD constructs pUIF4 and pUIF8 contained the *fpg* protein-coding region in-frame with the *lacZ'* translation initiation codon, under control of the *lac* promoter, any polypeptides produced would be fusion proteins. This was undesirable as important enzymatic and antigenic properties might be altered in the fusion products, and might complicate comparisons of the *fpg*4 gene product with wild-type FPG protein.

The 4.6 kb vector pKK223-3 (Brosius and Holy, 1984) contains the features necessary for the regulated over-expression of proteins in *E. coli* (Fig. 3.15). The *tac* promoter (P_{tac}) is a combination of the -35 region of the *trp* promoter and the -10 region and RBS of the *lac* UV-5 promoter, as such it provides a powerful stimulus to protein production (DeBoer, 1983). Immediately downstream of P_{tac} is the multiple cloning site (MCS) from pUC8 (Vieira and Messing,1982), followed by the *rrn*B sequence. This latter element contains an *E. coli* 5S rRNA gene and the T_1 and T_2 rRNA transcription terminators (Brosius, 1981), which have the effect of stabilising the host-vector system by inhibiting the expression of deleterious downstream plasmid-derived RNA or protein. The remainder of the vector is derived from pBR322, the β-lactamase gene conferring ampicillin resistance to the host.

Fig. 3.15
The *E. coli* expression vector pKK223-3



The *E. coli* expression vector pKK223-3 (Brosius and Holy, 1984). Key: P_{tac} = *tac* promoter, *rrn*B = *E. coli* 5S rRNA gene and T_1 + T_2 rRNA transcription terminators, ori = ColE1 origin of replication, → = direction of transcription. The plasmid is shown as if linearised at the unique *Sph* I site, the region containing nucleotides 2700 to 4300 has been omitted for clarity.

Overexpression of protein in *E. coli* with pKK223-3 is only achieved by the use of a suitable host strain. Within a *lac*I^Q strain, such as JM105, P_{tac} is repressed: the addition of isopropylthio-β-D-galactoside (IPTG) to the medium causes de-repression allowing overexpression of the inserted gene. For maximal protein production it is necessary for the translation initiation codon of the insert to be within 10 to 15 bp of the RBS. Using

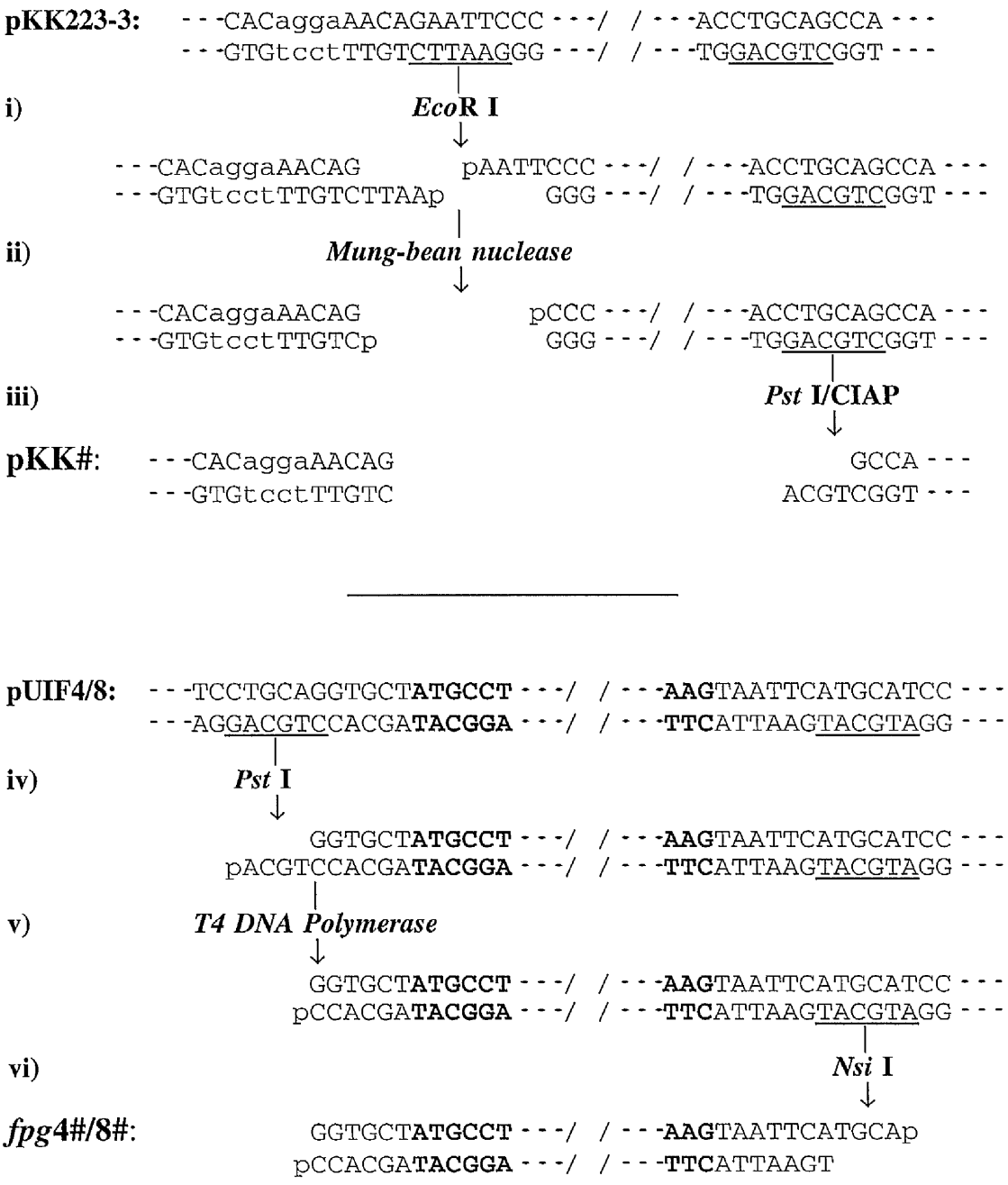
this system it is possible to achieve high levels of protein expression. In a related vector containing the P_{tac}/*rrn*B combination (pKK233-2; Amann, 1985) yields of 18-26% of total cellular protein have been observed in soluble cell extracts.

3.4.2 **Cloning of *fpg*4 and *fpg*8 in pKK223-3**

The DNA fragments pKK# and *fpg*4/8# were generated, from the plasmids pKK223-3 and pUIF4/8 respectively, by digestion with a series of nucleases (Fig. 3.16). The only way to fulfill the requirement for a 10 to 15 bp spacing between the RBS and start codon necessary for maximal protein production (3.4.1), was to first cut pKK223-3 at its unique *Eco*R I site (Fig. 3.16i) and the plasmids (pUIF4/8) containing the inserts (*fpg*4/8) at their unique *Pst* I sites (Fig. 3.16iv). Both restriction endonucleases cut to leave overhangs four nucleotides long at the ends of the dsDNA, *Eco*R I at the 5′ end and *Pst* I at the 3′ end. The 5′ overhang was removed by digestion with mung-bean nuclease (Fig. 3.16ii), which will digest such overhangs on DNA back to the first C:G base-pair encountered, and the 3′ overhangs were removed with T4 DNA polymerase (Fig. 3.16v), exploiting its 3′ to 5′ exonuclease activity in the absence of dNTPs (2.2.11/13 and 15). Subsequent ligation of the blunt ends of the two DNAs would give the required 11 bp spacing between the RBS and the start codon, as discussed above. In order to ligate the other end of the inserts into pKK223-3, the *Eco*R I/mung-bean nuclease-treated vector was then cut with *Pst* I, followed by a final treatment with calf intestinal alkaline phosphatase (CIAP) to 5′-dephosphorylate the DNA ends to minimise religation: this produced the fragment named pKK# (Fig. 3.16iii; 2.2.16). The inserts were released from their parent vectors by digestion with *Nsi* I, which generated *Pst* I-compatible cohesive ends (Fig. 3.16vi). A useful consequence of utilising this combination of blunt and cohesive ends was that the orientation of the inserts was constrained, i.e. an example of so-called "force-cloning".

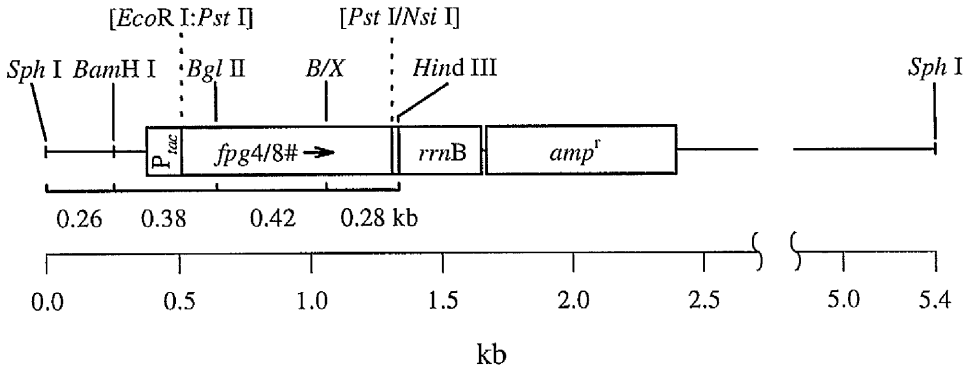
Ligation of *fpg*4/8# into pKK# would produce the expression vectors pKIF4 and pKIF8 (Fig. 3.17).

Fig. 3.16
Preparation of the DNA fragments pKK#, *fpg*4# and *fpg*8#



See methods: 2.2.11, 2.2.13, 2.2.15 and 2.2.16. The RBS sequence in pKK223-3 is shown in lower-case (i - iii); *fpg* protein-coding sequence is shown in bold (iv - vi); top strands are shown 5′ to 3′.

Fig. 3.17
The predicted structure of the pKIF4/8 constructs



The predicted structure of the pKK223-3-*fpg*4/8 constructs. Key: P_{tac} = *tac* promoter, *rrn*B = *E. coli* 5S rRNA gene and $T_1 + T_2$ rRNA transcription terminators, *fpg*4/8# = *fpg*4# or *fpg*8# insert (see Fig. 3.16 for details), *B/X* = position of *Bam*H I (*fpg*8#) or *Xho* II (*fpg*8#) site, [*Eco*R I:*Pst* I] = site of blunt-end ligation between respective modified sites, and [*Pst* I/*Nsi* I] = site of cohesive-end ligation between respective sites (see Fig. 3.16 for details), → = direction of transcription. The plasmid is shown as if linearised at the unique *Sph* I site, the region containing nucleotides 2700 to 4800 has been omitted for clarity.

The DNA fragments pKK#, *fpg*4# and *fpg*8# were purified by glasspowder adsorption after agarose gel electrophoresis (2.2.8). Ligation reactions were set up with pKK# alone, or with pKK# plus *fpg*4# or *fpg*8#, and the products used for transformation of *E. coli* JM105 (Table 3.3).

Table 3.3
Transformation efficiency of *E. coli* JM105 with pKK223-3/*fpg*4# and *fpg*8#

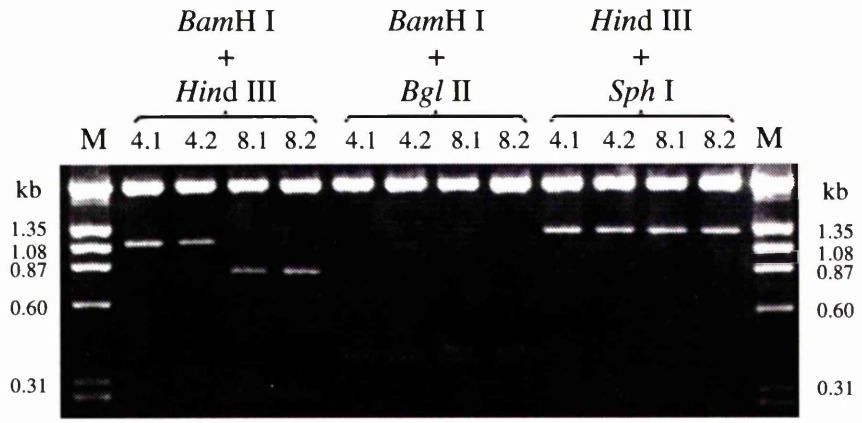
Vector	Efficiency[a]
pKK223-3 (uncut)	1.3×10^6
pKK#	0.4×10^3
pKK# + *fpg*4#	3.7×10^3
pKK# + *fpg*8#	1.6×10^3

[a] ampicillin resistant colonies per µg vector DNA.

The first two clones with each insert (i.e. pKIF4.1, 4.2, 8.1 and 8.2) were analysed by digestion with three different combinations of restriction endonucleases (Fig. 3.18). The *fpg*4# plasmids showed the expected pattern of 4.3 and 1.1 kb bands with

*Bam*H I + *Hin*d III, 5.0 and 0.38 kb bands with *Bam*H I + *Bgl* II and 4.0 plus 1.34 kb bands with *Hin*d III + *Sph* I. Likewise the *fpg*8# plasmids gave 4.3, 0.8 and 0.28 kb bands with *Bam*H I + *Hin*d III, 4.6, 0.42 and 0.38 kb bands with *Bam*H I + *Bgl* II and 4.0 plus 1.34 kb bands with *Hin*d III + *Sph* I. Analysis of the clones thus showed that all had incorporated insert as intended.

Fig. 3.18
Restriction analysis of plasmids pKIF4.1/4.2/8.1 and 8.2



DNA from small-scale preparations of the plasmids was digested with the various combinations of restriction endonucleases, and analysed by AGE. "4.1" = pKIF4.1, "4.2" = pKIF4.2, "8.1" = pKIF8.1, and "8.2" = pKIF8.2.

3.4.3 **FaPy-DNA glycosylase activity in the pKIF4 and pKIF8 clones**

Although the restriction patterns of the pKIF clones suggested that they had the same gross structure, minor differences between the clones, particularly in the critical RBS/start codon region, might cause the efficiency of translation to vary. Therefore, if differences between the clones were manifest, then those giving the greatest levels of over-expression could be chosen for large-scale culture and FPG extraction. By measuring the IPTG-induced FPG activity of the clones it might also be possible to gain some idea as to whether the multiple-mutant *fpg*4 insert did in fact encode a protein with FaPy-DNA glycosylase activity.

To this end *E. coli* JM105, JM105 containing pKK223-3 and the first six pKIF4 and pKIF8 clones were grown in small-scale cultures until mid-log phase ($A_{600\ nm} \approx 0.6$). IPTG was then added and incubation was continued for a further 3 hours, after which the cells were harvested (2.3.8). Crude cell lysates were then prepared as described (2.3.1) and the total protein (TP) and FaPy-DNA glycosylase activity assayed (2.3.2). The results are shown in Table 3.4.

Table 3.4
FPG activity after 3 h of IPTG induction in pKIF4 and 8 clones

Clone	FPG activity (μU/g TP)[a]
JM105	13.7 (1.0)[b]
JM105/pKK223-3	13.5 (1.5)
pKIF4.1	530 (170)
pKIF4.2	600 (160)
pKIF8.1	13.7 (1.1)
pKIF8.2	12.8 (1.5)

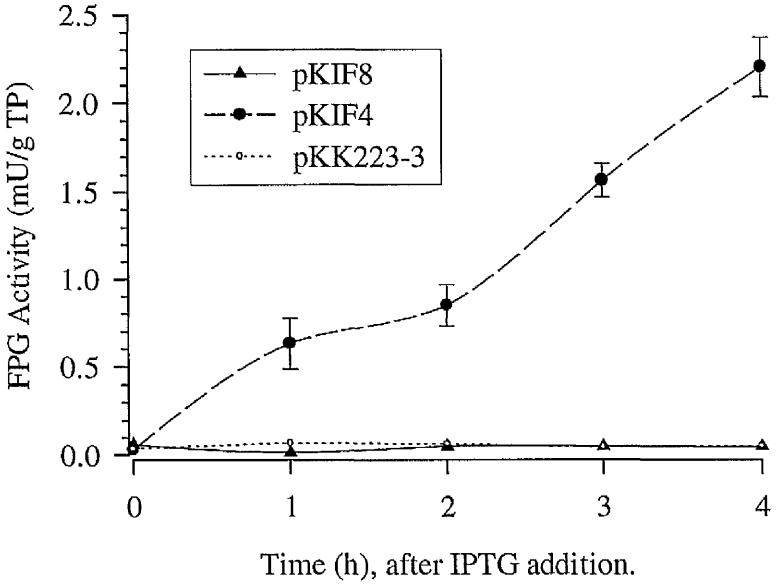
[a] by the DNA-based assay (2.3.2), [b] values given are mean (SEM) of three separate experiments.

The results showed that the presence of the vector pKK223-3 in *E. coli* JM105 did not affect the endogenous level of FPG activity. However, the most striking feature of the data was that although the mutant insert *fpg*4 clones gave 40 or 44-fold increased activity, the *fpg*8 clones' activities were no higher than the vector control.

To investigate whether the time-course of expression was different or the plasmid was unstable within the host on IPTG induction, it was decided to study one of each of the clones (pKIF4.1 and 8.1) during and after a more prolonged period of induction. Larger scale cultures (25 mL broth in 100 mL conical flasks) were therefore set-up with these clones together with JM105/pKK223-3 as a control. As before, IPTG was added when the $A_{600\ nm}$ was approximately 0.6 and 1 mL aliquots were taken immediately prior to induction and at hourly intervals up to 4 h, followed by a final sample at 20 h

post-induction. The results up to 4 h are shown in Fig. 3.19, the points on this graph are mean ± SEM (n = 3). It can be seen that pKIF4.1 gave steadily increasing levels of activity, while pKIF8.1 gave results no higher than pKK223-3. All three cultures had levels of activity less than 10 μU/g TP at 20 h post-induction. Restriction analysis of the plasmids isolated from the cultures after 20 h of induction gave no evidence of instability in the plasmids (data not shown). Overexpression of FPG in *E. coli* by a factor of 70-fold had been achieved by Boiteux *et al.* (1987). Thus there did not appear to be any reason why a similar system able to overexpress a novel protein containing two amino-acid changes (pKIF4.1), was not able to overexpress wild-type enzyme. A re-examination of the nucleotide sequences of *fpg*4 and *fpg*8 was therefore prompted.

Fig. 3.19
Induction of FPG expression in clones pKIF4 and pKIF8



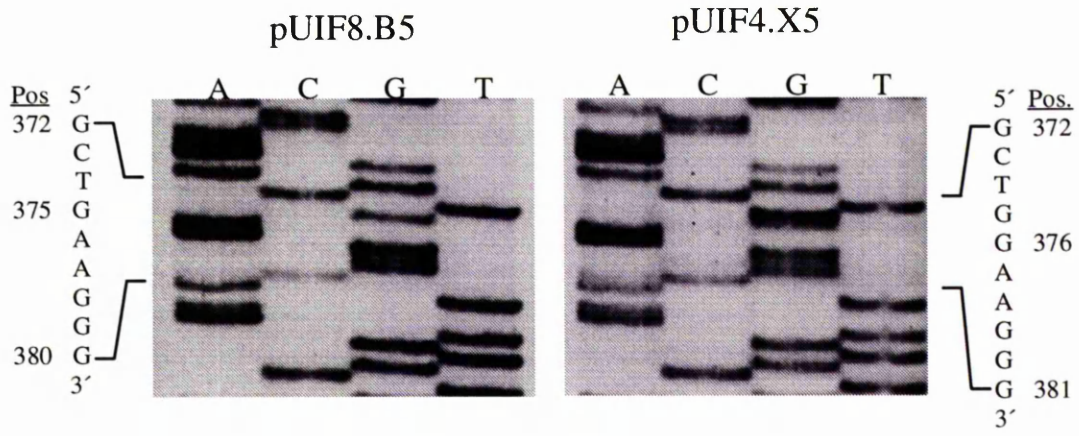
Cultures (25 mL) of *E. coli* JM105 clones containing the various plasmids had IPTG added to a concentration of 1 m*M* when the mid-log growth phase was attained (i.e. $A_{600\,nm} \approx 0.6$). Samples were taken at the times indicated and assayed in triplicate, the error bars represent SEM (which have been omitted for clarity when their magnitude did not significantly exceed the diameter of the symbol used).

3.5 FURTHER ANALYSIS OF pUIF CLONES

3.5.1 The nucleotide sequence of pUIF4/8

The four sub-clones of *fpg* in pUCD (3.3.4; pUIF4.X3 and X5, pUIF8.B3 and B5) were re-sequenced, as checking of the original autoradiographs did not reveal any unknown changes and there were some areas of G/C band compression impairing sequence reading. The (re-)sequencing autoradiograph of pUIF4.X5 and pUIF8.B5 is shown in Fig. 3.20.

Fig. 3.20
Sequencing autoradiographs showing the deletion of a single base in plasmid pUIF8.B5



The deletion of a guanine residue in pUIF8.B5 at position 375/6 was apparent when compared to the wild-type sequence of pUIF4.X5.

In the sequence of pUIF8.B5 (more easily seen when compared to the equivalent position in pUIF4.X5) there is a deletion of one of the two guanine residues at position 375 or 376. This loss of a base predicted a -1 bp frameshift which would result in termination of translation at codon 125, 16 bp after the deletion (Fig. 3.21). Four out of the five carboxyl-terminal amino-acids would also be changed from [GluGlyHisAsnVal] to [LysGlyIleMetCysEND]. With such a mutation the inability to induce FPG activity in the pKIF8 clones was no longer surprising.

Fig. 3.21
The frameshift mutation in *fpg*8

bp: 370 390

***fpg*:** ...TGGACCAAAGAGCTGGAAGGGCATAATGTGCTGACCCATCTT...

FPG: ...TrpThrLysGluLeuGluGlyHisAsnValLeuThrHisLeu...

***fpg*8:** ...TGGACCAAAGAGCTG_AAGGGCATAATGTGCTGACCCATCTT...

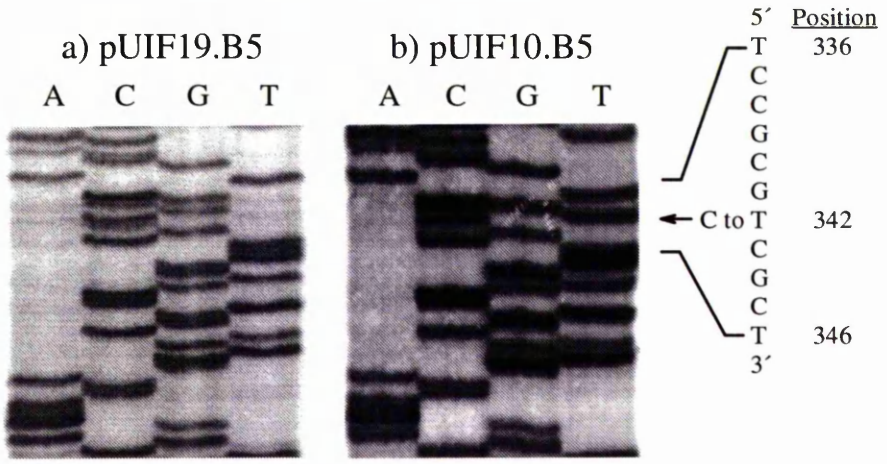
FPG8: ...TrpThrLysGluLeu LysGlyIleMetCysEND

Codon: 120 125

In comparison with wild-type *fpg* one of the guanine residues at position 375/6 in *fpg*8 is deleted. The resultant -1 bp shift in reading frame is predicted to cause premature termination of the FPG8 protein at codon 125.

It was, however, still necessary to produce *E. coli* overexpressing FPG protein in order to provide material for raising antibodies and allow a comparison with the mutant FPG4 protein (3.3.5 and Table 3.2); in addition, an insert known to be capable of producing active FPG protein was required for cloning into other expression vectors. The other two clones originally produced in pUCD, namely pUIF10 and pUIF19, were therefore sub-cloned as pUIF8 had been (3.3.4), except that DNA purification after agarose gel electrophoresis was achieved by adsorption onto glasspowder. The sub-clones were sequenced (2.2.19), revealing that pUIF10 had a single "silent" base-change at position 342 (codon 109: CGC → CGT; Fig. 3.22) and that pUIF19 had no mutations. The protein-coding inserts derived from these two clones were therefore designated *fpg*10 and *fpg*19 respectively. It was decided to use the *fpg*10 insert, rather than *fpg*19, for ligation into other expression vectors because its "silent" base change generated an extra *Hha* I restriction site as compared with the wild-type sequence. Thus, amplification of the *fpg*10 sequence by PCR could theoretically be easily distinguished from (possibly contaminating) wild-type sequence by what was, in effect, a serendipitous novel restriction fragment length polymorphism (RFLP).

Fig. 3.22
Sequencing autoradiographs of pUIF10.B5 and pUIF19.B5



Sequencing autoradiographs of the plasmids a) pUIF19.B5, and b) pUIF10.B5. The cytosine residue at position 342 in pUIF19.B5 had been mutated to a thymidine residue in pUIF10.B5.

3.5.2 *Taq* DNA-polymerase induced errors

The purified *Taq* DNA-polymerase used in the PCR generation of the *fpg* inserts lacks 3′ to 5′ exonuclease activity (Tindall and Kunkel, 1988) and thus is incapable of "proof-reading". Misincorporation or failure to incorporate a nucleotide is thus tolerated and mutations will be present in the final product, each one in amounts dependent on how early in the amplification process the original error occurred. If the PCR product is used en masse, the fact that at any one nucleotide a small proportion of the DNA has an incorrect base is usually negligible. However, if that product is subjected to clonal expansion, e.g. by ligation into a plasmid vector, then such mutations become significant. Estimates of *Taq* DNA-polymerase error rates vary. Saiki *et al.* (1988) used the formula in equation 2:

$$k_{mut} = \frac{2 \times f_{err}}{n_d} \quad (Eqn.\ 2)$$

Where k_{mut} = average mutation rate per cycle of PCR, f_{err} = observed error frequency, and n_d = number of DNA doublings.

to calculate the average rate at which errors occurred from the observed cumulative error frequency. From an f_{err} of 17 misincorporations per 6 692 nucleotides sequenced, their estimate of k_{mut} was approximately one in 5 900 nucleotides. For simplicity they assumed that the mutation rate through all cycles of the PCR was constant and this may not be so, also n_d is known not to equal the number of thermal cyclings, particularly in the later cycles of a PCR (the "plateau effect": 3.2.2). In addition, Innis *et al.*(1988) have found that a wrongly incorporated base can cause *Taq* DNA-polymerase to chain terminate, and it has also been reported that *Taq*-polymerase is sensitive to particular 3′-primer/template mismatches (i.e. C:C, A:G, G:A and A:A) which cause failure of chain extension (Gelfand, 1990), both of these effects may result in an under-representation of errors in the final product. A genetic fidelity assay was used by Tindall and Kunkel (1988) to estimate errors, which they reported as approximately one in 8 000 for misincorporations per cycle and approximately 1 in 30 000 for frameshifts per cycle. Subsequent studies, which used lower dNTP and magnesium concentrations than the earlier reports, have shown misincorporation error rates of less than one in 81 000 (Goodenow *et al.*, 1989) and less than one in 200 000 (Fucharoen *et al.*, 1989). In the experiments reported here the observed nucleotide misincorporation rate of 4 in 3 412, over 30-cycles of PCR, indicates an average mutation rate of one in every 12 795 nucleotides incorporated (using equation 2 and assuming n_d = 30). Similarly, the single base deletion observed (in *fpg*8) gives an average frame-shift mutation rate of one in 51 000. Within the error margins inherent with such small samples these estimates are consistent with those already reported. Therefore, although finding mutations in three out of four PCR-derived clones might have suggested an especially error-prone process it would appear not to have been exceptional, particularly considering the concentrations of magnesium and dNTPs used in the PCR. Subsequent developments in PCR technology have shown the utility, if using *Taq* DNA-polymerase, of keeping the number of cycles

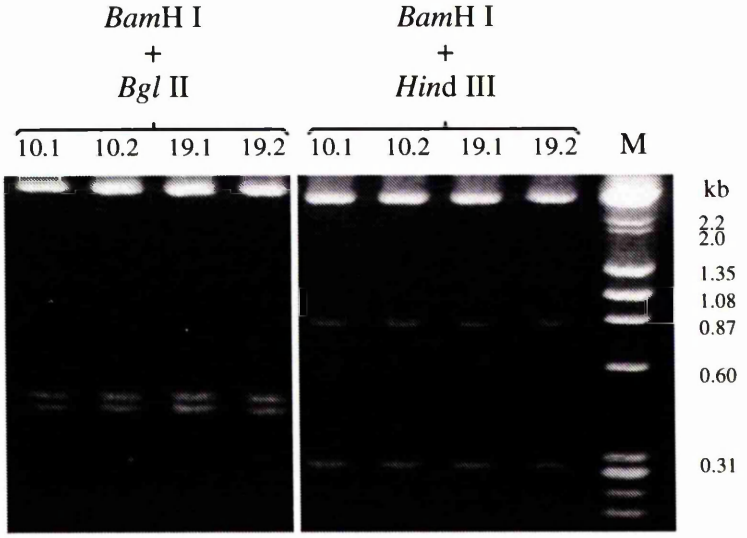
and the magnesium and dNTP concentrations as low as possible, or, better still, using proof-reading thermostable DNA-polymerases when cloning of a PCR product is required.

3.6 OVEREXPRESSION OF *fpg*4/8/10 AND 19 IN *E. coli*

3.6.1 Cloning of *fpg*10 and *fpg*19 in pKK223-3

Digestion of pUIF10 and pUIF19, as with pUIF4 and pUIF8 in Fig. 3.16 (3.5.2), provided *fpg*10# and *fpg*19# inserts for ligation into pKK# (Fig. 3.16); ligation and transformation of JM105 were performed as before (3.4.2) and two clones deriving from each insert were checked by restriction pattern analysis on agarose gel electrophoresis. All of the four clones chosen (designated pKIF10.1, pKIF10.2, pKIF19.1 and pKIF19.2) appeared to have successfully incorporated an insert (Fig. 3.23).

Fig. 3.23
Restriction analysis of plasmids pKIF10.1/10.2/19.1 and 19.2



Small-scale preparations of the plasmids was digested with the combinations of restriction enzymes shown, and analysed by AGE. "10.1" = pKIF10.1, "10.2" = pKIF10.2, "19.1" = pKIF19.1, and "19.2" = pKIF19.2.

3.6.2 Induction of pKIF10 and pKIF19 clones with IPTG

The clones were grown up as previously described (2.3.1, 2.3.2, and 3.4.3) and induced with 1 m*M* IPTG for 3 h (2.3.8; Table 3.5).

Table 3.5
FPG activity after 3 h of IPTG induction in pKIF10 and 19 clones

Clone	FPG activity (μU/g TP)
JM105	13.7 (0.9)[a]
JM105/pKK223-3	13.2 (1.5)
pKIF10.1	5 500 (580)
pKIF10.2	4 500 (600)
pKIF19.1	5 500 (400)
pKIF19.2	4 100 (350)

[a] values given are the mean (SEM) of three separate experiments.

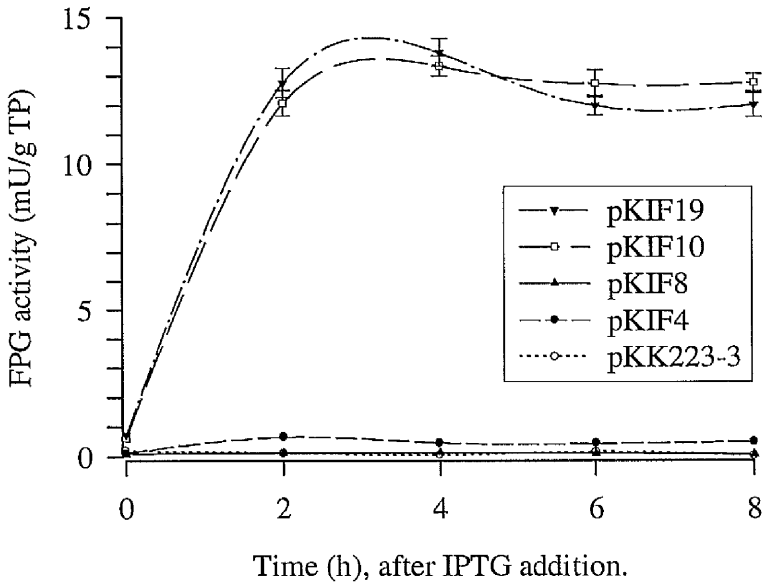
The levels of FPG activity observed with all of the pKIF10 and pKIF19 clones were similar to each other and substantially higher than the pKIF4 clones (3.5.3). Clones pKIF10.1 and pKIF19.1 were selected for further analysis, together with pKIF4.1 and pKIF8.1. For clarity the names of the various clones was simplified to pKIF4, pKIF8, pKIF10 and pKIF19.

3.6.3 The time-course of pKIF4/8/10 and 19 induction

In order to study the time-course of FPG induction 25 mL cultures of the four clones (pKIF4, 8, 10 and 19) plus JM105/pKK223-3 were set up as in section 3.4.3, although on a larger 25 mL scale in 100 mL conical flasks, and induced with 1 m*M* IPTG. Samples of the cultures (1 mL) were taken immediately prior to IPTG addition and at 2-hourly intervals thereafter, up to a maximum of 8 h. The time-course of FPG activity in the clones is shown in Fig. 3.24, points are mean ± SEM (n = 3). At no time did

clone pKIF8 exhibit any increase in FPG activity over the pKK223-3 control. The level of FPG activity seen with pKIF4, in contrast, showed only a small initial rise from 220 μU/g TP pre-induction, to 650 μU/g TP at 2 h, followed by lower levels at 4, 6 and 8 h (450, 430 and 500 μU/g TP). The activity in pKIF10 and pKIF19 on the other hand increased more rapidly and to considerably higher levels, between 11.7 to 13.3 mU/g TP, by 2 h after IPTG addition, and the level remained elevated at up to 8 h.

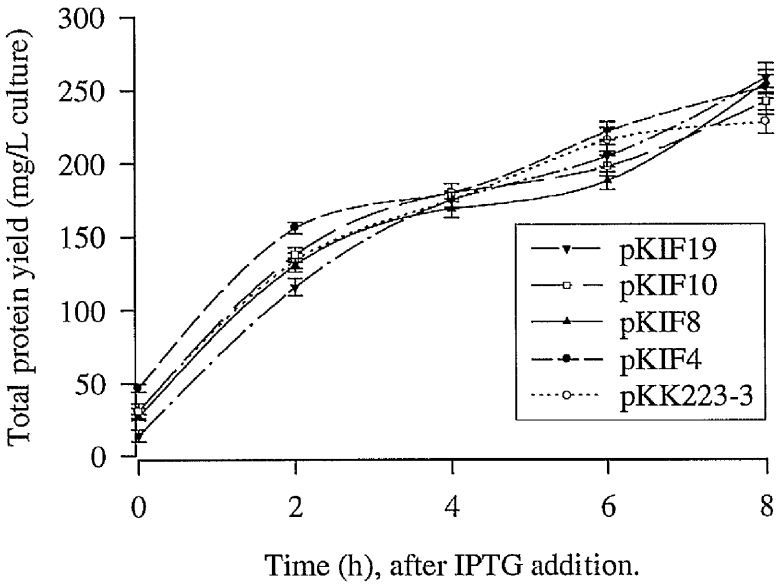
Fig. 3.24
Induction of FPG expression in clones pKIF4/8/10 and 19



Cultures (25 mL) of *E. coli* JM105 clones harbouring the various plasmids had IPTG added to 1 m*M* when the $A_{600\ nm} \approx 0.6$. Samples were taken at the times indicated and assayed in triplicate, the error bars represent the SEM (which have been omitted for clarity when their magnitude did not significantly exceed the diameters of the symbols used). The curves for pKIF10 and pKIF19 were effectively the same.

When the growth of the clones was assessed by plotting the total protein yield per litre of culture against time (Fig. 3.25) it was seen that there were no differences between the clones sufficient to explain the much reduced FPG activity in pKIF4.

Fig. 3.25
Total protein levels in clones pKIF4/8/10 and 19

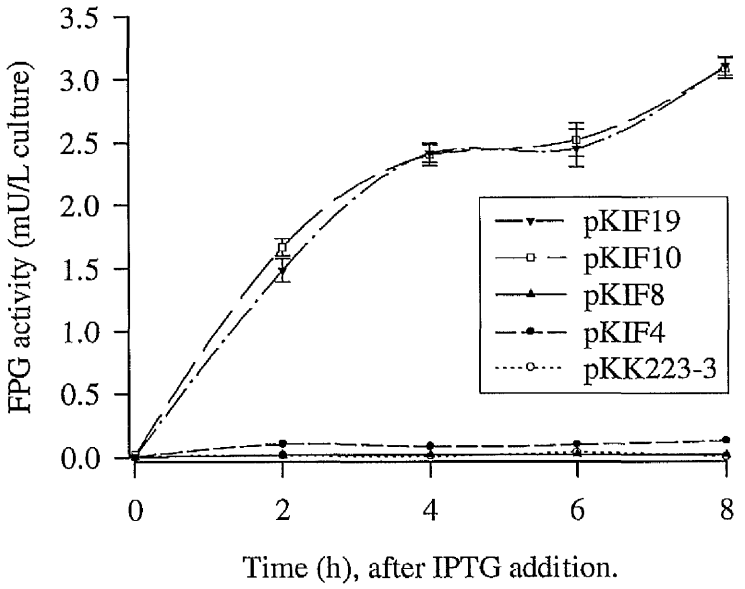


Cultures (25 mL) of *E. coli* JM105 clones harbouring the various plasmids had IPTG added to 1 m*M* when the $A_{600\ nm} \approx 0.6$. Samples were taken at the times indicated and assayed in duplicate for TP concentration, the error bars represent the SEM (which have been omitted for clarity when their magnitude did not significantly exceed the diameters of the symbols used).

Similarly, expressing the FPG activity in terms of volume of culture, instead of mass of total protein (Fig. 3.26), showed more clearly the differences between the clones, and that in pKIF10 and 19 the potential total yield of FPG protein was increasing all the time up to 8 h. These data suggested that the optimal time for harvesting pKIF10 would be between 4 to 6 h post-IPTG. In addition, as the FPG activity in pKIF4 appeared not to parallel the growth of the clone this suggested that FPG4 protein was unstable, or at least that large scale production of it would be problematical.

Fig. 3.26
Induction of FPG activity in clones pKIF4/8/10 and 19

(Expressed as enzyme activity per volume of culture.)



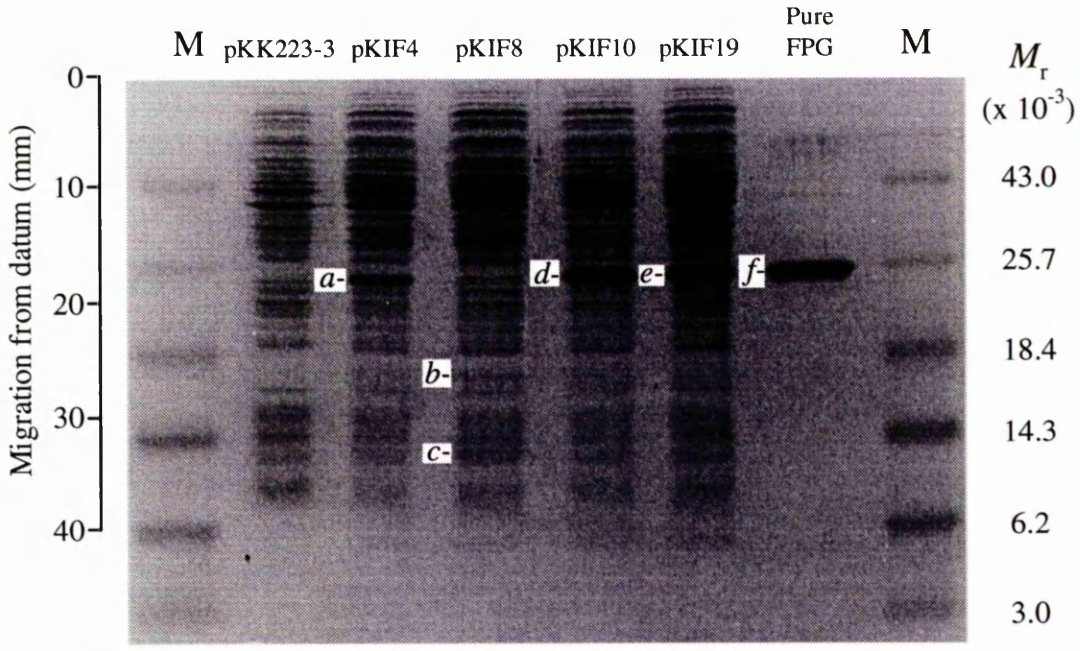
Cultures (25 mL) of *E. coli* JM105 clones harbouring the various plasmids had IPTG added to 1 m*M* when the $A_{600\ nm} \approx 0.6$. Samples were taken at the times indicated and assayed in triplicate, the error bars represent the SEM (which have been omitted for clarity when their magnitude did not significantly exceed the diameters of the symbols used). The FPG activity has been expressed in mU/L culture in order to determine the time at which to harvest the bacteria in the cultures to maximise yield.

3.6.4 **Protein analysis of clones pKIF4/8/10 and 19**

In the previous section induction of the various pKIF clones was described up to 8 h after IPTG addition. As well as samples being taken for FPG assay, aliquots were also taken at 8 h for SDS-PAGE analysis. By this means it was hoped that a semi-quantitative estimate of FPG yield in terms of the proportion of total protein could be gained, as well as possibly showing the presence of the prematurely terminated FPG protein (predicted M_r = 14 000) produced by pKIF8, in contrast to FPG itself (M_r = 30 200). SDS-PAGE analysis of the bacterial extracts, in addition to some purified FPG protein (from Dr S. Boiteux), was thus carried out (2.3.6). The gel (Fig. 3.27) clearly showed single bands, corresponding to the single band of pure FPG, in the tracks loaded with extract from pKIF4, 10 and 19. The marker proteins did not migrate

according to their true M_r because they were pre-stained with covalently bonded dye molecules (2.3.6) and hence their M_r indication is not accurate. In the pKIF8 track two bands (*b* and *c*) were just apparent, possibly indicating the position of the prematurely terminated FPG8 protein.

Fig. 3.27
SDS-PAGE analysis of *E. coli* JM105 clones harbouring plasmids pKK223-3, pKIF4/8/10 and 19

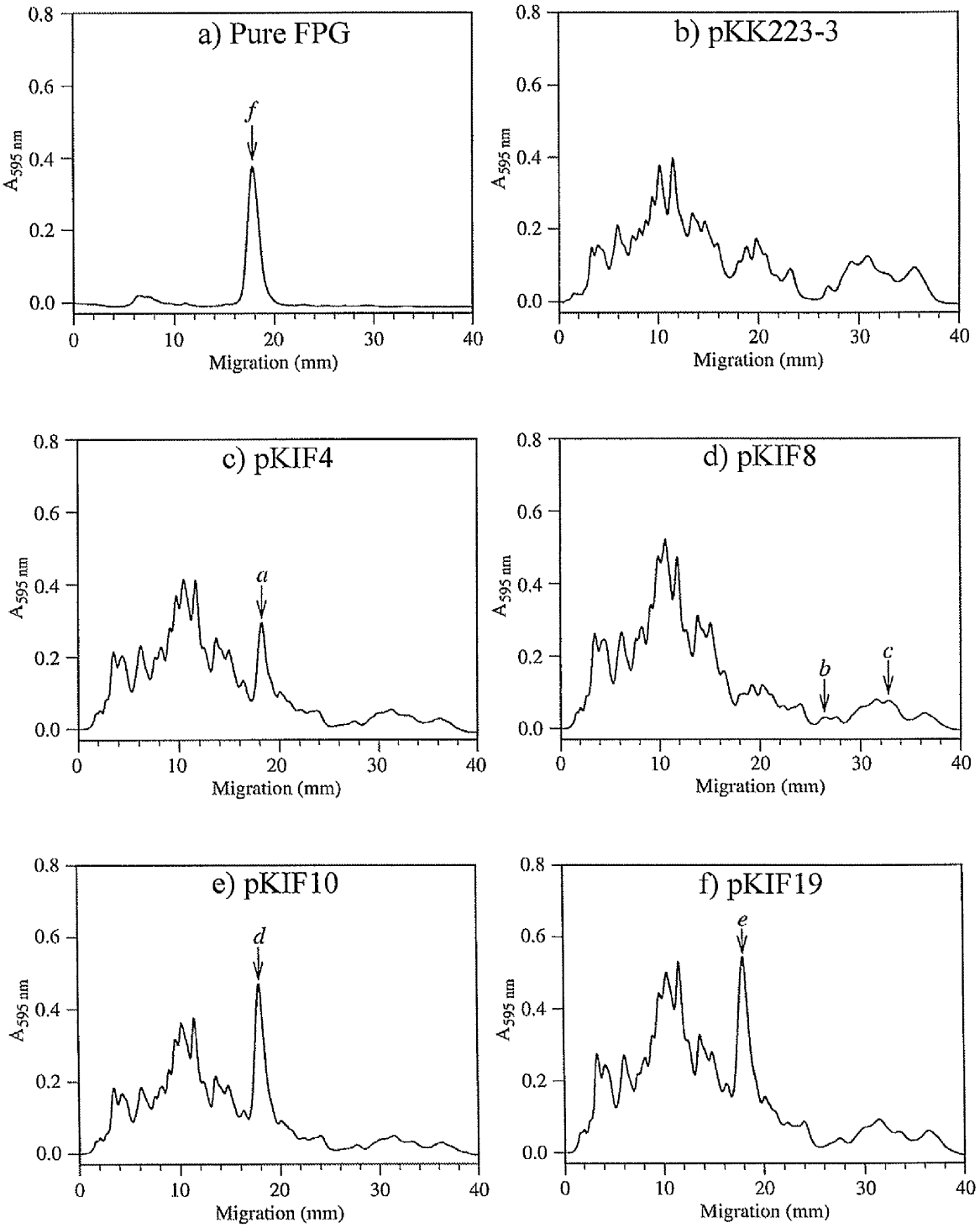


E. coli JM105 clones harbouring the respective plasmids were induced with 1 m*M* IPTG for 8 h. Total protein loaded in each lane = 10 µg, except for "Pure FPG" (4 µg TP). Marker proteins were pre-stained with a covalently-linked blue dye, and migrated with apparent M_r somewhat greater than the figures quoted for the un-labelled proteins (see 2.3.6).

In order to gain more information from the gel, densitometric scans were made of the various lanes (2.3.6; Fig. 3.28).

Fig. 3.28

Densitometric analysis of the SDS-PAGE of *E. coli* JM105 clones harbouring the plasmids pKK223-3, pKIF4/8/10 and 19



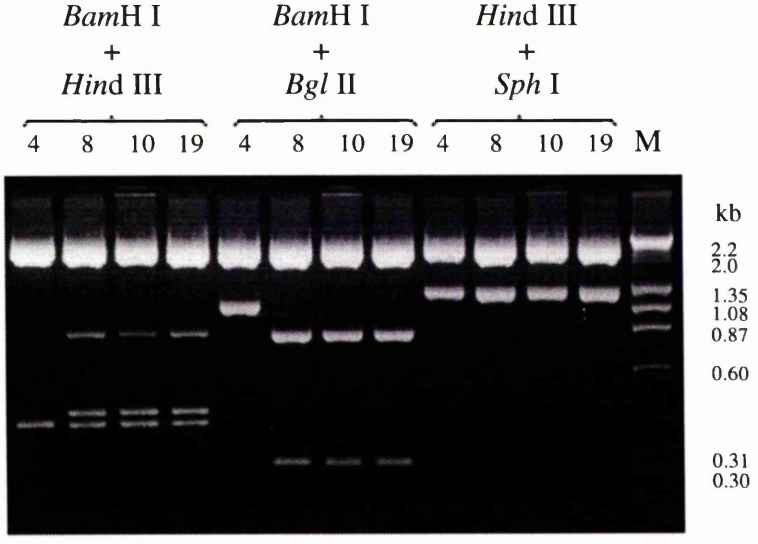
Densitometric scans were made of the lanes in the gel shown in Fig. 3.27 (2.3.6). See text for details and discussion of the labelled peaks (*a* to *f*).

This confirmed the single peak (*f*) seen in the pure FPG sample and the corresponding bands seen in pKIF4 (*a*), pKIF10 (*d*) and pKIF19 (*e*). Analysis of the densitometric scans gave estimates of the proportion of total cellular protein represented by the FPG peaks, these were 12% for pKIF4, 22% for pKIF10 and 15% for pKIF19. The two bands (*b* and *c*) in pKIF8 were more obvious in the scan; peak *c* migrated with similar mobility to the lysozyme marker (M_r = 13 400) and so was thought more likely to be the prematurely terminated product FPG8. To confirm these results it was decided to translate the various plasmids *in vitro* in a cell-free system.

3.6.5 ***In vitro* translation of clones pKIF4/8/10 and 19**

A cell-free *in vitro* transcription-translation system was used to determine the polypeptides produced by the pKIF clones (2.3.7). Large-scale preparations of the plasmids were made and their restriction patterns checked (Fig. 3.29).

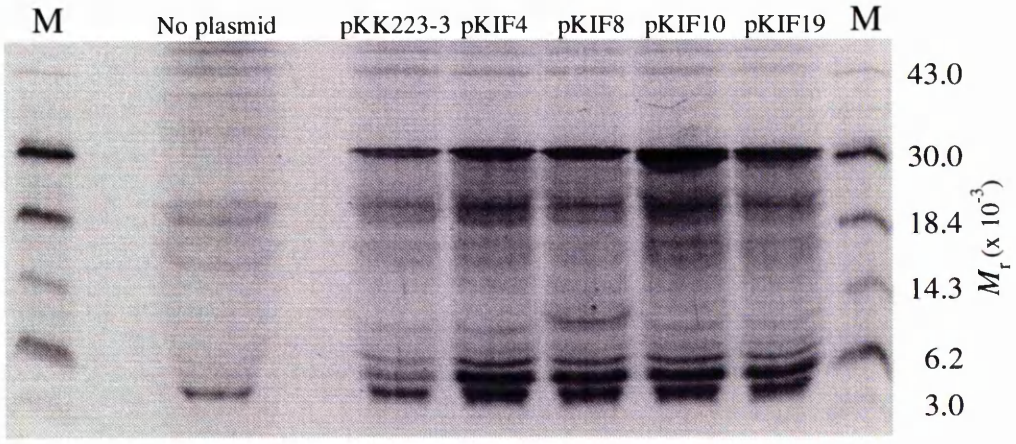
Fig. 3.29
Restriction analysis of large-scale preparations
of plasmids pKIF4/8/10 and 19



Caesium chloride gradient-purified preparations of the plasmids were made. Each plasmid (1 µg) was digested with each combination of restriction endonucleases (10 U of each enzyme) and analysed by AGE. "10.1" = pKIF10.1, "10.2" = pKIF10.2, "19.1" = pKIF19.1, "19.2" = pKIF19.2.

Estimates of the [^{35}S] incorporated into protein allowed loading of equal radioactivity in each lane of the SDS-polyacrylamide gel (Fig. 3.30).

Fig. 3.30
In vitro transcription and translation of plasmids pKK223-3, pKIF4/8/10 and 19



A commercially supplied extract of *E. coli* was used to transcribe and translate each plasmid *in vitro* (2.3.7). Each reaction contained 2.5 µg of plasmid and 20 µCi [^{35}S]-methionine; incorporation of radioactivity into protein was assayed and 60 000 cpm/lane of protein-bound [^{35}S] was loaded. Marker proteins were commercially labelled with [^{14}C] (16 000 cpm/lane).

It can be seen that a variety of labelled polypeptides, in low yield, were produced by the system without the addition of exogenous DNA. Addition of the parent plasmid, pKK223-3, resulted in a small number of additional bands, including one migrating with slightly slower mobility than the M_r = 25 700 marker. Close to this band in the pKIF10 and 19 tracks, an additional band was evident which migrated with the same mobility as the M_r = 25 700 marker, i.e. the same mobility as FPG in the Coomassie Blue-stained gel (Fig. 3.27). With pKIF8 a band was apparent migrating with a mobility between that of the M_r = 13 400 and 6 200 markers. This corresponded to band *b* in Fig. 3.28d, and is consistent with a polypeptide of M_r = 14 000. In contrast, no extra bands could be seen in the pKIF4 lane; either the yield of protein from pKIF4 was too low to be detected or the polypeptide (FPG4) was unstable in the cell-free transcription-translation reaction.

CHAPTER 4

PURIFICATION OF FPG PROTEIN AND POLYCLONAL ANTISERA PRODUCTION

4.1 CONSIDERATIONS OF PURIFICATION METHODS

Previous purification of DNA-alkyltransferases in the Department had exploited their dsDNA binding properties, being carried out by affinity chromatography on dsDNA-cellulose (Wilkinson *et al.*, 1990). Although DNA-cellulose is expensive it is possible to get protein of greater than 95% purity in a single step. Application of this method to FPG purification, however, proved difficult. Pilot experiments showed problems with backpressure, in spite of using sonicated and clarified bacterial cell lysates, and poor recovery of FPG protein, even with up to 1 *M* sodium chloride in the eluent buffer.

After a review of methods it was decided to try an alternative strategy, based on that of Boiteux *et al.* (1987). Anion exchange treatment of the primary extract with DEAE-cellulose would be utilised to remove nucleic acids (thereby reducing viscosity and hence backpressure), followed by cation-exchange/affinity chromatography on cellulose-phosphate, both materials were available in bulk and were considerably cheaper than DNA-cellulose. Final purification would be by gel-filtration.

4.2 PURIFICATION OF FPG PROTEIN

4.2.1 Maximisation of FPG protein overexpression by *E. coli*

Earlier experiments (3.6.3) had shown clone pKIF10 to overexpress FPG protein by a factor of approximately 250-fold to levels *ca.* 13 mU/g TP, representing some 10% of total cellular protein and equivalent to a total FPG activity of 3.18 mU/L of culture medium (3.6.3 and Table 4.1; experiment "A"). Under scaled-up conditions (Table 4.1;

experiment "B"), it was found that reducing the concentration of IPTG from 1.0 to 0.5 m*M* gave similar levels of bacterial wet weight, total protein and FPG activity (total and specific) to experiment "A", and effected a considerable reduction in cost. By then supplementing the LB medium with an extra carbon source (glycerol), zinc chloride and more tryptone extract, it was found that the bacterial wet weight could be increased 1.6-fold, the total protein ~~1.3~~ 2-fold, the total FPG activity 1.7-fold and the specific FPG activity 1.4-fold, as compared with experiment "B" (Table 4.1; experiment "C"). By increasing the oxygenation of the culture by supplying oxygen from a cylinder through a sintered glass frit, rather than relying on diffusion from the air, further improvements could be made: the yields of bacteria and total protein were increased 1.3-fold, the total FPG activity 1.5-fold and the specific FPG activity 1.14-fold, as compared with experiment "C" (Table 4.1; experiment "D"). The crude cell sonicate from experiment "C" was then used to provide material for pilot experiments, while the sonicate from experiment "D" (Fraction I) was used for the main purification.

Table 4.1
Maximisation of FPG production[a]

Expt	Conditions	Medium[b]	Yield of bacteria (g wet wt/L)[c]	Total protein (mg/L)[c]	Total FaPy-DNA glycosylase activity (mU/L)[c]	Specific FaPy-DNA glycosylase activity[d] (mU/g TP)
A:	100 mL flask, orbital shaker, 1 m*M* IPTG	25 mL LB	4.0	240	3.18	13.3
B:	1 L flask, orbital shaker, 0.5 m*M* IPTG	250 mL LB	4.5	215	3.40	15.8
C:	1 L flask, orbital shaker, 0.5 m*M* IPTG	250 mL LB + glycerol + tryptone + ZnCl	7.2	430	9.4	22.0
D:[e]	1 L flask, oxygenated, 0.5 m*M* IPTG	250 mL LB + glycerol + tryptone + ZnCl	9.2	550	13.8	25.0

[a] see methods 2.3.8 and 2.3.9, [b] nominal culture flask volumes were 4× that of the culture medium, [c](FaPy-DNA glycosylase activity)per litre of culture medium used (by DNA-based assay: 2.3.2), [d] of crude bacterial cell lysate, [e] experiment "D" was performed in quadruplicate, i.e. four flasks, each containing 250 mL of medium, the results quoted being obtained from the pooled cultures.

4.2.2 Extraction with DEAE-cellulose

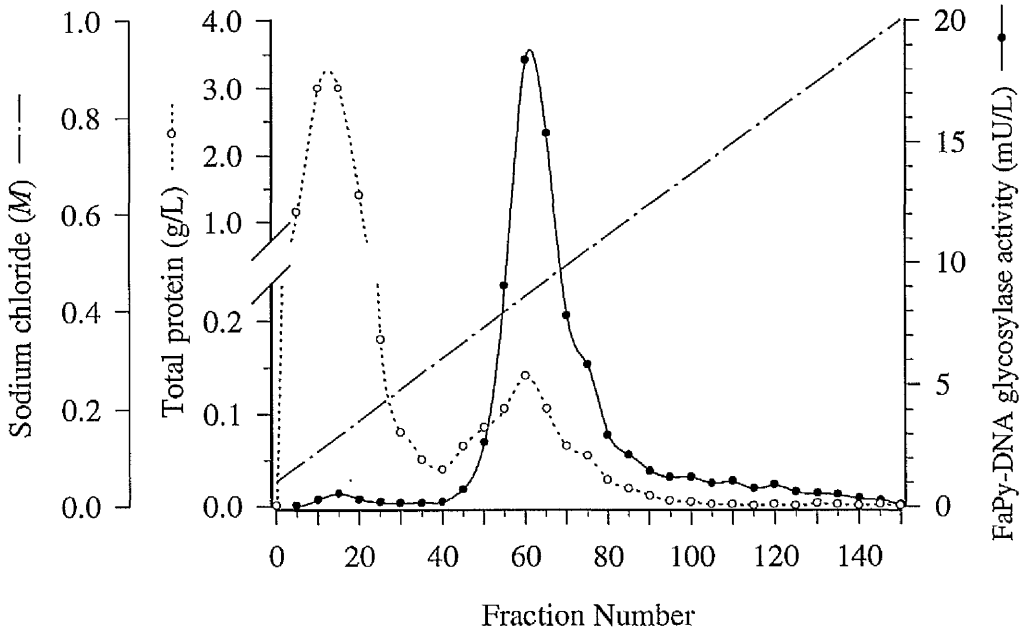
The first step in purification was to remove excess nucleic acid, as well as some protein, by anion exchange with diethylaminoethyl-cellulose (DEAE-cellulose) in bulk. The sonicate from experiment "C" (15 mL) was supplemented with sodium chloride to 250 m*M* and the DEAE-cellulose then added as a slurry (2.3.9). After 10 minutes chilled on ice, in order to bind nucleic acid and protein, the DEAE-cellulose was removed by centrifugation. This resulted in a modest rise in specific FPG activity from 22.0 to 25.3 mU/g TP, as well as a subjective reduction in viscosity as measured by the ease with which the extract could be drawn up into a 25 mL pipette. The process was then repeated using Fraction I (50 mL; 11.0 g TP/L), which generated 56 mL of supernatant

(Fraction II; 7.9 g TP/L). Although the total amount of FPG activity, compared to Fraction I, was reduced 3% by this step, the specific FPG activity was increased from 25.0 to 30.2 mU/g TP because of the removal of some 17% of the total protein.

4.2.3 **Chromatography on cellulose-phosphate**

The DEAE-cellulose treated extract from experiment "C" was dialysed against Buffer A + 50 m*M* sodium chloride so as to allow binding of protein to cellulose-phosphate (2.3.10). After loading of this dialysed sonicate onto the cellulose-phosphate, the column was eluted with a linear salt gradient of 50 to 1000 m*M* sodium chloride (Fig. 4.1). It can be seen that there was a clear distinction between the bulk of the cellular protein which eluted at low salt concentration (i.e. < 250 m*M* sodium chloride) and the peak of FPG which eluted at higher salt concentration (i.e. 300 to 500 m*M* sodium chloride).

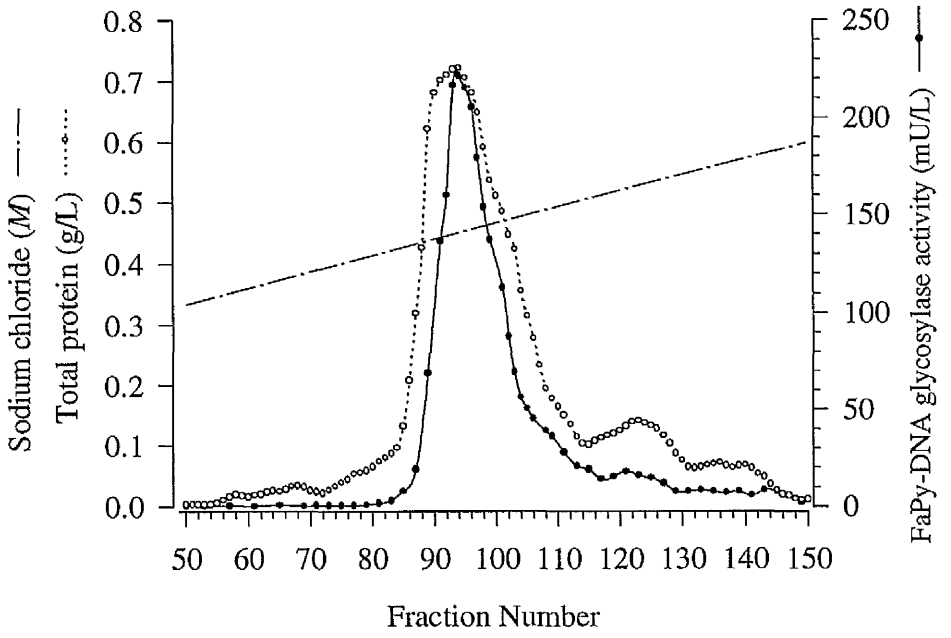
Fig. 4.1
FPG purification: Cellulose phosphate chromatography
50 to 1000 m*M* Sodium chloride gradient



Total protein and FaPy-DNA glycosylase activity (2.3.2) elution profiles of DEAE-cellulose-treated bacterial sonicate (Experiment "C") from cellulose phosphate, by an (assumed linear) 50 to 1000 m*M* sodium chloride (in Buffer A; 2.3.10) gradient. Fraction size = 1 mL. Note the discontinuity in the TP axis scale, necessary to show the peak around fraction 60, as well as the bulk of protein eluted up to fraction 30.

On the basis of this it was decided to omit the dialysis step and instead adjust the sodium chloride concentration in Fraction II from 250 to 200 m*M* by dilution with an appropriate volume of Buffer A. This slightly diluted Fraction II (440 mg TP) was then loaded onto a cellulose-phosphate column (4.5 cm diameter × 12 cm high; volume 190 cm^3), followed by washing with Buffer A + 200 m*M* sodium chloride. The first 85 mL of the wash had a total protein concentration of 3.8 g/L, the next 100 mL had 0.75 g TP/L, and the final 120 mL contained 0.11 g TP/L; the total washings (305 mL) therefore contained 410 mg of protein, showing that the bulk of the contaminating protein had been removed. The column was then eluted with a 200 to 600 m*M* sodium chloride gradient, and the FPG protein was found mostly in fractions 80 to 110, corresponding approximately to 400 to 500 m*M* sodium chloride (Fig. 4.2).

Fig. 4.2
FPG purification: Cellulose phosphate chromatography
200 to 600 m*M* Sodium chloride gradient



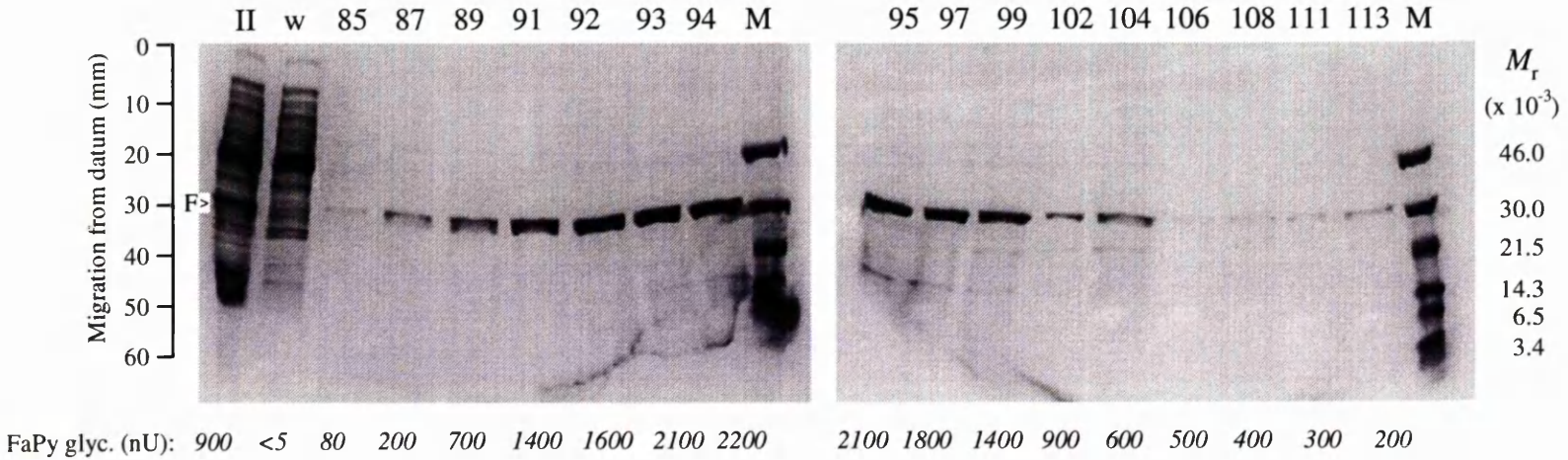
Total protein and FaPy-DNA glycosylase activity (2.3.2) elution profiles of DEAE-cellulose-treated bacterial sonicate (Fraction II) from cellulose-phosphate by an (assumed linear) 200 to 600 m*M* sodium chloride (in Buffer A) gradient. Fraction size = 4 mL, data not shown for fractions 1 to 49 as TP and FaPy-DNA glycosylase levels below limits of detection. Note that sodium chloride and TP concentration share the same axis scale.

In order to further assess the degree of purification achieved, Fraction II, the pooled column washings and a selection of the fractions comprising the peak of FPG activity were analysed by SDS-PAGE (2.3.6; Fig. 4.3a). It could be seen that a prominent band present in Fraction II, with an apparent M_r slightly greater than 30 000, was absent from the pooled column washings. In spite of there being a degree of distortion present in the gels, it was evident that this $M_r \approx 30\ 000$ species was responsible for the bulk of protein in those fractions comprising the FPG activity peak, and that the amount of it in each fraction varied according to the peak in total protein/FPG activity. It was concluded that this $M_r \approx 30\ 000$ species was in fact FPG protein. In fractions 92 to 104 (inclusive) there were also small amounts of lower molecular mass species apparent. In order to obtain an objective estimate of the proportion of FPG the lanes of Fraction II and fraction 91

were scanned densitometrically. This showed (Figs. 4.3b and 4.3c) that approximately 10% of the total protein of Fraction II, and greater than 95% of fraction 91, was FPG protein. The column fractions representing the FPG peak (fractions 88 to 98; 55 mL) were then pooled and concentrated to produce Fraction III (4 mL; 2.3.10).

Fig. 4.3
Further analysis of cellulose phosphate fractions

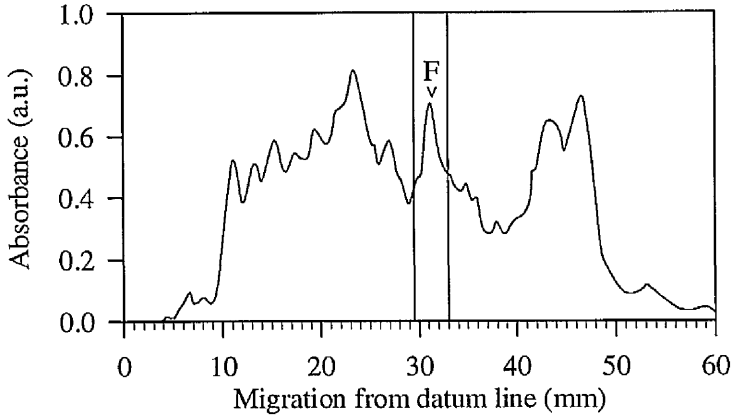
a) SDS-Polyacrylamide gel electrophoresis



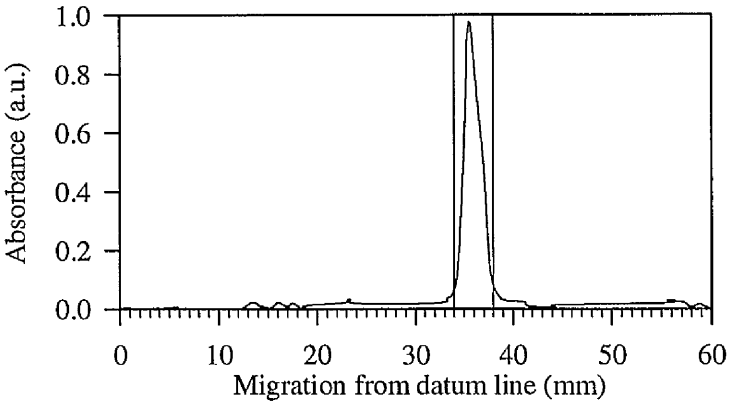
Analysis by SDS-PAGE of Fraction II ("II"; 30 µg TP), pooled column washings ("w"; 30 µg TP), and a selection of the fractions eluted from the column with highest FaPy-DNA glycosylase levels ("85"..."113"; 10 µL of each), numbered as in Fig. 4.2. The italicised values below the gels give the amount of FaPy-DNA glycosylase activity (to 2 s.f.) loaded in that lane. 12% polyacrylamide gels, stained with Coomassie Blue.

Fig. 4.3
Further analysis of cellulose-phosphate fractions

b) Densitometric scan of Fraction II



c) Densitometric scan of fraction 91

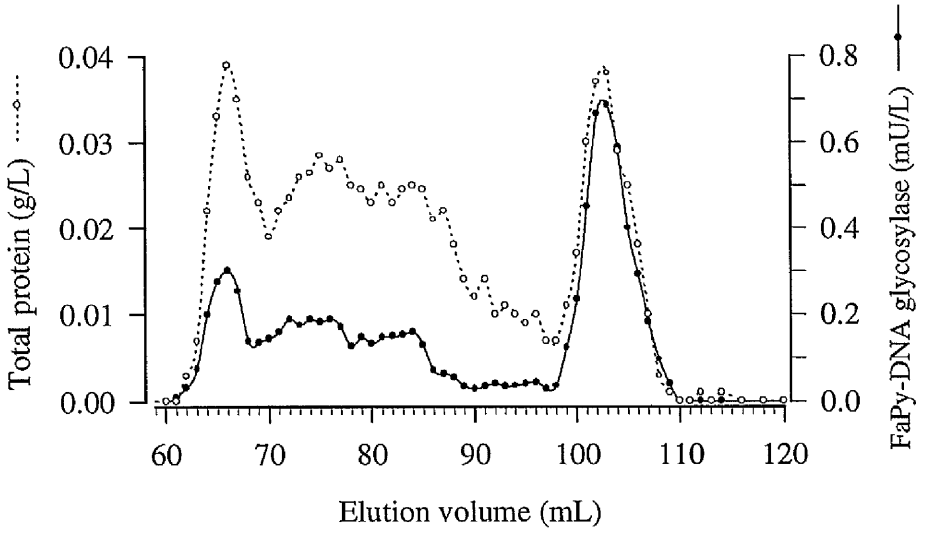


Densitometric scans, from the SDS-PAGE gels shown in Fig. 4.3a, generated by gel recording software (2.2.6). In b) the area under the curve between the dotted lines was 9% of the total area (i.e. from 0 to 60 mm from the datum line). The area similarly indicated in c) was 95% of the total. The difference in the position of the peak in b) marked "F" (31 mm from the datum line) and the major peak in c) (35.5 mm from the datum line) is due to post-electrophoresisartefactual distortion of the gel (see Fig. 4.3a).

4.2.4 **Gel filtration of Fraction III**

Final purification of FPG was carried out by gel filtration on Sephacryl HR-200; the column used was first checked and calibrated using a mixture of marker proteins and blue dextran (Fig. 2.2; 2.3.11). This showed that the column separated proteins in a linear fashion, on the basis of their size ($\log_{10}M_r$) in relation to their elution volume, over the range of M_r = 14 300 to 200 000. A trial experiment was next performed with a small portion of Fraction III (equivalent to 1 mg TP) which had been dialysed against Buffer A + 50 m*M* sodium chloride (Fig. 4.4). It showed two peaks of FPG activity eluting at volumes (83 and 103 mL) corresponding to $M_r \approx$ 110 000 and 31 000, as well as a broad spread in between. This suggested that under these conditions of low salt concentration the FPG protein was forming aggregates of higher molecular mass.

Fig. 4.4
Gel filtration of Fraction III
Low salt: Sodium chloride = 50 m*M*

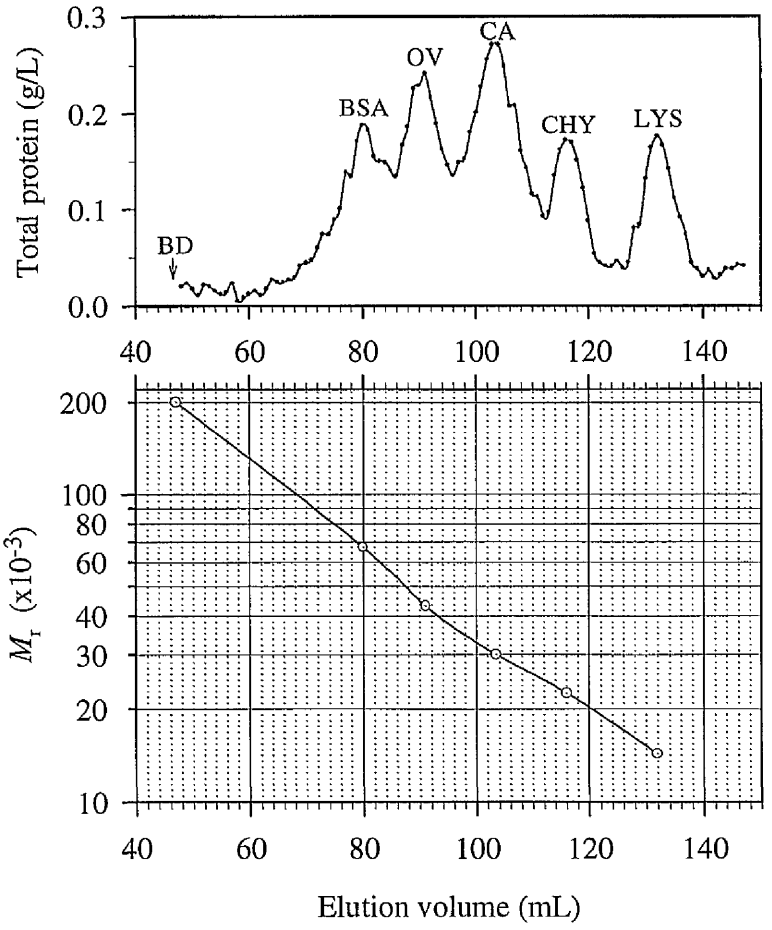


FaPy-DNA glycosylase activity (2.3.2) and TP elution profiles of Fraction III after gel filtration by Sephacryl HR-200 under low salt conditions (50 m*M* sodium chloride in Buffer A). Fraction size = 1 mL, data for fractions 1 to 57 and 121 to 200 not shown as TP concentration below limits of detection.

In order to counteract this tendency the Sephacryl column was re-equilibrated with Buffer A + 1 *M* sodium chloride and the remainder of Fraction III was similarly supplemented with sodium chloride to a final concentration of 1 *M*. Re-calibration of the

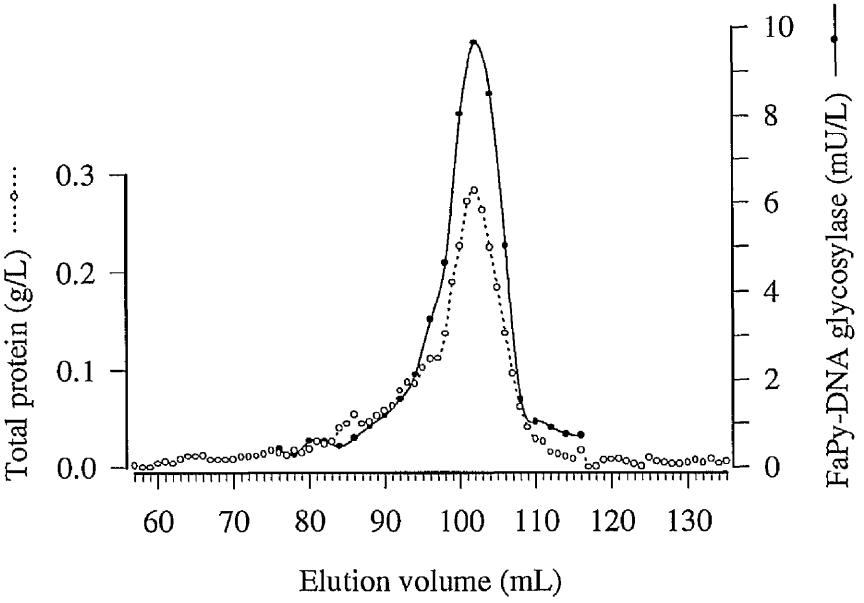
Sephacryl column (Fig. 4.5) showed little change from that with 50 m*M* sodium chloride (Fig. 2.2), although the void volume had been reduced slightly from 50.4 to 47 mL. Under these high salt concentration conditions the FPG activity now eluted as a single peak (Fig. 4.6). The column fractions (98 to 106; 9 mL) comprising the FPG peak were pooled and concentrated to produce Fraction IV (3 mL; 2.3.12). This was stored at -20°C after the addition of glycerol to 50% v/v.

Fig. 4.5
Gel filtration column re-calibration
High salt: sodium chloride = 1 *M*



Recalibration of Sephacryl HR-200 column under high salt conditions (1 *M* sodium chloride in Buffer A). Key: "BD" = position of Blue Dextran elution (= "Void volume"; determined visually), "BSA" = bovine serum albumin, "OV" = ovalbumin, "CA" = carbonic anhydrase, "CHY" = bovine chymotrysinogen, and "LYS" = lysozyme; M_r values are those quoted by the manufacturer (Sigma). Protein concentrations determined by Coomassie Blue G-250 assay (2.3.5). The void volume was reduced to 47 mL (cf. Fig. 2.2; where void volume = 50.4 mL with 50 m*M* sodium chloride), but the linear relationship between $\log_{10}M_r$ and elution volume remained.

Fig. 4.6
Gel filtration of Fraction III
High salt: sodium chloride = 1 *M*



FaPy-DNA glycosylase activity and TP elution profiles of Fraction III after gel filtration by Sephacryl HR-200 under high salt conditions (1 *M* sodium chloride in Buffer A). Fraction size = 1 mL, 160 fractions collected. Data for fractions 1 to 56 and 136 to 160 not shown as TP concentration below limits of detection. FaPy-DNA glycosylase activity (2.3.2) only assayed in even-numbered fractions from 76 to 116 inclusive.

4.2.5 Summary of FPG protein purification

The steps comprising the purification of FPG protein are summarised in Table 4.2.

Table 4.2
Steps in the purification of FPG protein

Fraction	Total protein (mg)	Total FaPy-DNA glycosylase activity (mU)[a]	Yield of FPG protein[b]	Specific FaPy-DNA glycosylase activity (mU/g TP)
I: Primary extract	550	13.8	100%[c]	25.0
II: DEAE-cellulose	440	13.4	97.1%	30.5
III: Cellulose-phosphate	27.5	6.0	43.5%	220
IV: Sephacryl HR-200	11.3	3.9	28.3%	350

[a] by DNA-based assay (2.3.2), [b] in terms of total FaPy-DNA glycosylase activity in Fraction I, [c] by definition.

It can be seen that in three steps a final yield of 28.3% of the initial total FPG activity was achieved; the specific FPG activity was increased nearly 14-fold from 25 to 350 mU/g TP. Assuming a 100% yield at each step in the process then 40 mg of FPG protein was present in Fraction I, representing 7.3% of the initial total protein.

4.3 PRODUCTION AND CHARACTERISATION OF FPG ANTISERA

4.3.1 Raising of antisera to FPG

Polyclonal antisera to FPG protein were raised by immunising two rabbits (Numbers 67 and 69; 2.5.1). After first obtaining blood for pre-immune sera (α67–pre and α69–pre), each animal was immunised by the repeated sub-cutaneous injection of 0.5 mg FPG protein (Fraction IV; 2.5.2). Each animal was given a total of four immunisations, with test bleeds being obtained 10-12 days after the second (α67–2 and α69–2) and third injections (α67–3 and α69–3). Twelve days after the final (fourth) injection the animals were bled out (α67–4 and α69–4; 2.5.1). The intervals between the various stages in the process are shown in Fig. 4.7. The antisera were stored for eventual use as probes in Western blotting of cell and tissue extracts (2.5.1).

Fig. 4.7
Rabbit immunisation schedule

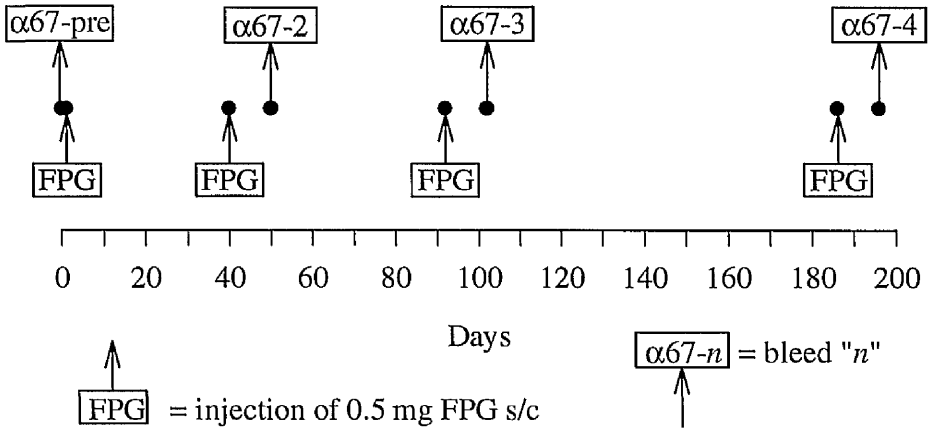


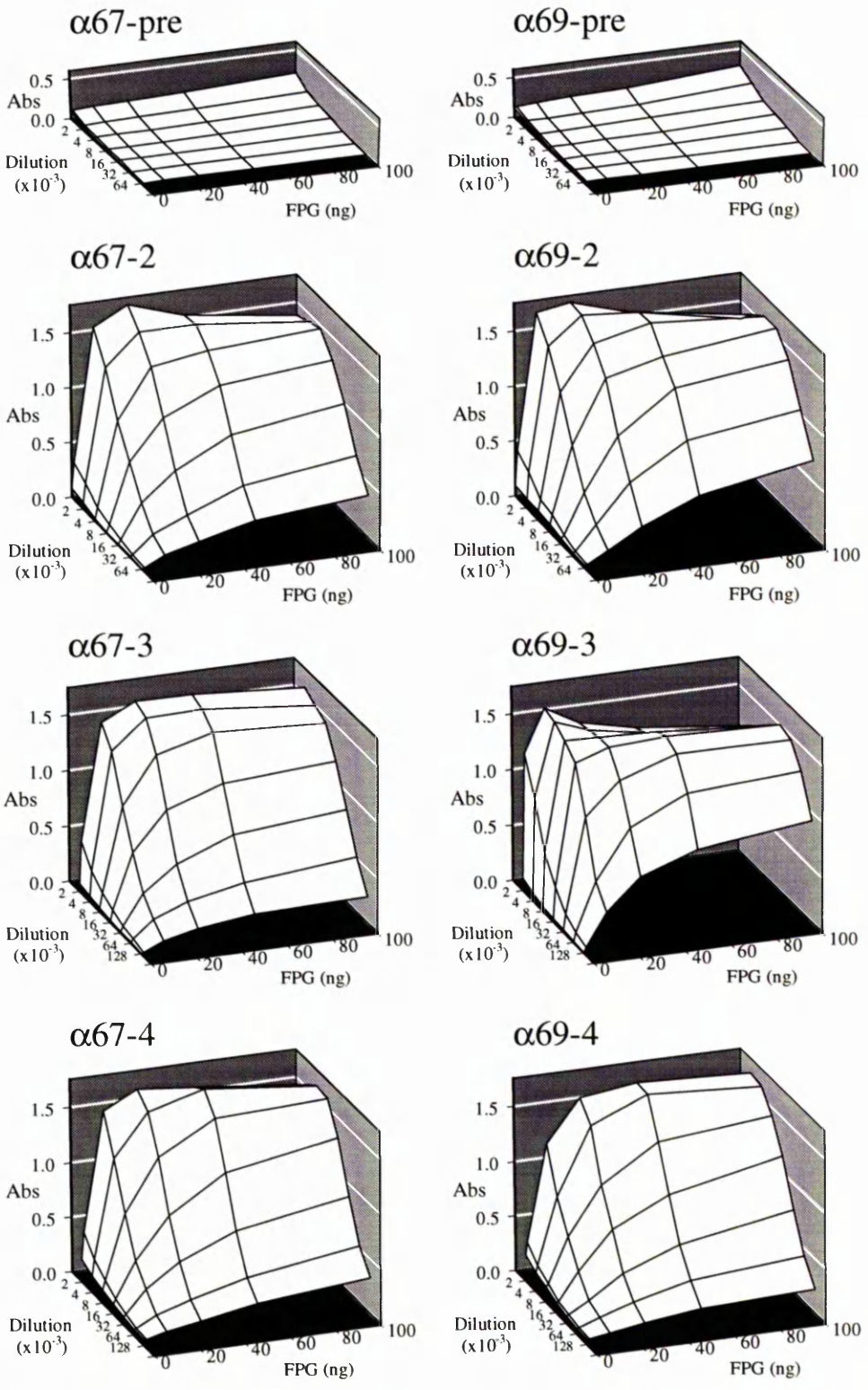
Diagram showing the timing of injections of FPG protein for immunisation, and bleeding for antisera, for rabbit 67; an identical schedule was used for rabbit 69. FPG protein was prepared for injection in incomplete Freund's adjuvant (2.5.1).

4.3.2 **Characterisation of antisera by ELISA**

The response of the animals was followed by measuring the anti-FPG titre of their antisera with an enzyme-linked immunosorbent assay (ELISA) using Fraction IV as the source of FPG protein (2.3.12). The results of this are shown in Fig. 4.8, where it can be seen that prior to immunisation both rabbits' sera had minimal reactivity towards FPG, whereas after immunisation both produced antibodies to FPG protein. At the second bleed rabbit 69's antiserum (α69–2) had the higher titre, as shown by its slightly greater response at the highest dilution (1:64 000). By the time of the third bleed rabbit 67's titre was hardly changed, but rabbit 69's was much higher, giving a response at a dilution of 1:128 000 similar to that of rabbit 67 at 1:32 000. At the final bleed-out rabbit 67's titre was much the same as its previous two bleeds, but that of rabbit 69 had now dropped to a level slightly below that of rabbit 67.

Fig. 4.8

Response of rabbits 67 and 69 to immunisation with FPG protein

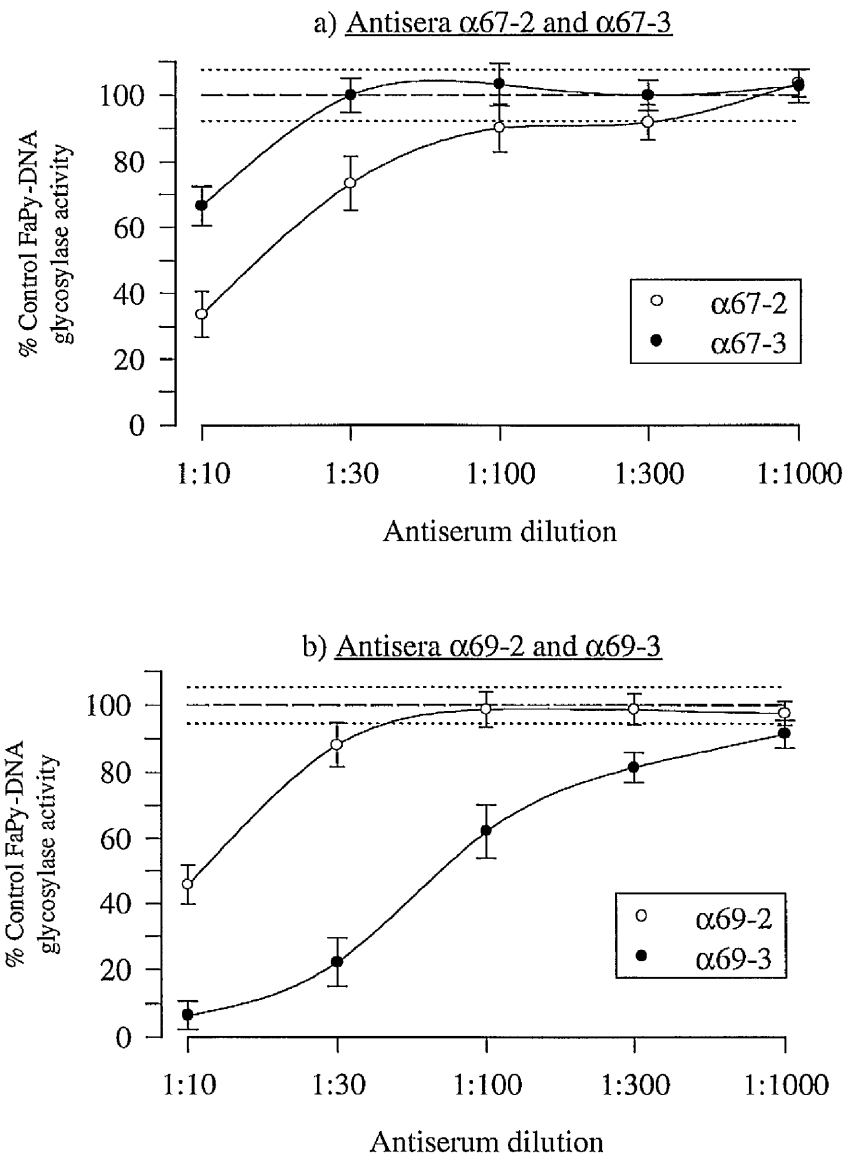


The anti-FPG antibody titres of the pre-immune sera ("α67/9-pre"), test sera after the second and third boosts ("α67/9-2/3"), and after the final (fourth) boost ("α67/9-4") were determined by ELISA (2.3.12). Response of the assay ("Abs") is given in $A_{450\ nm}$ absorbance units; wells were coated with 0, 1, 10, 25, 50 and 100 ng FPG protein; sera were diluted 1:1000 to 1:64 000, or 1:2000 to 1:128 000.

4.3.3 **Effect of antisera on FPG activity**

The ability of the second and third bleed antisera to inhibit the enzymic activity of FPG protein was investigated. The pre-immune and immune sera (α67–pre/2/3 and α69–pre/2/3) from both rabbits, at a range of dilutions, were incubated with a constant amount of purified FPG protein (70 nU; 0.2 μg TP) prior to assay of FaPy-DNA glycosylase activity (2.3.13). Neither of the pre-immune sera had any affect, at any dilution, on the FPG activity when compared to "protein only" controls. Therefore the levels observed with the immune sera were expressed as a percentage of the mean level seen with the respective pre-immune serum (Fig. 4.9). From the results of this it could be seen that α67–2 was partially inhibitory at lower dilutions, and that α67–3 was only inhibitory at the lowest dilution (1:10). In contrast, while α69–2 had an effect intermediate between that of α67–2 and α67–3, α69–3 showed the strongest effect, being inhibitory at all but the greatest dilution and at 1:10 being almost completely inhibitory (FPG activity 4 ± 2% of control).

Fig. 4.9
FPG inhibition by polyclonal rabbit antisera



Liquid-phase inhibition of the FaPy-DNA glycosylase activity of FPG protein was assessed by pre-incubation of a constant amount of FPG protein (70 nU; 0.2 μg TP) with varying amounts of sera (in duplicate; 2.3.13). As no effect was seen with either of the preimmune sera, at any dilution, when compared to FPG plus BSA controls, then the FaPy-DNA glycosylase levels were averaged (dashed lines) ± SEM (dotted lines). FaPy-DNA glycosylase activities with immune sera are expressed as a percentage of the mean activity in the samples containing the same dilution of preimmune serum.

4.4 DISCUSSION

4.4.1 Overexpression of FPG protein in *E. coli*

The results presented in section 4.2.1 show that production of FPG protein by *E. coli* JM105 containing plasmid pKIF10 could be scaled up successfully from the small volumes used in pilot experiments, while at the same time the concentration of IPTG could be reduced without affecting the levels of overexpression. By supplementation of the media with extra sources of energy (carbon source and oxygen), amino-acids (tryptone) and mineral (zinc chloride) it also proved possible to increase the level of overexpression of FPG protein by a factor of over 1.5-fold, from 15.8 to 25 mU/g TP (Table 4.1). The saving in costs achieved by reducing the amount of IPTG easily outweighed the small extra expense of the additional nutrients, thus the steps taken to increase FPG overexpression were worthwhile.

4.4.2 Purification of FPG protein

The problems encountered in purification of FPG protein using DNA-cellulose were somewhat unexpected given the previous experience and success within the Department of purification of DNA-alkyltransferases. This may be because the strain of *E. coli* used for overexpression was different (JM105 *v.* F26; 2.1.5), or that high levels of FPG protein are somehow injurious to DNA-cellulose. More likely, perhaps, there are significant amounts of oxidised and/or ring-opened purines in the DNA moiety of DNA-cellulose and these are acted upon by FPG protein to cause breakdown of the column material; such degradation of the matrix might have been responsible for the problems with excessive backpressure. Whatever the reason or reasons for this, however, it might well have taken a considerable amount of time to resolve the problems with no guarantee that the technique would ultimately have been successful, and it was therefore decided to use more conventional methods.

The use of ion-exchange treatment with DEAE-cellulose as a preliminary step to reduce viscosity by removing nucleic acids as well as some protein worked as intended. From Table 4.2 it can be seen that a modest increase in specific FPG activity was observed in spite of the loss of 3.2% of the FPG activity. This suggests that under the conditions used (in particular the *p*H and salt concentration) little active FPG protein is bound either to nucleic acid or DEAE-cellulose.

A single preliminary experiment defined the conditions, i.e. sodium chloride concentration, for elution of FPG protein from cellulose-phosphate. Once established, this step gave a clean separation of FPG protein from the bulk of other bacterial proteins, with a good yield (43.6%). The product had excellent purity (exceeding 95%), representing a greater than seven-fold increase in specific FPG activity. The strong binding of FPG protein to cellulose-phosphate may represent a more specific form of interaction, i.e. affinity chromatography, rather than a less specific ion-exchange process: FPG protein may bind strongly to cellulose-phosphate because its structure resembles the phosphate backbone of DNA. The fractions comprising that part of the main FPG peak eluting at higher salt concentrations did contain some contaminating species, as shown by the SDS-PAGE analysis, but these were present in low yield and were of lower relative molecular mass. The elution profile of the preliminary experiment (Fig. 4.1) when compared with the purification of the bulk of Fraction II (Fig. 4.2) shows the latter to have two smaller peaks of total protein, containing some FPG activity, eluting at higher salt concentrations (fractions 115 to 131 and 132 to 145, approximately). These extra peaks in the main experiment might be due to the much greater amount of protein separated, i.e. a function of some degree of overloading of the column. Alternatively, they may represent co-elution of FPG with other contaminating species, or an altered form or forms of (possibly denatured) FPG with greater cellulose-phosphate binding ability but reduced FaPy-DNA glycosylase activity.

The results of the subsequent gel-filtration of Fraction III (Fig. 4.4) strongly suggest that FPG protein tends to aggregate or oligomerise under the conditions of high protein concentration, low temperature, buffer and low ionic strength used, which are far from physiological. Also, while there is a smaller peak of FPG activity with an apparent M_r of approximately 110 000 (possibly representing FPG tri- or tetramers) there is a broad spread of activity between it and the main FPG peak (eluting with an apparent $M_r \approx 31\ 000$). The peak at higher relative molecular mass has a similar level of total protein but lower FPG activity, i.e. a lower specific FPG activity, and this may represent an enrichment for inactive FPG species. For these reasons this aggregation or oligomerisation probably does not represent a significant finding in terms of the mode of action of FPG, but it is known that a number of other DNA-binding proteins act by forming oligomers (Sugimoto *et al.*, 1994). The problem of aggregation was resolved by increasing the ionic strength of the buffer (Fig. 4.6). However, there is still a suggestion, given the asymmetry of the FPG activity peak (towards lower fraction numbers and hence higher M_r) that even at 1 *M* salt concentration some degree of aggregation is occurring.

The gel-filtration step did achieve a substantial increase in the specific FPG activity from 220 to 350 mU/g TP, albeit at the loss of one-third of the protein (yield = 28.3% of starting material for Fraction IV, compared to 43.5% of starting material for Fraction III; 28.6/43.6 = 0.65). Because the results of the SDS-PAGE (Fig. 4.3) showed that less than 5% of the total protein of Fraction III was not FPG this increase in specific activity by 1.6-fold suggests that the gel-filtration step separated out inactive FPG protein. This would be consistent with inactive FPG protein being present in an aggregated or oligomeric form, and the earlier finding that gel-filtration under low salt concentration gave higher relative molecular mass FPG species with lower specific activity.

The overall yield of FPG protein given by this method (28.3%; Table 4.2) was similar to that found by Boiteux *et al.* (1990) (41.4%), when they scaled up their original method

(Boiteux *et al.*, 1987). It is of note that in the fourth step (gel-filtration) of their five stage process they quote the buffer used as Buffer A plus 1 *M* sodium chloride, though without giving an explanation. The final yield by their method (Boiteux *et al.*, 1990) was 11.7 mg of FPG from 499 mg TP, figures almost identical to those found with the method presented here (11.3 mg of FPG from 550 mg TP starting material; Table 4.2).

4.4.3 Generation and characterisation of polyclonal antisera to FPG protein

The FPG protein purified from the *E. coli* clone pKIF10 (Fraction IV) was able to elicit an antibody response in both of the rabbits used (4.3). The ELISA assay (2.3.12, 4.3.2 and Fig. 4.8) showed that both pre-immune rabbit sera had very weak activity directed against FPG protein, but then only at the lowest titre (1:1000) and with the highest amount of FPG protein (100 ng), implying antibodies of low specificity and avidity. After two immunisations both rabbits had developed antisera showing maximal activity against FPG protein at a greater titre, i.e. 1:4000 with 100 ng of FPG. Rabbit 69's response at this dilution was highest with 40 ng of FPG suggesting antibodies of greater avidity. Both second bleed antisera showed a diminution in ELISA response at the lowest titre and greatest amount of FPG, indicative of antibody excess over antigen, which is characteristic of specific antibodies of high titre (Harlow and Lane, 1988). At the highest dilution (1:64 000) rabbit 69's antiserum gave an ELISA response with 100 ng of FPG approximately 50% greater than rabbit 67's antiserum, again indicating a stronger antibody response.

After the third immunisation rabbit 67's antiserum (α67–3) had its maximal response at lowest dilution (1:2000), albeit slightly greater at 20 *v.* 100 ng of FPG protein, and this showed that this animal's antibody response had become slightly weaker. In contrast, rabbit 69's antiserum (α69–3) was not only most reactive at a four-fold higher dilution (1:16 000) than the second bleed (α69–2; 1:4000), but its response was approximately

three times greater than either the same animal's second bleed, or rabbit 67's third bleed, with 1 ng of FPG protein at a dilution of 1:2000. Additionally, the profile of response of α69–3 at a dilution of 1:128 000 was similar to that of α67–3 at the four-fold lower dilution of 1:32 000.

The results with the final, fourth bleed, antisera showed rabbit 67's response to be very similar to that found after its second immunisation, with its highest response seen at a dilution of 1:4000 and 40 ng of FPG. Rabbit 69's final response stood in complete contrast to that found after its third bleed. It showed a greatly diminished titre, being maximally responsive only at the least dilution (1:2000) and most FPG protein (100 ng), with no signs of antibody:antigen excess. Progressive rechallenge of an animal with the same antigen is normally expected to produce antisera of steadily increasing titre and avidity (Harlow and Lane, 1988). Why this animal, which had earlier produced an antiserum with high titre and excellent avidity, now showed such a retrograde response is not clear. It may be that it had developed some underlying (but occult) infection or illness causing a form of immunosuppression, or that the effective dose of antigen was reduced by leakage of the FPG-adjuvant emulsion from the sub-cutaneous injection sites, possibly encouraged by the animal licking them. It is more likely that this is an example of immunological tolerance induced by the repeated presentation of antigen (Billingham *et al.*, 1953; Brent, 1991).

The results seen with liquid phase inhibition of FaPy-DNA glycosylase activity by the second and third bleed antisera (section 4.3.3 and Fig. 4.9) also showed a difference in the antibody responses of the two animals. Rabbit 67's antisera were only able to inhibit at the lowest dilutions (1:10 and 1:30 for α67–2; 1:10 for α67–3), and then only partially (by approximately 65% for α67–2, and by approximately 35% for α67–3 at 1:10). That α67–2 was more inhibitory and at greater dilution than α67–3 again indicated that rabbit 67's response had diminished. While rabbit 69's second antiserum showed an ability to

partially inhibit enzymic activity (*ca.* 55% at 1:10), slightly less than that of α67–2, the third bleed antiserum was partially inhibitory (*ca.* 20%) even at a dilution of 1:300, and at 1:10 was almost completely inhibitory (by 96 ± 2%). This fits with the relative ELISA responses for the two third bleed antisera, i.e. rabbit 69 having a more avid response at higher titre, but not for the two second bleed antisera, for which the ELISA results suggested a greater response in rabbit 69, not rabbit 67.

It is perhaps not surprising to find dissimilar results in antibody response to immunisation between two outbred animals. However, it is notable that one animal (rabbit 67) showed a generally consistent response, albeit with a slightly reduced penultimate response, while the other (rabbit 69) showed initially a rapidly increasing titre and avidity, with strongly inhibitory antibodies, which then failed to appear in the final bleed out. It also indicates that the greater diversity in response among outbred animals can be useful in generating antisera with contrasting properties, but that two animals should be the minimum number used in such experiments.

Having obtained FPG antisera then a number of possibilities are opened up. As originally envisaged (1.6) they may be of use as reagents in western blotting, or use as probes to detect mammalian homologues of FPG protein in expression libraries. They could also be used for immunostaining of *fpg*-transfected cells, or tissue sections from *fpg*-transgenic mice, to determine patterns of FPG expression. If they can be shown to inhibit mammalian FaPy-DNA glycosylase activity, then that might be exploited as another means of identifying eukaryotic homologues. The development of immunoassays to specifically measure FPG protein at low levels is also made possible.

CHAPTER 5

THE INTRODUCTION OF THE *fpg* GENE INTO MAMMALIAN CELLS *IN VITRO*

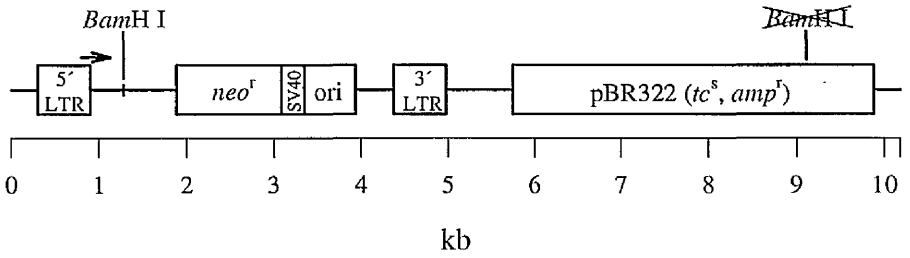
5.1 THE PRODUCTION OF RETROVIRAL-BASED EXPRESSION VECTOR CONSTRUCTS

5.1.1 Choice of vector

Eukaryotic expression vectors based on retroviruses have been in use for several years (Mulligan, 1983). They allow the expression of genes of interest, either pro- or eukaryotic, *in vitro* in cultured mammalian cells. The retroviral long terminal repeat sequences (LTRs) which they contain not only act as strong promoters, but also direct integration of the vector into the host genomic DNA *via* specific sites (Jaenisch, 1976). The 10.2 kb retroviral-based vector pZIP-NeoSV(X)1 (hereafter termed pZIP; Cepko *et al.*, 1984; Fig. 5.1) was available in the Department and there was considerable experience of its use. Initial ligation experiments using this vector, however, were far from satisfactory. Forty eight potential pZIP-*fpg* clones in *E. coli* DH5α were analysed, but none contained a plasmid with a restriction pattern wholly consistent with a single correctly-oriented insert without deletion or rearrangement of the vector, in spite of the host *E. coli* strain (DH5α) being *rec*A$^-$ (2.1.5). During this period, however, the alternative, and in some respects improved, vector pLJ became available (from Dr R.C. Mulligan, Whitehead Institute; Stair *et al.*, 1991; Schreiber *et al.*, 1993).

Fig. 5.1

The retroviral-based eukaryotic expression vector pZIP-NeoSV(X)1

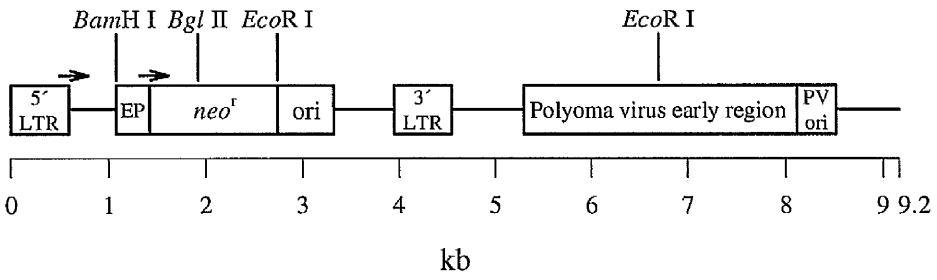


LTR = Long terminal repeat sequences (from the Moloney murine leukaemia virus), → = direction of transcription (the 5′ LTR drives expression of the *neo* gene and any gene inserted into the unique *Bam*H I site), SV40 = viral origin of replication, ori = ColE1 origin of replication (for replication within the prokaryotic host), tc^s = tetracycline resistance gene from which the *Bam*H I site has been removed (thus inactivating it), amp^r = β-lactamase, neo^r = neomycin phosphotransferase. Adapted from Cepko *et al.* (1994).

Like pZIP, the 9.2 kb mammalian cell expression vector pLJ is retroviral based (Fig. 5.2) and contains two Moloney murine leukaemia virus long terminal repeats (MMLV-LTR). The neomycin phosphotransferase (*neo*) gene permits selection of transfected eukaryotic cells with the antibiotic G–418 (Geneticin). In pLJ the *neo* gene has been put under the independent control of the SV40 virus early promoter, unlike pZIP where *neo* expression is driven by the 5′-LTR. Thus, in pLJ the 5′-LTR only drives expression of genes inserted in the unique *Bam*H I site. This has alleviated the problem with pZIP of some inserted sequences causing proviral genome rearrangement (Stuhlmann *et al.*, 1989), or even loss of the inserted gene from cell lines containing stably integrated vector (Dr L. Harris, personal communication). The colE1 origin (from pBR322) allows replication in the bacterial host, while expression of the *neo* gene allows selection of bacteria with kanamycin. The backbone of pLJ, derived from the polyoma virus early region, permits (if required) the rescue of cloned DNAs as infectious virus, under appropriate conditions (Mulligan, 1983; Stair *et al.*, 1991).

Fig. 5.2

The retroviral-based eukaryotic expression vector pLJ



LTR = Long terminal repeat sequences (from the Moloney murine leukaemia virus), → = direction of transcription (the 5´ LTR drives expression of any gene inserted into the unique *Bam*H I site), EP = SV40 virus early promoter, ori = ColE1origin of replication, PVori = polyoma virus origin of replication. Adapted from map provided by Dr R.C. Mulligan (Whitehead Institute).

5.1.2 **Production of pLJ/*fpg* constructs**

A complete digestion of pUIF4 and a partial digestion of pUIF8 with *Bam*H I was carried out to produce full-length 0.85 kb *fpg* protein-coding sequences (*fpg*4 and *fpg*8; 3.3.4). These were purified by agarose gel electrophoresis and isolated by glasspowder adsorption (2.2.8). When aliquots of the 0.85 kb inserts so obtained were analysed by agarose gel electrophoresis they were found to be homogeneous and of much improved yield, compared to the earlier method of LMP-agarose gel electrophoresis (2.2.7). The *fpg*4 and *fpg*8 *Bam*H I inserts were ligated into *Bam*H I/CIAP digested pLJ and the products used to transform *E. coli* DH5α, with selection on LB agar containing kanamycin (2.1.6 and 2.2.18), the efficiencies of this are given in Table 5.1.

Table 5.1

Transformation efficiency of DH5α with pLJ/*fpg*4 and *fpg*8

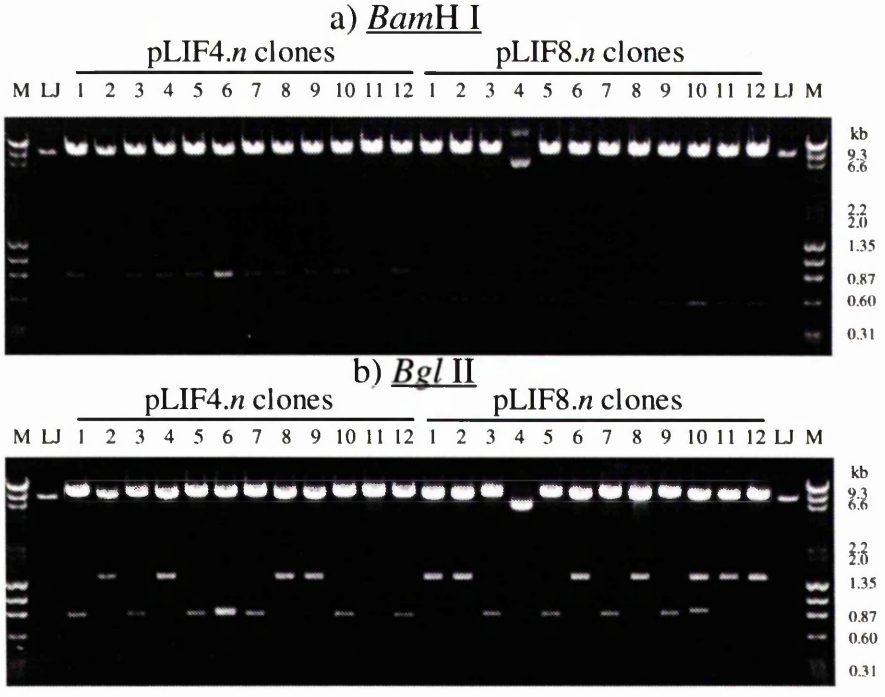
Vector	Efficiency[a]
pLJ (uncut)	4.0×10^6
pLJ/*Bam*H I	0.7×10^6
pLJ/*Bam*H I/CIAP	1.0×10^3
pLJ/*Bam*H I/CIAP + *fpg*4	45×10^3
pLJ/*Bam*H I/CIAP + *fpg*8	34×10^3

[a] kanamycin resistant colonies per μg vector DNA.

Twelve clones of each *fpg* version, named pLIF4.1 to 4.12 and pLIF8.1 to 8.12 were analysed by digestion of the respective plasmids with *Bam*H I and *Bgl* II, followed by agarose gel electrophoresis (Fig. 5.3).

Fig. 5.3

Restriction analysis of pLIF4 and pLIF8 plasmids

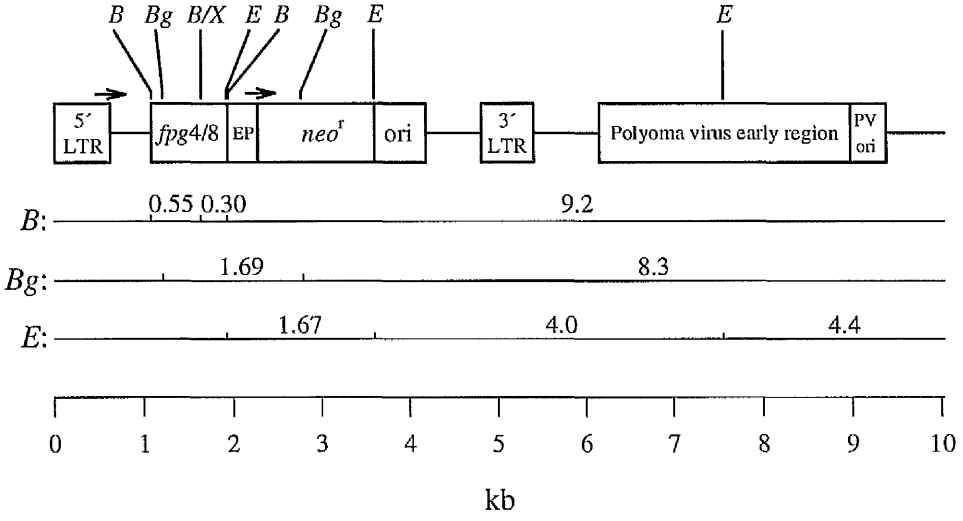


DNA from small-scale preparations of the plasmids was digested with a) *Bam*H I, or b) *Bgl* II. "LJ" indicates lanes loaded with pLJ digested with the respective enzyme, the numbers below the horizontal bars = *n* (as in "pLIF4.*n*" etc.).

With reference to the map of the intended pLIF constructs (Fig. 5.4), the *Bam*H I digestions indicated that 22 out of the 24 clones had a 0.85 kb insert, with two clones (pLIF4.6 and 8.10) appearing to have more than one insert, as judged by the greater intensity of the 0.85/0.55 kb bands seen with these clones relative to the others. Of the two clones without an insert, pLIF4.11 was simply religated vector while pLIF8.4 seemed to have undergone a deletion. The *Bgl* II digests showed that 10 clones had a single sense-oriented insert (pLIF4.2, 4.4, 4.8, 4.9, 8.1, 8.2, 8.6, 8.8, 8.11 and 8.12), while in 10 others it was anti-sense (pLIF4.1, 4.3, 4.5, 4.7, 4.10, 4.12, 8.3, 8.5, 8.7 and 8.9). They also confirmed that plasmids pLIF4.6 and 8.10 had two inserts, both anti-sense in pLIF4.6, but in pLIF8.10 both sense-oriented.

Fig. 5.4
The intended structure of the pLJ/*fpg* (pLIF) constructs

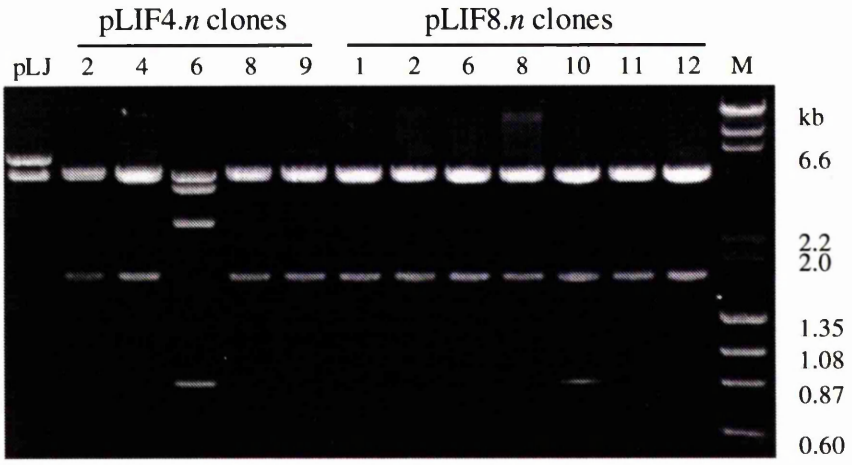
The predicted structure of a single *fpg*4 or *fpg*8 insert ligated, in the correct orientation, into the unique *Bam*H I site of pLJ:



LTR = Long terminal repeat sequences, → = direction of transcription, EP = SV40 virus early promoter, ori = ColE1 origin of replication (from pBR322), *fpg*4/8 = pUIF4 or pUIF8-derived fpg inserts, *B* = *Bam*H I, *B/X* = *Bam*H I (if *fpg*8) or *Xho* II (if *fpg*4), *Bg* = *Bgl* II, *E* = *Eco*R I, the numbers above the three scales give the size (kb) of the fragments predicted to be produced by the respective restriction endonucleases. N.B. with insert *fpg*4 *Bam*H I digestion should give only 0.85 and 9.2 kb fragments.

As a further check (after the experience with pZIP) the sense-oriented constructs were digested with *Eco*R I together with the two double-insert-containing plasmids (Fig. 5.5). The *Eco*R I digest of pLJ produced the two bands of 5.2 and 4.0 kb predicted from the map (Fig. 5.4). The sense-oriented clones all give the same result with *Eco*R I (i.e. 4.4, 4.0 and 1.67 kb bands) proving that they all had a single sense-oriented *fpg*4 or *fpg*8 insert (Fig. 5.4). The two plasmids containing double-inserts each gave a 0.85 kb band showing that in each one the two inserts were oriented in the same sense; the pattern (4.4, 4.0 and 1.67 kb) confirming the sense orientation of the inserts in pLIF8.10, and anti-sense orientation in pLIF4.6 (4.0, 3.5 and 2.5 kb).

Fig. 5.5
*Eco*R I restriction analysis of selected pLIF4 and pLIF8 plasmids



DNA from small-scale preparations of those plasmids in Fig. 5.3 which appeared to contain either single sense-oriented inserts, or double-inserts, was digested with *Eco*R I. "pLJ" indicates lane loaded with pLJ digested with *Eco*R I.

These results were a great improvement over those seen with pZIP, given twenty out of twenty four clones with a single insert and only one (pLIF8.4) aberrant. As a consequence of this and the expected greater stability of pLJ constructs in mammalian cells, it was decided to use the pLIF4 and pLIF8 constructs to transfect eukaryotic cells

in vitro. Thus large-scale preparations were made of pLIF4.2 and 8.1 and these plasmids were re-designated pLIF4 and pLIF8.

5.2 INTRODUCTION OF *fpg*4 INTO MAMMALIAN CELLS

5.2.1 Transfection of RJKO cells with pLJ constructs

Chinese hamster lung fibroblasts (RJKO cells) were grown and transfected with pLJ, pLIF4 and pLIF8 (2.4.1). During this time the revised sequence and FPG activity data for *fpg*8 became available (3.3.5, 3.4.3 and 3.6.3 *et seq.*), showing it to be inactive because of a mutation causing a -1 bp frameshift. Hence four colonies from the pLJ (J1 to J4) and twelve from the pLIF4 (4L1 to 4L12) transfection were picked. A line of mock-transfected RJKO cells were maintained in G–418-free medium for comparison of FaPy-DNA glycosylase activity.

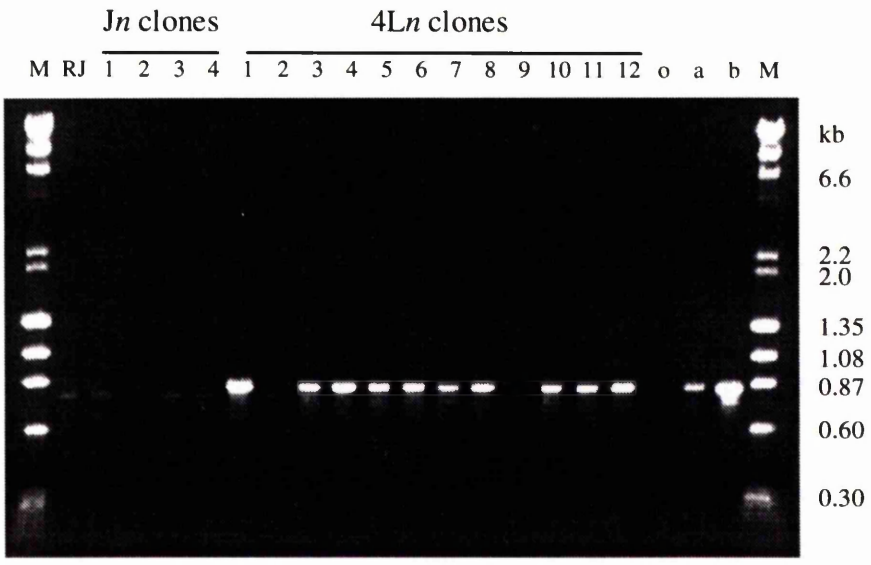
5.2.2 Analysis of RJKO clones

The G–418-resistant RJKO cell clones (J1 to 4 and 4L1 to 12) were grown to confluence, harvested and then extracts made for assay of FPG activity (2.3.3). Extracts of RJKO and mock-transfected RJKO lines were also made. The mean (SEM) of each of the controls using the DNA-based FaPy-DNA glycosylase assay (2.3.2) was 4.0 (2.0) μU/g TP for the RJKOs and 3.2 (1.8) μU/g TP for the mock-transfected line (n = 12 in each case). None of the sixteen G–418-resistant lines (J1 to J4 and 4L1 to 4L12) had detectably increased levels of FPG, all having mean levels between 2.7 and 4.7 μU/g TP (n = 3 in each case). Although these results appeared conclusive the lines were also analysed by PCR.

Aliquots of the cell lines taken at passage were analysed by PCR (2.2.4 and 2.2.5) for the presence of *fpg* sequence (Fig. 5.6). It can be seen that the positive controls (pLIF4 and RJKO + pLIF4) resulted in a 0.85 kb band while no bands were apparent in the negative controls (no added DNA and RJKO cell extracts). All extracts, except those from lines 4L2 and 4L9, gave a 0.85 kb band as in the positive controls. These results suggested that ten of the twelve pLIF4 transfected lines contained the bacterial *fpg*4

sequence. Southern analysis of genomic DNA from the clones was therefore warranted and the respective clones were therefore grown up to provide enough DNA for this. While this was in progress, however, it had become apparent (3.6.3) that the mutations in *fpg*4 had severely reduced the FaPy-DNA glycosylase activity of the protein encoded by it (FPG4), and therefore (in spite of the poor precision of the DNA-based FaPy-DNA glycosylase assay at low-levels of activity) the inability to detect raised FaPy-DNA glycosylase levels in the 4L1 to 12 clones appeared to be accounted for. The fact that both *fpg*4 and *fpg*8 were compromised by mutations meant that it would be necessary to use *fpg*10 instead.

Fig. 5.6
PCR analysis of pLJ- and pLIF4-transfected RJKO clones



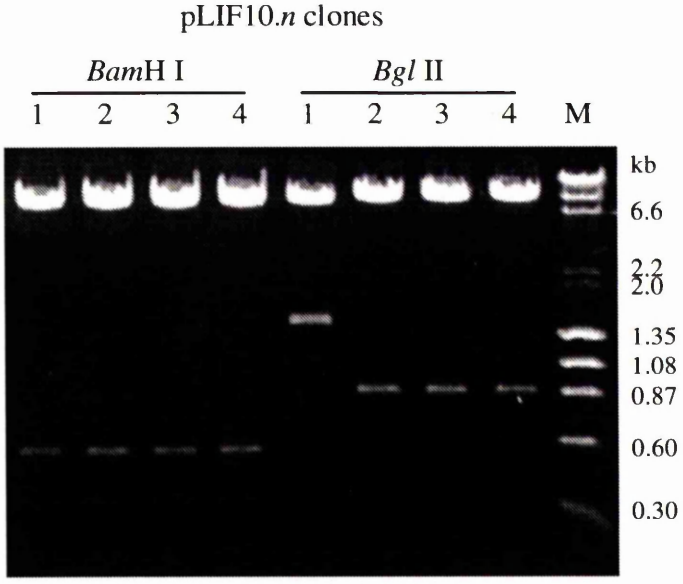
Samples were taken from the various cell lines and analysed by PCR using oligonucleotides 184 and 185. "RJ" = RJKO cells, "o" = no DNA, "a" = pLIF4, and "b" = pLIF4 + RJKO controls, numbers below bars refer to clone numbers ("*n*").

5.3 PRODUCTION OF MAMMALIAN CELL EXPRESSION VECTORS CONTAINING *fpg*10

5.3.1 Construction of the pLJ-*fpg*10 expression vector

The *Bam*H I 0.85 kb insert from pUIF10 (*fpg*10) was ligated into pLJ and used to transform *E. coli* DH5α. Following the previous success in this respect (5.1.2) with *fpg*4 and *fpg*8 the plasmids from just four clones were analysed by restriction pattern analysis with *Bam*H I and *Bgl* II (Fig. 5.7). All of the clones contained a single insert (one sense and three anti-sense); the sense-oriented clone pLIF10.1 (redesignated pLIF10) was therefore selected for large-scale preparation and transfection of RJKO cells.

Fig. 5.7
Restriction analysis of pLIF10.*n* plasmids



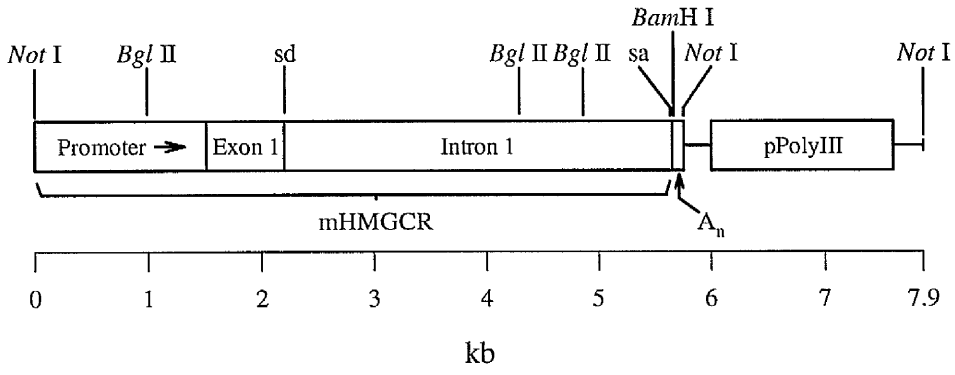
DNA from small-scale preparations of the various plasmids was digested with either *Bam*H I or *Bgl* II and analysed by AGE. The numbers below the bars refer to "*n*" (as in "pLIF10.*n*").

5.3.2 The mammalian cell expression vector pHMG

The 7.9 kb plasmid vector pHMG is based on the promoter of the mouse hydroxymethylglutaryl-Coenzyme A reductase gene (HMGCR; Gautier *et al.*, 1989; Fig. 5.8); the enzyme product of this gene catalyses the initial rate-limiting step in the

biosynthesis of cholesterol, and as the gene is constitutively expressed in every cell type is described as a "housekeeping" gene. Such genes are characterised not only by their ubiquitous expression but also their high degree of conservation between species. Putting a foreign gene under the control of a mammalian "housekeeping" gene promoter offers the potential of a most useful vector. Both cell lines and transgenic mice have been produced containing constructs of pHMG (Gautier *et al.*, 1989). The cell lines gave levels of expression between 94 and 104% of that obtained with the vector pRSV-L, based on the SV40 early promoter. Analysis of the transgenic mice revealed expression in a wide range of tissues, including all major internal organs, integument and musculo-skeletal tissue (Gautier *et al.*, 1989; De Wet *et al.*, 1987).

Fig. 5.8
The mammalian cell expression vector pHMG



mHMGCR = region derived from the mouse Hydroxymethylglutaryl-Coenzyme A Reductase gene, sd = splice donor site, sa = splice acceptor site, A_n = SV40 polyadenylation signal, pPolyIII = region derived from pPolyIII, containing ColE1 originof replication and *amp*^{r} (β-lactamase), → = direction of transcription (the transcription start site is at the promoter/exon 1 boundary; exon 1 is untranslated). The plasmid is shown as if linearised at one of the two *Not* I sites; digestion with *Not* I can be used to remove the pPolyIII-derived fragment. Genes containing start codons can be ligated into the unique *Bam*H I site: intron 1 is spliced out in the mRNA. Adapted from Gautier *et al.* (1989).

The structure of pHMG is derived from a fusion of pPolyIII (providing a bacterial origin of replication and β-lactamase gene for ampicillin selection) with the mouse HMGCR gene (Fig. 5.8). Intron I of HMGCR is linked to the SV40 polyadenylation signal by a short segment containing a unique *Bam*H I site, into which may be inserted

protein-coding sequences containing their own start and stop codons. Correct transcript processing occurs, as predicted, between the major HMGCR splice donor site (sd) and the invariant splice acceptor site (sa; Gautier *et al.*, 1989).

5.3.3 Cloning of *fpg*10 in pHMG

The *fpg*10 0.85 kb *Bam*H I insert (as used with pLJ; 5.3.1) was ligated into *Bam*H I/CIAP digested pHMG (2.2.14; Fig. 5.9) and used to transform DH5α to ampicillin resistance (2.2.18; Table 5.2).

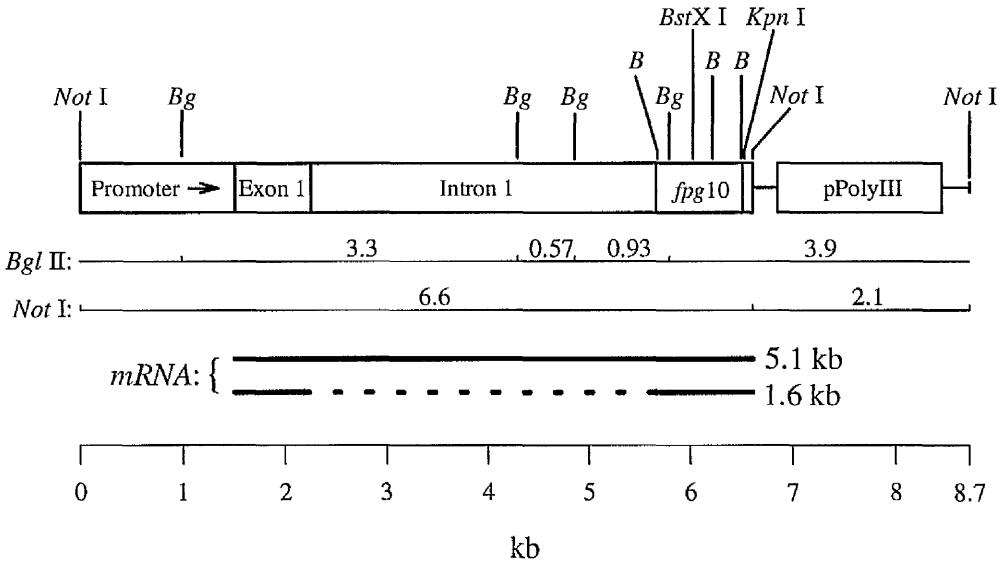
Table 5.2
Transformation efficiency of DH5α with pHMG-*fpg*10

Vector	Efficiency[a]
pHMG (uncut)	9.5×10^6
pHMG/*Bam*H I/CIAP	45×10^3
pHMG/*Bam*H I/CIAP + *fpg*10	400×10^3

[a] ampicillin resistant colonies per μg vector DNA.

Fig. 5.9
The intended structure of the pHMG/*fpg*10 (pHIF10) construct

The predicted structure of a single *fpg10* insert ligated, in the correct orientation, into the unique *Bam*H I site of pHMG:

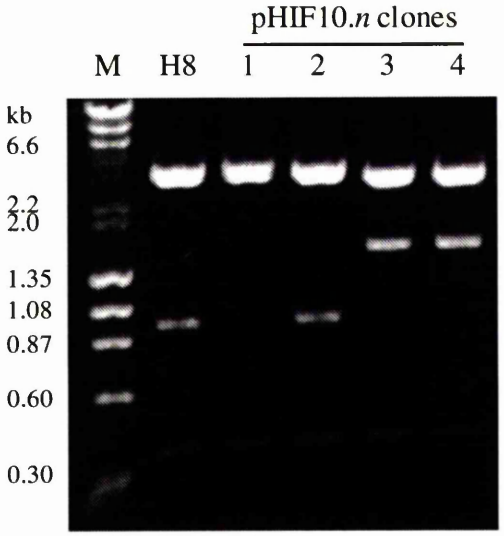


"*fpg*10" = *fpg*10 *Bam*H I insert, *B* = *Bam*H I, *Bg* = *Bgl* II, *mRNA*: { = predicted RNA transcripts (with and without splicing of intron 1), sizes on right. The numbers above the upper two scales (below the plasmid map) give the size in kb of the DNA fragments predicted to be produced by the respective restriction endonucleases.

Digestion with *Bgl* II of plasmids from ampicillin resistant clones showed that, out of the four tested, one had a correctly-oriented insert (pHIF10.2), two had an incorrectly-oriented insert (pHIF10.3 and 10.4) and one had no insert ((pHIF10.1; Fig. 5.10). The correctly-oriented 8.7 kb plasmid (pHIF10.2, renamed pHIF10) was digested with *Not* I to release the 6.6 kb HMGCR-*fpg*-polyA unit from the 2.1 kb pPolyIII-derived fragment (as in Fig. 5.9). These were then separated by electrophoresis (using LMP agarose) and the 6.6 kb fragment purified by adsorption onto glasspowder (2.2.8). Analysis of the purified 6.6 kb fragment by agarose gel electrophoresis provided an estimate of its concentration and confirmed its purity (Fig. 5.11). No contaminating bands being present, the DNA was prepared for microinjection into fertilised mouse ova (2.5.3), in order to produce *fpg*10-transgenic mice (Chapter 7).

Fig. 5.10

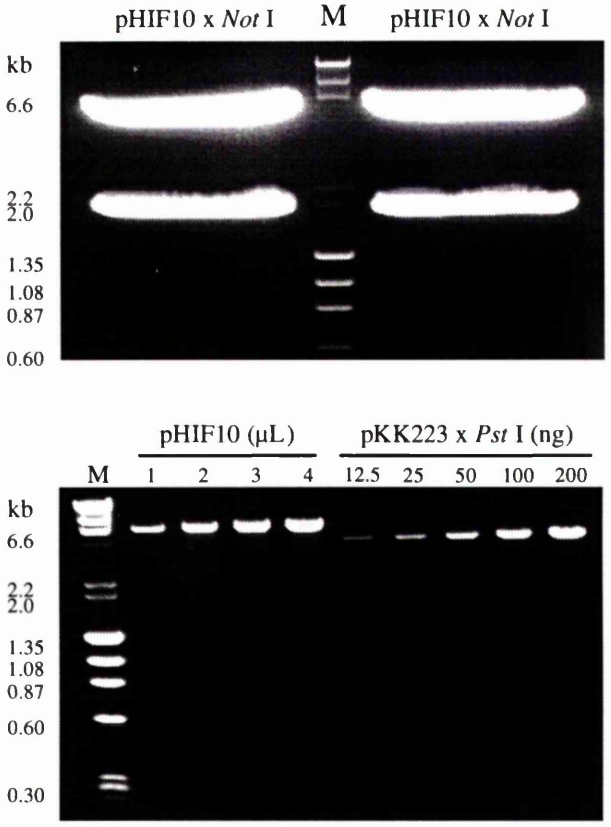
Bgl II restriction analysis of pHIF10.*n* plasmids



DNA from small-scale preparations of the various plasmids was digested with *Bgl* II and analysed by AGE. "H8" = pHIF8 digested with *Bgl* II, the numbers below the bars refer to "*n*" (as in "pLIF10.*n*").

Fig. 5.11

Isolation of the 6.6 kb *Not* I fragment from plasmid pHIF10



Caesium chloride gradient-purified plasmid pHIF10 (66 µg) equivalent to 50 µg of 6.6 kb *Not* I fragment was digested with *Not* I (100 U), half was loaded in each lane (upper gel). After electrophoresis in LMP agarose with TAE buffer the gel containing the 6.6 kb fragment was excised and the DNA purified by adsorbtion on glassmilk (2.2.8). It was then eluted into 400 µL of sterile water. Portions were analysed by AGE (lower gel), in comparison with known amounts of pKK223-3 digested with *Pst* I, to provide an estimate of the concentration: 1 µL appeared to contain 50 ng DNA, giving a yield of 20 µg (i.e. 40% of the starting material).

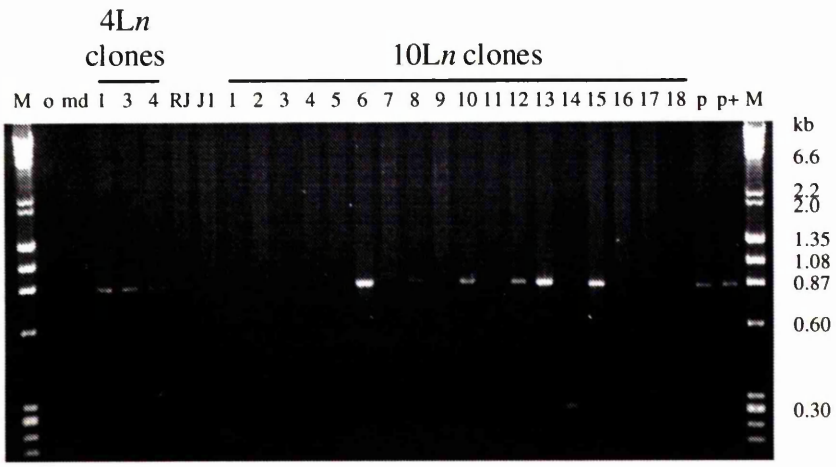
5.4 TRANSFECTION OF RJKO CELLS WITH pLIF10 AND pHIF10

5.4.1 Initial attempts to introduce pLIF10 into RJKO cells

As in section 5.2.1 RJKO cells were transfected with pLIF10 (2.4.2). Geneticin-resistant clones were selected and samples for PCR analysis were taken as early on in the process as possible, i.e. at the time of clone/colony picking (2.2.4, 2.4.2). Of the eighteen clones picked (designated 10L1 to 10L18) only six (10L6, 10L8, 10L10, 10L12, 10L13 and 10L15) were PCR-positive for the *fpg* sequence (Fig. 5.12). It was

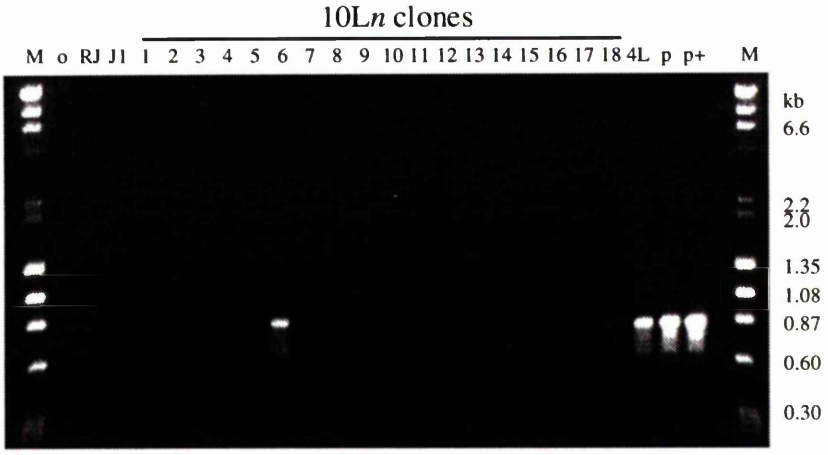
thought that this minority of positive clones was due to inhibition of the PCRs by the crude cell extracts used. So the clones were expanded to produce enough DNA for Southern analysis, and some of this purified DNA was analysed by PCR. Somewhat surprisingly, it now only proved possible to get a PCR-positive result with clone 10L6 (Fig. 5.13).

Fig. 5.12
PCR analysis of RJKO cell lines 10L1 to 10L18 at cloning



Samples were taken from the 10L*n* cell lines at the time of cloning and analysed by PCR using oligonucleotides 184 and 185. Similar samples were taken from RJKO cells and lines J1, 4L1, 4L3 and 4L4 as controls. "o" = no DNA, "md" = cloning medium, "p" = pLIF10 and "p+" = pLIF10 + RJKO extract controls.

Fig. 5.13
PCR analysis of RJKO cell lines 10L1 to 10L18
from purified DNA after expansion of clones

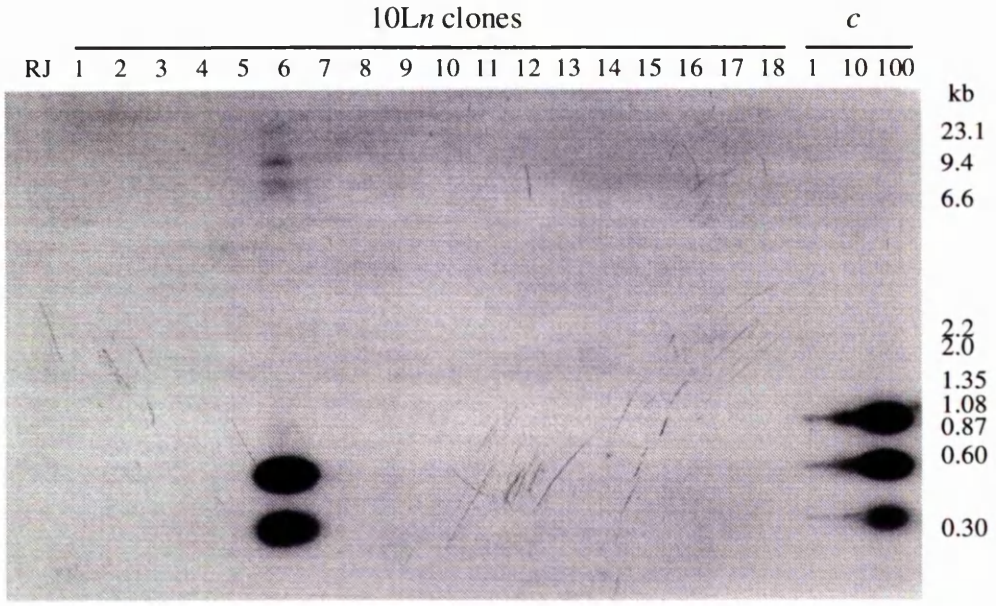


The 10L*n* cell lines were expanded to provide purified DNA for Southern analysis, and this DNA was analysed by PCR using oligonucleotides 184 and 185. DNA from RJKO ("RJ") cells and lines J1 ("J1") and 4L1 ("4L") were included as controls. "o" = no DNA, "p" = pLIF10 and "p+" = pLIF10 + RJKO DNA controls

5.4.2 **Southern analysis of RJKO clones 10L1 to 10L18**

The RJKO-clones transfected with pLIF10 (10L1 to 18) were checked for the presence of *fpg* sequence by Southern analysis, after restriction of their DNA with *Bam*H I and probing with the 0.85 kb *Bam*H I *fpg*4 fragment ([^{32}P]-labelled; 2.2.20). Positive controls were included, consisting of *Bam*H I digested pLIF4 and pLIF8 combined with RJKO DNA in amounts corresponding to copy-numbers (of *fpg* per RJKO haploid genome) of 1, 10 and 100. An autoradiograph of the subsequent blot is shown in Fig. 5.14. This confirmed the second PCR findings (Fig. 5.13) in that the only positive clone was 10L6, and in addition, showed that in this clone the *fpg*10 copy-number was in excess of 100/haploid genome.

Fig. 5.14
Southern analysis of RJKO clones 10L1 to 10L18



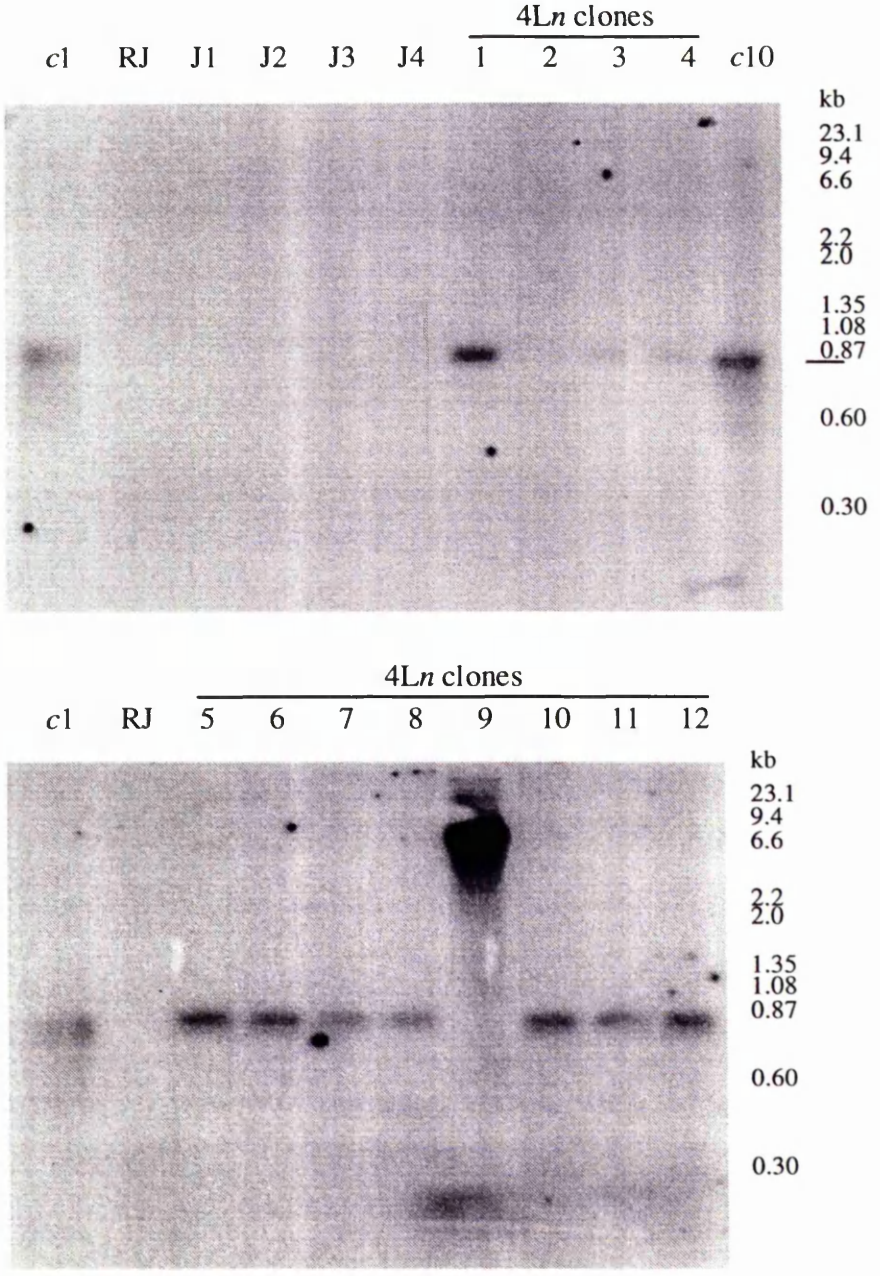
DNA from each cell line (10 µg) was digested with 50 U *Bam*H I and subjected to Southern analysis (2.2.20; capillary blotted). "RJ" = RJKO DNA. "*c*" = *fpg* copy-number controls: a mixture of pLIF4 and pLIF10 was digested with *Bam*H I and mixed with similarly digested RJKO DNA in ratios simulating *fpg* copy numbers (*c*) of 1, 10 and 100 per haploid genome.

5.4.3 **Further analysis of the RJKO clones 4L1 to 4L12**

The RJKO clones produced by transfection with pLIF4 (i.e. 4L1 to 4L12) had earlier been analysed by PCR and all but two (4L2 and 4L9) found to give positive results. Southern analysis of these clones' DNA was now performed, together with some DNA from clone 10L6 and the 1- and 10-copy number controls as before (5.4.2). The autoradiographs from this are shown in Fig. 5.15. Of the 4L clones it can be seen that, apart from 4L2 and 4L9, they all gave the expected 0.85 kb *Bam*H I band and had copy-numbers from 1 to 10. Of the two PCR-negative clones, 4L2 did not give a hybridisation signal, and 4L9, whose copy-number appeared to be substantially greater than 10, gave an aberrant restriction pattern with a broad band (*ca.* 6 to 10 kb) of substantially greater size than expected. Thus, as in the case of the 10L clones, the PCR result had agreed with the Southern analysis. It was concluded from this that, at least for

the initial analysis of clones, PCR testing for the presence of intact *fpg* sequence was as efficient as Southern blotting, and as it was considerably quicker would be the method of choice in further experiments.

Fig. 5.15
Southern analysis of RJKO clones J1 to J4 and 4L1 to 4L12



DNA from each cell line (10 µg) was digested with 50 U *Bam*H I, and subjected to Southern analysis (2.2.20; capillary blotted). "RJ" = RJKO DNA. "c" = *fpg* copy-number controls: pLIF4 was digested with *Bam*H I and mixed with similarly digested RJKO DNA in ratios simulating *fpg* copy numbers (*c*) of 1 and 10 per haploid genome.

5.4.4 Clone 10L6: FaPy-DNA glycosylase activity

Southern analysis of clone 10L6 had shown it to contain *fpg*10 at a high copy-number of about 100 (5.4.2), and so its FaPy-DNA glycosylase activity was measured to see if expression of FPG could be detected. Using the DNA-based FaPy-DNA glycosylase assay (2.3.2), the mean (SEM) FaPy-DNA glycosylase activity of clone 10L6 in three separate experiments was found to be 4.5 (1.8) μU/g TP, a level which was not significantly different by *t*-test from that of the RJKO cells measured in parallel (3.3 (2.7) μU/g TP). Thus no striking difference in FaPy-DNA glycosylase activity could be demonstrated at this stage. However, because the DNA-based assay had insufficient precision at these low levels (5.2.2) there may have been a small but statistically significant difference. This prompted the introduction of the potentially more sensitive polyGC-based FaPy-DNA glycosylase assay (2.3.4), as used by Boiteux *et al.* (1987).

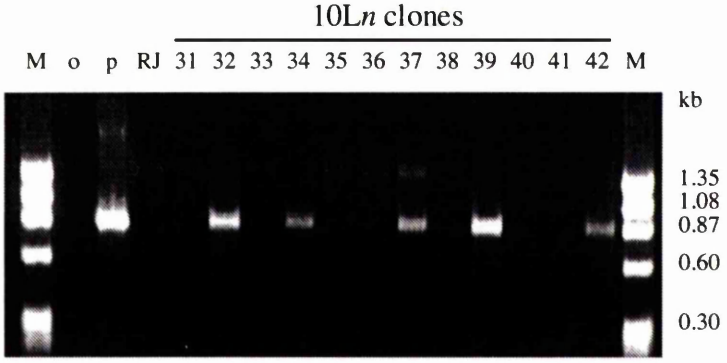
5.4.5 The results of further transfections of RJKO cells with pLIF10

As it was considered possible that some error had been made in the original RJKO/pLIF10 transfection, which could explain the relatively poor results obtained in efficiency of successfully transfected clones compared with pLIF4, a second transfection was set up using the same methods as before. In this repeat experiment a total of twelve clones were picked (10L19 to 10L30). At no stage did any of these prove to be positive by PCR, and neither did any prove to be positive on Southern analysis (data not shown).

Although the selection conditions (G–418 concentration and the length of time cells were exposed; 2.4.2) were no more rigorous than those used at other times in the Department to isolate RJKO clones containing other DNA-repair genes (Mrs A.J. Watson, personal communication), it was considered possible that the conditions were too severe for selecting *fpg*-containing clones. A third transfection experiment was therefore carried out, this time reducing the time spent at the highest concentration of G–418 from seven

to four days (2.4.2). A further twelve clones (10L31 to 10L42) were selected under these revised conditions and analysed by PCR as before (2.2.4; 2.2.5; Fig. 5.16). A total of five clones proved positive (10L32, 10L34, 10L37, 10L39 and 10L42) and together with 10L6 these were expanded for additional analyses.

Fig. 5.16
PCR analysis of RJKO clones 10L31 to 10L42



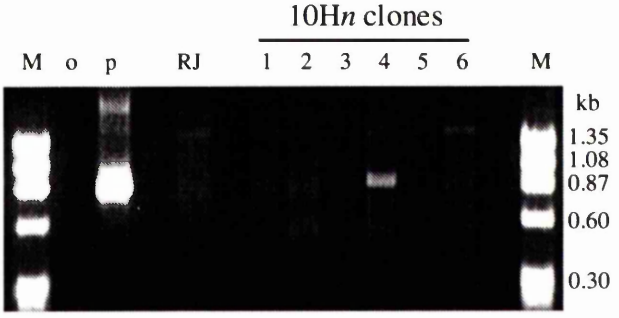
Crude extracts of the clones were analysed by PCR using oligonucleotides 184 and 185 (2.2.4, 2.2.5). "RJ" = RJKO cells, "o" = no DNA, "p" = pLIF10 + RJKO controls.

5.4.6 **Transfection of RJKO cells with pHIF10**

Simultaneously with the third pLIF10 transfection (5.4.5) it was decided to try transfection of RJKO cells with the 6.6 kb *Not* I pKIF10 fragment (5.3.3). The pHMG vector had been successfully used in the Department to achieve over-expression of human ATase (hAT) in XP cells; co-transfection of a selectable marker (i.e. the *Not* I pHMG-hAT fragment together with a ten-fold molar excess of pLJ) had allowed isolation of clones with G–418 (Mrs A.J. Watson, personal communication). A co-transfection experiment of RJKO cells was therefore carried out with the 6.6 kb *Not* I pHIF10 fragment and an eight-fold molar excess of pLJ. Selection with G–418 followed the revised protocol as in section 5.4.5 (2.4.2) and six clones were generated, named 10H1 to 10H6. Only one of these, 10H4, proved positive on PCR analysis for the *fpg* sequence (Fig. 5.17). It was therefore grown up in parallel with the 10L*n* clones (5.4.5) for additional analyses.

Fig. 5.17

PCR analysis of RJKO clones 10H1 to 10H6



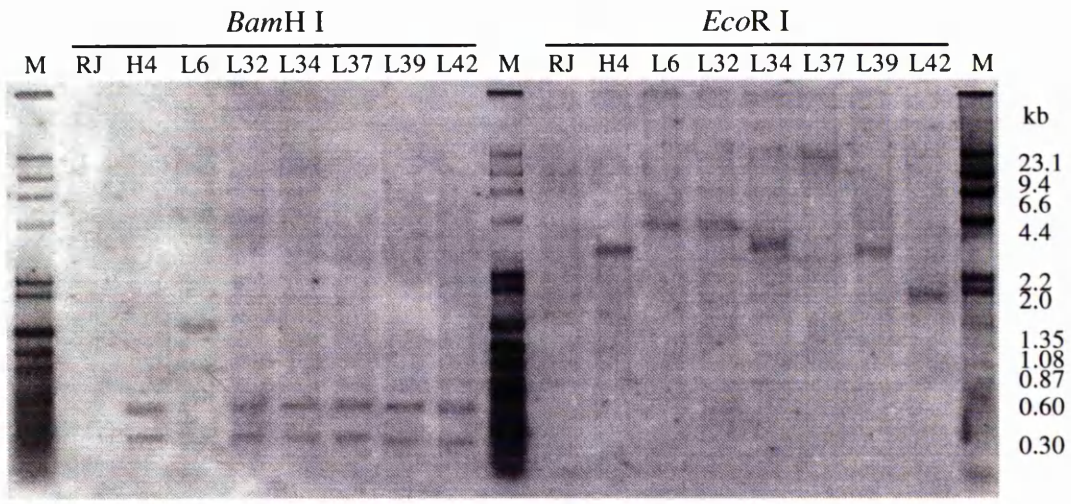
Crude extracts from the clones were analysed by PCR using oligonucleotides 184 and 185 (2.2.4, 2.2.5). "RJ" = RJKO cells, "o" = no DNA, "p" = pHIF10 + RJKO controls.

5.4.7 Analysis of RJKO clones containing *fpg*10

5.4.7.1 Southern analysis

DNA from each of the *fpg*10-containing clones was digested with *Bam*H I and *Eco*R I and used in Southern analysis, as with the earlier clones (5.4.2, 5.4.3 and 2.2.20). However, the DNA from clone 10L6 was a repeat purification from a later passage than that used in section 5.4.2, as insufficient DNA was left from the original extraction. The resultant autoradiographs (Fig. 5.18) showed that clones 10H4, 10L32, 10L34, 10L37, 10L39 and 10L42 produced 0.55 and 0.30 kb *Bam*H I bands as expected. Somewhat surprising however was the result with 10L6, which had previously given a very strong signal at 0.5 and 0.35 kb (5.4.2, Fig. 5.14), but which now gave a single band of weak intensity at 1.5 kb. The *Eco*R I digests gave hybridisation bands in all *fpg*10-containing clones. Amongst the pLIF10-derived clones a variety of patterns were seen with *Eco*R I, no two clones having the same pattern, although 10L6 and 10L32 appeared to share a *ca.* 4.4 kb band.

Fig. 5.18
Southern analysis of RJKO clones 10H4, 10L6/32/34/37/39 and 42



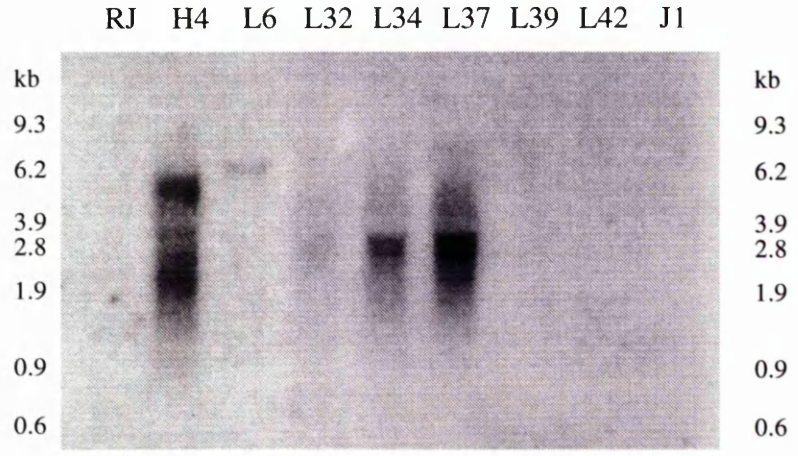
DNA from each cell line (10 μg) was digested with 50 U enzyme and subjected to Southern analysis (2.2.20; vacuum blotted). "RJ" = RJKO, "H4" = 10H4, "L*n*" = 10L*n*, "M" = [^{35}S]-end-labelled DNA markers.

5.4.7.2 **Northern analysis**

Total RNA was extracted from the various clones (2.2.21) at the same time as extracts were made for the first comparative FaPy-DNA glycosylase assay (using the polyGC-based assay; 2.3.4). These RNAs were then electrophoresed under denaturing conditions in agarose gel, which was then vacuum blotted and probed with [^{32}P]-labelled *fpg*4 (2.2.22; Fig. 5.19). All the clones, except 10L39 and 10L42, gave positive hybridisation signals. The pLIF10-based clones 10L32, 34 and 37 all gave a single diffuse band at *ca.* 2.5 kb, rather weak with 10L32, intermediate with 10L34, and quite strong with 10L37, whereas 10L6 gave a single weak, but discrete, band at *ca.* 5.5 kb. In contrast, the pHIF10-derived clone 10H4 gave two strong bands, one of 5 kb and the other slightly smaller than 1.9 kb.

Fig. 5.19

Northern analysis of RJKO clones 10H4, 10L6/32/34/37/39 and 42



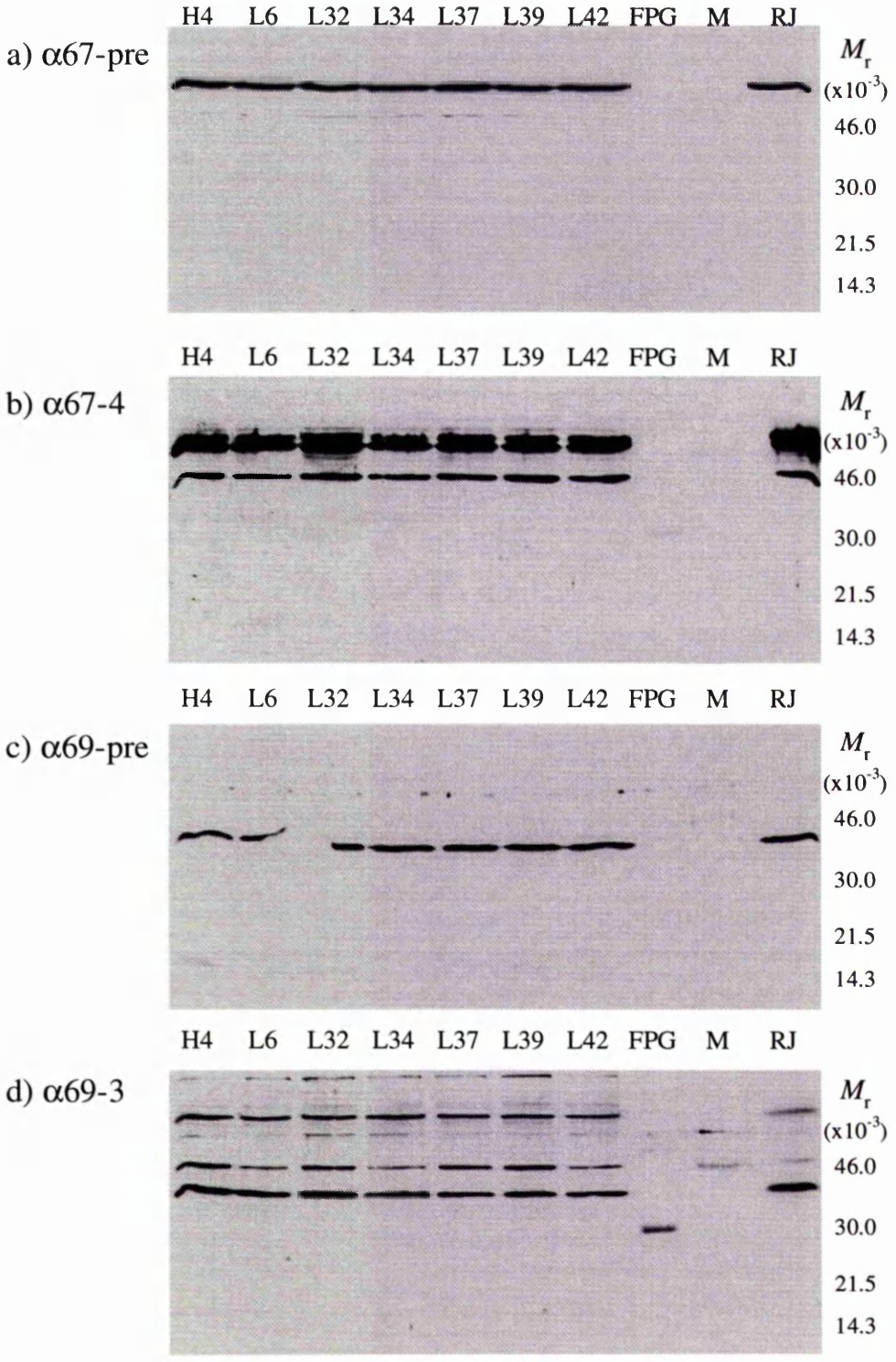
Total RNA extracted from the various cell lines (10 µg; 2.2.21) was separated by agarose gel electrophoresis under denaturing conditions (1× MOPS buffer, 2.2 *M* formaldehyde) and subjected to northern analysis (2.2.22; vacuum blotted).

5.4.7.3 **Western analysis**

Extracts made for FaPy-DNA glycosylase assay from the *fpg*10-containing clones, at the same time as the total RNA extractions (5.4.7.2), were separated by SDS-PAGE and electroblotted onto nitrocellulose membranes. These blots were then probed with the anti-FPG antisera α67–4 and α69–3 (2.3.14; 4.3; Fig. 5.20). The pre-immune sera detected different sized bands which were also present in the respective immune sera probed blots. The probing with α67–pre showed two bands, a major one with a M_r of about 60 000 (by extrapolation of a plot of the migration distance of the markers *v.* their M_r) and a minor one showing a weak signal at a M_r of 46 000. In contrast, α69–pre hybridised with a single band with a M_r of 40 000. All these bands were evident in the immune sera-probed blots, with similar intensities except for the M_r 46 000 band in α67–4 which had increased in intensity to a level similar to the other major bands. In addition, with α67–4 an extra band of similar strength to that M_r 60 000 was apparent with an M_r of 56 000 (by extrapolation). In α69–3 the band pattern was more complex with, in addition to the M_r 40 000 band, two other major bands with M_r of 46 000 and 65 000 (by extrapolation), plus three or four other minor bands. Antiserum α69–3 also appeared to react with the M_r 46 000 marker protein (dye-labelled ovalbumin). Thus, both of the antisera detected specific bands in all extracts, whether from the parent RJKO line or the *fpg*10-transfected clones, but none corresponded with that from purified FPG protein.

Fig. 5.20

Western analysis of RJKO clones 10H4, 10L6/32/34/37/39 and 42

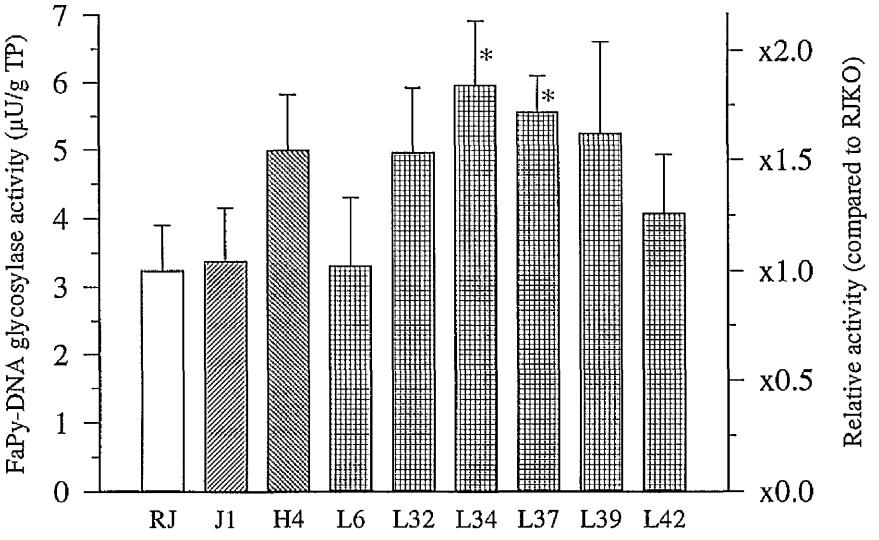


Extracts from the various cell lines (30 µg TP), as used for glycosylase assay, were separated by SDS-PAGE, and then subjected to western analysis, with detection of secondary HRP-labelled antibody by enhanced chemiluminescence (2.3.14). "M" = coloured dye-labelled marker protein mix, "FPG" = purified FPG protein (5 ng),"RJ", "H4", "L6" etc. = cell lines, as in Fig. 5.18.

5.4.7.4 **FaPy-DNA glycosylase activity**

All of the *fpg*10-containing RJKO cell clones (10L6/32/34/37/39/42 and 10H4) together with the parent RJKO cells and one of the pLJ-containing clones (J1) were grown up in bulk and assayed for their FaPy-DNA glycosylase activity (2.4.3). This time, however, the polyGC-based assay of Boiteux *et al.* (1987) was used (2.3.5), as its precision at measuring the lower levels of FaPy-DNA glycosylase found in eukaryotic cells had proved superior to that of the DNA-based assay (2.3.3). Five separate experiments were carried out, using the same incubator, the same batch of medium and same batch of flasks throughout, in addition to the length of time the cells were allowed to grow after sub-culturing (2.3.3; 2.3.4; 2.4.1). The average FaPy-DNA glycosylase levels found in the various clones are shown in Fig. 5.21. It can be seen that the standard errors of the mean specific activities were large, corresponding to CVs of 21 to 68%. This was a reflection of the variation between experiments within individual clones (despite the care taken to keep variation between separate experiments to a minimum), rather than due to day-to-day variation in assay performance, as the inter-assay CV across the five experiments for the same sample (some female BDF_1 mouse liver extract, kept frozen in aliquots at -20°C, and included in each assay as a control) was only 6.0% (and this at a mean level of 2.7 μU/g TP, i.e. somewhat lower than that in RJKO cells). Only two clones had significantly raised levels of activity when compared to RJKO cells, 10L34 (1.83-fold higher; *t*-test: $p = 0.03$) and 10L37 (1.71-fold higher; $p = 0.02$), however clone 10H4's level at 1.54-fold higher did approach statistical significance ($p = 0.11$). All the other clones, and in particular J1 and 10L6, were not significantly different from RJKO.

Fig. 5.21
FaPy-DNA glycosylase activity in RJKO-*fpg*10 clones



Results given are means of 5 independent experiments, with error bars indicating SEM, * indicates $p < 0.05$, when levels compared with RJKO by *t*-test. "RJ" = RJKO, "J1" = J1, "H4" = 10H4, "L*n*" = 10L*n*.

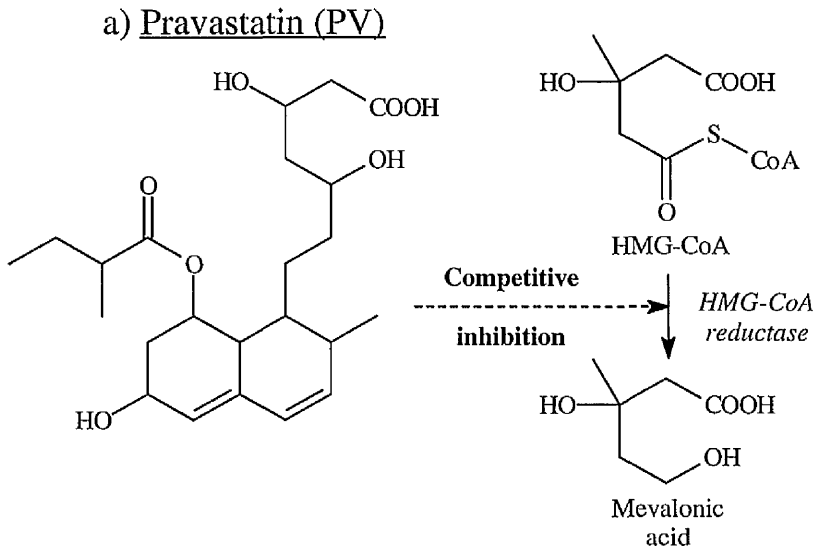
5.4.8 **Manipulation of FaPy-DNA glycosylase activity in clone 10H4**

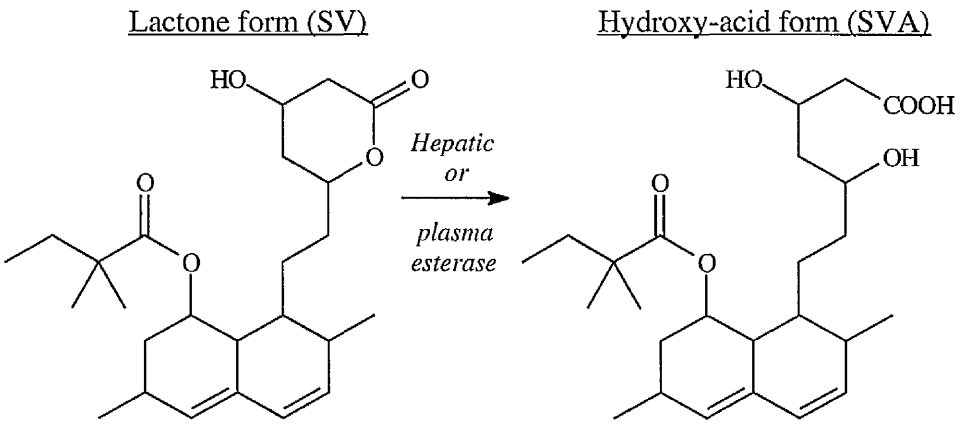
5.4.8.1 **HMGCR inhibitors as potential up-regulators of pHMG constructs**

As described in section 5.3.2, the vector pHMG used to construct pHIF10 is based on the murine hydroxymethylglutaryl-Coenzyme A reductase (HMGCR) gene; HMGCR catalyses the rate-limiting step in cholesterol biosynthesis. The introduction of HMGCR inhibitors, such as Pravastatin and Simvastatin, into clinical practice (in the late 1980's) for the control of hypercholesterolaemia, prompted the idea that they might stimulate compensatory up-regulation of the HMGCR gene. Both agents are derived from a class of fungal metabolites which are reversible competitive inhibitors of HMGCR: Pravastatin (PV) is direct-acting, whereas Simvastin is a pro-drug and does not become active as an HMGCR inhibitor until it has been enzymatically hydrolysed *in vivo* from the cyclic lactone form (SV) to an open-ring hydroxy-acid form (SVA; Todd and Goa, 1990;

Fig. 5.22). In the majority of mammalian species investigated this hydrolysis occurs by means of hepatic esterase activity, however in some species (e.g. rats) a plasma esterase rapidly converts SV to SVA (Vickers *et al.*, 1990). As the cyclic lactone form is more lipid soluble this means that, *in vivo*, SV is taken up by cells and accumulates there in its SVA form (Todd and Goa, 1990). Pravastatin, in contrast, is lipophobic and therefore partitions relatively poorly into cells, except in the liver where it appears to be selectively taken up by hepatocytes: thus *in vivo* PV is a more organ-selective agent than SV (Koga *et al.*, 1990; Koga *et al.*, 1992). A study of the literature also showed that rats and mice treated with these inhibitors paradoxically increase their hepatic HMGCR levels (Kita *et al.*, 1980; Liscum *et al.*, 1983), strongly suggesting that up-regulation of the gene might take place, at least in these species. Because their HMGCR levels increase in this manner, rodents are unsuitable for use in pharmacological studies on the inhibition of cholesterol biosynthesis by this class of agents (Vickers *et al.*, 1990). Indeed, for these purposes it is necessary to use rabbits or dogs which, like humans, do not up-regulate their HMGCR expression in response to these inhibitors. Additionally, there was some evidence that the mechanism by which HMGCR up-regulation occurred in mice was by a 6– to 10–fold increase in transcription (Kita *et al.*, 1980). Thus, as clone 10H4 had been generated by transfection of a Chinese hamster-derived cell line (RJKO) with a mouse HMGCR promoter-*fpg* construct, it was considered worthwhile to investigate whether these HMGCR inhibitors could be used to up-regulate expression of FPG protein in these cells.

Fig. 5.22
Hydroxymethylglutaryl-Coenzyme A inhibitors



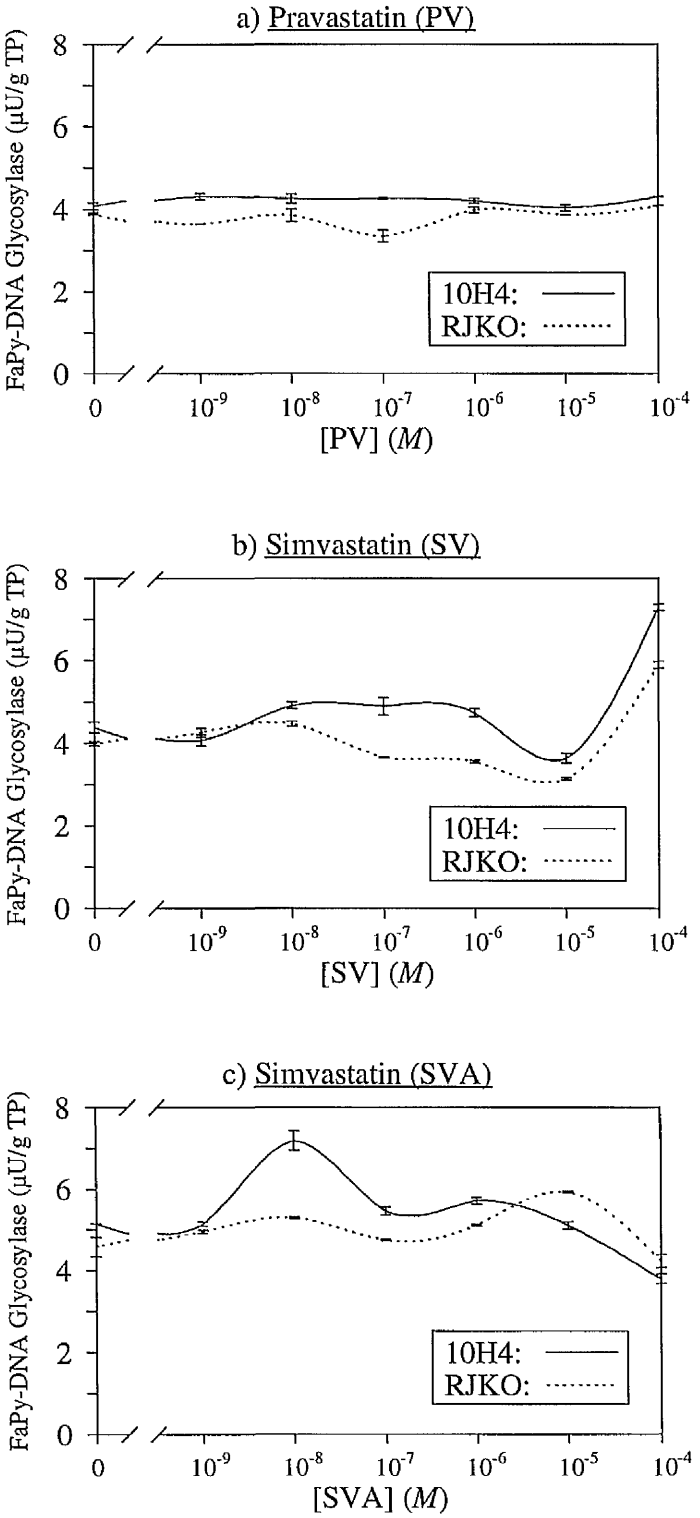


5.4.8.2 **The effect of Pravastatin and Simvastatin on clone 10H4**

The clone 10H4 was treated with Pravastatin and Simvastatin, in parallel with RJKO cells under identical conditions (2.4.3). Simvastatin was used both in pro-drug form (SV) and hydroxy-acid form (SVA), the latter being achieved by incubation of SV with normal rat serum prior to treatment of cells (2.4.3; Vickers *et al.*, 1990). Concentrations from 10^{-4} *M* to 10^{-9} *M* were used to establish dose-response curves over as wide a range as

possible. The results of this first experiment can be seen in Fig. 5.23.

Fig. 5.23
The effect of HMGCR inhibitors on
the FaPy-DNA glycosylase activity in 10H4 cells



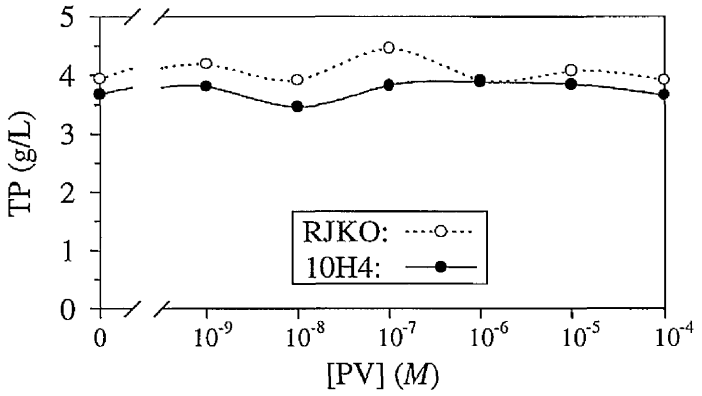
All points are means of triplicate FaPy-DNA glycosylase assay determinations carried out on single samples, i.e. each point represents a single flask of cells, error bars indicate SEM.

All three agents appeared to increase the levels of FaPy-DNA glycosylase activity in clone 10H4 when compared to the parent RJKO cells, but to different extents and at different concentrations. Pravastatin gave consistently higher levels in clone 10H4 at all concentrations, with somewhat greater effect from 10^{-9} to 10^{-7} *M*, but the differential was never greater than 1.2-fold. Simvastatin (as SV or SVA) had a greater effect on the FaPy-DNA glycosylase activity, but in the 10^{-8} to 10^{-6} *M* range, with SVA giving the widest differential of 1.4-fold higher at 10^{-8} *M*. However it also had a marked inhibitory effect on cell growth at the highest two concentrations. This was evident by the subjective observation of diminished cell-pellet size at the highest levels of SV, which was consistent with the objective estimate of cell mass given by the total protein concentration of the extracts made for enzyme assay, which were made in a constant volume of buffer (Fig. 5.24). The high levels seen with SV at 10^{-4} *M*, in both 10H4 and RJKO, contrasted with the low levels seen with SVA at the same concentration.

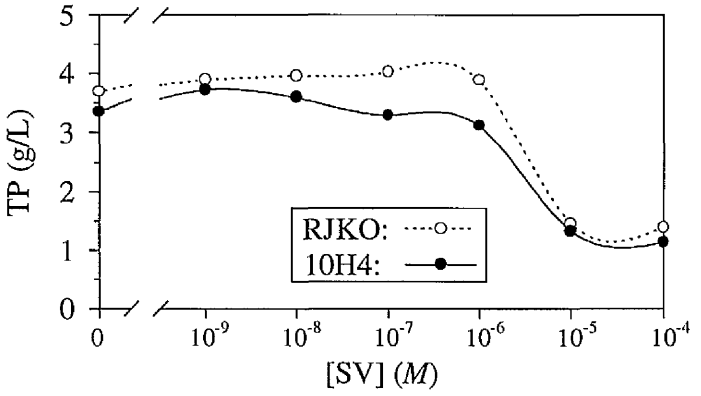
Fig. 5.24

The effect of HMGCR inhibitors on the total protein level in 10H4 cells

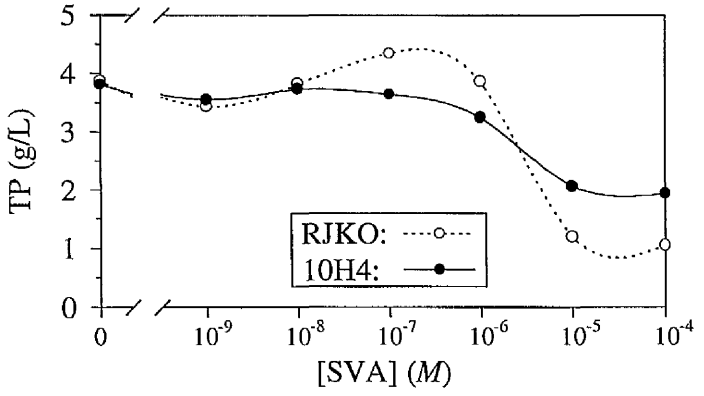
a) Pravastatin (PV)



b) Simvastatin (SV)



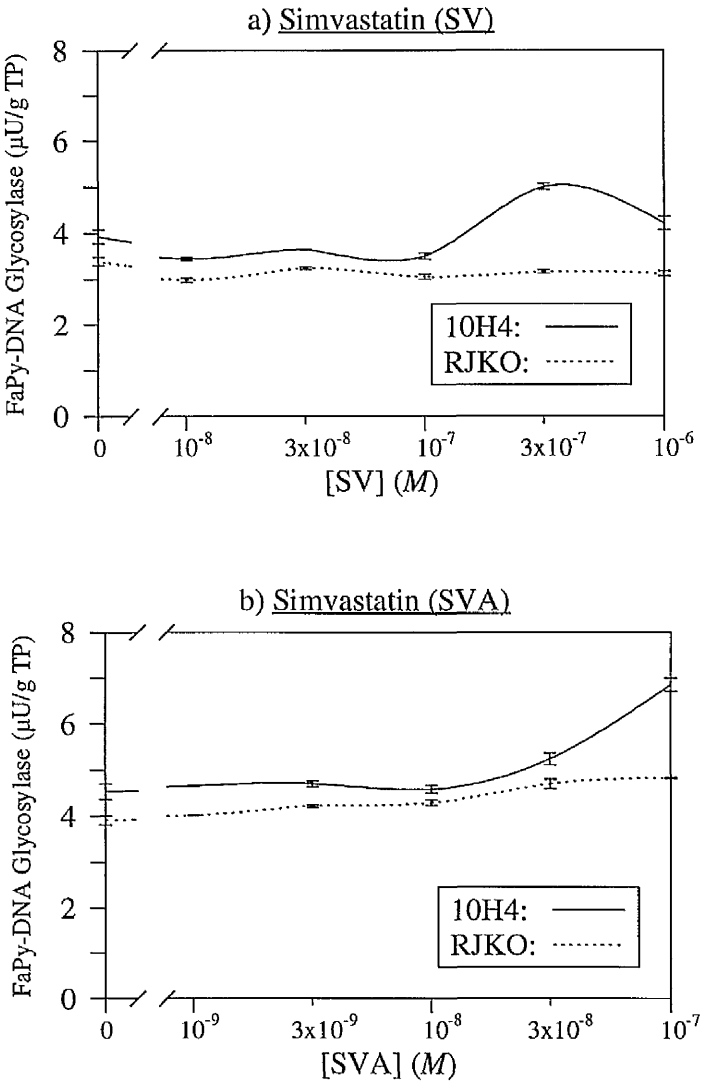
c) Simvastatin (SVA)



All points are single determinations of total protein (2.3.5) carried out on single samples, i.e. each point represents a single flask of cells. All extracts were made in the same volume of buffer and therefore the TP concentrations are a measure of the TP per flask.

On the basis of this, Simvastatin seemed to be more promising as a means of increasing FaPy-DNA glycosylase activity in 10H4 and a second experiment was therefore performed, this time with SV at 10^{-8} to 10^{-6} *M* and SVA at 10^{-9} to 10^{-7} *M* as well as intermediate concentrations. Figure 5.25 shows the somewhat different results that were obtained compared to the first experiment (Fig. 5.23). The differential in activity with SV between 10^{-8} to 10^{-7} *M* had been lost and the previous peak of activity with SVA at 10^{-8} *M* appeared to have shifted to 10^{-7} *M* or higher. The greatest effect, however, with 1.6-fold higher FaPy-DNA glycosylase activity at 3×10^{-7} *M* was no greater than the difference seen previously, without treatment with PV or SV (5.4.7.1).

Fig. 5.25
The effect of Simvastatin on
the FaPy-DNA glycosylase activity of 10H4 cells



All points are means of triplicate FaPy-DNA glycosylase assay determinations carried out on single samples, i.e. each point represents a single flask of cells, error bars indicate SEM.

5.5 DISCUSSION

5.5.1 Production of expression vector-*fpg*4/8 constructs

The initial choice of the vector pZIP was largely based on its previous successful use in the Department to express *ada* and *ogt* in RJKO cells. However, the cloning experiments with pZIP-*fpg*4 were wholly unsatisfactory, with deletions or rearrangements of the vector, although the host (DH5α) was not recombination proficient. This contrasted with the ease with which it proved possible to generate pLJ–*fpg*4/8- and, later, *fpg*10-containing plasmids, in the same strain of *E. coli*. One explanation for this is that the experiments with pLJ used *fpg*4/8/10 insert which had been purified by adsorption on glassmilk, rather than phenol extraction. However, while this might be expected to improve the efficiency of ligation it is difficult to envisage how it might have caused aberrant constructs, especially in view of the phenol purification method having been used successfully in earlier experiments (e.g. 3.3.4). In addition, pZIP was maintained in DH5α without any problems of deletion or rearrangement of the vector.

It would therefore appear that the problems with pZIP were a function of the addition of the *fpg* insert to the combination of vector and host. Switching to the vector pLJ allowed the satisfactory generation of a retroviral-based expression vector-*fpg* construct. One possible reason why the pZIP/*fpg* combination was problematical is that an excess of the FPG protein was toxic, or otherwise deleterious to the host. However, this seems unlikely because the experiments using pKK223-3-*fpg* constructs achieved over-expression of FPG protein by over 1000-fold, without any evidence of instability in the plasmids (albeit that the *E. coli* host was JM105 and the vector pKK223-3). More likely, perhaps, is that the combination of *fpg* ligated into pZIP strongly stimulated either deletion or rearrangement by a mechanism independent of *rec*A-mediated recombination.

The generation of a pHMG-*fpg* construct was, as with pLJ, achieved without problems. From the resultant plasmid (pHIF10) the 6.6 kb *Not* I fragment, required for the

generation of transgenic mice, was purified and its concentration established for such experiments (5.3.3).

5.5.2 Transfection of RJKO cells with pLIF4 and analysis of clones

While the transfection experiments with pLIF4 and pLIF8 were being carried out it became evident that the *fpg*8 insert would not produce active FPG protein, and therefore only pLIF4-transfected clones were picked. Analysis of these with the DNA-based FaPy-DNA glycosylase assay was unable to demonstrate a significant increase in glycosylase level. However, as is evident from the magnitude of the standard errors compared to the means themselves, the DNA-based FaPy-DNA glycosylase assay was at its limit-of-detection (defined as being within 2.8× the SD at zero, i.e. when assaying blanks; Varley *et al.*, 1980) and hence inherently unable to demonstrate differences at this level.

Testing by PCR showed that 10 out of 12 pLIF4-transfected RJKO clones contained intact *fpg*4 sequence, which was later confirmed by Southern analysis with 100% concordance. This suggested that initial testing by PCR might be an efficient way of screening cell lines, or even transgenic animals, thus allowing Southern analysis to be reserved for PCR-positive clones.

The concurrent experiments with pKIF4 showed that FPG4 had only a very low level of FaPy-DNA glycosylase activity and this was taken as the reason why an increase in FaPy-DNA glycosylase activity could not be detected in the 4L*n* clones. As a result of both *fpg*4 and *fpg*8 being adversely affected with mutations subsequent experiments were carried out with the *fpg*10 insert, known to contain a silent mutation but proven to produce active FPG protein (3.6.1 and 3.7.3).

5.5.3 Transfection of RJKO cells with pLIF10/pHIF10 and their analysis

The decision to carry out an initial screening of clones by PCR provided useful data when pLIF10-transfected RJKO clones were analysed at the earliest possible opportunity (5.4.1). Six out of the first eighteen 10L*n* RJKO clones picked were PCR-positive, compared to 10/12 of the 4L*n* clones. This showed that a lower proportion of G–418-resistant pLIF10-derived clones contained the *fpg* sequence compared to the experiment with pLIF4. Subsequent PCR analysis could only demonstrate that a single clone was positive (10L6), suggesting that perhaps the PCR assay was variable in performance, but Southern analysis of the clones was consistent with these second PCR findings. Thus, the proportion of positive clones diminished from 6/18 to 1/18 after a further period of growth in G–418-containing medium. Taken together, this implied that there was selection occurring against the maintenance of pLIF10-containing RJKO clones. The finding by Southern analysis that clone 10L6 had an *fpg*10 copy-number of *ca.* 100 possibly argued against this. However, it is perhaps not inconsistent: lines with a low copy number, and hence a low level expression of neomycin phosphotransferase, would be eliminated by high levels of G–418; lines with a high copy number, giving high expression of FPG, might be disadvantaged in some way, possibly due to a cytotoxic or cytostatic effect of FPG in a eukaryotic environment, and thus liable to evolve *via* a reduction in copy number into lines with reduced expression. The measurement (by the DNA-based assay; 2.3.2) of the FaPy-DNA glycosylase activity in 10L6 failed to show an increase and therefore a further transfection experiment was carried out.

This second experiment proved even less successful than the first, and prompted a change to less severe selection conditions in a third experiment. This resulted in five 10L*n* clones, i.e. 10L32/34/37/39 and 42, testing positive by PCR. These were further subjected to Southern analysis (Fig. 5.18), which showed that while all clones gave hybridisation bands with the *fpg*4 probe, only one pLIF10-derived clone (10L32) had a

restriction pattern consistent with an intact vector/insert construct (5.4.7.1). The clone 10L6, which had earlier been found to have a 0.55 + 0.30 kb band pattern with *Bam*H I had also undergone a change in its pattern to a single band at 1.5 kb, with a considerable drop in the intensity of the hybridisation signal, implying a reduction in *fpg* copy-number as well as rearrangement(s) in the integrated construct.

The results seen with northern analysis (5.4.7.2) showed that out of the three clones from the third experiment which showed hybridisation with the *fpg*4 probe (10L32/34 and 37), 10L32 gave the weakest signal, whereas the others (which had both shown aberrant patterns with *Eco*R I) gave stronger signals. Clone 10L6 gave a weak signal, similar in intensity to 10L32, but with a different pattern from that of the other three 10L*n* clones. Western analysis (5.4.7.3) was unable to detect expressed FPG protein in any of the clones, using antiserum from either rabbit. Assay of the FaPy-DNA glycosylase activity in the clones with the polyGC-based method (2.3.4, 5.4.7.4) was able to demonstrate a significantly higher level in clones 10L34 and 10L37, compared to the parent RJKO cells. In the western analysis experiment 5 ng of recombinant FPG protein had been detected by the better of the two antisera (α69-3). A detection limit of 1 ng of FPG protein would be equivalent to detecting 350 pU of FPG (from the specific activity of Fraction IV in Table 4.2) per 30 μg TP (the amount of TP loaded per lane), i.e. a rise in FaPy-DNA glycosylase level of 12 μU/g TP. The greatest such rise detected was with clone 10L34, which had a glycosylase level 2.8 μU/g TP higher than that of the parent RJKO cells, less than one quarter of the likely limit of detection of western analysis. This probably accounts for its failure to detect FPG protein in the clones.

On Southern analysis the single clone (10H4) from the dual pLJ/pHIF10 6.6 kb *Not* I fragment transfection experiment had a satisfactory restriction pattern with *Bam*H I and, as far as was ascertained, with *Eco*R I. It gave two major bands on northern analysis, consistent with correct splicing of the transcript, but, like the 10L*n*

clones, no FPG protein could be demonstrated by western analysis. Measurement of the FaPy-DNA glycosylase activity in clone 10H4 suggested that it had a higher level than RJKO cells, but this did not quite reach statistical significance. These findings for the 10L*n* clones, as well as clone 10H4, are summarised in Table 5.3.

Table 5.3
A summary of the analysis of clones 10L6/32/34/37/39/42 and 10H4

Clone	Southern[a]		Northern[b]	Activity[c]
	*Bam*H I	*Eco*R I		
10H4	+	+	+++	× 1.54 (0.11)
10L6	+/(+)[d]	+	+[e]	× 1.02 (0.96)
10L32	+	+	+	× 1.53 (0.18)
10L34	+	(+)	++	× 1.84 (0.047)
10L37	+	(+)	+++	× 1.72 (0.026)
10L39	+	(+)	-	× 1.62 (0.22)
10L42	+	(+)	-	× 1.26 (0.46)

[a] results expressed qualitatively: "+" = hybridisation signal as predicted; "(+)" = hybridisation signal but not as predicted; "-" = no hybridisation signal, [b] results expressed semi-quantitatively, [c] FaPy-DNA glycosylase activity levels expressed relative to that of RJKO cells (with, in parentheses, the *p* value by *t*-test with respect to RJKO cells), [d] hybridisation signal observed to change both in pattern and reduce in intensity, [e] pattern different from the other three pLIF10-derived clones 10L32, 10L34 and 10L37.

Taken together the data in Table 5.3 show a degree of correlation between the results for northern analysis and glycosylase activity, in that the three clones with some evidence of raised enzyme activity (10L34, 10L37 and 10H4) also gave the strongest hybridisation signals. Of the four clones which showed no evidence of significantly raised glycosylase activity, three had abnormalities on Southern analysis, with one (10L6) also having a different (perhaps aberrant) pattern on northern analysis (compared with the other three clones derived from the same vector which all gave the same signal on northern analysis, which was different from that of 10L6). Clone 10L32 gave hybridisation bands on Southern analysis of the predicted size and also a signal on northern analysis, but it did

not have a significantly raised glycosylase level. Therefore all of the 10L*n* clones appeared to have undergone some change in the predicted structure of the *fpg*10-construct and/or failed to express a significantly raised level of FPG protein. If the clones did express FPG protein, it was either in an inactive form, which did not hybridise under the conditions used in western analysis and was also inactive in the FaPy-DNA glycosylase assay (unlikely), or was being turned over so rapidly it was unable to reach a significant level (possible).

Two features are apparent from comparing this data with that from the pLIF4 and pLIF8 transfection experiments: 1) a lower proportion of *fpg*-containing G–418-resistant clones were obtained with *fpg*10-containing constructs, and 2) those *fpg*10-containing clones that were obtained either did not express FPG protein, or showed only a minimal rise in FaPy-DNA glycosylase level. This suggested that, at least in RJKO cells, there is a strong selection pressure against expression of *E. coli* FPG protein. The results with clone 10L6 in which, after a period of growth in G–418-containing medium, a change in restriction pattern and, probably, copy number were observed with Southern analysis is perhaps additional evidence of the selection process in progress.

There are a variety of reasons why such a phenomenon might be observed. The first is that the process is one of selection *for* improved expression of the *neo* gene in an environment containing G–418, and that this happens when *fpg* is lost from the constructs. However, there was no evidence of outright loss of *fpg* sequence, and pLJ was designed with separate promoters to minimise loss of inserted genes (5.1.1). Also, though it might occur with the pLIF10 construct, it is difficult to see how the process might apply to clone 10H4, in which pLJ and the pHIF10 construct were separate. The alternative explanation that it was selection *against* FPG expression could take two forms, related or un-related to the DNA repair functions of FPG protein. If un-related, it suggests that expression of this prokaryotic protein deleteriously interferes with eukaryotic

cell metabolism. If related, it could be a function of any one, or any combination, of the activities of the FPG protein as either a FaPy/8oxoPurine-DNA glycosylase, or AP-lyase. In Chapter 1 it was discussed how any one DNA repair function cannot be considered in isolation, i.e. that a cell with DNA glycosylase activities generating AP-sites also requires AP-endonuclease, phosphatase, DNA-polymerase and DNA-ligase activities to facilitate complete DNA repair. It might well be harmful for a cell to have an increased DNA glycosylase activity if this produces more AP-sites or DNA single-strand nicks than can be repaired by the other components; it would predict a chronic increase in unscheduled DNA synthesis. The continuous generation of oxidative DNA damage within cellular DNA, particularly in cultured cells kept in relatively hyperoxic conditions, could be sufficient in combination with FPG protein to have this effect. In turn this could cause delaying or blocking of the entry of cells into S-phase, thereby lowering the overall proliferation rate.

There are a number of alternative strategies that could be followed to investigate the effect of FPG within a eukaryotic cell, that might also help to address the hypothesis that FPG expression in such cells is selected against. Given the purification of large quantities of active FPG protein (Chapter 4) it could be microinjected into cells *in vitro* and the, possibly cytotoxic, effects monitored. However, this would require expensive equipment, and the acquisition of the necessary expertise. Electroporation could also be tried as a way of introducing FPG protein into cells. The use of an expression vector with aninducible promoter would address the question of whether the observed instability of the constructs within RJKO cells was a function of the DNA sequence of *fpg*-constructs, or related to the expression of transcript and FPG protein. Few, if any, promoters are entirely switchable on/off, however, which might be problematical. Rather than try and achieve stable integration of the *fpg*-construct into the host cells' genome, the transient expression of FPG protein could be attempted by infection of suitable cells,

e.g. cos7, with SV40 virus constructs of *fpg*. This has been successfully carried out with ATase genes (Dr G.P. Margison, personal communication).

5.5.4 Treatment of clone 10H4 with HMGCR inhibitors

Northern analysis of clone 10H4 had shown transcription occurring as predicted, with apparent correct splicing of the mRNA (Figs. 5.9 and 5.19). Assay of the FaPy-DNA glycosylase activity suggested that there might be expression of FPG protein, but this did not quite reach statistical significance at the $p = 0.05$ level. HMGCR inhibitors, known to up-regulate transcription of the endogenous rodent HMGCR gene, were therefore used to investigate whether up-regulation of transcription from the pHIF10 construct could be achieved to give a significant rise in FaPy-DNA glycosylase level. Two compounds were tried, Pravastatin (PV) and Simvastatin, the latter in two forms: the pro-drug lipophilic lactone form (SV), and the active lipophobic hydroxy-acid form (SVA).

In the first experiment (5.4.8.2 and Fig. 5.23) the level of FaPy-DNA glycosylase observed in clone 10H4, untreated, (4.38 (0.13) μU/g TP; mean (SEM), $n = 3$) was significantly higher ($p = 0.049$, by *t*-test) than that in RJKO cells grown under identical conditions (3.99 (0.05) μU/g TP; mean (SEM), $n = 3$), however this was by a much reduced margin (1.1-fold compared to the 1.54-fold value already determined (5.4.7.4)). The was not just due to a lower absolute level in 10H4 cells (4.38 *v.* 4.98 μU/g TP), but also a greater value for the activity in RJKO cells (3.99 *v.* 3.23 μU/g TP). Consistently higher activity was seen in 10H4 cells treated with PV (Fig. 5.23a), but the dose-response relationship was flat, without any suggestion that an increase in FaPy-DNA glycosylase was taking place. Any increase in the differential between 10H4 and RJKO cells appeared to be a function of inconsistent drops in glycosylase level in the RJKO cells. The total protein levels (Fig. 5.24a) showed a consistently higher value in RJKO compared with 10H4 cells, sufficient to account for the difference in specific FaPy-DNA

glycosylase activities. It was therefore concluded that PV was of no use in increasing the FaPy-DNA glycosylase activity of 10H4 cells.

The results with Simvastatin (Fig. 5.23b and c) differed from those with PV in that there was a suggestion of some effect on increasing FaPy-DNA glycosylase levels, with SV between 10^{-8} and 10^{-6} *M*, and SVA at around 10^{-8} *M*. With SV this appeared to be a result of a difference in total protein levels, but with SVA it appeared to be as a result of a definite increase in FaPy-DNA glycosylase activity, presumably due to FPG expression. At the highest concentrations used both forms of Simvastatin had an adverse effect on cell growth, as estimated by the total protein levels. With the highest concentration (10^{-4} *M*) SV caused an increase in specific FaPy-DNA glycosylase in both RJKO and 10H4 cells, and this could be accounted for by the decrease in TP. In contrast, treatment with SVA caused a drop in the specific glycosylase activity at the highest concentration, in spite of a reduction in the TP level. These effects are evidence of a strong influence on cellular metabolism; a reduction in the supply of cholesterol, an essential component of all cell membranes, would be expected to inhibit cell growth; but, because the effects were seen to an equal extent in both RJKO and 10H4 cells they cannot be the specific result of the pHIF10 construct. Thus the possibility that the HMGCR inhibitors at the highest concentrations were stimulating FPG expression, which had a cytostatic effect, can be discounted. In the second experiment (Fig. 5.25), investigating more closely over a restricted range of concentrations whether Simvastatin in either form could effectively increase FaPy-DNA glycosylase activity, the levels in clone 10H4 cells were consistently higher than in RJKO cells, whether treated or un-treated. The greatest differential in activity (1.6-fold) seen with SVA at 10^{-8} *M* was of the same degree as seen previously without treatment with Simvastatin. It was therefore concluded that although there might be some stimulatory effect with Simvastatin in either form it was

inconsistent, and not of a sufficient magnitude to be useful.

A similar agent (Lovastatin) has been shown to cause 2- to 70-fold increases in HMGCR activity *ex vivo* in the rat and 6- to 10-fold increases in HMGCR transcription *in vivo* in the mouse (Liscum *et al.*, 1983; Kita *et al.*, 1980). The greatest effect seen in these experiments was 1.6-fold, much smaller than expected. This class of HMGCR inhibitors are highly protein bound (>95%), in human plasma at least (Todd and Goa, 1990), and this might have reduced their availability *in vitro*, but SV and SVA both had an adverse effect at the highest concentrations suggesting that this was not a limiting factor. It was also assumed that the rat serum treatment was adequate to convert SV to SVA. As this was not independently tested it is not possible to be sure. However Vickers *et al.* (1990) had shown the conversion of SV to SVA to be complete under the conditions used, and different dose-response curves were seen with SV and SVA implying that conversion had occurred. The slightly greater FaPy-DNA glycosylase levels seen in RJKO and 10H4 cells in the SVA experiments (compared to the same cells treated with SV) is likely to have been due to the effect of the rat serum (20% v/v) in the medium (2.4.3) in addition to the FCS (10% v/v), causing a change in the TP content of the cells.

The smaller-than-anticipated effect on FaPy-DNA glycosylase levels was not necessarily due to a failure to increase transcription from the pHIF10 construct. An increase in transcription may have occurred but been cancelled out by a concomitant increase in message turnover. Or it could be that FPG protein was being degraded as fast as it was being produced in spite of an increase in transcription and translation, and hence no net increase was observed. Another possibility is that the Chinese hamster is unlike other rodents and does not up-regulate transcription of HMGCR when treated with HMGCR inhibitors (alternatively RJKO cells may have lost this ability). The promoter and/or other controlling sequences of the Chinese hamster and

mouse HMGCR genes may have critical differences, such that regulation of the mouse gene does not occur in Chinese hamster (RJKO) cells. Another explanation is that the region of the mouse HMGCR gene present in the vector pHMG does not include all of the necessary controlling sequences. It must contain at least those regions necessary for the construct to behave as a "housekeeping" gene and to give normal constitutive levels of expression (Gautier *et al.*, 1989), but perhaps pHMG is missing those elements required for up-regulation.

While some of these possibilities might have been amenable to investigation by studying mRNA levels, e.g. by northern analysis, or to check for up-regulation of the endogenous HMGCR activity, e.g. by assay of incorporation of [^{14}C]-acetate into cholesterol (Henwood and Heel, 1988), these would have been expensive experiments in terms of time and reagents, in the latter case particularly the supply of normal rat serum used for the conversion of SV to SVA. Treatment with any of the agents (PV, SV or SVA) did not give a differential increase in FaPy-DNA glycosylase levels of at least 2-fold between RJKO and 10H4 cells. In view of this and the inconsistency between experiments, the cause or causes of which could have taken considerable time to identify, it was decided not to pursue *in vitro* treatment of cells with these inhibitors any further.

CHAPTER 6

STUDIES ON MAMMALIAN FaPy-DNA GLYCOSYLASE ACTIVITY

6.1 ATTEMPTS TO INCREASE THE LEVEL OF FaPy-DNA GLYCOSYLASE IN RODENTS AND IN MAMMALIAN CELLS *IN VITRO*

The levels of many DNA-repair enzymes can be up-regulated in both pro- and eukaryotes. One of the best characterized of these phenomena is the adaptive response in bacteria, where methylating agents cause a rapid increase in ATase and MAG (Cairns, 1980; Karran *et al.*, 1982; McCarthy *et al.*, 1983; Sedgwick, 1983; Pegg and Byers, 1992). This happens by the ada protein becoming methylated in the process of removing methyl groups from methylphosphotriesters in DNA. The methylated ada protein, with an altered conformation, then acts on the "*ada* box" promotor to up-regulate the *alk*A and other genes. In eukaryotes the phenomenon whereby rodent liver ATase and MAG are up-regulated by γ-irradiation has been well described, albeit that the mechanism is less well characterized (Laval and Laval, 1984; Frosina and Laval, 1987; Hall *et al.*, 1990; von Hofe and Kennedy, 1991; Fritz and Kaina, 1992; Lefebvre *et al.*, 1993). The paradoxical nature of this increase, in an apparently inappropriate DNA-repair protein (i.e. one that does not act on γ-irradiation-induced DNA damage) may be explained by ATase being just one of a suite of repair enzymes which are all up-regulated together in mammalian cells in response to xenobiotics. This method of raising ATase levels has been exploited as a means of increasing the amount of protein produced from a given amount of liver (Wilkinson *et al.*, 1990). Induction of expression, if it involved an increase in mRNA levels, could be exploited in a strategy involving proportional subtractive hybridisation, such as described by Hampson *et al.* (1992). For these reasons it was considered worthwhile to see if the level of endogenous FaPy-DNA glycosylase in

rodent liver could be similarly up-regulated, as this would be useful in any attempt to purify the enzyme, which could provide peptide sequences from which oligonucleotide probes might be designed. Additionally, it might provide an improved understanding of the normal physiological regulation of this DNA repair activity.

Another method of increasing the expression of a particular gene product is to repeatedly select cells *in vitro* by exposure to selective agents that are cytotoxic in a way that is related to the function of the protein of interest. The up-regulation that occurs in such circumstances may due to amplification of the gene, i.e its copy-number is increased, rather than simply a temporary increase in transcription/translation; it is thus a relatively stable change in the cells. The continuous exposure of RJKO cells to, at first low but progressively increased, levels of the methylating agent mitozolamide has allowed the derivation of clones with a stable increase in ATase levels, but without evidence of gene amplification (Morten and Margison, 1988; Morten *et al.*, 1992).

It would be reasonable to expect, therefore, that by repeated selection of cells with an agent that generated FaPy-DNA glycosylase-repaired lesions, clones could be produced which had increased their levels of endogenous FaPy-DNA glycosylase. Some clones might have acquired this through gene amplification, but by whatever precise mechanism the change had occurred, so long as the increase in activity was stable, the clone(s) could be used to test whether other agents acted *in vitro* by the formation of FaPy-DNA glycosylase-repaired lesions. The generation of such clones by a particular agent would, in itself, be good *a priori* evidence that the agent acted *via* FaPy-DNA glycosylase-repaired lesions. One problem with this sort of approach though is that few, if any, DNA-damaging agents produce a "clean" spectrum of lesions (1.2.1.1, 1.3.1), also clones might be derived whose resistance to the agent used was not mediated by a change in a DNA repair mechanism. However, reports

available at the time of this study suggested that a particular and relevant lesion could be cleanly generated in DNA, *vis* 8oxoG (Floyd *et al.*, 1989; Floyd *et al.*, 1990; Muller *et al.*, 1990; Schneider *et al.*, 1990). An added benefit was that the photodynamic treatment of cells (1.3.6) was simple to perform and easily manipulated: cells could be treated with photodynamic agents, with and without simultaneous exposure to visible light. It was therefore decided to investigate whether repeated selection of cells by this method could produce clones with a stable increase in endogenous FaPy-DNA glycosylase activity.

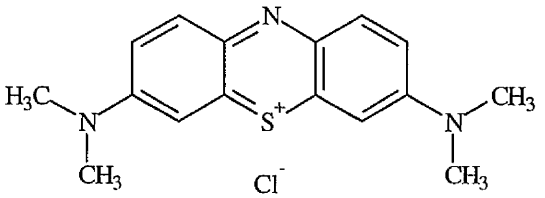
6.2 CELL CLONE SELECTION BY THE USE OF THE PHOTODYNAMIC EFFECT

6.2.1 Initial experiments

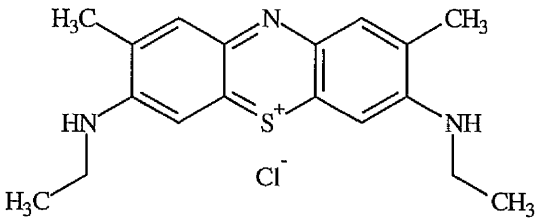
The three essential components of the photodynamic process: the cells, the photodynamic agents and the light source were chosen in the following manner. As RJKO cells were already being used in transfection experiments, and their endogenous level of FaPy-DNA glycosylase was reasonably well characterised, it was decided to use them for repeated selection by photodynamic treatment. The first choice of photodynamic agent was clear cut as previous studies had all used the thiazine dye Methylene Blue (MB; Fig. 6.1a), a well characterised compound available in pure form, which was an important consideration as some impurities in MB may be cytotoxic in the absence of light. It was also decided to test two other dyes: New Methylene Blue N (NMB; Fig. 6.1b) and Bengal Rose B (BRB; Fig. 6.1c).

Fig. 6.1
The dyes used for photodynamic treatment of RJKO cells:
Chemical structures

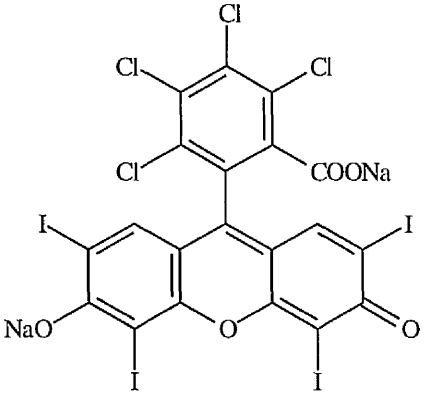
a) Methylene Blue (MB)



b) New Methylene Blue N (NMB)



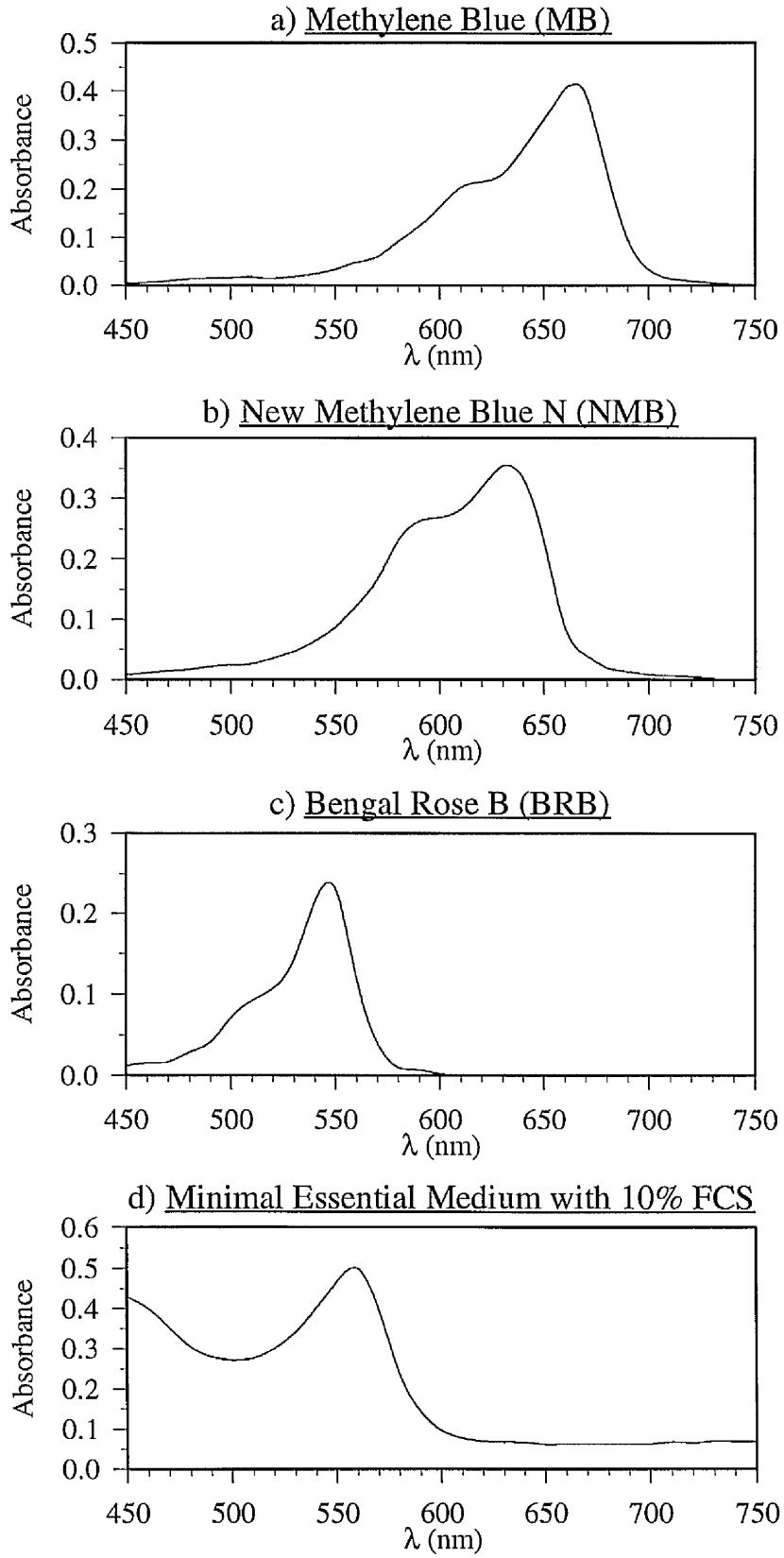
c) Bengal Rose B (BRB)



It can be seen that all three compounds have a (tricyclic) structure composed of alternate single and double bonds, accounting for their strong absorption in the visible part of the spectrum. Their absorption spectra show similar profiles, comprised of a major peak of absorption with a subsidiary peak at a shorter wavelength (Fig. 6.2a, b and c).

Fig. 6.2

Absorption spectra of the dyes used for photodynamic treatment



The absorbance spectra of the dyes (MB, NMB and BRB at 10 μ*M* in EBSS), and the medium used for culture of RJKO cells (MEM + 10% FCS), were measured with a dual-beam scanning spectrophotometer, using EBSS (or water for MEM+FCS) as the reference solution.

Both MB and BRB are known to be efficient generators of singlet oxygen when irradiated with visible light (Bonneau *et al.*, 1975; Shinkarenko and Aleskovskii, 1981; Lamberts and Neckers, 1983). The absorbance spectrum of the medium used for culture of RJKO cells (MEM with 10% FCS), which contains Phenol Red (50 mg/L; 140 μ*M*), was also measured (Fig. 6.2d), showing that it absorbed strongly in the region 450 to 590 nm. This suggested that although only slight interference might occur with MB, this would be greater with NMB, and probably complete with BRB. In addition, the presence of other components in this complex culture medium, such as riboflavin, which is known to have a photodynamic effect, and BSA, which binds small molecules, was considered undesirable. The requirement, therefore, was for a simple, colourless, medium in which to treat the cells. Phosphate-buffered saline (PBS) was considered, but its absence of an energy source was unacceptable as this would cause redox stressing of the cells and MB, at least, can act as a redox agent *in vitro* and *in vivo* at high concentrations (Hunter *et al.*, 1967). (This effect is exploited clinically to treat the condition of methaemoglobinaemia, MB converting the inactive oxidised Fe(III)-haemoglobin to Fe(II)-haemoglobin.) It was therefore decided to use Earle's Buffered Salt Solution (EBSS), which contains 10 g/L glucose (2.1.8).

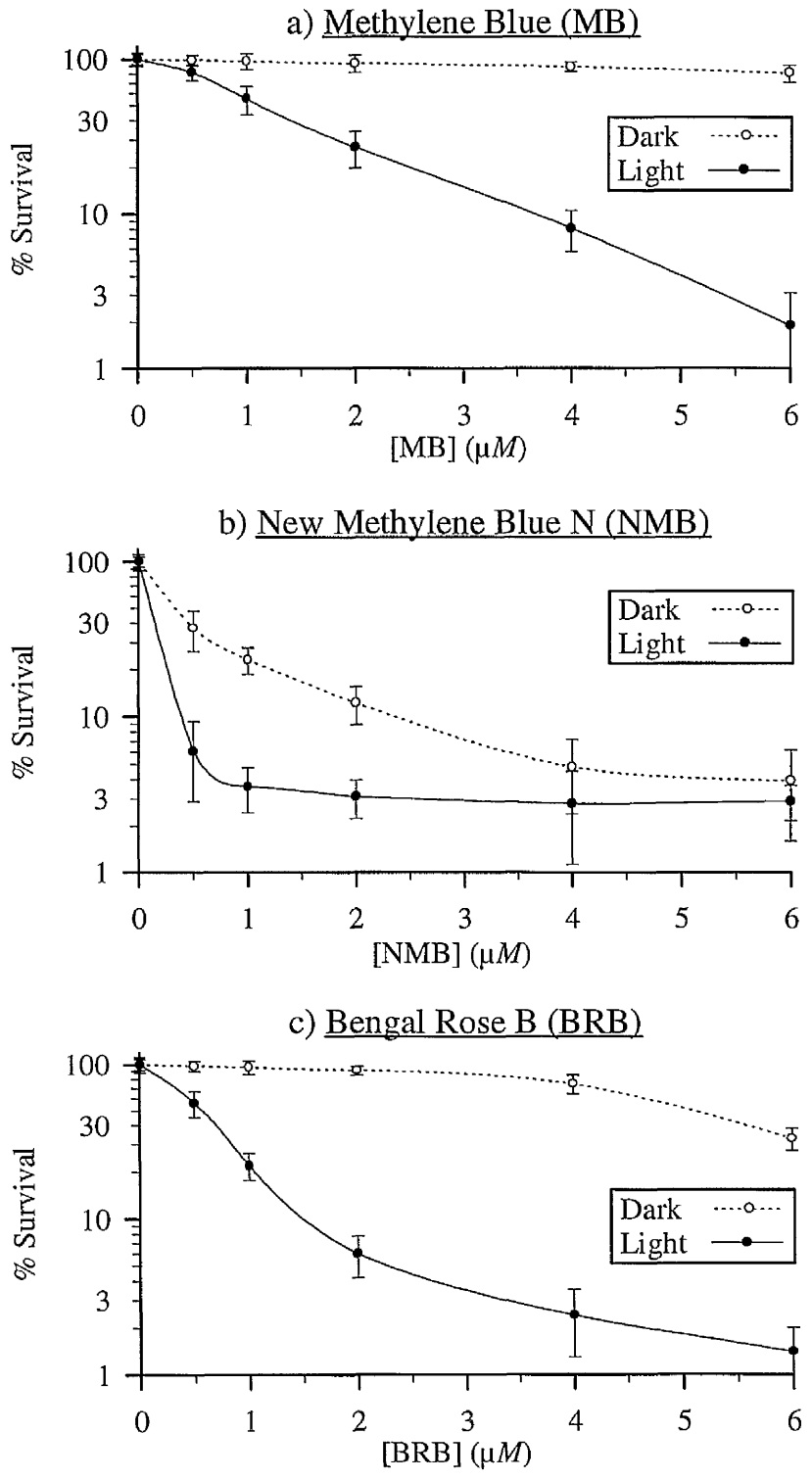
The third component of the system, a source of visible light, needed to fulfil certain parameters. These were: the emission of diffuse light of suitable wavelength (for MB i.e. 580 to 680 nm) at constant luminance (so as to irradiate evenly and simultaneously as many culture flasks as possible), and a filter to absorb the infra-red part of the spectrum (to minimise the heating effect). Experiments performed by others, looking at the photodynamic effect of MB on DNA *in vitro*, had used ordinary tungsten-filament light bulbs (100 to 150 W) with an infra-red filter in the form of a dish of water (Floyd *et al.*,1989; Schneider *et al.*, 1990). A photographic enlarger

available in the Institute fulfilled all these requirements (2.4.4). It produced light of the desired spectrum from a pair of tungsten-halogen bulbs (50 W each) and had a diffuser screen, with an integral infra-red filter, which allowed even illumination of a wide area (approximately 30 × 20 cm). The control unit had a voltage-stabilised output and had the added benefit of an accurate timer which directly controlled the lamps.

This system was then used to examine the effect of the variables involved. Clonal survival experiments were carried out with each dye, at a range of concentrations, with and without exposure to light (2.4.4). Varying the concentration of dye (Fig. 6.3) showed, a) that MB had minimal killing effect in the dark and a satisfactory killing effect with light, which progressively increased with dye concentration, b) that NMB had a strong killing effect in the dark and was also highly toxic with light at all concentrations tested, and c) that BRB had an effect intermediate between that of MB and NMB, both under dark and light conditions. The effect of varying the exposure time was also investigated, using the dyes at the (highest) concentration of 6 μM (Fig. 6.4), and dark and light dose-response curves were obtained with each agent. In view of the excessive toxicity of NMB in the dark the choice of agent was between MB and BRB: MB was chosen as it had a smaller dark effect and was the better characterised of the two compounds in terms of the spectrum of DNA lesions it was known to generate (Floyd *et al.*, 1990; Muller *et al.*, 1990; Schneider *et al.*, 1990).

Fig. 6.3

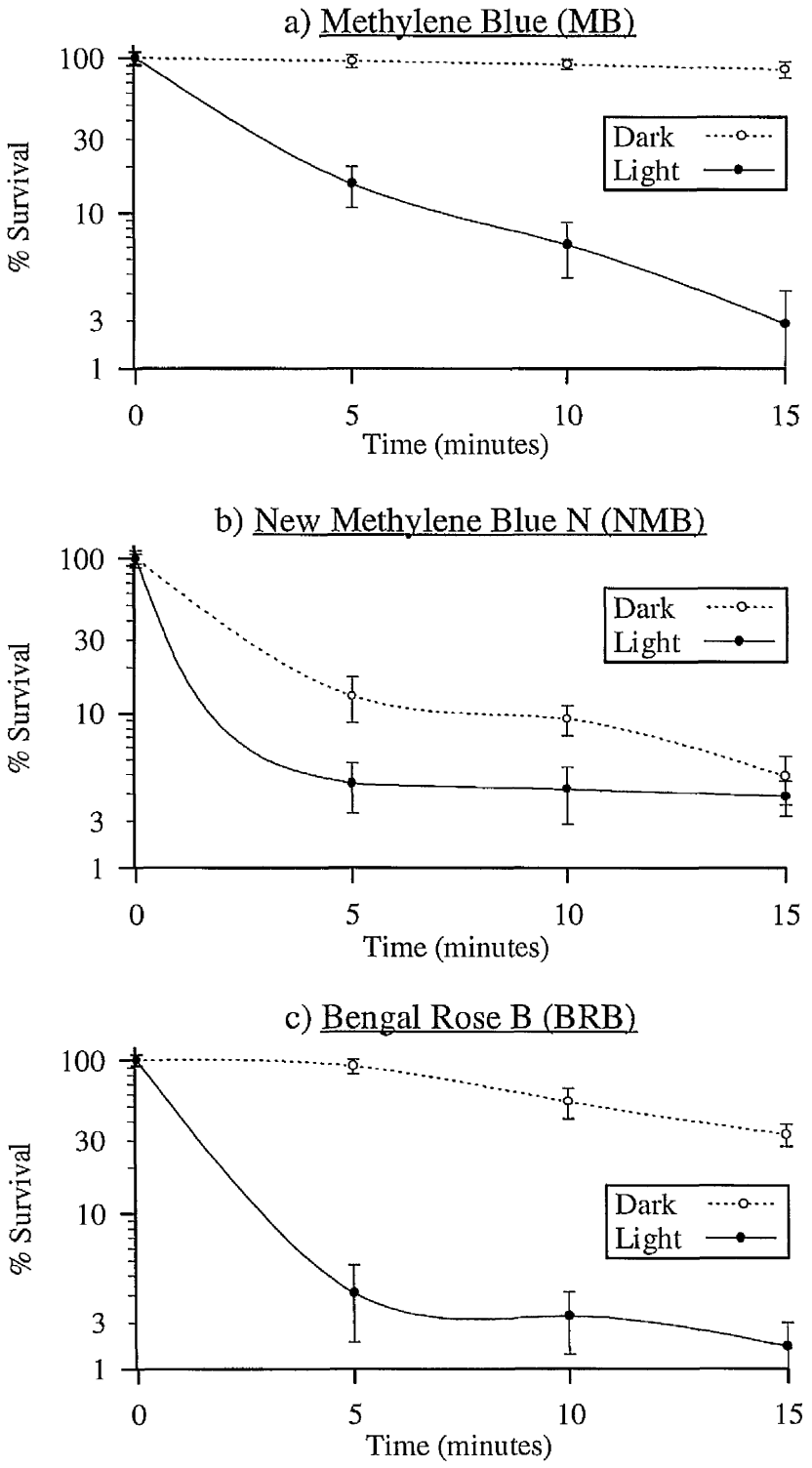
Clonal survival of RJKO cells after treatment with photodynamic agents: The relationship with dye concentration



RJKO cells were plated at 200 cells/6 cm dish, allowed to adhere and then washed prior to the addition of MB, NMB or BRB in EBSS; the dishes were either exposed to visible light ("Light"), or kept in a light-proof box ("Dark"), for 15 minutes at RT. The EBSS was replaced with medium, and colonies were counted 5 days later. Each point is the mean of duplicates, error bars = SEM.

Fig. 6.4

Clonal survival of RJKO cells after treatment with photodynamic agents: The relationship with length of treatment



RJKO cells were plated at 200 cells/6 cm dish, allowed to adhere and then washed prior to the addition of 6 μM MB, NMB or BRB in EBSS; the dishes were either exposed to visible light ("Light"), or kept in a light-proof box ("Dark"), for 5, 10 or 15 minutes at RT. The EBSS was replaced with medium, and colonies were counted 5 days later. Each point is the mean of duplicates, error bars =

SEM. The zero and 15 minute time points on these graphs correspond to those at zero and 6 μM in Fig. 6.3.

6.2.2 Selection of RJKO clones with photodynamic treatment

On the basis of the initial experiments, described in the previous section, it was decided to select surviving clones after irradiating RJKO cells in the presence of 5 μ*M* MB (estimated from the curve in Fig. 6.3a to give about 4% survival). Some RJKO cells (10^3) were therefore treated accordingly and a week later 70 colonies (7% survival) had grown to sufficient size to be picked. Six colonies, of various sizes, were selected and grown up in separate flasks; they were named MPR1.1 to 6.1 (for Methylene blue-Photodynamic treatment-Resistant clones 1 - 6, selection number 1). These six clones were then subjected to another round of selection with 5 μ*M* MB plus light and a single colony picked, from the survivors of each original clone, to give clones MPR1.2 to 6.2. This process of selection and clone picking was repeated a further four times resulting in clones MPR1.6 to 6.6. The survivability of clones MPR1.6 to 6.6 was assessed by treatment with 5 μ*M* MB plus light and the results are shown in Table 6.1.

Table 6.1
Survival rate of RJKO clones MPR1.6 to 6.6 after treatment with 5 μ*M* methylene blue plus light

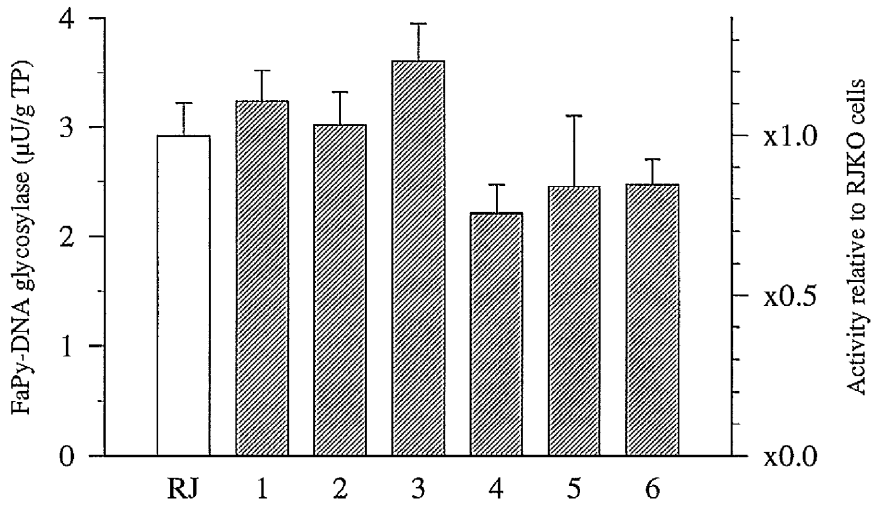
Cells	Clonal Survival Rate (%)[a]
RJKO	4.5 (1.5)
MPR1.6	5.8 (1.0)
MPR2.6	7.0 (1.6)
MPR3.6	5.5 (1.5)
MPR4.6	2.4 (1.2)
MPR5.6	2.6 (1.8)
MPR6.6	6.9 (2.2)

[a] values are means (SEM) of triplicate estimates, no value for an MPR clone was significantly different (by *t*-test) from that for RJKO cells.

In addition to clonal survival testing the clones also had their FaPy-DNA glycosylase activity measured by means of the polyGC-based assay (2.3.4). The levels that were found are illustrated in Fig. 6.5a and, though they varied by a factor of between × 0.75 (MPR4.6) and × 1.23 (MPR3.6) when compared to RJKO cells, the levels were not significantly different (by *t*-test) from the parental cell line. Neither was the mean (SEM) activity in the MPR*n*.6 clones (2.83 (0.20) μU/g TP) significantly different to that in RJKO cells (2.92 (0.30) μU/g TP). It appeared that after six rounds of selection, under the conditions given above, neither the clonal survival rate nor the FaPy-DNA glycosylase levels of RJKO cells could be increased. To determine whether there might be a relationship between FaPy-DNA glycosylase activity and clonal survivability the best-fit linear regression was plotted for the clones plus RJKO cells (Fig. 6.5b). From this it appeared that there might be a positive correlation between the two parameters ($r = 0.52$), however this value was not significantly different from an r value of zero ($p = 0.23$, $n = 7$) and therefore it was concluded that there was no correlation. Taken together, these data strongly suggested that further selection of the clones (at least under the conditions being used) was unlikely to produce an increase in FaPy-DNA glycosylase levels, and it was therefore decided to discontinue the approach.

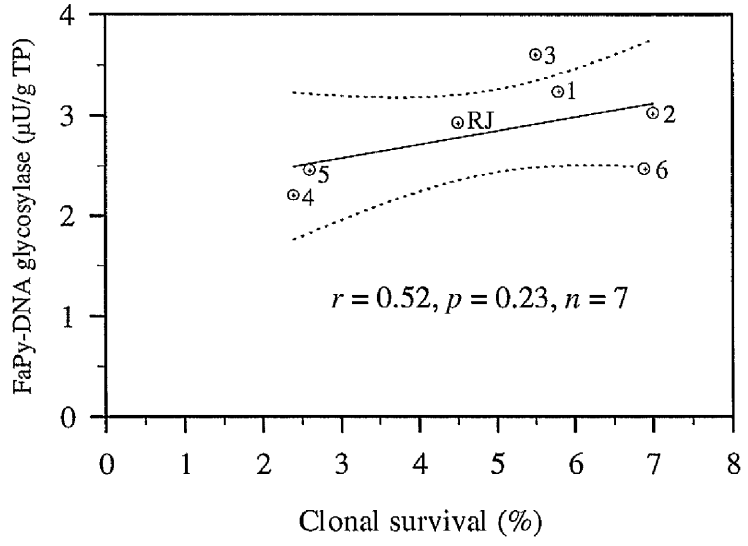
Fig. 6.5
FaPy-DNA glycosylase activity in MPR*n*.6 clones

a) Absolute levels (relative to RJKO cells)



Six MPR clones ("MPR1-6") were derived from RJKO cells by repeated selection with methylene blue plus light: the "MPR*n*.6" clones (the numbers below the hatched bars represent *n*) were the result of six such selections. "RJ" = RJKO cells. FaPy-DNA glycosylase levels were measured in triplicate: the means are shown with the error bars giving the SEM.

b) Correlation with clonal survival



The clonal survival values (from Table 6.1) were plotted against the FaPy-DNA glycosylase levels for RJKO cells and the MPR*n*.6 clones and the best-fit linear regression plotted (solid line) with 95% confidence limits (dotted lines). Points are labelled with the identity of the cell lines they represent: "RJ" = RJKO and numbers = *n* (as in MPR*n*.6). Although there appeared to be a positive correlation ($r = 0.52$), as is evident from the confidence limits of the regression, r was not significantly different from zero ($p = 0.23$, $n = 7$).

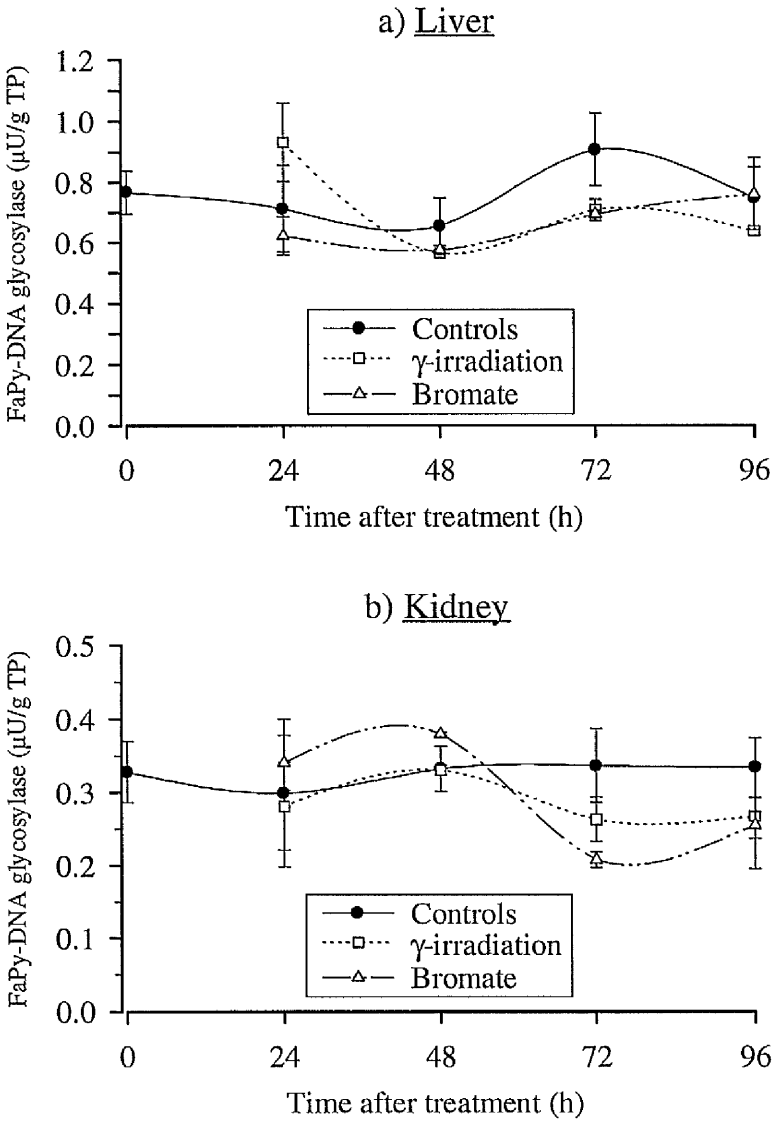
6.3 THE EFFECT OF DNA-OXIDISING AGENTS ON ENDOGENOUS RODENT FaPy-DNA GLYCOSYLASE LEVELS

6.3.1 FaPy-DNA glycosylase levels in rats treated with γ-irradiation or potassium bromate

Two agents known to generate lesions in DNA repaired by FaPy-DNA glycosylase are γ-irradiation and potassium bromate (1.3.1). Rats given a substantial dose of γ-irradiation, of the order of 5 Gy, exhibit a rise in hepatic ATase levels 24 to 48 hours later (Schmerold and Wiestler, 1986). As mentioned in section 6.1 this may be part of a generalised rise in DNA-repair activity in response to DNA damage. The oxidant potassium bromate is a renal carcinogen in the rat (Kurokawa *et al.*, 1983). When rats are given bromate, either parenterally or in their diet, the level of 8oxoG in the DNA of the renal cortical cells increases (Kasai *et al.*, 1987); homogenates of rat renal cortex, unlike other organ extracts, generate singlet oxygen upon the addition of bromate (Sai *et al.*, 1992).

An experiment was therefore carried out to see if the FaPy-DNA glycosylase levels in rat tissues were increased by γ-irradiation or potassium bromate. A group of male OBW rats were selected, fourteen controls and sixteen test subjects. Eight of the test subjects were given 4.5 Gy of γ-irradiation from the Institute's [^{60}Co] source, at a rate of 0.5 Gy/min (2.1.10), while the other eight were given a single dose of potassium bromate (70 mg/Kg body weight) by i/p injection. The controls were sham treated, i.e. lightly anaesthetised but neither irradiated or injected, and six were sacrificed immediately to provide samples for the zero time points. Two animals from each of the three groups were then killed at 24 h intervals, starting at 24 h after the initial irradiation/injection, and the hepatic and renal FaPy-DNA glycosylase levels assayed (Fig. 6.6).

Fig. 6.6
FaPy-DNA glycosylase levels following
treatment of rats with γ–irradiation or potassium bromate



Male OBW rats (225-245 g) were lightly anaesthetised (Enflurane) and then either allowed to recover (Controls), given 4.5 Gy γ-irradiation ([^{60}Co]; 0.5 Gy/min), or injected with 70 mg/Kg b.w. of potassium bromate. All points are means (±SEM) of results for 2 animals, except controls at zero hours (6 animals).

At autopsy those animals treated with potassium bromate were found to have methaemoglobinaemia, evident by their blood having a brown rather than scarlet colour, even up to 96 h after dosing, but they otherwise gave no evidence of being clinically affected by the treatment. In the control group the levels in liver and

kidney, at any time point, were not significantly different to those at zero hours (by *t*-test). Similarly, the FaPy-DNA glycosylase levels, in liver and kidney, at any time point, in animals treated with either agent were not significantly different from either the respective control level at zero hours or the control group at the same time point.

6.3.2 FaPy-DNA glycosylase levels in mice treated with γ-irradiation

The observation that mice, like rats, up-regulate their levels of ATase in response to γ–irradiation (Wilson *et al.*, 1993) had prompted a further study (Dr G.P. Margison, personal communication). This had shown that mice with severe combined immune deficiency (SCID), a heritable condition due to a defect in DNA double-strand break repair and associated with hypersensitivity to ionizing radiation, showed a greater rise in ATase levels following γ–irradiation than congenic non-SCID control (BALB/c) mice. This response had been determined 48 h after irradiation, the time previously found at which ATase induction was maximal, with a range of doses from 0.25 Gy to 5 Gy. The greatest differential in hepatic ATase induction was found at the 2 Gy dose: 2.3 ± 0.2 fold in SCID animals (compared to un-irradiated controls) *v.* 1.5 ± 0.1 fold in the BALB/c mice (similarly compared to un-irradiated controls). Frozen tissues were available in the Department from the animals in this study and it was considered worthwhile to investigate whether FaPy-DNA glycosylase levels were increased in mice, and if so whether SCID mice showed a greater response. Extracts for FaPy-DNA glycosylase assay were therefore made from the frozen liver samples from the un-irradiated control and 2 Gy-irradiated animals (2.3.4); each was assayed in duplicate (2.3.5), the means of each group of mice are shown in Table 6.2.

Table 6.2

Hepatic FaPy-DNA glycosylase levels in γ-irradiated BALB/c and SCID mice

Treatment	BALB/c[a] (μU/g TP)	SCID[a] (μU/g TP)
Un-irradiated	1.51 (0.02)[b]	1.43 (0.11)
2 Gy[d]	1.13 (0.15)	1.64 (0.23)[c]

[a] BALB/c and SCID mice were congenic, [b] $n = 2$ (animals) in each case, except for [c] where $n = 3$, SEM are shown in parentheses, [d] [^{60}Co] source (dose rate = 0.5 Gy/min).

These data showed that the hepatic FaPy-DNA glycosylase levelswere no different in BALB/c *v.* SCID mice (controls: $p = 0.55$, or irradiated animals: $p = 0.21$; *t*-test), nor was the level changed by γ-irradiation in either strain (BALB/c: $p = 0.13$, SCID: $p = 0.54$; *t*-test). While these figures were based on small group sizes, which would not have given the experiment sufficient statistical power to show small differences, the group sizes had been adequate to show differences of 1.5-fold in BALB/c ATase levels (Dr G.P. Margison, personal communication). It was therefore concluded that γ-irradiation did not increase FaPy-DNA glycosylase levels in mice, at least not to the extent seen with ATase.

6.4 INVESTIGATION OF ENDOGENOUS MAMMALIAN FaPy-DNA GLYCOSYLASE ACTIVITY WITH RABBIT ANTI-FPG ANTISERA

6.4.1 The effect of anti-FPG antisera on mammalian FaPy-DNA glycosylase activity

As described previously (4.3.3), some of the rabbit polyclonal antisera had been shown to inhibit the FaPy-DNA glycosylase activity of FPG protein. If similar inhibitory activity was shown towards endogenous mammalian FaPy-DNA glycosylase activity, then this might allow isolation of eukaryotic homologues of *fpg*, and would

also make possible attempts to clone eukaryotic homologues of *fpg*, using a cDNA expression library in *fpg*⁻ *E. coli*. For this purpose antisera directed at epitopes present in native (FPG) protein would be more appropriate than those suited to detecting denatured polypeptides (as e.g. in western analysis). To this end it was decided to test whether the FPG antisera could inhibit the FaPy-DNA glycosylase activity in extracts from mouse liver or human kidney (2.3.13), and the results of this are given in Table 6.3.

Table 6.3
The effect of the rabbit antisera to FPG on FaPy-DNA glycosylase levels in mouse liver and human kidney

Antiserum[a]	BALB/c liver (% Control)	Human kidney[b] (% Control)
Nil (Control)	100.0 (2.4)[c]	100.0 (2.7)[d]
α67-pre	104.8 (3.7)	98.5 (2.5)
α67-2	102.2 (3.2)	97.0 (3.1)
α67-3	99.1 (432)	98.3 (3.0)
α69-pre	99.8 (3.1)	102.5 (3.4)
α69-2	96.7 (2.1)	102.2 (3.0)
α69-3	101.3 (2.9)	99.8 (2.4)

[a] all antisera were used at a dilution of 1:10 (2.3.13), [b] extract from snap-frozen normal tissue obtained at nephrectomy for a renal carcinoma in a 35 y male (kindly supplied by Dr S.P. Guy, St Mary's Hospital, Manchester), [c] each sample was assayed in duplicate (2.3.4), SEM are shown in parentheses; absolute value = 1.28 nU, [d] absolute value = 1.26 nU.

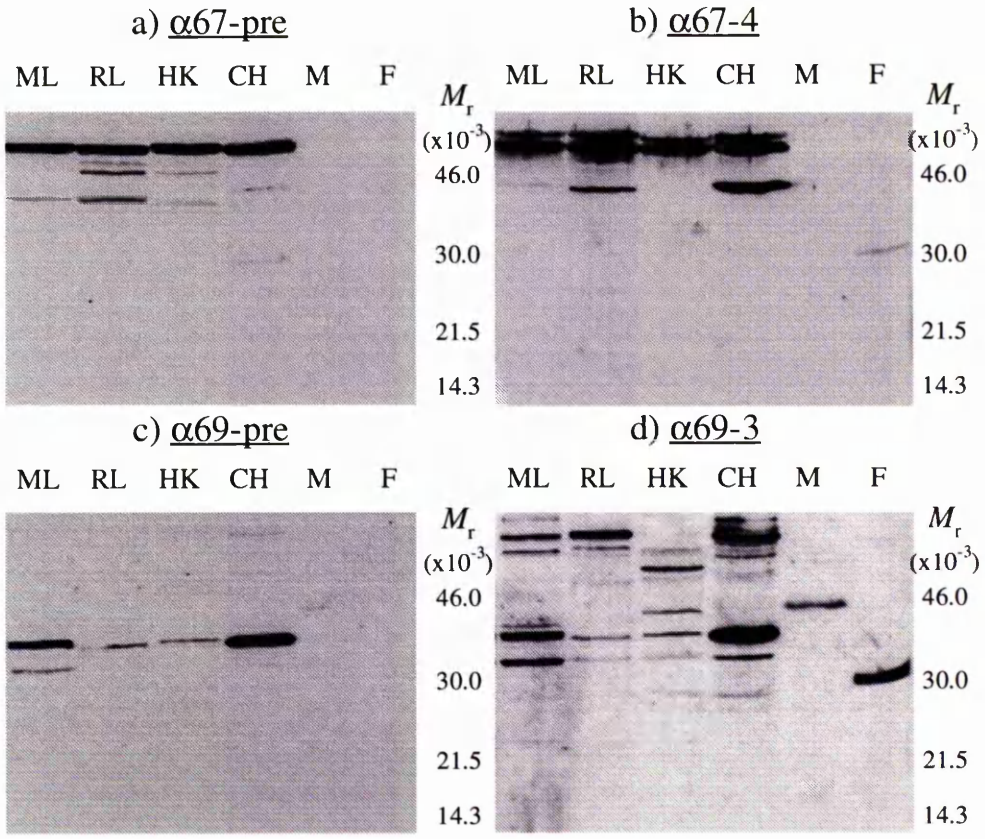
It was evident from this that under the conditions used none of the antisera had any effect on the FaPy-DNA glycosylase activity in either the mouse or human extracts.

6.4.2 Western analysis of mammalian cell extracts with anti-FPG antisera

The antisera α67-4 and α69-3 had already been used to probe western blots of extracts from RJKO cells and clones derived from the *fpg*10-derived expression

vectors (5.4.7.3). To investigate if these antisera might be of use in probing for mammalian homologues of *E. coli* FPG protein, western analysis was carried out using extracts from a number of mammalian species (2.3.4; Fig. 6.7). The hybridisation band patterns obtained are summarised in Table 6.4.

Fig. 6.7
Western analysis of mammalian cell extracts



A variety of mammalian cell extracts (30 µg TP) were separated by SDS-PAGE, and then subjected to western analysis (2.3.14) with detection by chemiluminescence. "M" = coloured dye-labelled marker protein mix, "ML" = C57BL/6J mouse liver, "RL" = OBW rat liver, "HK" = human kidney (for details see Table 6.3), "CH" = Chinese Hamster cells (RJKO), and "F" = FPG protein (5 ng).

Table 6.4

Hybridisation bands seen on western analysis of mammalian cell extracts[a]

Antiserum	Band ($M_r \times 10^{-3}$)	Mouse liver	Rat liver	Human kidney	Chinese hamster	FPG protein
a) **α67-pre**	*57*	4[b]	4[c]	4	4	
	53		1			
	51		3	1		
	44	1	3		1	
	42			1		
	32				1	
b) **α67-4**	*60*	**3**[d]	**3**		**3**	
	58			**3**		
	57	(4)[e]	(4)	(4)	(4)	
	53		(3)	**1**		
	51		(2)			
	49		**1**			
	44	(1)	(3)		(4)	
	32					**1**
c) **α69-pre**[f]	*62*					
	39	3	2	2	3	
	35	2				
d) **α69-3**[g]	*67*	1			2	
	64				2	
	62	**3**	**4**		**4**	
	59		2		1	
	58	2		**1**	1	
	57		1	1	2	
	53			**3**		
	51			1	1	
	49		**1**			
	44			3		
	41	2				
	39	(4)	(3)	(3)	(5)	
	35	(3)	1	1	3	
	32					**4**
	30			1	1	

[a] as in Fig. 6.7, [b] M_r as estimated by extrapolation or interpolation from positions of marker proteins, [c] semi-quantitative estimate of hybridisation signal: "5" = very strong, "4" = strong, "3" = moderate, "2" = weak, "1" = very weak, [d] those bands detected by both antisera (but *not* the corresponding pre-immune sera) shown in bold, [e] bands found with immune as well as pre-immune sera shown in parentheses, [f] also gave very weak, and [g] moderate strength hybridisation signals with the M_r 46 000 marker protein (ovalbumin).

It can be seen that with antiserum α67-4 six bands were detected in the mammalian extracts which had not been detected by the preimmune serum α67-pre. This

antiserum only gave a weak hybridisation signal with purified FPG protein, at an apparent M_r of 32 000. In contrast, antiserum α69-3 gave a strong signal with purified FPG protein at the same apparent M_r, and detected 24 bands in the mammalian extracts, which had not been detected by the preimmune serum α67-pre, mostly in the range M_r = 41 000 to 67 000. Of these bands, three were in common with those found with α67-4 (i.e. at M_r = 49 000 in rat liver, plus M_r = 58 000 and 53 000 in human kidney). In addition, given the inaccuracy inherent in extrapolation from the range of marker proteins used, the bands with α67-4 at M_r = 60 000 in mouse, rat and Chinese hamster and M_r = 58 000 may be equivalent to those seen with α69-3 at M_r = 62 000.

6.5 DISCUSSION

6.5.1 Selection of RJKO clones using MB plus light

The initial RJKO cell clonal survival experiments in which MB, NMB and BRB were investigated as to their use as selective photodynamic agents demonstrated differences between the three agents. Methylene Blue showed the smallest dark effect, and its light effect was related in a similar fashion to either the concentration of dye or length of exposure. The related compound NMB had a greater light effect than MB, at least at low concentrations/exposure times, but it also had a considerably greater dark effect. This suggested that it was toxic in the absence of light, however, the preparation of cells for the dark treatment was not completely devoid of light (2.4.4) and therefore this may represent much more efficient generation of singlet oxygen by NMB with visible light. The structures of the two compounds (Fig. 6.1a and b) are similar, but it is of note that the by-product of MB synthesis with only three N-methyl groups is considered cytotoxic in the absence of light (Bergmann and O'Konski, 1963). It could therefore be that NMB is strongly dark toxic because both

amine groups are incompletely alkylated. Bengal Rose B had both light and dark effects intermediate between those of MB and NMB. It was considered that NMB had too excessive a dark effect, and that because MB was better characterized than BRB in terms of the DNA lesions it was known to produce, that MB would be used to select RJKO cells.

The results presented in section 6.2.2 show that repeated selection of RJKO cells, a total of six times, did not result in clones with either increased FaPy-DNA glycosylase activity or greater resistance to MB photodynamic treatment. There are a number of possible reasons why this was so. Firstly, the number of selections may have been inadequate. But there was no suggestion of an upward trend in FaPy-DNA glycosylase levels, in particular no evidence of an increase in the mean of the six clones, which it might be reasonable to expect after this time. Secondly, the severity of selection might not have been sufficient. The concentration of MB chosen (5 μM) was a compromise to try and maximise the differential between the dark and light effects, as it was considered undesirable to have too large a dark effect. However, this may have been erring on the side of caution and a higher concentration might have been used. Thirdly, it may be a consequence of the generation by MB of both 8oxoG and FaPy-G in DNA, the ratio of which, at least *in vitro*, depends on the oxygen tension (Fuchiarelli *et al.*, 1990). Under the conditions used the DNA lesions produced may have been mostly 8oxoG with little FaPy-G, but this must remain speculation in the absence of a direct assay for these. It is noteworthy that vascular endothelial cells are considered to be the major target of photodynamic therapy to tumours (Cowled and Forbes, 1989), and out of all cell types in a tumour, or other tissues, they will be exposed to the greatest oxygen tension. Lastly, it has been assumed that the cytotoxicity of MB plus light is mediated by the effect of singlet oxygen on DNA and this may not be entirely valid. Other cellular macromolecules,

such as enzymes, can also be the target for this form of reactive oxygen species (1.3.6). Thus the MPR clones may not have developed a rise in FaPy-DNA glycosylase levels, however they would have been expected to show an increase in survival rate, which they did not.

6.5.2 The effect of DNA-oxidising agents on rodent FaPy-DNA glycosylase levels

No significant changes could be demonstrated in OBW rat FaPy-DNA glycosylase levels, in either liver or kidney, up to 96 h after γ-irradiation or treatment with potassium bromate (6.3.1). The dose of γ-irradiation given (5 Gy) was well established to cause a significant rise in hepatic ATase levels with this (Dr G.P. Margison, personal communication) and other strains of rat (Schmerold and Wiestler, 1986). Treatment with potassium bromate has not been reported to cause a rise (or fall) in the level of a DNA repair enzyme, however the dose used (70 mg/Kg b.w.) has been found to generate a significant rise in the 8oxoG level of rat kidney cortex DNA (Kasai *et al.*, 1987; Sai *et al.*, 1992). In addition, the toxicity in the form of methaemoglobinaemia observed in the bromate treated rats suggests that they received a significant dose. This is supported by a comparison of known LD50 values (Material Safety Data Sheet, 1992), as a rat (i/p) LD50 is not available: mouse (oral) LD50 = 289 mg/Kg b.w., mouse (i/p) LD50 = 177 mg/Kg b.w., and rat (oral) LD50 = 321 mg/Kg b.w., by proportion (which it is accepted may not be entirely valid) this implies a rat (i/p) LD50 = 197 mg/Kg b.w. Thus, they received a dose approximately one-third that of the estimated LD50. While it cannot be assumed that bromate treatment induced sufficient DNA damage (of the appropriate kind) to cause up-regulation of FaPy-DNA glycosylase, it would appear that the animals received a dose near the safe maximum.

The experiment with mice gave similar results to those seen with rats: neither BALB/c nor congenic SCID mice showed any evidence of an increase in hepatic FaPy-DNA glycosylase levels after γ-irradiation. In this instance the liver ATase levels had been measured and were found to have significantly increased, 1.5-fold in BALB/c and 2.3–fold in SCID mice (Dr G.P. Margison, personal communication). This greater differential in SCID mice suggests that DNA double-strand breaks are involved in the mechanism whereby ATase is up-regulated.

This lack of an increase, in rats and mice, in an enzyme which is involved in the primary repair of oxidative DNA lesions is curious, particularly when it is considered that other enzymes such as ATase and MAG do show a response, and yet they are involved in the repair of alkylation damage. It is unlikely that the treatment given, at least the γ–irradiation, was insufficient to have elicited an increase in FaPy-DNA glycosylase activity, if in fact there is one. This suggests that the rodent FaPy-DNA glycosylase gene is not subject to control in the same fashion as ATase. These experiments do not rule out the possibility that there is more than one enzyme responsible for the repair of the lesions generated other than FaPy-DNA glycosylase, and that this/these other enzyme(s) are inducible. There is a precedent, albeit in prokaryotes, for the existence of two DNA repair enzymes fulfilling the same function, one constitutively expressed and the other inducible by increased DNA damage, i.e. *ada* and *ogt* in *E. coli* (Rebeck *et al.*, 1988). However, the report of a separate mammalian 8oxoG-DNA glycosylase lacking FaPy-DNA glycosylase activity, which might have added weight to this argument, is open to doubt (Chung *et al.*, 1991b; 1.3.4). As far as is known, there is only one report looking specifically at mammalian FaPy-DNA glycosylase levels and their possible up-regulation by DNA-oxidising agents (Hall *et al.*, 1993). With the recent (1995) availability of 8oxoG-phosphoramidite making feasible a quantitative oligonucleotide-based assay for

8oxoG-DNA glycosylase, it could be possible to determine whether this strategy produces a rise in 8oxoG-DNA glycosylase, rather than FaPy-DNA glycosylase.

6.5.3 Probing mammalian cell extracts with polyclonal anti-FPG antisera

Two contrasting effects were explored in the experiments using the rabbit polyclonal antisera to investigate mammalian FaPy-DNA glycosylase activity. The first (6.4.1) was looking for evidence that the antisera α67-2, α67-3, α69-2 or α69-3 could inhibit mouse or human FaPy-DNA glycosylase activity. As was seen none could be demonstrated, although all the antisera (in particular α69-3) had been shown to inhibit the FaPy-DNA glycosylase activity of FPG protein (4.3.3). Liquid-phase inhibition of an enzyme by a polyclonal antiserum is presumably due to the presence of antibodies directed against epitopes, present in the native enzyme, connected with its catalytic activity. It was therefore concluded that FPG protein and mouse or human FaPy-DNA glycosylase are unlikely to share common epitopes, at least those directly related to their glycosylase functions. It is possible that the FPG-directed antibodies bind to analogous mammalian enzymes, at the pre-incubation temperature used (0°C; 2.3.13), but that the bonds involved dissociate on warming up to the assay temperature of 37°C: it might be possible to test this by carrying out the FaPy-DNA glycosylase assay at a lower temperature.

The second experiment was carried out to investigate whether the antisera α67-4 and α69-3, which had been used previously to probe western blots of the transfected RJKO clones (5.4.7.3), might be of use in detecting mammalian homologues of FPG protein. This could be exploited by screening mammalian cDNA expression libraries contained in an *fpg*⁻ *E. coli* host (Boiteux and Huisman, 1989). Certainly both antisera detected various bands in all of the extracts tested, however many of these bands were also present in the blots probed with the respective preimmune sera, some giving

strong hybridisation signals. Antiserum α67–4 gave only a weak signal with purified FPG protein, in contrast to α69-3 which gave a strong signal. The strong non-specific signal seen with α67-4 at M_r = 57 000 in all extracts rules this antiserum out as useful for screening libraries. Similarly the strong/very strong non-specific signals seen at M_r = 39 000 in mouse and Chinese hamster extracts with α69-3 would be expected to cause problems with frequent false-positive detection of cDNA clones. In contrast, the rat and human extracts only gave moderate strength signals at M_r = 39 000, as well as a number of specific bands. It might therefore be possible to use antiserum α69-3 for screening human or rat cDNA libraries, but further testing would be required, such as checking whether it also detected bands in the *fpg*⁻ *E. coli* host. If this occurred the antiserum could be adsorbed first with an extract from the *fpg*⁻ *E. coli* strain (Dr S. Boiteux, personal communication). It is of note that antiserum α69-3 had a) the highest anti-FPG titre by ELISA, b) the greatest anti-FPG DNA glycosylase effect, and c) the strongest anti-FPG signal on western analysis. It was also the best candidate for use in screening a cDNA expression library.

Enhanced excision endonuclease repair activity has been observed in monkey CV-1 and mouse Krebs II cells 12 h after 5 Gy [^{137}Cs] ionizing radiation (Bases *et al.*, 1992). Size-fractionation by gel filtration of crude extracts of both unirradiated and irradiated cells showed an EDTA-resistant activity cleaving 8oxoG, from a duplex oligonucleotide substrate, in the range M_r = 60 000 to 70 000. In this respect it is of note that a consistent band was seen on western analysis (Fig. 6.7; Table 6.4) in rodent extracts at $M_r \approx$ 60/62 000, and in human kidney extract at M_r = 58 000. Although no increase was observed in FaPy-DNA glycosylase levels following γ-irradiation (6.3), this was only sought at time points 24 h and later. Further experiments might be carried out, therefore, on cultured cells or animals, with samples taken in the 0 to 24 h range. These could then be assayed for FaPy-DNA (and

8oxoG-DNA) glycosylase activity, together with western analysis looking for any increase in intensity of particular bands.

CHAPTER 7

THE GENERATION AND CHARACTERISATION OF *fpg*-TRANSGENIC MICE

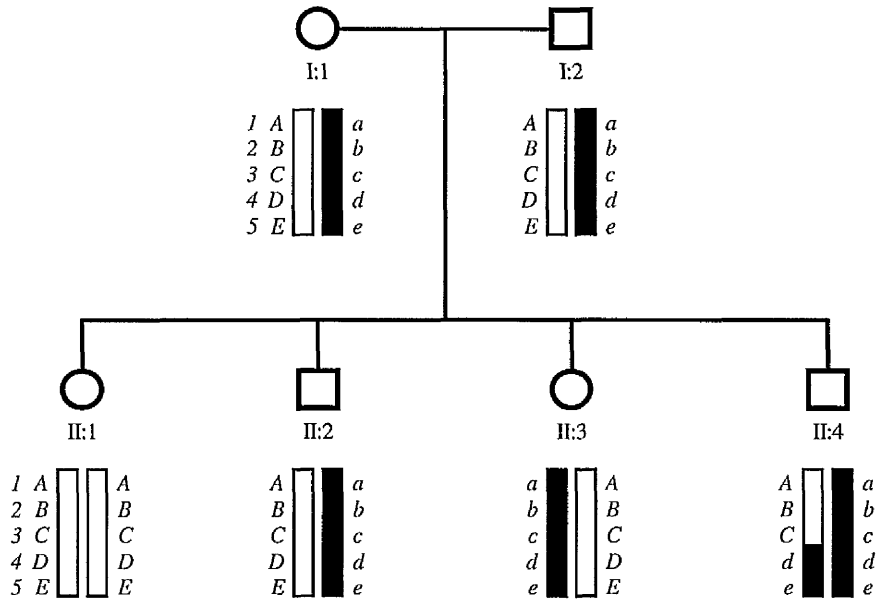
"Breeders habitually speak of an animals's organization as something quite plastic, which they can model almost as they please." (Darwin, 1866)

7.1 BACKGROUND CONSIDERATIONS

The phenotype of an organism is a function of its genotype and environment. The phenotypic variation between individuals complicates the use of animals in the study of toxicity and carcinogenesis. Genotypic variation occurs in all organisms as a result of the generation of novel alleles of genes by mutation. In eukaryotes the diploid state has permitted an extra dimension to this by allowing different alleles of the same gene to act together in the same individual. Chromosomal recombination (also known as "crossing over") during the process of sexual reproduction has taken this a step further: the exchange of segments of homologous chromosomes during meiosis ensures that combinations of alleles do not remain constant (Fig. 7.1). At least one crossover must occur on every chromosome arm during meiosis (Sturt, 1976). Somewhat simplified, the likelihood of recombination occurring between any two loci is a function of their distance; the further apart they are, the more likely it is that they will become separated by recombination. Loci far enough apart so that the probability of recombination would be >50% (i.e. $\theta > 0.5$) always appear to be inherited as if they were on different chromosomes (i.e. $\theta = 0.5$), and are described as unlinked. Loci close enough together such that $\theta < 0.5$ are said to be genetically linked, two adjacent loci may be so close as to appear to be linked with θ = zero. While θ is a measure of genetic distance, the relationship between the two is complicated by the possibility of multiple crossover events. The unit of measurement of genetic distance

is the centimorgan (cM), and over small distances 1 cM is equivalent to $\theta = 0.01$, i.e. a recombination fraction of 1%. The probability of recombination is not uniform along a chromosome's length and also varies depending on whether the meiosis occurred in the male or female (Ayala and Kiger, 1984). However, in the laboratory mouse 1 cM is equivalent to approximately 2 Mb of DNA, a region sufficient to contain 30 genes (Hogan *et al.*, 1986). The 40 chromosomes in the diploid mouse genome average 80 Mb (40 cM) in length (Hogan *et al.*, 1986).

Fig. 7.1
Chromosomal recombination in eukaryotes



Hypothetical pedigree illustrating the process of chromosomal recombination which takes place during meiosis in eukaryotes. The parents (mother I:1 and father I:2) each have a pair of homologous chromosomes with five loci *1, 2, 3, 4* and *5*, having alleles *A/a*, *B/b*, *C/c*, *D/d* and *E/e* respectively. Offspring II:1, II:2, and II:3 have inherited the parental chromosomes intact, but II:4 has inherited a novel chromosome formed by recombination between loci *3* and *4*. Recombination might have occurred between any two loci, the frequency per generation with which this occurs gives a measure of their genetic distance (θ).

The variation inherent in a population of animals, e.g. rodents, can be minimised by various means. The environment can be controlled ~~such~~ so that the diet, temperature, humidity, ventilation, lighting levels and day length, amongst other parameters, are constant and standardised. It is more difficult, perhaps impossible, to standardise the

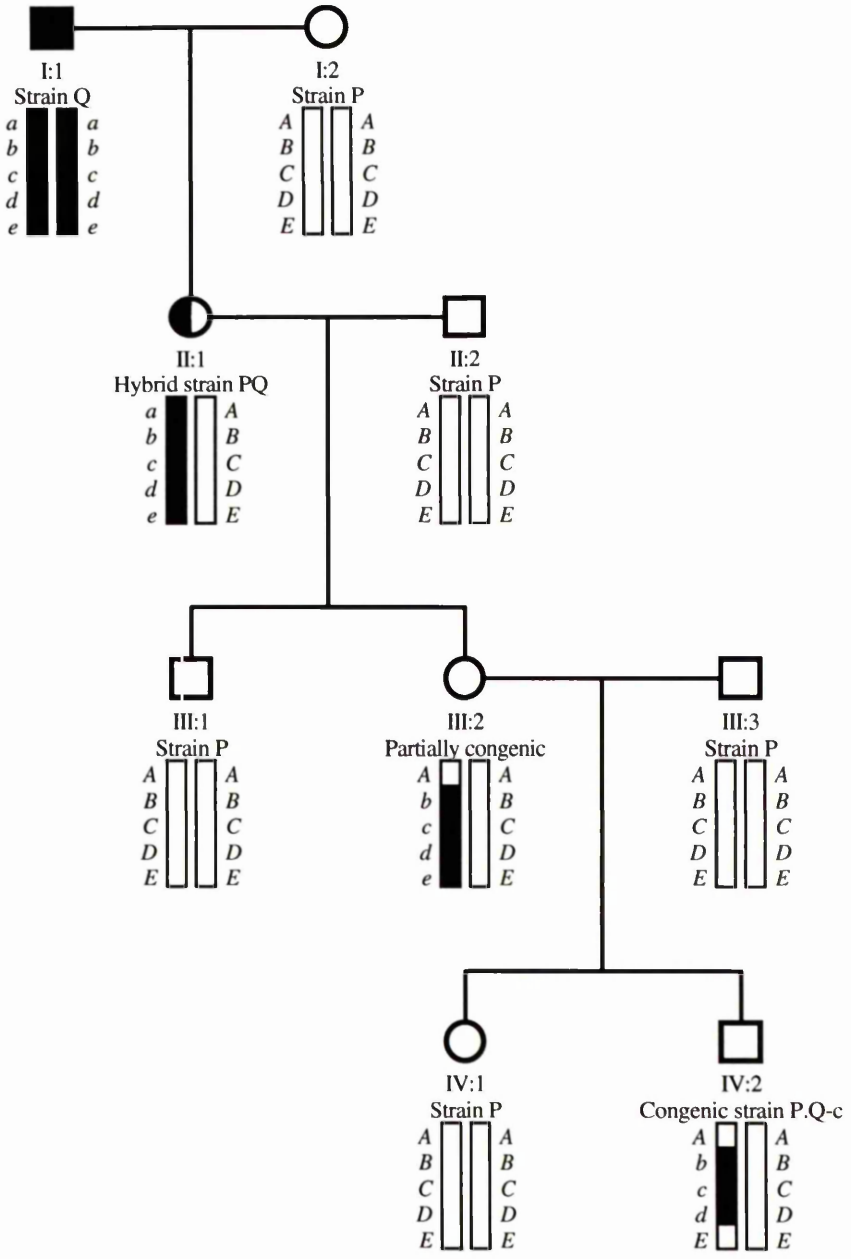
commensal organisms of the gut or skin, but these may be assumed to be reasonably consistent between individuals of a population kept in the same environment. The presence of pathogens, e.g. viruses, bacteria, fungi or parasites, can mostly be checked for and thus controlled appropriately. Genetic variation can be controlled by inbreeding. In a wild population of individuals many, if not most, genes will have different alleles. By repeatedly breeding brother with sister, over at least 20 generations, an inbred strain can be derived which effectively will be homozygous at all loci (Altman and Katz, 1979), and in the process all recessive lethal alleles are eliminated. Any inbred strain can never be homozygous at every locus all of the time because of the small but finite spontaneous mutation rate, estimated to be approximately 10^{-5} per locus per generation (Morse, 1981). This rate is sufficiently low ~~such~~ so that its effect can usually be ignored.

There are now approximately 200 different inbred strains of the laboratory mouse, most of which can be traced back to a relatively small number of founding mice (Hogan *et al.*, 1986; Altman and Katz, 1979). Much of the demand for inbred mouse strains was driven by studies on the genetics of cancer susceptibility, in particular studies on the problems of tumour transplantation (Hogan *et al.*, 1986). Analysis of mitochondrial and nuclear DNA has shown the modern laboratory mouse to be derived from *Mus musculus domesticus*, with many strains (e.g. BALB/c, C57BL/6J (abbreviated B6) and DBA/2J (abbreviated D2)) having the Y chromosome from *Mus musculus musculus* (Ferris *et al.*, 1982; Bulfield, 1985). Physical, and hence genetic, isolation of colonies of the same strain will in time result in those colonies becoming different ~~to~~ from the original strain (~~due~~ owing to the finite spontaneous mutation rate). This may or may not be a disadvantage, but it has to be borne in mind that separate colonies of what is ostensibly the same strain may not be genetically identical. Such, usually slight, differences between colonies are probably outweighed by the advantages of a

long-term constant environment, particularly in terms of commensal organisms and pathogens.

Another type of laboratory mouse has been developed, originally to study the genes of the histocompatibility locus, but the process is applicable to the study of any set of closely linked genes. This is the production of congenic inbred strains, a term coined by Snell (1978) who used the process to investigate histocompatibilty genes. For example, if it is desired to study the effect of the strain Q allele *c* (at locus *3*) on the genetic background of strain P, which has allele *C*, then the hybrid PQ is backcrossed with strain P and those offspring having allele *c* are selected. (In this example alleles in upper-case, e.g. *C*, are not necessarily dominant, nor those in lower-case recessive.) These animals are then crossed with strain P and the offspring selected for the presence of allele *c*, and so on. Thus, in time, the genetic background of this strain becomes almost entirely of strain P (and homozygous) but with locus *3* heterozygous *C/c*. Let locus *3* be flanked on either side by the loci *1, 2* and *4, 5*. Recombination around locus *3* will probably have replaced the strain Q alleles at the loci furthest away (i.e. *1* and *5*) with strain P alleles, however, those loci more closely linked to *3*, i.e. *2* and *4*, are more likely to be present as strain Q alleles (Fig. 7.2). This phenomenon has implications for the strategy used to generate transgenic mice, as will be discussed later.

Fig. 7.2
Derivation of a congenic strain of laboratory animal



Hypothetical pedigree illustrating the process of deriving a congenic strain of laboratory animal. If it is desired to put allele *c* of locus *3* (from strain Q) onto the genetic background of strain P (which is homozygous *C/C* at the same locus), then a PQF_1 hybrid (individual II:1) is backcrossed with strain P (II:2). The offspring are selected for allele *c* (III:2), and then backcrossed with strain P (III:3), to produce the congenic strain P.Q-*c* (IV:1): this process is usually repeated for at *ca.* 20 generations, but for clarity has been reduced to just 2 generations. Alleles in upper-case, e.g. *C*, are not necessarily dominant, nor those in lower-case recessive. Recombination occurs between the strain P and Q chromosomes eliminating the strain Q alleles *a* and *e*, but those strain Q alleles (*b* and *d*) closest to locus *3* remain. The same process can occur when repeatedly backcrossing transgenic animals (originally generated in a hybrid genetic background) with an inbred strain, and selecting the offspring for the transgene.

Because of the necessity to minimise the variables involved in studies of toxicity and carcinogenesis the use of defined animal stocks is required. These may either take the form of inbred strains (e.g. B6 or D2), or first generation (F_1) hybrids obtained by crossing two inbred strains (e.g. BDF_1, from the crossing of B6 × D2). Hybrid strains may be better suited for some purposes, either because of the characteristic "hybrid vigour" resulting from their heterozygosity at many loci, or the combination of other characteristics. However, it may be difficult to determine whether a particular phenomenon, e.g. susceptibility to one class of xenobiotic, is due to a gene or genes from one or other of the parent strains, or their combination in the hybrid. With the use of inbred animals this problem is avoided, resulting in greater confidence in the interpretation of data.

The production of transgenic animals depends to a considerable degree on the manipulation of reproductive physiology. It is well recognised that inbred strains of animals show diminished reproductive efficiency, and that hybrids exhibit increased fertility and fecundity (Brinster *et al.*, 1985). The generation of most transgenic mice has involved hybrid strains, as in many cases their heterozygosity at other than one defined locus is not significant, e.g. the two parent strains only need to be identical at a particular histocompatibility locus. Matings between F_1 hybrid males and females are used to produce F_2 hybrid zygotes for microinjection. This implies a direct conflict if transgenic animals are to be used in experiments in carcinogenesis. A transgenic mouse derived from such a hybrid zygote will have as its genetic background a random mixture of chromosomes from the two parent strains. In addition, chromosomal recombination will have generated novel chromosomes unique to that individual. It may be thought that by repeated backcrossing of the transgenic animals to the strain of interest the desired genetic background can be reestablished. To a large extent this may be true, but by repeated backcrossing and selection for a

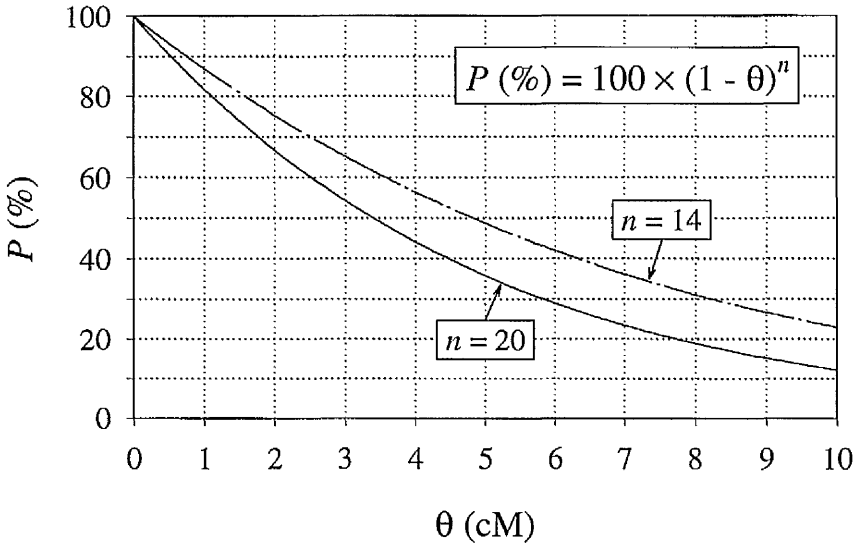
novel locus, in effect a congenic strain is derived. For example, it may be desired to make a transgenic derivative of strain A, but it is necessary to use ABF_2 hybrid zygotes for microinjection. After microinjection the transgene may integrate in one of the strain A chromosomes. Backcrossing this animal to strain A should therefore eventually result in the intended outcome, always assuming that recombination does not occur between the strain A chromosome containing the transgene and the homologous strain B chromosome when it is passed on to the first generation, otherwise strain B alleles of genes will be linked to the transgenic locus. There are two other ways in which this might occur if the transgene a) integrates directly into a strain B chromosome, or b) integrates into a novel chromosome formed by recombination between the homologous strain A and B chromosomes. In any of these events it is most unlikely that all the strain B alleles will be eliminated by repeated backcrossing. The probability (P) that two linked loci will become separated by chromosomal recombination is related to their genetic distance apart (θ), for example two loci 10 cM apart stand a 10% chance per generation that recombination will occur between them, that is a 90% chance per generation that they will *not* become separated. Over short genetic distances P is related to θ and the number of generations (n) by the equation (Terwilliger and Ott, 1994):

$$P = (1-\theta)^n \quad (Eqn.\ 3)$$

The function of P *v.* θ for $n = 20$ is illustrated in Fig. 7.3, which shows that there is a 95% probability of the genes within 0.25 cM of the selected locus still being present, and a 50% probability of the genes within 3.4 cM, a region likely to contain 100 genes.

Fig. 7.3

The probability that after n generations two linked genetic loci will not become separated by chromosomal recombination



These curves give the probability P (%) that two linked genetic loci, separated by a given genetic distance θ (cM), will not become separated by chromosomal recombination after n generations, where n = 14 or 20.

This would not be a problem were it not for the great strain-specificity in susceptibility to tumours, with or without the effect of exogenous xenobiotics. Any experiment in carcinogenesis requires control animals. If the experiment uses an inbred or F_1 hybrid strain then the control animals will be genetically identical. However, if performing an experiment using transgenic animals which are only congenic with the control strain then, although the control and transgenic animals will be matched closely, it cannot be ensured that they do not carry an allele of an endogenous gene having a major influence on the experiment. Ideally, therefore, transgenic animals produced for use in experiments in carcinogenesis should be generated in an inbred strain, if possible one for which data is available on tumour susceptibility. An additional advantage of producing any transgenic line in an inbred

strain is that they can be compared immediately with perfectly matched control animals of the parent strain.

Transgenic founder animals are properly termed hemizygous for the transgene, because the new transgene locus only exists on one of a pair of homologous chromosomes. This locus will be segregated according to the Mendelian rules of inheritance, i.e. half of that animal's offspring will carry the transgene. If these transgenic offspring are mated together, or with the founder animal, then it would be expected that homozygous transgenic, hemizygous transgenic and non-transgenic offspring would result, in the ratio 1:2:1. Thus, it should be possible with this strategy to produce homozygous transgenic animals in two generations. This compares most favourably with a strategy using hybrid zygotes (F_1 or F_2), which would necessitate backcrossing for some 20 generations, followed by breeding to homozygosity. Assuming an 8 week life-cycle for the laboratory mouse this amounts to over three years breeding to produce usable animals, compared theoretically to only four months for the approach using inbred animals. Once a homozygous transgenic strain has been obtained, by either strategy, it could be used to produce hemizygous transgenic hybrid animals.

It has already been stated that most experiments in producing transgenic mice have been carried out with hybrid zygotes (usually F_2), in view of the superior fertility and fecundity of the animals being used to produce the zygotes (Brinster *et al.*, 1985) and, presumably, the hybrid vigour of the transgenic embyos and animals generated as a result. However, a number of transgenic lines had been derived using inbred mice, usually of the strain B6 (Hogan *et al.*, 1986). This is a strain widely available and for which a considerable bank of data and experience in carcinogenesis exists, in addition, the hybrid strain BDF_1 on which similarly much work has been carried out is produced from B6 animals. In particular, BDF_1 mice had been used in the Paterson

Institute for three decades, and a colony of B6 animals, kept for stock breeding purposes, were available. It was therefore decided to produce *fpg*-transgenic B6 mice. Transgenic mice previously produced in the Institute had been generated in BDF_2 zygotes, using either constructs of the *E. coli ogt* ATase gene with the murine mammary tumour virus promoter (Harris, 1990) or, more successfully, the *E. coli ada* or human ATase gene (*hAT*) in the vector pMThGH (Palmiter, 1982; Fan *et al.*, 1990; Fan, 1991). Because of the great propensity of D2 animals to develop thymic lymphosarcoma after treatment with NMU these *hAT*-transgenic animals were backcrossed 14 times to strain D2 to put the *hAT*-transgene on a D2 background (Dr J. Rafferty, personal communication). However, it is of note that in spite of this their coat colour was not that of D2 animals, i.e. dilute, brown, agouti (genetically *dba*), which appears a pale brown, but rather was dilute, black, agouti (*dBa*), which appears a pale grey. The dominant *B* allele at the *Tyrp* (Tyrosine related protein 1) locus which gives black pigmentation must have come from the B6 component of the founder animal's genome, and it would therefore appear that the *hAT*-transgene integrated in the B6 chromosome near enough to the *Tyrp* locus (on mouse chromosome 4; Rinchik, 1994) to still be genetically linked to it after 14 generations (see Fig. 7.3). This illustrates the potential problems associated with attempting to derive congenic transgenic strains.

7.2 TRANSGENIC MOUSE PRODUCTION

Because transgenic mouse production in the Paterson Institute had previously involved hybrids or strains other than B6, it was necessary first to redefine some of variables involved.

7.2.1 Optimisation of superovulation conditions

In order to produce a supply of zygotes for microinjection female mice can either be allowed to come into oestrous naturally, or they can be treated with gonadotrophins to cause superovulation. The latter procedure is considered the more efficient, as a regular supply of zygotes can be ensured, although the fertilisation rate may not be as high as with natural oestrous (Hogan *et al.*, 1986). The gonadotrophins used are a preparation of pregnant mare's serum (PMS), which acts as follicle stimulating hormone (FSH), and human chorionic gonadotrophin (hCG), which acts as luteinising hormone (LH). There are a number of factors involved in achieving satisfactory superovulation and fertile zygote production. These are: strain of mouse, age/weight (which depends on weaning age and diet), timing and length of the light-dark cycle in the animal house, as well as timing, dose and spacing of the gonadotrophin injections. The males used to mate with the superovulating females are also subject to the same factors, with the exception of the gonadotrophin treatment.

The conditions for achieving superovulation of BDF_1 females had already been established as: 10 ± 2 weeks old (22 to 25 g b.w.), 5 IU FSH i/p between 12:00 and 15:00 (on a 12 h light-dark cycle: 07:00 to 19:00), followed by 5 IU LH i/p 46 to 48 h later (2.5.3). Hogan *et al.* (1986) had found superovulation of B6 females to be maximal using 25 day old mice (12.5 to 14 g; 1978), with 5 IU of each hormone. Experiments were therefore carried out to determine the age of B6 females at which maximal superovulation could be achieved, and the results are summarised in Table 7.1.

Table 7.1

The effect of age on superovulation efficiency in C57BL/6J female mice

Strain[a]	Age (weeks)[b]	Total eggs[c]	Eggs/ female[d]	Females ovulating	Eggs/ ovulating female
B6	5[e]	164	16.4	6/10	27.3
B6	6	192	19.2	8/10	24.0
B6	8	124	12.4	9/10	13.8
B6	10	71	7.1	9/10	7.9
B6	12	55	5.5	9/10	6.1
BDF_1	10	492	32.8	15/15	32.8
D2[f]	10	20	20	6/6	20

[a] females were mated with males of the same strain in all cases, [b] mice born during the same calendar week were given the same age, thus their individual ages might vary ±3.5 days, all animals were given 5 IU of FSH between 12:00 and 14:00, with 5 IU of LH 47 h later (2.5.3), [c] this figure includes unfertilised eggs and fertile zygotes, [d] for B6 mice ten animals were used of each age, for BDF_1 15 were used, [e] it was originally intended to test 4 week old mice but they were considered to be too small and so 5 week old animals were used, [f] DBA/2J data with permission from Fan (1991).

The B6 males used in these experiments were animals which had proven themselves able to consistently plug BDF_1 females in natural oestrous (Hogan *et al.*, 1986). It was notable that failure to superovulate only occurred with strain B6, at least one female in each age group failing to superovulate, with the proportion increasing with decreasing age. It can be seen that because of this the overall efficiency of zygote production fell off at the youngest age in B6 mice, in spite of the number of zygotes per superovulating animal being highest in this group. The greater number of zygotes harvested from the hybrid BDF_1 mice, compared to the inbred strains, is also evident. Overall it appeared that 6 week old B6 mice would be the most efficient for producing inbred zygotes.

7.2.2 **The effect of other variables**

In order to minimise the effect of variation in reagents, batches of the various reagents were validated by pilot experiments (2.5.2). Having produced microinjected zygotes it is necessary to implant the resulting embryos into pseudopregnant females in sufficient numbers to achieve pregnancy and a sufficiently large litter. A litter size which is too small, i.e. less than about 4 pups, is likely to result in the mother ignoring the young and allowing them to die (Hogan *et al.*, 1986). Implanting too many embryos is wasteful, as the maximum litter size is regulated by resorption of the excess (Hogan *et al.*, 1986). The strain of females into which embryos are implanted is also important because, especially with first litters, mice from inbred strains are more liable to ignore their young than with mothers of hybrid stock. Hybrid BDF_1 mice were therefore used to provide both the pseudopregnant females, as well as the vasectomised males (2.5.2). Experiments were performed in order to ascertain the relative efficiency with which pregnancies could be established using B6 or BDF_2 embryos in BDF_1 pseudopregnant females (Table 7.2).

Table 7.2

The efficiency of establishing successful pregnancy using C57BL/6J or BDF_2 embryos

Strain[a]	Total eggs	Total zygotes[b]	Total embryos[c]	Embryos implanted/ female[d]	Litters	Pups/ litter[e]
BDF_2 (5)	152	138	126	21 - 24	5/6	14 - 18 (86)
B6 (10)	198	146	121	23 - 26	3/5	13 - 16 (44)

[a] strain is that of the resultant embryos in each case (number of females used in parentheses), [b] number of eggs exhibiting pronuclei, [c] number of viable 2 or 3-cell embryos 24 h after harvesting eggs, [d] BDF_1 females used, [e] range, with the total number of pups in parentheses.

From this it was evident that a smaller proportion of zygotes and embryos were obtained with B6 compared to BDF_2. There was also a lower proportion of successful

pregnancies and pups per litter with strain B6, although the number of embryos implanted per female was slightly greater.

7.2.3 **Microinjection of C57BL/6J zygotes**

Having shown that a sufficient number of zygotes could be obtained from 6 week old B6 females, and that the embryos from them could be successfully implanted into pseudopregnant females, experiments were next carried out to validate the conditions for microinjection. A batch of microinjection buffer (2.5.2, 2.5.4) was made up from selected components and frozen in aliquots, this was used to dilute the highly purified 6.6 kb *Not* I pHIF10 DNA previously prepared for microinjection (5.3.3). A preliminary experiment was carried out to test this microinjection buffer: out of 118 BDF_2 zygotes injected 48 (41%) were immediate survivors, with 42 (36%) developing into 2 or 3-cell embryos by the next day. On implantation into two BDF_1 pseudopregnant females the resultant embryos gave rise to two litters totalling 24 pups. An experiment was next carried out using 165 B6 zygotes, 28 (17%) of which immediately survived microinjection to give rise to 22 (13%) two-cell embryos, which were implanted in to a single BDF_1 female, resulting in a litter of 11 pups. This suggested that B6 zygotes were less able to withstand microinjection. An incidental finding was that, whereas Fan (1991) had found D2 zygotes to have indistinct pronuclei making microinjection difficult, the pronuclei of B6 zygotes were distinct, like those in BDF_1 or BDF_2 zygotes.

The microinjection of zygotes with the 6.6 kb *Not* I pHIF10 fragment could now take place. Work by Brinster *et al.* (1985) had established that optimal integration efficiency was obtained with DNA concentrations of 1 and 2 mg/L, lower concentrations giving reduced integration efficiencies and higher ones being toxic to zygotes. An initial experiment was carried out with 125 BDF_2 embryos being injected

with the DNA at a concentration of 2 mg/L, giving a total of 39 (31%) zygotes that immediately survived the injection. The 32 (26%) survivors that had developed into 2 or 3-cell embryos by the next day were implanted into two BDF_1 females, who delivered a total of 11 pups. When these were old enough (5 weeks), tail biopsies were taken from them and DNA prepared (2.5.6, 2.2.23). Testing by both PCR and Southern analysis after *Bam*H I digestion, as with the transfected RJKO cell clones (5.2.2, 5.4.1 and 5.4.2), failed to show any evidence of integration of the 6.6 kb *Not* I pHIF10 fragment (data not shown). This failure to generate a BDF_2 transgenic animal was ascribed to the small total number of offspring produced. Although the survival figures in this experiment were somewhat less than those in the previous experiment, they were considered satisfactory enough to procede to microinjection of B6 zygotes.

To provide a sufficient number of potentially transgenic B6 offspring, so as to have a reasonable chance of obtaining founder mice, five separate experiments were carried out. In each experiment 15 six week old B6 females were superovulated: from the 65 females thus used 1080 eggs were obtained, 755 of which were fertile zygotes. These were microinjected with the 6.6 kb *Not* I pHIF10 fragment at a concentration of 2 mg/L, and from the 145 immediate survivors (over the five experiments) 120 2-cell embryos were obtained, which were implanted into six pseudopregnant females (Table 7.3).

Table 7.3

Results of implantation of BDF_1 pseudopregnant females with C57BL/6J embryos microinjected with the 6.6 kb *Not* I pHIF10 fragment

Experiment	Female[a]	Embryos implanted/ female[b]	Pregnancy established[c]	Litter	Pups
1	#1	17	Yes	Yes	6
	#2	15	No	(n/a)[d]	(n/a)
2	#3	25	Yes	No	(n/a)
3	#4	20	Yes	Yes	6
4	#5	23	Yes	Yes	8
5	#6	20	Yes	No	(n/a)

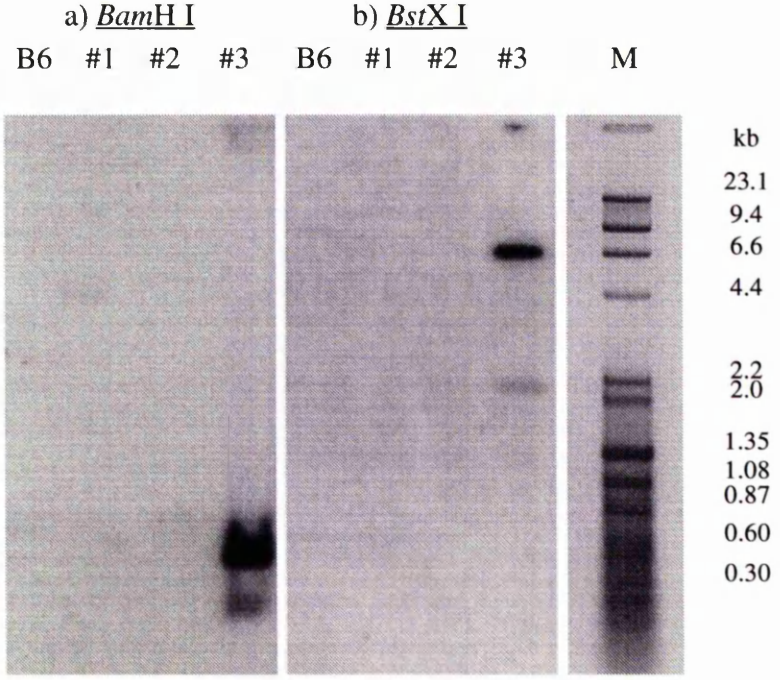
[a] psueodopregnant females were of strain BDF_1, [b] 2-cell B6 embryos from zygotes microinjected with the 6.6 kb *Not* I pHIF10 fragment at a concentration of 2 mg/L, [c] ascertained by abdominal examination 10 to 12 days post-implantation, [d] n/a = not applicable.

It can be seen that in the first experiment two pseudopregnant females had embryos implanted, in one of which pregnancy was not established. In experiments 2 and 5 no litters were produced in spite of successfully established pregnancies being evident by clinical examination at 10 to 12 days post-implantation. It was noted that during the third (final) week of pregnancy the abdominal girth of these animals became progressively reduced to that of an unpregnant female, and it was concluded that resorption of the foetuses had occurred. Females #1, #4 and #5 all littered successfully producing a total of 20 potentially transgenic pups, and the analysis of these is described in the next section.

7.2.4 **Southern analysis of potentially transgenic mice**

Tail-biopsies were obtained from the 20 potentially transgenic pups at 5 weeks of age and DNA extracted (2.5.6, 2.2.23). The DNAs were then digested with *Bam*H I and, separately, with *Bst*X I (2.2.11); the 6.6 kb *Not* I pHIF10 fragment contains three *Bam*H I sites, plus a unique *Bst*X I site located in the centre of the *fpg*10 insert (see Fig. 5.9). Southern analysis was carried out on the digested DNAs, using [^{32}P]-labelled *fpg*4 0.85 kb DNA as the probe (2.2.20): one male give a positive hybridisation signal (Fig. 7.4). The single 6.6 kb band with *Bst*X I was indicative of a multiple head-to-tail array of 6.6 kb *Not* I pHIF10 fragments, the significance of the minor band at *ca.* 2.1 kb is discussed later (7.4.2).

Fig. 7.4
Southern analysis of *fpg*-transgenic founder mouse



A total of 755 B6 zygotes were microinjected with the 6.6 kb *Not* I pHIF10 fragment, the 120 surviving embryos resulted in 20 pups. Tail tip biopsies were taken, and DNA extracted. These DNAs were digested with a) *Bam*H I, and b) *Bst*X I, and then subjected to Southern analysis (2.2.20). Results are only shown for B6 male control DNA ("B6") and pups "#1" to "#3", "M" = [^{35}S]-labelled DNA markers. Mouse #3 (a male) was the only animal to give a positive hybridisation signal.

In view of the effort expended in producing this single founder animal, it was considered too hazardous to take any further biopsies, in particular by invasive procedures requiring general anaesthesia such as partial hepatectomy. It was therefore decided to breed from the founder animal and analyse the offspring. If it proved impossible to obtain transgenic offspring then further analysis of this founder would be warranted.

7.3 FURTHER ANALYSIS OF HEMIZYGOUS *fpg*-TRANSGENIC MICE

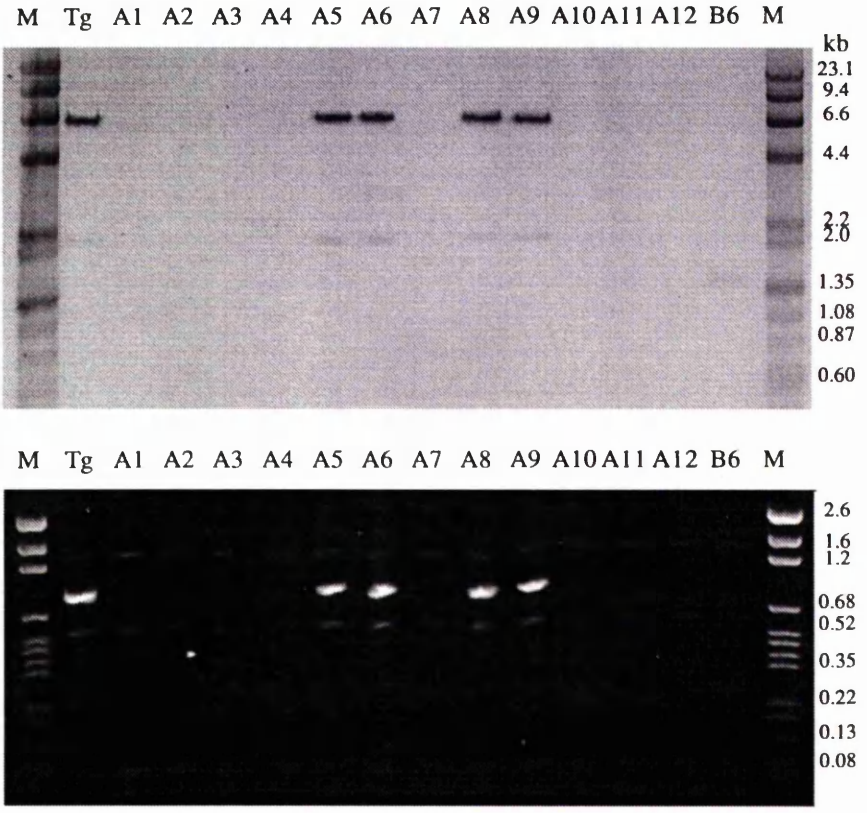
7.3.1 Initial breeding from the *fpg*-transgenic founder mouse

The male transgenic founder mouse was mated with a total of eight 10 week old B6 females, in every case resulting in a pregnancy successfully carried to term. Tail-tip biopsies for DNA purification were obtained from the first two litters at 5 to 6 weeks of age. These were then used to test whether, following on from the experience with the initial analysis of cell clones (5.2.2), analysis by PCR might also be used to rapidly and easily check for the presence of the *fpg*-transgene.

7.3.2 Southern and PCR analysis of hemizygous offspring

Southern analysis of the DNAs obtained from the first 12 offspring of the transgenic founder was carried out after digestion with *Bst*X I. Digestion with *Bst*X I was more consistent than *Bam*H I, being apparently less sensitive to impurities in tail-tip DNAs. Analysis by PCR of the same DNAs was also carried out, and the results compared (Fig. 7.5). This demonstrated concordance between the results of the two methods, just as with the analysis of transfected cell clones.

Fig. 7.5
Southern *v.* PCR analysis of *fpg*-transgenic offspring



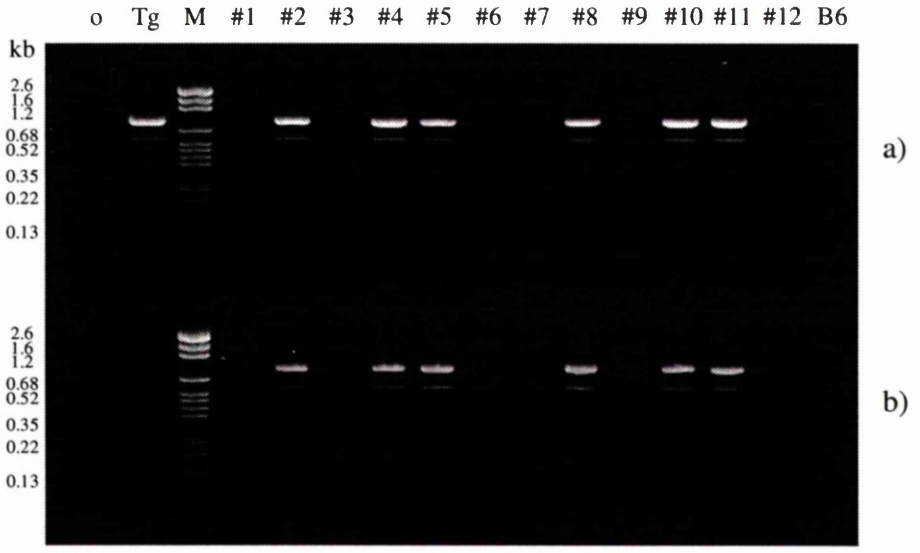
The *fpg*-transgenic founder male was crossed with B6 females and the offspring were tested for the presence of the *fpg*-transgene, by both Southern blotting (upper panel; 2.2.20; *Bst*X I; vacuum blotted) and PCR (lower panel; 2.2.5), using DNA made from large (1 to 2 cm) tail tip biopsies. "Tg" = transgenic founder male, "A1 to A12" = offspring from Tg × B6 female cross, "B6" = male B6 control, "M" = [^{35}S]-labelled DNA markers. The same five samples (Tg, A5, A6, A8 and A9) tested positive by both Southern and PCR analysis.

Analysis by PCR was quicker, taking one working day, compared with the 3 to 5 days necessary to get a result by Southern analysis. However, preparing purified DNA from the large (≈1 cm) tail-tip biopsies, required to provide sufficient material for Southern analysis, also took a significant amount of time, typically 3 or 4 days. In addition, the animals had to be at least 5 weeks old so that they were large enough. It was considered that given the sensitivity of PCR and its less stringent requirement for purified DNA, it should be possible to analyse smaller biopsies with little if any purification needed. Analysis by PCR of crude extracts from blood had been carried

out (Kawasaki, 1990) and it was decided to investigate whether this would be a feasible approach to testing. Younger (smaller) animals could be tested, and sample preparation would be quicker and simpler than for Southern analysis.

Two litters of 4 week old potentially transgenic mice were therefore tested in the following manner. Approximately 20 to 50 μL of blood was collected after removing the extreme tip of the tail (not more than 0.5 mm) with a sterile scalpel (new blade for each subject). As this was being undertaken it was realised that the small piece of tissue removed from the tail tip probably contained sufficient DNA for analysis, and if it could be processed easily for PCR might prove to be a better biopsy than blood. Collecting blood required a few minutes per animal, while collecting just the extreme tail tip would require less than one minute. It was found that preparation of the tail tip biopsies for PCR involved fewer steps than processing blood (2.2.24, 2.2.25). The paired blood and tissue samples were processed and analysed by PCR (Fig. 7.6), which showed that both blood and tissue samples gave the same results, i.e. the same 6 animals tested positive regardless of the nature of the biopsy. In view of the tail tip procedure being simpler, quicker, but apparently equally as efficient as the blood-based test, it was decided to use it for all future screening of potentially transgenic mice.

Fig. 7.6
PCR analysis of *fpg*-transgenic mice:
comparison of blood *v.* small tail tip biopsies

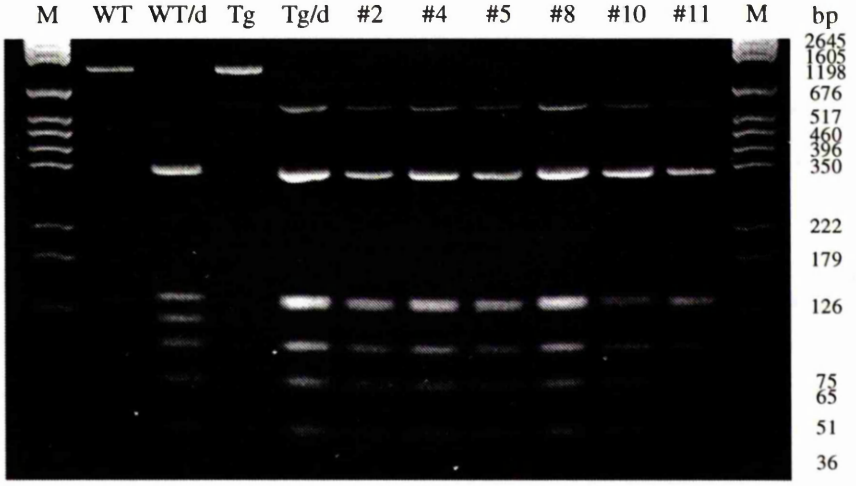


PCR analysis (2.2.5) for the presence of the *fpg*-transgene was carried out on crude extracts of a) small (*ca.* 0.5 mm) tail tip biopsies (upper lanes; 2.2.24), and b) blood (lower lanes; 2.2.25), from 12 potentially transgenic animals. "Tg" = PCR from purified DNA from founder male, "#1 to #12" = samples from mice being tested, "B6" = biopsies from B6 control male, "o" = no DNA control. The same six animals (#2, #4, #5, #8, #10 and #11) tested positive on both blood and small tail tip biopsies.

However, it was considered that because of the smaller amount of DNA in these tail tip biopsies and the sensitivity of PCR there might be a problem with contamination by bacteria, in particular *E. coli*, which the mice were known to harbour as a commensal organism in their gut (Mr T. Bruder, personal communication). This might result in false-positive testing. This was readily testable, as the *fpg*10 insert in the 6.6 kb *Not* I pHIF10 fragment contained a single base change, compared to the wild-type sequence (3.6.1), which generated a *Hha* I RFLP. Thus, the PCR products derived from wild-type *fpg* and *fpg*10 sequence would be distinguishable by restriction analysis with *Hha* I. The PCR products from the small tail tip biopsies were therefore digested with *Hha* I and analysed by AGE (Fig. 7.7).

This showed that all the PCR products were compatible with amplification from *fpg*10 and not wild-type *fpg*, and that the method of testing by PCR from small tail tip biopsies was probably not liable to significant contamination by *E. coli*. It also confirmed that the silent mutation detected in *fpg*10 (C to T at nucleotide 342; section 3.6.1) disrupted the *Hha* I restriction site at that point. It was subsequently found that biopsies could be taken at weaning (i.e. 3 to 4 weeks old), further improving on the age at which transgenic status could be determined.

Fig. 7.7
Hha I restriction analysis of *fpg* PCR products from small tail tip biopsies

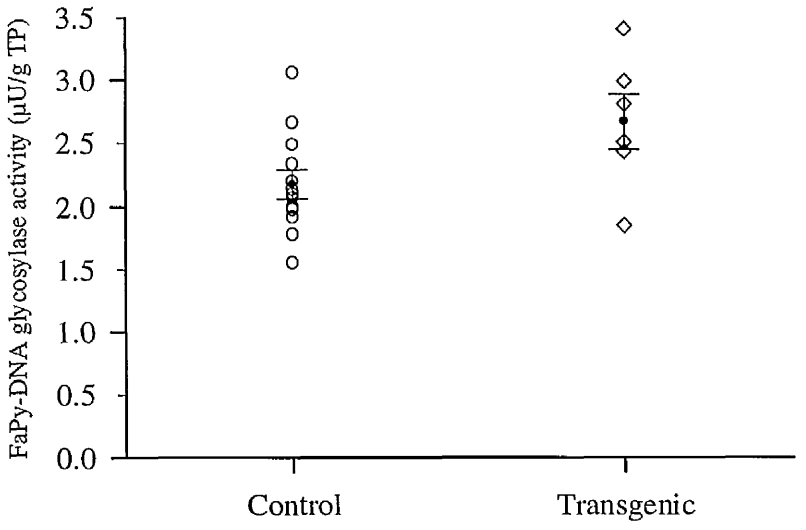


Because of the possibility that contamination of small tail tip biopsies with *E. coli* might cause false-positive PCR results, the PCR products obtained (see Fig. 7.6a) were subjected to restriction analysis with *Hha* I. Digestion of PCR product derived from wild-type *fpg* would be predicted to generate restriction fragments of 329, 135, 116, 98, 75 and 51 bp plus two fragments of 15 bp, whereas digestion of PCR product derived from the recombinant *fpg*10 version would be predicted to produce fragments of 329, 135, 131, 98, 75, 51 and 15 bp, because of the C→T mutation in *fpg*10 at nucleotide 342. "WT" = *fpg* PCR product from *E. coli*, "WT/d" = this product digested with *Hha* I, "Tg" = PCR product from founder male's purified tail tip DNA, "Tg/d" = this product digested with *Hha* I, "#2, #4 etc." = *Hha* I digested PCR products from mice (as indicated in Fig. 7.6a), "M" = pGEM DNA markers, 4% agarose gel. The absence of a band at 116 bp, in addition to a somewhat more intense band at *ca.* 130 bp in the products derived from the tail tip biopsies indicates their origin from *fpg*10, rather than wild-type *E. coli*. The 15 bp products were not seen on this gel, and the weak band at *ca.* 600 bp is a side-reaction product observed in this PCR when carried out using mouse DNA, it would appear to be resistant to digestion with *Hha* I.

7.3.3 FaPy-DNA glycosylase levels in hemizygous mice

Six 8 week old hemizygous transgenic males, and 13 age-matched B6 males had liver biopsies taken by partial hepatectomy (2.5.7). The liver samples were taken in this manner, in preference to being taken *post mortem*, as the animals could then be used to investigate the effects of the HMGCR-inhibitors Pravastatin and Simvastatin, with the advantage of being able to compare paired results for each individual (7.5.1). The biopsies were processed and the FaPy-DNA glycosylase activities measured (Fig. 7.8). The mean of the transgenic group was significantly greater than that of the controls ($p = 0.040$, by t-test), but only by a factor of 1.2-fold. There was also considerable overlap in the spread of individual values in the two groups. However, this did provide preliminary evidence that the *fpg*-transgene was being expressed, albeit at an apparently low level.

Fig. 7.8
Liver FaPy-DNA glycosylase activity in male B6 and *fpg*-transgenic mice



Six 8 week old male hemizygous *fpg*-transgenic mice, and thirteen 8 week old male B6 control mice, had liver biopsies taken, by partial hepatectomy, for FaPy-DNA glycosylase assay. The mean activities of the groups are shown by the solid circles, with error bars indicating SEM. Values for individual animals are shown by open circles (controls) or diamonds (transgenics). The mean of the transgenic group was significantly greater than that of the controls ($p = 0.04$; by t-test).

7.3.4 **Genetic analysis of hemizygous offspring**

The sex and transgene status of the first 89 offspring, from 14 litters, resulting from the mating between the founder *fpg*-transgenic male and B6 females are shown in Table 7.4.

Table 7.4

Offspring of the *fpg*-transgenic founder male crossed with B6 females: sex and transgene status

Status[a]	Males	Females	Total
Non-transgenic	32	20	52
Transgenic	18	19	37
Total	50	39	89

[a] determined by PCR analysis of small tail tip biopsies.

If the transgene had integrated at two or more different (i.e. genetically un-linked) sites the ratio of transgenic to non-transgenic animals would be expected to be 3:1 or greater. If the transgene were present in the X/Y chromosome then it would be transmitted exclusively to females/males (if it were also present at an autosomal locus then a preponderance of transgene-positive females/males would be expected). The value of χ^2 for the data in Table 7.4 (with Yates's correction, as recommended for all 2×2 tables: Armitage and Berry, 1987) is 0.982, which with one D.F. gives a value of P between 0.50 and 0.25, thus showing that there was no significant difference in the number or sex of transgenic or non-transgenic animals. It was therefore concluded that the *fpg*-transgene was present at a single autosomal (i.e. non-sex linked) locus.

7.4 GENERATION AND ANALYSIS OF HOMOZYGOUS *fpg*-TRANSGENIC MICE

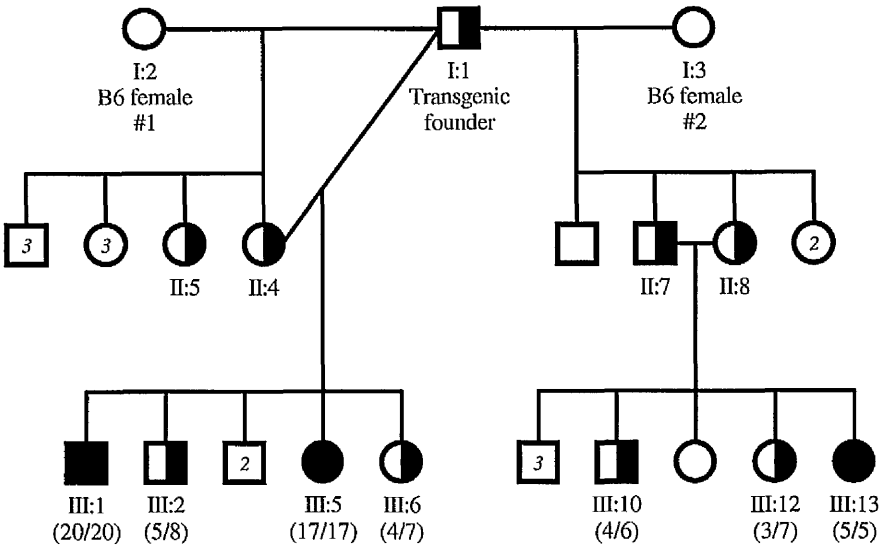
7.4.1 Breeding to generate *fpg*-transgenic homozygous animals

As was discussed earlier (7.1), one of the advantages of producing a transgenic founder animal in an inbred strain is that it should be possible to generate homozygous transgenic animals of that strain in two generations. This assumes that the transgene is transmissible from the founder to its offspring and that the transgene is not present at more than one locus. If it were present at more than one locus, the different loci would need to be separated first by breeding from offspring carrying only a single transgenic locus. It had been demonstrated in the previous section (7.3.4) that the *fpg*-transgene was both transmissible and segregated as a single autosomal locus. A breeding program was therefore undertaken to generate animals homozygous for the *fpg*-transgene. Such mice would provide both an opportunity for investigating the effect of (trans)gene dosage, and also the possibility of crossing them with strain D2 to provide transgenic BDF_1 offspring. In addition, the *fpg*-transgenic strain could be maintained as a fully homozygous inbred line, if required, by repeated brother × sister mating.

Two separate matings were therefore set up, one between the founder male and the first transgenic F_1 female (II:4 in Fig. 7.9), and a second between the first F_1 transgenic male (II:7) and the transgenic female (II:8) from the same litter.

Fig. 7.9

The breeding program used to generate *fpg*-transgenic homozygotes

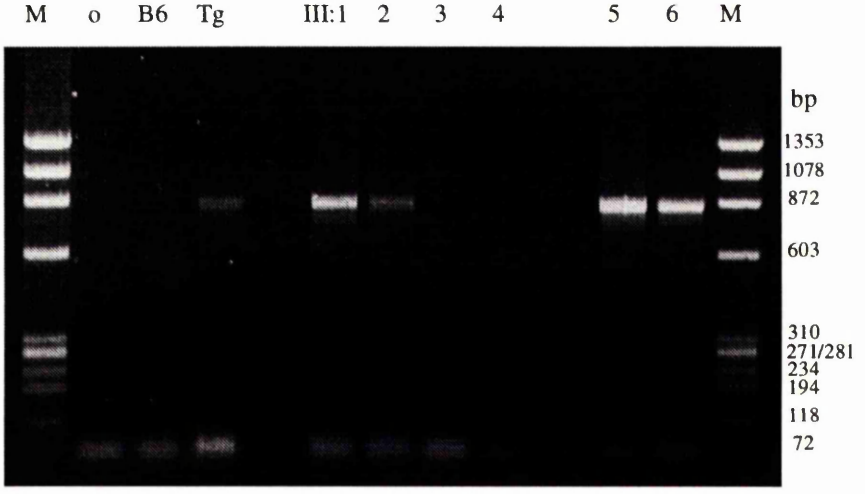


This shows breeding program carried out to generate *fpg*-transgenic animals. The transgenic founder male (I:1) was mated with two B6 females and the resulting offspring (generation II) tested for the presence of the transgene by Southern analysis (see Fig. 7.5). The first (hemizygous) transgenic female (II:4) was backcrossed with the founder to produce a litter of six animals, four of which tested positive for the transgene by PCR (III:1, III:2, III:5 and III:6). Similarly, a brother (II:7) × sister (II:8) mating was set up and the five offspring from that tested by PCR, three proving positive (III:10, III:12 and III:13). These transgenic mice in generation III were then mated with B6 animals (not shown) and the offspring tested for the presence of the transgene by PCR (the number of transgenic offspring compared with the number tested is given in parentheses below each individual, e.g. 5 positive out of 8 tested for III:2). Any animals that gave non-transgenic offspring (i.e. III:2, III:6, III:10 and III:12) were rejected as being hemizygous, and the remainder were mated again with B6 animals, but female III:13 was not able to produce a second litter. Three mice (III:1, III:5 and III:13) were thus found that gave only transgenic offspring. The probability that III:1 and III:5 were not homozygous was extremely low (see text for details).

These matings resulted in two litters, comprising 8 males and 5 females. Two males (III:1 and III:2) and two females (III:5 and III:6) from the first litter tested positive for the *fpg*-transgene by the PCR assay using small tail tip biopsies (Fig. 7.10), plus one male (III:10) and two females (III:12 and III:13) from the second litter. It was of note that the concentration of *fpg* PCR product generated from three of the animals in the first litter (III:1, III:5 and III:6) appeared greater than that from a known hemizygous animal (used as the positive control) and the other transgenic littermate (III:2).

Although there are a great many confounding variables determining product yield in PCR, and it had not been established that the yield from the PCR used was quantitative, it was tempting to speculate that these animals giving stronger bands were homozygous for the *fpg*-transgene.

Fig. 7.10
PCR analysis of offspring from mating between transgenic founder and first F_1 transgenic female



PCR analysis of the litter resulting from the mating between the transgenic founder male and the first F_1 hemizygous transgenic female, using small tail tip biopsies, (see Fig. 7.9). "o" = no DNA control, "B6" = B6 control mouse, "Tg" = sample from hemizygous transgenic mouse, "1 to 6" = individuals III:1 to III:6 as in Fig. 7.9, "M" = DNA markers.

Formal testing of these mice for homozygosity was carried out by crossing them with non-transgenic animals: any giving non-transgenic offspring would be assumed to be hemizygous, those giving universally transgenic offspring would likely be homozygous. Separate matings were therefore set up with non-transgenic B6 animals, and the offspring tested for the transgene by the PCR assay. (This was also a convenient way to produce hemizygous offspring for other experiments.) The number of offspring, from each animal, testing positive compared with the number tested (e.g. 5/8 for male III:2) are shown in Fig. 7.9. Four mice (III:2, III:6, III:10 and III:12)

each gave rise to first litters containing non-transgenic pups: thus they could be eliminated as potential homozygotes. The other three (III:1, III:5 and III:13) all gave rise to litters of exclusively transgenic pups and were re-mated with B6 animals. The female III:13 failed to become pregnant, in spite of being caged with a number of B6 males, all of which had successfully fathered litters with other females. The male III:1 and female III:5 both gave rise to second and third litters all testing positive. From the total number of offspring tested (n) the probability (p) of the transgenic parent being hemizygous can be calculated from equation 4:

$$p = \frac{1}{2^n} \quad (Eqn.\ 4)$$

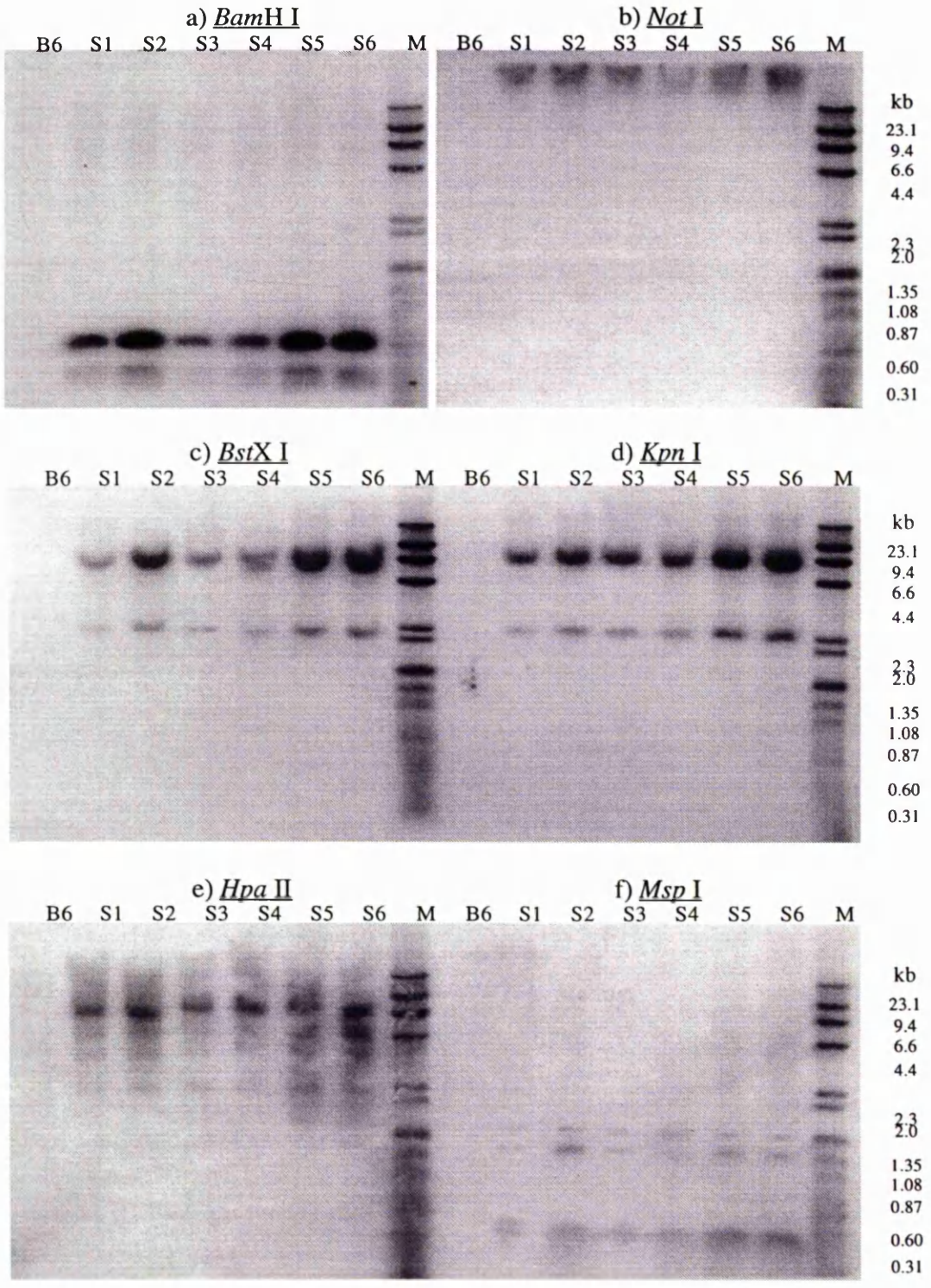
For the female III:13 $p = 0.031$ (i.e. 1/32), which was reasonable evidence that this animal was homozygous, but as she failed to breed again she was, unfortunately, of no further use. However, for the male III:1 $p = 0.95 \times 10^{-6}$ and the female III:5 $p = 7.6 \times 10^{-6}$, figures which could confidently be taken to indicate that both animals were homozygous. Therefore III:1 and III:5 were mated to produce homozygous offspring for further analysis (7.4.3 and 7.4.4).

7.4.2 Southern analysis of homozygotes

As well as being used to breed homozygotes, the male (III:1) was also mated with his hemizygous sibling (III:6), to produce offspring which would be predicted to consist of homo- and hemizygotes in equal proportion. A litter was produced of three males (S1 - 3) and three females (S4 to 6), which were sacrificed at 5 weeks of age and DNA purified from the livers. After digestion of these DNAs with a number of restriction endonucleases the products were subjected to Southern analysis (Fig. 7.11).

Fig. 7.11

Southern analysis of offspring from mating between a homozygous *fpg*-transgenic male and hemizygous female



Three male mice ("S1 to S3") and three female mice ("S4 to S6") were produced from a mating between the homozygous *fpg*-transgenic male (III:1) and his hemizygous sibling (female III:6). Liver DNA was digested with: a) *Bam*H I, b) *Not* I, c) *Bst*X I, d) *Kpn* I, e) *Hpa* II, and f) *Msp* I, and subjected to Southern analysis (2.2.20). "B6" = male B6 DNA, "M" = [^{35}S]-labelled DNA markers.

This showed that all transgenic animals gave the predicted pattern of hybridisation bands after digestion with *Bam*H I (0.55 and 0.30 kb; Fig. 7.11a). A single, diffuse, band (≫ 23.1 kb) was seen with *Not* I (Fig. 7.11b), while both of the single-cutting enzymes, *Bst*X I and *Kpn* I (see Fig. 5.9), gave a major band at 6.6 kb (Fig. 7.11c and d). This showed that multiple copies of the transgene had integrated in a head-to-tail array, but with disruption of the *Not* I sites at the ends of the copies. The significance of the minor bands seen at 2.1 kb with *Bst*X I and *ca.* 2.5 kb with *Kpn* I lies in the possible arrangement of adjacent restriction sites in the mouse chromosome (to generate "junction fragments", as termed by Hogan *et al.* (1986)) and/or the presence of aberrant copies of the transgene in the array. The *Bst*X I site in the 6.6 kb *Not* I pHIF10 fragment lies 6.0 kb from the 5′ end within the *fpg* insert, whereas the *Kpn* I site is 17 bp 3′ of the *fpg* insert, i.e. 6.5 kb from the 5′ end of the same fragment (Fig. 5.9). Thus, if there were a site in the chromosome with *Bst*X I and *Kpn* I sites juxtaposed, located 5′ of the transgene array, this would have to be at least 6.5 kb from the *Kpn* I site in the most 5′ copy of the transgene array. Minor bands larger than 6.6 kb with both enzymes would be predicted from this. As *Bst*X I cuts within the *fpg* insert, 0.6 kb from the 3′ end of the 6.6 kb *Not* I pHIF10 fragment, then it would also be predicted that another minor band at least 0.6 kb in size should be seen, in addition to that from the 5′ end. However, because the *Kpn* I site is 3′ of the *fpg* insert only one minor band would be revealed by probing with *fpg*, i.e. that corresponding to the junction fragment from the 5′ end of the array. One explanation for the pattern of minor bands seen is that at least one copy of the transgene within the array was truncated, such that approximately 4.3 kb was missing from the 5′ end. If two copies in the array were present in a head-to-head orientation then a novel fragment *ca.* 12 kb in length would be predicted with *Bst*X I (with a corresponding

band at *ca.* 13 kb with *Kpn* I), similarly, if two copies were oriented tail-to-tail this would predict a novel band at *ca.* 1 kb with *Bst*X I, without a corresponding *Kpn* I band. This doubt over the precise origin of the minor bands from junction fragments precluded their use to directly assess the number of copies of the transgene, within the array, from the ratio of the signal strengths of the major *v.* minor bands. Finally, the pattern seen with *Hpa* II (5mC-methylation sensitive) consisted of a major band at *ca.* 6.6 kb plus minor bands at 4.4, 3.4 and 2.2 kb (Fig. 7.11e). This contrasted with the pattern seen with *Msp* I, the 5mC-insensitive isoschizomer of *Hpa* II, which was of bands at 1.4, 1.1 and, indistinctly, *ca.* 0.3 kb (Fig. 7.11f), showing that the *Msp* I/*Hpa* II sites in the transgene array were methylated. The significance of this is discussed later (7.5.2 and 7.6.2).

With those enzymes which gave discrete and consistent bands (i.e. *Bam*H I, *Bst*X I and *Kpn* I), the strength of the hybridisation signals for animals S2, S5 and S6 appeared greater than for S1, S3 and S4. It had been intended to scan the autoradiographs densitometrically to quantitate the signals from the individual mice, because if the signals from the mice were bimodally distributed, with some giving twice the signal of the others, this would provide extra evidence that transgene homozygotes had been generated. At the time this experiment was carried out a phosphorimager became available in the Institute, which gave the opportunity to directly and accurately quantitate the signal strengths, rather than use densitometry (which is liable to errors due to the inherent non-linearity of response of X-ray film to radiation, and ideally requires calibration (Sambrook *et al.*, 1989)). Absolute measurement of the hybridisation signals by phosphorimaging was therefore carried out for the bands given by *Bam*H I, *Bst*X I and *Kpn* I, and the results are given below (Table 7.5).

Table 7.5

Quantitation of the hybridisation signals seen on Southern analysis of potentially homozygous *fpg*-transgenic mice

Enzyme	Total radioactivity (Counts × 10^{-3})[a]							Ratio (S2+S5+S6:S1+S3+S4)
	C57	S1	S2	S3	S4	S5	S6	
*Bam*H I[b]	3.8	91	145	67	84	172	153	1.94
*Bst*X I[c]	6.2	49	108	62	80	133	153	2.06
Kpn I[d]	7.4	63	88	73	77	135	169	1.84
All	17.4	203	358	202	241	440	475	1.94

[a] as determined by analysis of the hybridisation membranes used in Southern analysis (see Fig. 7.11) by phosphorimager; individual test counts (for S1 - S6) are given as corrected by subtraction of counts in the C57 control lane, [b] 0.55 and 0.30 kb bands quantitated, [c] 6.6 and 2.1 kb bands quantitated, [d] 6.6 and 2.5 kb bands quantitated.

Applying the Poisson heterogeneity test (Armitage and Berry, 1987) to the results with each enzyme gave χ^2 values of 78.8 (for *Bam*H I), 85.8 (for *Bst*X I) and 86.8 (for *Kpn* I): with 5 D.F. this gives $P < 0.001$ in all cases, showing that the variability was in excess of that expected if the individual counts were unimodally distributed. Therefore, it appeared justified to assume a bimodal distribution in total radioactivity for animals S2, S5 and S6 (Group 1) *v.* S1, S3 and S4 (Group 2). The measurement of the individual counts was prone to a number of experimental inaccuracies, for example in quantitation of DNA concentration, and the variation in efficiency of transfer of DNA to the hybridisation membrane, both between lanes and with DNAs of different sizes. In view of this, the ratios for the individual enzymes of between 1.84 and 2.06 (and the ratio for all together of 1.94) were remarkably close to the predicted ratio of 2.0, and provided independent evidence that mice homozygous for the *fpg*-transgene had been generated. Taken together with the evidence from breeding (7.4.1) this was considered conclusive proof that homozygotes had been generated.

7.4.3 FaPy-DNA glycosylase levels in hemi- and homozygotes

The litter resulting from the mating between the two homozygous animals, III:1 and III:5, consisted of four males and four females. These were sacrificed at 5 weeks of age, so that the liver and kidney samples taken were matched for age with the samples taken from the hemi-/homozygous litter described in the previous section (7.4.2). In addition, three male and three female 5 week old hemizygous animals were sacrificed, together with six B6 control mice of each sex. The FaPy-DNA glycosylase activity was assayed in the extracts made from these samples and the results are shown in Table 7.6.

Table 7.6

Liver and kidney FaPy-DNA glycosylase activity in hemi- and homozygous *fpg*-transgenic mice

Group	Sex	*n*	Liver		Kidney
B6 Controls[a]	♂	6	2.37 (0.064)[b]	2.33 (0.114)	3.11 (0.089)
	♀	6	2.28 (0.215)		4.52 (0.121)
Tg-Hemizygous	♂	6	2.38 (0.112)	2.56 (0.143)	2.93 (0.142)
	♀	6	2.65 (0.184)		4.58 (0.105)
Tg-Homozygous	♂	6	2.48 (0.106)	2.36 (0.114)	3.12 (0.096)
	♀	5	2.24 (0.122)		4.20 (0.174)

[a] all animals 5 weeks old, [b] mean (SEM) FaPy-DNA glycosylase activities given in μU/g TP.

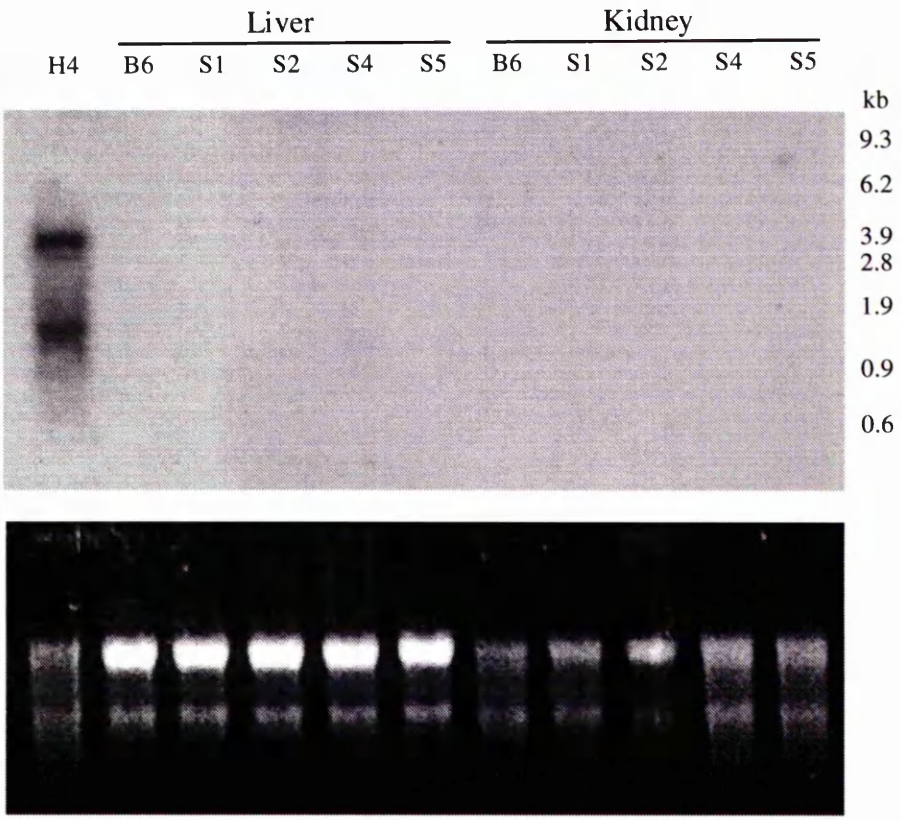
Considering first the levels seen in liver: there were no significant differences (by *t*-test) between the males and females within each group, and therefore the mean levels for both sexes combined are also given. There were also no significant differences between the groups, either for animals of the same sex, or for both sexes combined. The levels seen in kidney were significantly greater than those in liver ($p \leq 0.01$ in all cases, by *t*-test), and also showed a marked difference between the sexes in all groups,

being 1.4 to 1.6-fold greater in females ($p < 0.001$ in all cases; *t*-test). However, as with liver, there were no significant differences between the groups for animals of the same sex. It appeared from this that the transgene did not confer enhanced FaPy-DNA glycosylase activity on those animals carrying it. In particular, increased levels were not seen in the homozygous mice, implying that transgene dosage had no effect. This contrasted with the results obtained for liver from 8 week old male transgene hemizygotes, as described in section 7.3.3 (Fig. 7.8), where a small but significant difference in FaPy-DNA glycosylase levels had been seen: this is discussed later. However, the inability to show higher activity did not necessarily imply failure of the transgene to be transcribed or, if mRNA was produced, a failure to be expressed as protein (e.g. if the protein were being turned over at a high rate). Other analyses were therefore carried out, as described in the next section.

7.4.4 **Northern and western analysis**

Total RNA was purified from portions of the liver and kidney samples from one B6 control male, and four transgenic animals: one hemizygous and one homozygous of each sex (S1, S2, S4 and S5; 7.4.3). The RNA samples, plus one from the RJKO cell clone 10H4 (5.4.7.2), were then subjected to northern analysis as shown in Fig. 7.12. Only the sample from the 10H4 cells gave a positive signal (as previously, see Fig. 5.19), with none of the mouse RNAs showing any evidence of hybridisation with the 0.85 kb *fpg*4 probe.

Fig. 7.12
Northern analysis of *fpg*-transgenic mice

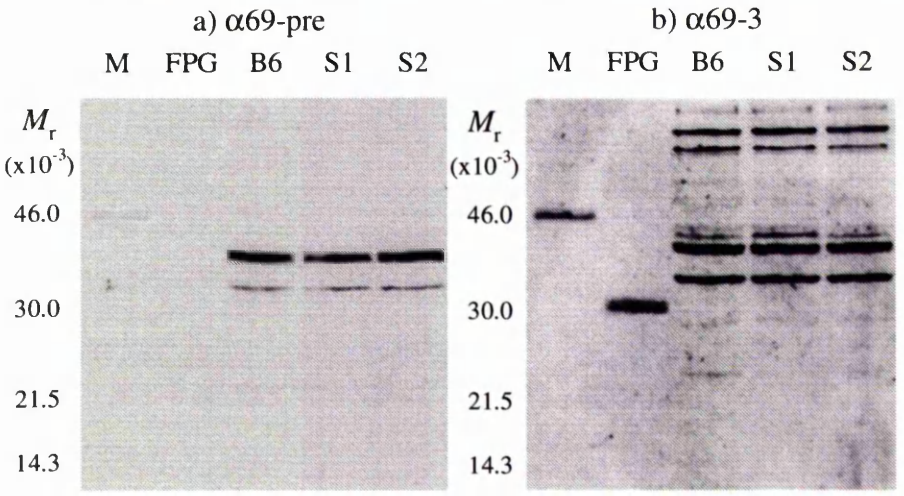


Upper panel: RNA was purified from liver and kidney samples obtained at *post mortem* from one hemizygous and one homozygous *fpg*-transgenic mouse of each sex, and subjected to northern analysis (2.2.22; vacuum blotted). "H4" = RNA from the cell line 10H4, "S1" = hemizygous male, "S2" = homozygous male, "S4" = hemizygous female, "S5" = homozygous female. The position of MB-stained RNA size markers is indicated. The lower panel shows the UV-transilluminated denaturing agarose gel prior to vacuum blotting.

Western analysis was also carried out, on the extracts made for assay of liver FaPy-DNA glycosylase activity from the same mice as were used for northern analysis. It was decided to probe only with the antiserum α69-3, as the other (α67-4) had shown itself to be relatively insensitive at detecting FPG protein in the western analysis of the RJKO cell clones (5.4.7.3) and mammalian cell extracts (6.4.2). The results of this are shown in Fig. 7.13. When probed with the preimmune serum the same pair of bands was seen in all extracts, at $M_r \approx 35\,000$ and 39 000, as had been noted before in B6 liver extract (6.4.2). With the antiserum α69-3 the same pattern of bands was

seen in all extracts, again as had been noted before in B6 liver extract (6.4.2 and Table 6.4), in addition some very weak bands were seen at $M_r \approx 51\,000$, 47 000, 28 000 and 25 000. However, no bands were seen at $M_r = 32\,000$ in the transgenic extracts corresponding to FPG protein.

Fig. 7.13
Western analysis of liver extracts from *fpg*-transgenic mice



Extracts from liver samples (30 µg TP), as used for FaPy-DNA glycosylase assay, were subjected to western analysis (2.3.14). Blots were probed with, a) pre-immune (α69-pre), or b) immune serum (α69-3). "M" = coloured dye-labelled marker proteins, "FPG" = purified FPG protein (5 ng), "B6" = B6 male, "S1" = hemizygous *fpg*-transgenic male, "S2" = homozygous *fpg*-transgenic male.

When considered in conjunction with the data on FaPy-DNA glycosylase activity, these results from northern and western analysis indicated that the *fpg*-transgene was inactive; all the assays for an *fpg* gene product of any sort had proven negative. Southern analysis had also shown that the transgene was present as a multi-copy array with extensive cytosine methylation, a phenomenon associated with down-regulation of genes (Cedar *et al.*, 1982; Stein *et al.*, 1983; Costello *et al.*, 1994). Efforts were therefore taken to achieve expression of FPG protein by treatment with HMGCR inhibitors (to increase/induce transcription) or 5-azacytidine (to decrease cytosine methylation), and these are described in the next section.

7.5 ATTEMPTS TO ACHIEVE EXPRESSION OF FPG PROTEIN IN *fpg*-TRANSGENIC MICE

7.5.1 Treatment with HMGCR inhibitors

It has been described (7.3.3) that six 8 week old hemizygous transgenic male mice, plus 13 age-matched B6 control animals, underwent partial hepatectomy to obtain liver biopsies. This provided preliminary evidence that a modest, but statistically significant, increase in FaPy-DNA glycosylase had been obtained (Fig. 7.8), implying that the *fpg*-transgene was expressed, at least at a low level. These animals could now be used in a pilot experiment to determine whether the HMGCR inhibitors, Pravastatin (PV) and Simvastatin (SV), might be capable of causing up-regulation of the HMG-*fpg* transgene (5.4.8.1). Treatment of clone 10H4 cells with PV, SV and SVA (the active metabolite of SV) had given variable and inconsistent results in their effects on FaPy-DNA glycosylase levels (5.4.8.2), possibly related to the HMGCR sequence in the pHIF10 construct being derived from the mouse gene whilst the parent cells (RJKO) originated from the Chinese hamster. Neither could it be certain that effective conversion of SV to SVA had been achieved by the treatment with rat serum *in vitro*. The use of transgenic mice, however, would avoid both of these problems.

Previous studies, in which mice were treated with the related compound Lovastatin (LV), had used doses of 33 to 40 mg/Kg b.w., given by twice daily injection for 3 days (40 mg LV = 0.1 mmol). Increases of 6 to 10-fold had been observed in endogenous hepatic HMGCR activity (Henwood and Heel, 1988; Kita *et al.*, 1980; Liscum *et al.*, 1983). It was therefore decided to give both PV and SV to mice by twice daily injection at a dosage of 0.1 mmol/Kg b.w. (2.1.11). Pravastatin could be given dissolved in saline (2 μmol/100 μL), but SV being lipophilic was made up in vegetable oil (2.1.11) so that the required dose was delivered in a 20 μL volume, using a calibrated automatic delivery device. The partially hepatectomised animals

(7.3.3) were allowed two weeks convalescence to permit liver regeneration to occur and otherwise fully recover from the effects of the operation. They were then treated for 4 days with PV, or SV, or sham treated with vehicle alone. Two hours after the final injection the animals were sacrificed and liver samples taken for FaPy-DNA glycosylase assay, the results of which are shown in Table 7.7.

Table 7.7

Liver FaPy-DNA glycosylase activity in hemizygous *fpg*-transgenic male mice treated with HMGCR inhibitors

Animals	Group	Treatment[a]	*n*	Pre-treatment[b]	Post-treatment[c]	Change (%)
B6	Control	Saline	2	2.53 (0.53)	2.57 (0.23)	+1.6
		Oil	2	1.85 (0.29)	1.97 (0.12)	+6.5
		Oil or saline	**4**	**2.19 (0.32)**	**2.27 (0.20)**	**+3.7**
	Test	PV[d]	4	2.25 (0.16)	2.15 (0.05)	-4.4
		SV[e]	3	2.19 (0.15)	2.17 (0.09)	-0.9
		PV or SV	**7**	**2.23 (0.10)**	**2.16 (0.04)**	**-3.1**
Transgenic	**Control**	**Oil**	**2**	**2.34 (0.48)**	**2.24 (0.05)**	**-4.3**
	Test	PV	2	2.76 (0.24)	1.86 (0.08)*	-32.6
		SV	2	2.93 (0.49)	1.96 (0.30)	-33.1
		PV or SV	**4**	**2.85 (0.20)**	**1.91 (0.11)****	**-33.0**

[a] commencing 2 weeks after partial hepatectomy, [b] FaPy-DNA glycosylase activities in liver biopsies obtained by partial hepatectomy, given in μU/g TP, [c] liver samples obtained *post mortem*, [d] 0.1 mmol Pravastatin (PV)/Kg b.w. in 100 μL saline, given by i/p injection b.d. × 4 days, [e] 0.1 mmol Simvastatin (SV)/Kg b.w., in 20 μL vegetable oil, by i/p injection b.d. × 4 days, * $p = 0.07$ by *t*-test compared with pre-treatment level, ** $p = 0.006$ by *t*-test compared with pre-treatment level.

There were no significant differences (by *t*-test) between pre-treatment FaPy-DNA glycosylase levels in the B6 sub-groups, i.e. B6 controls treated with saline *v.* oil, or B6 tests treated with PV *v.* SV. Neither was there a significant difference between pre- and post-treatment levels in any of these sub-groups. The B6 control data ~~was~~ were therefore combined in a single "Oil or saline" group, and similarly the B6 test data

were
~~was~~ combined in a single "PV or SV" group: there was no significant difference between the pre- and post-treatment levels in these combined groups. This was taken to indicate that neither the procedure, nor the HMGCR inhibitors, had a significant effect on the levels of endogenous hepatic FaPy-DNA glycosylase levels in B6 mice. Because of the small number of animals available the transgenic control group were treated with oil, as it was considered that this agent was more likely to have a significant physiological effect than saline: there was no significant difference between the pre- and post-treatment glycosylase levels in the transgenic control group. Surprisingly, however, the levels seen post-treatment in both of the transgenic test sub-groups (PV and SV) were lower than pre-treatment values. When formally checked with the *t*-test the difference was not significant in the SV-treated group ($p = 0.23$), but did approach significance in the group treated with PV ($p = 0.071$). There were no significant differences between the pre-treatment, or post-treatment, transgenic test sub-groups, and so the data were combined in a single transgenic test (PV or SV) group. When the pre- and post-treatment values for this combined group were compared by *t*-test there was a highly significant reduction in the post-treatment level ($p = 0.006$). In view of the apparently paradoxical result of this pilot experiment it was decided to repeat it, but with a larger number of animals and using PV alone to maximise the group sizes. The treatment with SV had not been well tolerated by the animals, and at *post mortem* examination those animals which had received oil injections (whether containing SV or not) were found to have free fluid in the peritoneum, chylous in nature. This indicated that probably not all of the material injected had been absorbed and/or that an inflammatory peritoneal reaction had been provoked.

A total of 16 hemizygous transgenic males (8 to 9 weeks old; the product of the test matings to prove homozygosity - see 7.4.1) and 8 age-matched B6 controls had

partial hepatectomies performed, to provide liver biopsies for FaPy-DNA glycosylase assay and, if required, RNA extraction. The animals were given 2 weeks to recover, and were then treated, as before, with twice daily i/p injections of 0.1 mmol PV/Kg b.w. in 100 μL saline for 4 days. Liver samples were taken *post mortem* for glycosylase assay and RNA extraction. The results of the FaPy-DNA glycosylase assays are shown in Table 7.8.

Table 7.8

Liver FaPy-DNA glycosylase activity in hemizygous *fpg*-transgenic male mice treated with Pravastatin

Animals[a]	Group	Treatment[b]	*n*	Pre-treatment[c]	Post-treatment[d]	Change (%)
B6	**Control**	Saline	4	2.49 (0.22)	2.32 (0.11)	-6.8
	Test	PV[e]	4	2.12 (0.07)	2.44 (0.18)	+15.1
Transgenic	**Control**	Saline	8	2.33 (0.11)	2.59 (0.06)*	+11.2
	Test	PV	8	2.13 (0.09)	2.37 (0.12)	+11.3

[a] 8 to 9 weeks old, [b] commencing 2 weeks after partial hepatectomy, [c] FaPy-DNA glycosylase activities in liver biopsies obtained by partial hepatectomy, given in μU/g TP, [d] liver samples obtained *post mortem*, [e] 0.1 mmol Pravastatin (PV)/Kg b.w. in 100 μL saline, given by i/p injection b.d. × 4 days, * $p = 0.025$ by paired *t*-test compared with pre-treatment levels.

It was of note that the pre-treatment mean (SEM) for all of the transgenic animals was 2.23 (0.104) μU/g TP ($n = 16$), a level which was not significantly different from that of the B6 group (2.31 (0.128) μU/g TP; $n = 8$). The B6 control group showed a slight drop in FaPy-DNA glycosylase level, whereas the B6 test group showed a rise of just over 15%, however these changes were not significant when the differences for the individual animals were compared using the paired *t*-test. Both of the transgenic groups appeared to show almost identical rises in activity of just over 11%. However, when the pre- and post-treatment levels amongst the individual animals in each group were compared, using the paired *t*-test, the transgenic test group did not show a

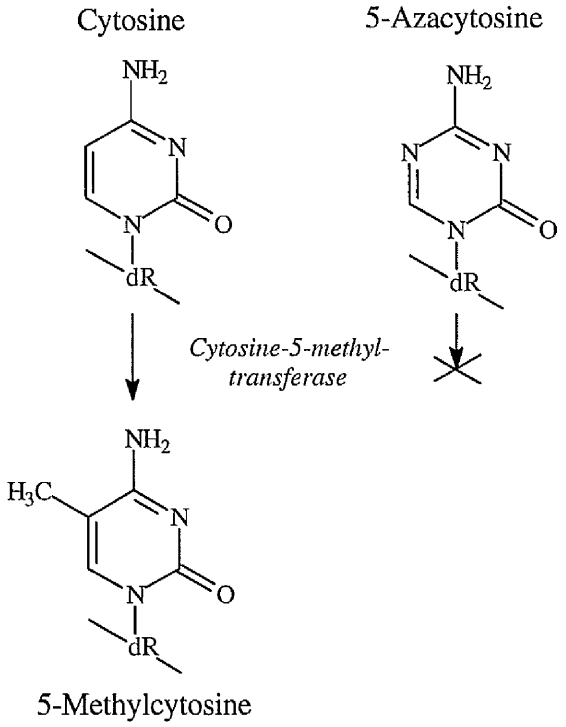
significant difference ($p = 0.12$), but the control group did ($p = 0.025$). These data demonstrated that the hepatic FaPy-DNA glycosylase activity in hemizygous male transgenic mice was no different from that in age-matched B6 animals. It further showed that treatment with PV did not change the level of activity in transgenic animals, either up or down, suggesting that the result of the pilot experiment in which a statistically significant drop had been observed was fortuitous. The modest, but significant, rise in activity seen in the transgenic control animals was baffling and is discussed later.

7.5.2 **Treatment with 5-azacytidine following partial hepatectomy**

Southern analysis of both hemi- and homozygous *fpg*-transgenic animals after digestion of genomic DNA with *Hpa* II and *Msp* I had demonstrated that the multi-copy transgene array was extensively cytosine-methylated. This phenomenon has been associated with modulation of gene expression, with generally an inverse correlation between the levels of 5mC and the level of gene expression (Cedar *et al.*, 1982; Stein *et al.*, 1983; Costello *et al.*, 1994). It had also been observed in *ada*-transgenic mice previously produced in the Department which exhibited low or undetectable levels of expression of the transgene (Fan, 1991; Dr G.P. Margison, personal commununication). Methylation of cytosine residues in vertebrate genomic DNA takes place shortly after cell division, with some 2 to 7% of cytosine residues being converted, mostly in the sequence context 5′...CG...3′ (Bird, 1986). Treatment with 5-azacytidine (5azaC) has been used in a variety of systems both *in vitro* and *in vivo* to subvert cytosine 5-methylation, and thus alter gene expression (Detke, 1984; Kroger *et al.*, 1984; Wagner *et al.*, 1988; Constantoulakis *et al.*, 1991). This compound is a synthetic analogue of cytidine which can be incorporated into DNA, but cannot then be methylated at the 5 position of the cytosine moiety (Fig. 7.14). There is also evidence that 5azaC in DNA acts by irreversibly combining with the active site of cytosine-5-methyl transferase (Friedman, 1985). But by whatever mechanism it does act, sufficient 5azaC must be incorporated into DNA, which is more likely to be achieved when rapidly dividing cells are treated, e.g. bone marrow. However for cells or tissues with a lower rate of turnover, e.g. liver, a different strategy must be employed. Stimulation of hepatic cell division is readily obtained by performing a partial hepatectomy. This proliferative response to surgical resection has been known since the early part of this century and has been extensively studied (Wright, 1984). In mouse liver, mitotic activity commences 30 to 36 h after partial hepatectomy, with

peak values occurring a few hours later. Values of between 50 and 70% in the proportion of cells in *S* phase (as determined by [^{3}H]-TdR labelling), and 3 to 7% in mitosis have been described, thus a prolonged effect on proliferation is evident: *ca.* 1% of cells may still be mitotically active up to 14 days post-operatively (Wright, 1984).

Fig. 7.14
The use of 5-azacytidine to prevent cytosine 5-methylation



Some cytosine residues in eukaryotic DNA are methylated at the 5 position by a specific enzyme, cytosine-5-methyltransferase. This phenomenon is considered to be related to the control of gene expression. The synthetic compound 5-azacytidine is, like 2′-deoxycytidine, incorporated into DNA but, unlike its natural analogue, 5-azacytosine does not allow methylation at the 5 position. It may also act by irreversibly combining with cytosine-5-methyl transferase. Thus it can be used to alter patterns of gene methylation if given during DNA replication.

If treatment of animals with 5azaC is carried out during this period of hepatic cell proliferation which follows partial hepatectomy, then a sufficient proportion of 5–azacytosine residues may become incorporated in the new DNA such that cytosine 5–methylation is prevented. If a gene has previously been down-regulated by such

methylation then after such treatment it may become active (Wagner *et al.*, 1988; Constantoulakis *et al.*, 1991). An experiment was therefore carried out to assess the effect of 5azaC treatment as a means of hypomethylating the *fpg*-transgene and thereby stimulating expression of FPG message and hence protein. Previous studies on the effect of 5azaC in altering gene expression in mice had used either a single i/p injection of 25 mg/Kg b.w. (Wagner *et al.*, 1988), an i/p injection of 4 mg/Kg b.w. followed 12 h later by another of 12 mg/Kg b.w. or a regime of 10 mg/Kg b.w./day × 4 days (Constantoulakis *et al.*, 1991). This latter study had shown treatment with 5azaC to induce expression in adult mice of a transgene (human γ-globin) normally only expressed during foetal life. It was decided to cover the period of maximum cell proliferation with high doses, which would then be tailed off as proliferation diminished, in deference to the quoted toxicological data for 5azaC (Material safety data sheet, 1992). Twice daily i/p injections of 5azaC were to be given over a course of 4 days, starting 26 h after partial hepatectomy. The doses on day 1 would be 20 mg/Kg b.w., diminishing in 5 mg/Kg b.w. increments per day to 5 mg/Kg b.w. on day 4; it was envisaged that at these doses the animals would show some signs of toxicity, probably in the form of lethargy (Material safety data sheet, 1992).

Thirteen 8 to 9 week old hemizygous female *fpg*-transgenic mice, and 6 age-matched B6 control animals were selected and underwent partial hepatectomies, at which liver samples were obtained for FaPy-DNA glycosylase assay. Treatment with 5azaC commenced 26 h later. The following day, when the dose of 5azaC was due to be reduced from 20 to 15 mg/Kg b.w., it was noted that those mice which had received 5azaC did not appear to be exhibiting any ill effects and therefore it was decided not to reduce the dose, and to observe the effects the next day. On the third day the mice appeared to be tolerating the treatment well and again the dose was not reduced. The final dose was given on the morning of the fourth post-operative day

and the animals sacrificed 4 h later, with liver samples obtained for glycosylase assay, DNA, and, if required, RNA extraction. The FaPy-DNA glycosylase levels, before and after 5azaC treatment, are shown in Table 7.9.

Table 7.9

Liver FaPy-DNA glycosylase activity in hemizygous *fpg*-transgenic female mice treated with 5-azacytidine following partial hepatectomy

Animals[a]	Group	Treatment[b]	*n*	Pre-treatment[c]	Post-treatment[d]	Change (%)
B6	Control	Saline	3	1.63 (0.37)	1.65 (0.28)	+1.2
	Test	5azaC[e]	3	1.17 (0.09)	1.11 (0.15)	+5.1
Transgenic	Control	Saline	5	1.62 (0.21)	1.23 (0.22)	-24.1
	Test	5azaC	8	1.53 (0.18)	1.07 (0.07)*	-30.1

[a] 8 to 9 weeks old, [b] commencing 26 h after partial hepatectomy, [c] FaPy-DNA glycosylase activities in liver biopsies obtained by partial hepatectomy, given in μU/g TP, [d] liver samples obtained *post mortem*, [e] 20 mg 5azaC/Kg b.w., given by i/p injection b.d. × 4 days, * $p = 0.040$ by paired *t*-test compared with pre-treatment levels.

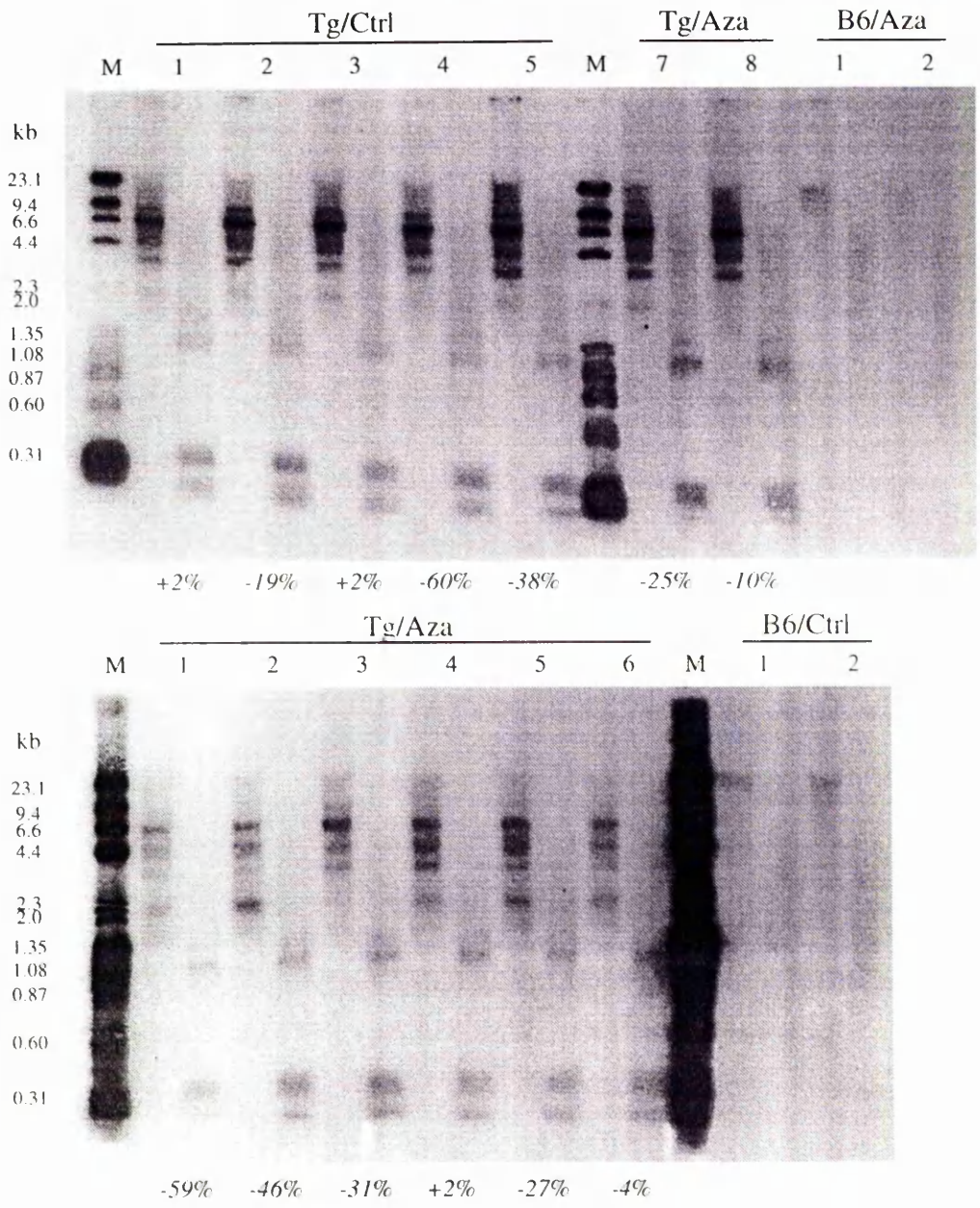
The pre- and post-treatment values for the animals in each group were compared using the paired *t*-test, which showed that the small differences observed in the B6 control and test groups were not statistically significant. Of the large, and somewhat surprisingly, negative changes in enzyme activity in the transgenic groups, that in the control animals proved not to be significant, however, that in the test group did prove significant ($p = 0.040$). It was of note that the mean (SEM) level of activity in the transgenic mice pre-treatment (1.56 (0.133) μU/g TP; $n = 13$) was no different ~~to~~ from that in the controls (1.40 (0.200) μU/g TP: $p = 0.51$; *t*-test). These values were also lower than those seen in 5 week old female mice or male mice of a similar age and this is discussed later.

In order to establish if the 5azaC treatment had had an effect on the methylation pattern of the *fpg*-transgene, DNA was extracted from the post-treatment liver samples from all of the transgenic animals plus two each from the B6 groups. These were

then digested with *Hpa* II and, separately, *Msp* I, and subjected to Southern analysis as carried out previously (7.4.2) except that a higher percentage of agarose was used in the gels to improve resolution of the smaller DNAs. The resultant autoradiographs are shown in Fig. 7.15, with the changes in FaPy-DNA glycosylase activity for the individual transgenic 5azaC-treated animals (Tg/Aza) given below their respective pair of lanes.

Fig. 7.15

Southern analysis of 5-azacytidine-treated female *fpg*-transgenic mice



B6 and hemizygous *fpg*-transgenic mice were treated with 5azaC following partial hepatectomy ("B6/Aza" and "Tg/Aza"), controls were treated with saline ("B6/Ctrl" and "Tg/Ctrl"), see text for details (and Table 7.9). DNA was extracted from liver samples and digested with the 5mC-sensitive restriction enzyme *Hpa* II (left-hand lane of each pair) and its 5mC-insensitive isoschizomer *Msp* I (right-hand lanes), and subjected to Southern analysis (2% agarose gel; 2.2.20; vacuum blotted). The percentage changes in individual FaPy-DNA glycosylase levels of the transgenic animals are shown beneath each respective pair of lanes. "M" = [^{35}S]-labelled markers.

The B6 mice, whether treated (B6/Aza) or untreated (B6/Ctrl), all showed the same pattern of diffuse, weak, hybridisation signals, between approximately 23 and 9 kb with *Hpa* II (5mC-methylation sensitive) and between *ca.* 9 and 1 kb with *Msp* I (5mC-methylation insensitive); no discrete bands were visible. The untreated transgenic mice (Tg/Ctrl) showed the same patterns as previously observed: with *Hpa* II a major band at 6.6 kb, plus minor ones at *ca.* 4.4, 3.5, 2.3 kb, and very weakly *ca.* 1.6 kb, plus a previously unobserved band at 9 kb; with *Msp* I major bands at 1.1, 0.3 and 0.2 kb plus a weak band at *ca.* 1.4 kb. These patterns were generally seen with the 5azaC-treated mice (Tg/Aza), except that some differences were noted in individual animals. An extra weak band was seen with *Hpa* II in Tg/Aza numbers 1 and 5 at *ca.* 1.5 kb, whereas the band at *ca.* 9 kb was weak or absent in numbers 2 and 4 to 6, the 2.3 kb band was absent in number 3 but prominent in number 2. With *Msp* I the band at 1.2 kb appeared generally more intense in all animals (more easily seen in the blot (Fig. 7.15a) where numbers 7 and 8 were run out next to the Tg/Ctrl samples), otherwise the patterns were no different when compared to the Tg/Ctrl samples. Thus, there was some evidence that treatment with 5azaC had altered the methylation pattern of the *fpg*-transgene, but it was inconsistent, small in extent and did not appear to correlate with the observed changes in FaPy-DNA glycosylase activity. It was therefore concluded that the treatment had been largely unsuccessful in altering the methylation pattern of the transgene. This is discussed later in relation to the statistically significant drop observed in the FaPy-DNA glycosylase activities of this group.

7.6 DISCUSSION

7.6.1 Transgenic mouse production

It was first necessary to establish the conditions for superovulation in the inbred strain B6. The experiment to investigate the optimum age at which B6 females should be superovulated (Table 7.1; 7.2.1) showed that the yield of eggs diminished with increasing age, especially at and after puberty (around 8 weeks). However, as the age decreased, so too did the proportion of animals failing to ovulate at all, and thus the process was most efficient with 6 week old females. Hogan *et al.* (1986) has divided mouse strains into two groups, "high responders" which ovulate 40 to 60 eggs per mouse, and "low responders" which ovulate 15 eggs or less per mouse, however this classification was made on the basis of the response of prepubescent animals. Strain B6 is classed as a high responder, while strains D2 and BDF_1 are classed as low responders. On the basis of the data presented here prepubescent B6 mice would appear to be intermediate in response. Adjustment to the precise timing of the gonadotrophin injections might have improved on these figures; the PMS injection needs to be given 2 to 3 h before the endogenous LH release (Gates, 1971), except where the amount of this is insufficient to induce ovulation on its own. The "all or nothing" nature of the response of these young B6 mice is perhaps evidence of a threshold effect, related in some way to the precise timing and/or dose of the gonadotrophins, factors which were not explored in this work. It is of note that pubescent or young adult females show a different pattern of response compared with that reported for prepubescent animals. Comparing the figures for 10 week old females, B6 mice showed a very poor response (6.1 eggs/animal), with D2 being somewhat greater (20/animal; data from Fan (1991)), and the hybrid BDF_1 greatest of all at 32.8/animal, presumably an effect of hybrid vigour which is not manifest before puberty.

At all subsequent stages of the process of generating transgenic mice, strain B6 proved to be less efficient than hybrid animals, thus confirming the experience of Hogan *et al.* (1986). A lower proportion of the eggs collected were fertile, they were less able to withstand the rigours of microinjection, and fewer zygotes surviving microinjection developed into 2-cell embryos. Thus, three times as many B6 females were required per microinjection experiment than with BDF_1 mice (15 *v.* 5), even then the number of eggs obtained was often only just sufficient. The survival figures for B6 *v.* BDF_2 zygotes microinjected either with buffer or the 6.6 kb *Not* I pHIF10 fragment are summarised in Table 7.10.

Table 7.10

Summary of survival figures for microinjected zygotes of strains BDF_2 and C57BL/6J

Solution injected	Strain	Total number of zygotes micro-injected	Immediate survivors[a]	2-cell embryos[a]	Pups[a]
Buffer	BDF_2	118	40.7%	35.6%	20.3%
	B6	165	17.0%	13.3%	6.7%
Buffer + DNA[b]	BDF_2	125	31.2%	25.6%	8.8%
	B6	755	19.2%	15.9%	2.6%[c]

[a] as a percentage of the total number of zygotes, [b] the 6.6 kb *Not* I pHIF10 fragment, [c] litters only obtained from 3 out of 6 pseudopregnant females implanted with 2-cell embryos.

This shows that a lower proportion of BDF_2 zygotes survived microinjection with DNA than with buffer alone. Also, that the survival figures for B6 zygotes are about half those for hybrid BDF_2 zygotes. The results with B6 zygotes either immediately surviving or developing to 2-cell embryos show minimal differences for buffer *v.* DNA. The figure for pups ultimately produced are lower for the group injected with DNA, but it has to be borne in mind that only three out of the six females implanted

were delivered of litters, if all had carried their pregnancies to term then this figure would likely have been *ca.* 5%. Of the three females who did not carry to term, #2 (see Table 7.3) failed to become pregnant, possibly because the number of embryos implanted was marginally sufficient. The other two (#3 and #6) became pregnant, but did not deliver litters, there being evidence that the pregnancies failed with resorption of the foetuses. It is possible that this was because of an infection, however, there was no evidence of new pathogens having been introduced into the Paterson mouse colonies during this time (Mr T. Bruder, personal commununication). Alternatively, it could be that expression of the *fpg*-transgene in mouse embryos becomes toxic during the middle week of foetal life, but it seems most unlikely (though not impossible) that all of the embryos, in two litters, were transgenic and expressed FPG protein, (especially given that out of the 20 pups born only one turned out to be transgenic.)

The fact that only a single *fpg*-transgenic founder animal was obtained out of the 20 mice born suggests a low integration frequency. A number of investigators have found integration frequencies of 10 to 40% for DNAs introduced by pronuclear microinjection (Brinster *et al.*, 1981; Costantini and Lacy, 1981; Gordon and Ruddle, 1981; Wagner E. *et al.*, 1981; Wagner T. *et al.*, 1981; Lacy *et al.*, 1983). For the transgenic mice produced previously in the Department Fan observed an integration frequency of 17% (8 transgenic animals in 46 born) using a construct of the mouse metallothionein-1 gene (mMT-1) with the *E. coli ada* ATase gene, and 31% (11 transgenics out of 35 pups) using a mMT1-human ATase construct (Fan *et al.*, 1990; Fan, 1991). Harris had obtained a single founder animal (from 14 born) using a construct containing the *E. coli ogt* ATase gene (Harris, 1990). Both Fan and Harris had used BDF_2 zygotes, however, so their results are not directly comparable with those using strain B6. In view of this, the apparently low integration frequency of 5% observed in the work presented here could be a characteristic of B6 zygotes, just as in

many other respects they were found to be less efficient than hybrids, as discussed above. An alternative reason could be one of operator effect, given a technique critically dependant on manual skills, though this might also be expected to be manifest as a reduced immediate survival efficiency of microinjected zygotes. In the experiment where BDF_2 zygotes were microinjected with buffer alone an immediate survival rate of 40.7% was seen (Table 7.10), a figure very close to the rates seen by Fan (also with BDF_2 zygotes) of 39.9% with the *ada* construct and 36.5% with the hAT construct (Fan, 1991), which would argue against this.

7.6.2 **Analysis of *fpg*-transgenic mice**

Initial investigation of the founder male by Southern analysis indicated that the *fpg*-insert in the transgene construct ~~were~~ was intact (as shown by the results with *Bam*H I) and that the transgene was probably present as a multi-copy head-to-tail array (as shown by the major 6.6 kb band with *Bst*X I) integrated into the host's genome. Analysis of F_1 offspring of the founder by Southern blotting showed that the transgene was transmissible, while PCR analysis provided extra evidence that the *fpg*-inserts were intact, or at least appeared full-length. The PCR method was further validated by checking the *Hha* I restriction pattern of the products, which in all cases were consistent with having been derived from *fpg*10 rather than, possibly contaminating, wild-type *fpg* in *E. coli*. It also showed that PCR testing was a suitable method for screening potentially transgenic offspring, there being concordance in the results obtained by Southern analysis. The PCR analysis method was facilitated by being able to use crude extracts from small tail tip biopsies, rather than from blood, or the purified DNA from large tail-tip biopsies.

The analysis of homozygous *fpg*-transgenic mice will be discussed in the next section, however, the Southern analysis of a litter containing, potentially, both

hemizygotes and homozygotes provided additional information about the transgene array (7.4.2). Digestion with *Not* I showed that the head-to-tail junctions of the copies of the transgene in the array had not reformed *Not* I sites, otherwise a single 6.6 kb band would have resulted. The enzyme *Not* I otherwise cuts extremely rarely in mammalian DNA, as its 8 bp recognition site is composed entirely of guanine and cytosine residues (i.e. 5′...GCGGCCGC...3′), thus the hybridisation signal at ≫23 kb is also evidence that the transgene was integrated into the host's genome, rather than present episomally.

The significance of the patterns seen with the two restriction enzymes known to cut once within the 6.6 kb *Not* I pHIF10 fragment (*Bst*X I and *Kpn* I) has already been discussed in section 7.4.2, and above, from which it could be concluded that the members of the transgene array were oriented head-to-tail. Because of the sizes of the junction fragments it was considered that at least one copy of the transgene in the array was probably truncated at the 5′ end of the injected fragment. It is possible that some other more complicated arrangement might be responsible for the patterns seen, but there was in any event felt to be sufficient doubt so such that the ratio of counts between the possible junction fragments and the main bands could not be used to absolutely determine the copy-number of the transgene. The ratio of the counts as determined by analysis with the phosphorimager was approximately 10:1 (major:minor bands, for *Bst*X I and *Kpn* I; data not shown). While this might indicate an absolute copy-number (per haploid genome) of 10, it could mean that for every 10 full-length members of the array there was one truncated copy and therefore this represents a relative copy-number. If this were so then the absolute copy-number might be 10, or a multiple thereof, i.e. 20 or 30, or even a more complicated combination giving overall a ratio of 10:1. No formal comparison of hybridisation strength with bands of predictable intensity was carried out to independently estimate the transgene copy-

number (as had been carried out with the cell lines; 5.4.2 and 5.4.3), however, the intensity of the bands seen on Southern analysis of the transgenic animals did not appear as strong as, for example, the original bands seen with cell clone 10L6 (5.4.2 and Fig. 5.14). Such a comparison is inevitably somewhat subjective and prone to many different errors, so any conclusions drawn must be circumspect. Assuming the simplest situation would imply that the transgene was present as an approximately 10-copy array, one of which was truncated, but this is by no means a definite conclusion.

It was possible, however, to come to some firm conclusions regarding the 5mC-methylation status of the transgene array. Digestion with the 5mC-insensitive restriction enzyme *Msp* I (recognition site 5′...CCGG...3′) would predict (from Fig. 3.8) the generation of six bands, four of defined size: 285 bp, 156 bp, 142 bp and 132 bp, plus two others, one >122 bp, the other >22 bp (in effect junction fragments of the *fpg*10 sequence with the flanking pHMG sequences). The pattern seen of a broad band at *ca.* 0.3 kb, with two others at 1.4 and 1.1 kb would be consistent with this. The most 5′ *Msp* I fragment of the *fpg*10-insert is that of >122 bp, with adjacent to it the 285 bp fragment (≈ 0.3 kb), it is therefore possible that the 1.1 kb band is the >122 bp fragment, with the 1.4 kb band representing incomplete digestion of the 1.1 + 0.3 kb fragments. In subsequent analyses (Fig. 7.15) the 1.4 kb band was generally absent, or much less distinct, than in this first blot, which would be compatible with this. Alternatively, the broadness of the *ca.* 0.3 kb hybridisation band in the later experiment (Fig. 7.15), in spite of a higher percentage agarose gel being used to improve the resolution of such small DNAs, could be compatible with a 5′ junction fragment of approximately 0.3 kb. The junction fragment from the 3′ end of the *fpg*10-insert would only have 22 bp of homology with the probe (the *Bam*H I-*fpg*4 fragment) and would be unlikely to remain hybridised under the reasonably stringent conditions used to wash the blot (0.6× SSC, 0.1% SDS, 65°C). The pattern seen with

Hpa II, the 5mC-sensitive isoschizomer of *Msp* I, was of a major band at 6.6 kb plus others, less intense, at 9, 4.4, 3.5, 2.3 and 1.6 kb. This clearly indicated that most, if not all, of the *Msp* I/*Hpa* II sites within the *fpg*10-inserts, as well as those within 9 kb of this region, were subject to cytosine methylation. (The attempt to alter this methylation pattern is discussed later in section 7.6.4.) The presence of a major band with *Hpa* II at 6.6 kb strongly suggests, given the head-to-tail multi-copy array of the transgene, that the same *Msp* I/*Hpa* II site within each copy of the array was not methylated. This is discussed below, in the context of the evidence for and against expression of the transgene.

Integration of transgenes as multi-copy head-to-tail arrays has been observed by many groups (Brinster *et al.*, 1985; Costantini and Lacy, 1981; Gordon, 1993). Brinster *et al.* (1985) concluded that mouse eggs contain large amounts of enzyme activities able to modify linear DNA ends and ligases able to join such molecules. Gordon (1993) has suggested that homologous recombination between microinjected fragments occurs prior to integration, with possible reduplication of some pieces of the DNA during integration itself. This latter point is of interest as Gordon (1993) comments on occasionally finding aberrant restriction fragments within transgene arrays; also that an array which did not contain at least one complete expressible copy of the transgene has never been observed in his laboratory, and, in view of this, extensive Southern analysis of new transgenic lines is no longer carried out once a multi-copy array has been demonstrated.

Analysis of the pattern of inheritance of the transgene, from the founder mouse to its F_1 offspring, showed that it was present at a single autosomal locus. Hogan *et al.* (1986) has observed that when multiple copies of a transgene are present, they are usually found at a single chromosomal locus. It has been noted, however, that transgenes may integrate at more than one site, often on different chromosomes

(Gordon, 1993; Lacy *et al.*, 1983). Again, the copies of the transgene at each site have been found to be arranged in head-to-tail arrays (Brinster *et al.*, 1981; Costantini and Lacy, 1981). Integration is generally considered to take place at the one-cell stage (Gordon, 1993), thus the transgene is present in all cells of the founder, including the germ cells. It is possible for the transgene to be integrated after the first cell division, which results in animals which are mosaic, and the frequency of this amongst founder animals has been noted to be approximately 20% (Costantini and Lacy, 1981). A side-by-side comparison of the copy number of the *fpg*-transgene in the founder animal with its offspring was performed (Fig. 7.5), and, although the intensities of the bands were not compared densitometrically, the intensity of the band observed with the founder was very similar to the bands from the offspring showing that the copy-numbers were certainly not grossly different. The genetic analysis of the *fpg*-transgenic offspring (7.3.4) showed that the ratio of transgenic to non-transgenic animals was not significantly different from unity, implying that the founder was not a mosaic.

The first experiment comparing FaPy-DNA glycosylase activity in *fpg*-transgenic mice with B6 animals demonstrated statistically significantly higher activity in the liver in a group of six 8 week old transgenic animals (7.3.3), although this was only by *ca.* 1.2–fold. However, subsequent experiments (7.4.3, 7.5.1, and 7.5.2) failed to demonstrate greater FaPy-DNA glycosylase activity in liver or kidney, in males or females, at a range of ages (5 to 9 weeks). Assay of activity in other tissues was not carried out, and it is possible that in some other tissue(s) the transgene may have been expressed. However, while HMGCR is considered a "house-keeping" gene in view of its constitutive expression in almost all cell types, it is expressed to its highest level in liver, where cholesterol synthesis is greatest (Henwood and Heel, 1988). Northern analysis failed to show any evidence of message production, implying that the

transgene array was transcriptionally inactive (7.4.4). In addition, western analysis (7.4.4) was unable to show expression of FPG protein (which may not necessarily have been detectable by enzyme assay), although the previous comments about the limits of detection of FPG protein by western blotting being rather greater than the FaPy-DNA glycosylase assay (5.5.3) are also valid here. Taken together, these various lines of evidence indicated that the *fpg*-transgene was not expressed. The first experiment did give a statistically significant result at the $p = 0.05$ level, but without any other corroborating data this must be regarded as presumably an instance of "regression to a mean"; such a result is to be expected once in every twenty experiments simply because of random statistical variation. While absence of evidence is not necessarily evidence of absence, it could be concluded from this that even if the *fpg*-transgene did express it was at a low level, and unlikely to have much, if any, influence in toxicity and carcinogenesis experiments involving these animals.

It has already been mentioned how cytosine methylation has been correlated with the transcriptional activity of endogenous genes, as well as transgenes (7.5.2; Cedar *et al.*, 1983; Stein *et al.*, 1983; Costello *et al.*, 1994). There are a number of reasons why the *fpg*-transgene array might have become methylated. Firstly, the site of chromosomal integration of the transgene might be one naturally prone to methylation. Secondly, the presence of a large tandem array of the same (trans)gene is an un-natural arrangement and might prevent proper regulation of transcription (Hogan *et al.*, 1986). Thirdly, the specific presence of prokaryotic sequence(s) might stimulate methylation in and around that region, and this effect might be stronger given their organisation in a multi-copy array. The extensive methylation of the *fpg*-transgene is probably related to its failure to be expressed. Tissue-specific genes have been found to be heavily methylated in most tissues except those in which the gene is expressed (Yisraeli and Syzf, 1984). In contrast, the CpG islands in house-keeping genes are

unmethylated in every tissue, at all stages of development, the only exception being such genes on the inactivated X-chromosome (Stein *et al.*, 1983). It is most interesting, therefore, that an array of transgenes, based on a mouse house-keeping gene, is heavily methylated and shows no signs of transcriptional activity, in the very mouse tissue where expression might be expected to be maximal. The two different strategies used to try and accomplish expression of the *fpg*-transgene are discussed in section 7.6.4.

7.6.3 **Homozygous *fpg*-transgenic mice**

The advantages in breeding a transgenic line to homozygosity have already been reviewed in sections 7.1 and 7.4.1. Successful breeding of the *fpg*-transgenic line to homozygosity was demonstrated by a combination of two techniques: test breeding with wild-type mice for genetic proof, and molecular analysis to test transgene dosage. Other techniques have been applied to this issue, such as measurement of hybridisation signal on dot- or slot-blots, with normalisation of the signal by probing a second filter, equivalent to the first, for a single-copy endogenous mouse gene (Hogan *et al.*, 1986; Gendron-Maguire and Gridley, 1993; Krulewski, 1989), however, these authors note that it is better to prove potential homozygotes by test breeding with non-transgenic animals. Curiously, there are no guidelines given as to the number of offspring from such matings which all need to test positive before homozygosity can be confidently established, although Bronowski (1960) has commented: "*The reason is that we can never discount a run of luck of the kind which Dr Johnson once observed when a friend of his was breeding horses. 'He has had', said Dr Johnson, 'sixteen fillies without one colt, which is an accident beyond all computation of chances.'* (from equation 4: $p = 0.000\ 015$, but an X-linked mutation causing lethality *in utero* seems a more likely explanation than bad luck). The use of a probabilistic model,

such as equation 4, seems obvious, and its use is suggested for those wishing to have objective evidence of proof. One of the reasons it was decided to prove homozygotes by breeding was the facility with which offspring could be tested, and at a young age, by PCR, if testing had only been possible by e.g. Southern blotting then it may well have been more efficient to test first by a dot/slot-blotting-based method. A useful product of the test breeding program was the large number of hemizygous animals it made available for other experiments. The use of dot/slot-blotting methods to quantitate transgene dosage is usually advocated because they circumvent the problems of DNA quantitation and its unequal transfer to filters associated with Southern blotting. The experiment performed here (7.4.2) shows that Southern blotting can provide a satisfactory result, admittedly with the use of a phosphorimager to quantitate the signals. It is by no means guaranteed that a hemizygous transgenic line can be bred to homozygosity. A significant proportion of transgenes integrate into endogenous mouse genes, in the process producing a recessive lethal mutation, which can be recognized by finding reduced litter sizes when inbreeding (Gordon, 1993; Hogan *et al.*, 1986). Meisler has reviewed this area (1992), and a number of studies have found the insertional mutation rate to be between 5% to 10% (Palmiter *et al.*, 1986; Soriano *et al.*, 1987; Weiher *et al.*, 1990). While a formal study looking for evidence of insertional mutagenesis in the *fpg*-transgenic line was not carried out, the animals, whether hemi- or homozygous, showed no gross anatomical abnormalities, or exhibited any obvious physiological or behavioural disturbances. However, it will be appreciated that systematic screening for abnormalities which are not obvious is potentially extremely time-consuming.

It is also possible for such mutations to affect fertility, thus making it impossible to breed from one sex or the other when they are homozygus for the transgene (Gordon, 1993). In the experiment carried out and described here, one male and two female

hemizygotes were produced, but one female was only able to have a single litter. As the other female was able to produce three, apparently full-size, litters when mated with B6 males, and also produced a litter of eight homozygotes when crossed with her homozygous brother, it was concluded that the reason the first female was unable to have further litters was peculiar to her as an individual, and not related to being homozygous for the *fpg*-transgene.

7.6.4 **Attempts to achieve expression of FPG protein in *fpg*-transgenic mice**

It had originally been envisaged that it might be possible, through the use of the HMGCR inhibitors PV and SV, to up-regulate the expression of the *fpg*-transgene. Following on from the realisation that the *fpg*-transgene was almost certainly not expressed, at least to a significant level, it was recognised that any attempts to produce such up-regulation would be starting from a baseline of zero. The pilot experiment with these agents (7.5.1) showed that they might be capable of causing a paradoxical reduction in FaPy-DNA glycosylase activity in transgenic animals but not in the B6 controls, which might be construed as evidence that the transgene was active, but then, for some reason, down-regulated by the treatment with SV/PV. It was possible that the region of the mouse HMGCR gene contained in the vector pHMG did not contain all of the *cis*-acting promoter/enhancer elements necessary for proper physiological regulation (and this may have been one reason for the inconsistent results seen *in vitro* with clone 10H4; 5.4.8.2 and 5.5.4). A second experiment with greater numbers of individuals, for increased statistical power, was therefore carried out to address this preliminary finding. The results of this (Table 7.8) demonstrated that PV had no significant effect on B6 animals nor on transgenic animals. However, an unexpected and somewhat puzzling finding was that the transgenic control (i.e. saline treated) mice showed a significant rise in FaPy-DNA

glycosylase activity. This was to the same average extent as observed in the treated transgenic group (i.e. *ca.* +11%), but the testing for significance had been carried out with the paired *t*-test. It is possible that this does not represent any actual biological effect; it is perhaps equally possible that this modest rise in activity represents a natural ageing effect in *fpg*-transgenic mice. No tests were carried out on animals older than these so no firm conclusions can be made, but this might be an area worthy of further study.

The extensive cytosine-methylation of the transgene array has already been discussed (7.6.2). This methylation appeared to include *Hpa* II/*Msp* I sites throughout the 6.6 kb transgene construct and therefore may well account for not only the failure of the transgene to be expressed, but also its failure to be up-regulated by agents likely to cause such an effect. Although it can not be certain that a sufficient dosage of the agents was given to cause up-regulation of the endogenous (hepatic) HMGCR activity, because no assays were performed of hepatic HMGCR activity, the dosing regime was based on previous published experiences (Henwood and Heel, 1988; Kita *et al.*, 1980) and also given for a somewhat longer period (4 *v.* 3 days). Regulation of endogenous HMGCR activity might involve the methylation of specific cytosine residues within promoter/enhancer elements of the *mHMGCR* gene, and the inactivity of this transgene might provide an opportunity to investigate this. Alternatively, if this does not represent a normal physiological process, it might still provide data as to how promoter/enhancer activity is altered by methylation at certain critical sites. Approaches to studying methylation based on the PCR have been described (Zuccotti *et al.*, 1993), and in conjunction with the presence nearby of the *fpg* sequence it might be possible to devise a strategy to study methylation in the transgene without interference from the endogenous gene.

The second strategy devised to achieve up-regulation of the *fpg*-transgene involved treatment with 5azaC. This was given during the expected period of active cell division in the liver following partial hepatectomy, to try and alter the methylation pattern and, hence, realise expression of the transgene. Analysis of liver DNA by Southern blotting, after 5azaC treatment, showed little, if any, difference in the pattern of methylation, suggesting that the treatment had been ineffective. There were slight differences seen between different animals, but these were not consistent. Neither was any significant difference seen in hepatic FaPy-DNA glycosylase activity, nor any correlation of this with the Southern blot pattern. This inability to change the methylation pattern could have been due to a number of reasons. The dosage of 5azaC may have been insufficient, however, the doses given, although based on expected toxicity, were larger than originally intended. As the animals tolerated even these higher doses, then the planning of any future such experiments should take this into account; different mouse strains may show differences in their susceptibility to treatment with 5azaC, for whatever reason, which would be useful to know for practical purposes, and might throw some light on this aspect of gene regulation. The Southern analysis-based method of testing may not have been sufficiently sensitive to detect changes at certain sites, which may also not necessarily have occurred at the same site in all copies of the transgene in the array. A PCR-based method to look at methylation of particular sites might be more appropriate (as discussed above). It is of note that Fan (1991) had attempted to gain expression of mMT-1-*ada* transgenes by 5azaC treatment without success, but the different transgenic lines tested did not have their methylation status assessed, before or after treatment, and the 5azaC was given in drinking water *ad libitum*, rather than by injection. One way of investigating the phenomenon further might be to establish, for example, primary cultures of neonatal

skin fibroblasts or an immortalised cell line, from the *fpg*-transgenic mice, which could be manipulated *in vitro*.

CHAPTER 8

SUMMARY

It might well be concluded from the results presented here that expression of FPG protein in mammalian cells is difficult, if not impossible, to achieve, certainly to levels at which protection from the effect of relevant xenobiotics might be expected. Summarising the various findings allows an overview to be made:

1) Overexpression of active FPG protein was achieved in *E. coli* by a factor of over 1800-fold (Chapter 4), without any apparent ill effects on the organism, even when the cultures were aerated with pure oxygen. In fact, this had the effect of improving the yield of bacteria and protein. A similar degree of overexpression has been obtained by Boiteux *et al.* (1990). This demonstrates that, at least in *E. coli*, a gross excess of FPG DNA–glycosylase activity, when compared to the rest of base excision repair pathway, is not lethal.

2) Transfection of RJKO cells with G-418-selectable constructs of different versions of the *fpg* gene had differing results (Chapter 5). The proportion of clones (4L*n*) containing the *fpg*4 gene construct in an intact state was 10 out of the 12 obtained; it was later determined that *fpg*4 encoded a protein with much diminished FaPy-DNA glycosylase activity. Transfection of RJKO cells with the pLIF10 construct, shown to have an *fpg* insert encoding a fully active FaPy-DNA glycosylase by the bacterial overexpression experiments, gave only a single clone (10L6; 5.4.2) containing the *fpg* gene, out of 30 clones picked in two experiments. This was using selection conditions which had allowed other DNA repair genes to be introduced into RJKO cells. There was also evidence of loss of *fpg* sequence from clones during the isolation process,

and/or other rearrangements of the integrated retroviral-based constructs, a phenomenon observed by others (Stuhlmann *et al.*, 1988; Dr L. Harris, personal communication).

A subsequent experiment with milder G-418 selection conditions allowed the isolation of five more clones in addition to 10L6, which appeared to have undergone a reduction in its *fpg* gene copy number (5.4.5). Of the six clones three showed evidence of transcription of *fpg*–containing mRNA. Two clones had significantly raised FaPy-DNA glycosylase activity, but only by a factor of *ca.* 1.8-fold (5.4.7.4). Co-transfection of RJKO cells with pHIF10 plus pLJ allowed the isolation of six clones, only one of which contained the *fpg* gene (10H4; 5.4.6). This clone had evidence of message production, but the level to which its FaPy-DNA glycosylase activity was raised (*ca.* 1.5-fold) was only of borderline significance. Furthermore, it was not possible to achieve expression of the *fpg* gene in this line by treatment with HMGCR inhibitors.

3) Of 20 potential *fpg*-transgenic founder mice born, only one actually carried the pHIF10 *fpg* construct. This low proportion may have been a function of the use of the inbred B6 strain. The *fpg*-transgene was shown to have integrated as a multi-copy head-to-tail array, and was transmitted to offspring as a single autosomal locus. Subsequent Southern analysis revealed the array to be highly methylated, a phenomenon associated with down-regulation of genes. In addition, there was no evidence of transcript production on northern analysis, nor evidence of FPG protein on western analysis. An initial experiment suggested a small (*ca.* 1.3-fold) increase in hepatic FaPy-DNA glycosylase levels, but this finding was never repeated and was probably fortuitous. Treatment of the mice, either with HMGCR inhibitors, or 5azaC following partial hepatectomy, did not result in any increase in FaPy-DNA glycosylase

levels. Thus, it appeared that perhaps only a small proportion of pups could be obtained which were transgenic, but certainly the only one which was had no evidence of FPG protein expression.

Taken together this suggested that there was strong selection against the expression of FPG protein in mammalian cells. This might have been related to the fact that FPG protein is prokaryotic in origin, rather than to a specific function of its activities as a DNA repair enzyme. If it were a consequence of its enzymatic behavior, this would suggest that a gross imbalance in the various enzyme activities, which together constitute the base excision repair pathway, is not tolerated by mammalian cells. The endogenous level of 8oxoG in DNA (1 per 10^5 G) is much higher than the endogenous level of other lesions formed by alkylation (1 per $10^7/10^8$ G; Kasai and Nishimura, 1991b); excessive DNA nicking at sites of endogenous 8oxoG might be too much for the capacity of the other enzymes responsible for repairing such nicks, resulting in lethality as a consequence of un–repaired strand breaks. Overexpression of 3mA-DNA glycosylase in Chinese hamster ovary cells has recently been shown to render them liable to undergo chromosomal aberrations and DNA breaks when treated with methylating agents (Coquerelle *et al.*, 1995). Thus, there is a precedent for imbalance in base excision repair causing increased sensitivity to xenobiotics, rather than enhancing protection against them. Alternatively, FPG might be acting to directly subvert a normal DNA repair function. There is a precedent for this too, in that expression of the human (hMSH2) and other divergent homologues of the *E. coli* mismatch repair enzyme mutS in wild-type *E. coli* confers on them a mutator phenotype like that of *mutS*⁻ strains (Fishel *et al.*, 1993).

However, this has to be interpreted in the light of the findings of Laval (1994), who was able to achieve expression of FPG protein in both V79 and CHO (Chinese

hamster ovary) cells, causing an increase in FaPy-DNA glycosylase activity in such clones by factors of 25- to 40-fold. This provided no protection to the cytotoxic effects of γ–irradiation or hydrogen peroxide treatment, but it did reduce by two-thirds the frequency of mutations induced by γ–irradiation at the HPRT locus causing 6-thioguanine resistance. These results suggested that FPG when expressed in mammalian cells can indeed protect them from the mutagenic effects of a γ–irradiation-induced DNA lesion, presumably 8oxoG. The expression vector used in her experiments was pSV2, based on the SV40 virus early promoter. Another vector based on the SV40 early promoter (pRSV-L) has given similar levels of expression as pHMG (Gautier *et al.*, 1989; 5.3.2), but this was determined in a rat fibroblast cell line transformed by the *ras* oncogene (FR3T3ras) and therefore may well not be representative of the situation in V79/RJKO cells. Thus, the results presented here may not be incompatible with Laval's findings, suggesting that the degree of FPG expression is of critical importance, a factor which might be explored by the use of expression vectors with inducible promoters.

Finally, given that it is possible to obtain FPG expression in at least some mammalian cells it would be worth attempting to obtain an *fpg*-transgenic strain of mouse, either by pro-nuclear microinjection, or by transfection of a mouse embryonic stem cell line, using a different/inducible promoter. With the eventual cloning of mammalian homologues of the *fpg* gene, it will be possible to introduce targeted mutations into the mouse *fpg* gene in embryonic stem cell lines, thus allowing investigations into the effect(s) of deficiency in this DNA repair function. With such experimental models available, it should be possible to determine the significance of FaPy and 8-oxopurine lesions, and their repair, in the acute and long-term biological effects of xenobiotic agents.

REFERENCES

Abbott, P.J. and Saffhill, R. *Biochim. Biophys. Acta* (1979) **562**: 51-61.

Aida, M. and Nishimura, S. *Mutat. Res.* (1987) **192**: 83-89.

Akiyama, M., Horiuchi, T. and Sekiguchi, M. *Mol. Gen. Genet.* (1987) **206**: 9-16.

Altman, P.L. and Katz, D.D. Part I: Mouse and rat. In: *Inbred and Genetically Defined Strains of Laboratory Animals* (1979) Federation of American Societies for Experimental Biology, Bethesda, MD.

Amann, E. *Dev. Biol. Stand.* (1985) **59**: 11-22.

Ames, B.N. *Science* (1983) **221**: 1256-1263.

Armitage, P. and Berry, G. *Statistical Methods in Medical Research,* 2nd edn. (1987) Blackwell Scientific Publications, Oxford.

Aruoma, O.I., Halliwell, B. and Dizdaroglu, M. *J. Biol. Chem.* (1989) **264**: 13024-13028.

Aruoma, O.I., Halliwell, B., Gajewski, E. and Dizdaroglu, M. *Biochem. J.* (1991) **273**: 601-604.

Asahara, H., Wistort, P.M., Bank, J.F., Bakerian, R.H. and Cunningham, R.P. *Biochemistry* (1989) **28**: 4444-4449.

Ayala, F.J. and Kiger, J.A. (1984) *Modern Genetics*. Benjamin, Menlo Park, CA.

Bailly, V., Derydt, M. and Verly, W.G. *Biochem. J.* (1989) **261**: 707-713.

Barrows, L.R., Borchers, A.H. and Paxton, M.B. *Carcinogenesis* (1987) **8**: 1853-1859.

Bases, R., Franklin, W.A., Moy, T. and Mendez, F. *Int. J. Radiat. Biol.* (1992) **62**: 427-441.

Beranek, D.T., Weis, C.C., Evans, F.E., Chetsanga, C.J. and Kadlubar, F.F. *Biochem. Biophys. Res. Commun.* (1983) **110**: 625-631.

Berenblum, I. and Shubik, P. *Br. J. Cancer* (1947) **1**: 383-386.

Bergmann, K. and O'Konski, C.T. *J. Phys. Chem.* (1963) **67**: 2169-2177.

Bhatnagar, S.K. and Bessman, M.J. *J. Biol. Chem.* (1988) **263**: 8953-8957.

Bickersteth, R.A. *St Bartholomew's Hospital Reports* (1890) **26**: 299-301.

Billingham, R.E., Brent, L. and Medawar, P.B. *Nature* (1953) **172**: 603-606.

Bird, A.P. *Nature* (1986) **321**: 209-213.

Blakely, W.F., Fuciarelli, A.F., Wegher, B.J. and Dizdaroglu, M. *Radiat. Res.* (1990) **121**: 338-343.

Bodmer, W. F., Bailey, C. J., Bodmer, J., Bussey, H. J., Ellis, A., Gorman, P., Lucibello, F. C., Murday, V. A., Rider, S. H. and Scambler, P. *Nature* (1987) **328**: 614-616.

Boiteux, S. and Laval, J. *Biochem. Biophys. Res. Commun.* (1983) **110**: 552-558.

Boiteux, S., Belleney, J., Roques, B.P. and Laval, J. *Nucl. Acids Res.* (1984) **12**: 5429-5439.

Boiteux, S., O'Connor, T.R. and Laval, J. *EMBO J.* (1987) **6**: 3177-3183.

Boiteux, S., Bichara, M., Fuchs, R.P. and Laval, J. *Carcinogenesis* (1989) **10**: 1905-1909.

Boiteux, S. and Huisman, O. *Mol. Gen. Genet.* (1989) **215**: 300-305.

Boiteux, S., O'Connor, T.R., Lederer, F., Gouyette, A. and Laval, J. *J. Biol. Chem.* (1990) **265**: 3916-3922.

Boiteux, S., Gajewski, E., Laval, J. and Dizdaroglu, M. *Biochemistry* (1992) **31**: 106-110.

Bonicel, A., Mariaggi, N., Hughes, E. and Teoule, R. *Radiat. Res.* (1980) **83**: 19-26.

Bonneau, R., Pottier, R., Bagno, O. and Joussot-Dubien, J. *Photochem. Photobiol.* (1975) **21**: 159-163.

Bos, J.L., Fearon, E.R., Hamilton, S.R., Verlaan de Vries, M., van Boom, J.H., van der Eb, A.J. and Vogelstein, B. *Nature* (1987) **327**: 293-297.

Boutwell, R.K. *Prog. Exp. Tumour Res.* (1964) **4**: 207-250.

Boylan, J.F., Jackson, J., Steiner, M.R., Shih, T.Y., Duigou, G.J., Roszman, T., Fisher, P.B. and Zimmer, S.G. *Anticancer Res.* (1990) **10**: 717-724.

Bradford, M.M. *Anal. Biochem.* (1976) **72**: 248-254.

Breimer, L.H. and Lindahl, T. *Nucl. Acids Res.* (1980) **8**: 6199-6211.

Breimer, L.H. and Lindahl, T. *J. Biol. Chem.* (1984) **259**: 5543-5548.

Breimer, L.H. *Nucl. Acids Res.* (1984) **12**: 6359-6367.

Breimer, L.H. *Mol. Carcinog.* (1990) **3**: 188-197.

Brennand, J. and Margison, G.P. *Carcinogenesis* (1986a) **7**: 2081-2084.

Brennand, J. and Margison, G.P. *Proc. Natl. Acad. Sci. U. S. A.* (1986b) **83**: 6292-6296.

Brent, L. *Transplant. Proc.* (1991) **23**: 2056-2060.

Brent, T.P. *Biochemistry* (1979) **18**: 911-916.

Brinster, R.L., Chen, H.Y., Trumbauer, M.E., Senear, A.W., Warren, R. and Palmiter, R.D. *Cell* (1981) **27**: 223-231.

Brinster, R.L., Chen, H.Y., Trumbauer, M.E., Yagle, M.K. and Palmiter, R.D. *Proc. Natl. Acad. Sci. U. S. A.* (1985) **82**: 4438-4442.

Bronowski, J. Chapter 6: The idea of chance, p.91. In: *The Common Sense of Science* (1960) Pelican Books, London.

Brosius, J., Dull, T.J., Sleeter, D.D. and Noller, H.F. *J. Mol. Biol.* (1981) **148**: 107-127.

Brosius, J. *Gene* (1984) **27**: 151-160.

Brosius, J. and Holy, A. *Proc. Natl. Acad. Sci. U. S. A.* (1984) **81**: 6929-6933.

Brothwell, D.R. The evidence for neoplasms, pp.320-345. In: *Diseases in Antiquity*. (1967) Springfield.

Brown, T., Leonard, G.A., Booth, E.D. and Chambers, J. *J. Mol. Biol.* (1990) **212**: 437-440.

Bulfield, G. *Trends Genet.* (1985) **1**: 39.

Cabrera, M., Nghiem, Y. and Miller, J.H. *J. Bacteriol.* (1988) **170**: 5405-5407.

Cairns, J. *Nature* (1980) **286**: 176-178.

Cedar, H., Stein, R., Gruenbaum, Y., Naveh Many, T., Sciaky Gallili, N. and Razin, A. *Cold Spring Harb. Symp. Quant. Biol.* (1982) **47 Pt 2**: 605-609.

Cepko, C.L., Roberts, B.E. and Mulligan, R.C. *Cell* (1984) **37**: 1053-1062.

Cerutti, P.A. *Science* (1985) **227**: 375-381.

Chetsanga, C.J. and Lindahl, T. *Nucl. Acids Res.* (1979) **6**: 3673-3684.

Chetsanga, C.J., Bearie, B. and Makaroff, C. *Chem. Biol. Interact.* (1982a) **41**: 217-233.

Chetsanga, C.J., Polidori, G. and Mainwaring, M. *Cancer Res.* (1982b) **42**: 2616-2621.

Chetsanga, C.J. and Frenette, G.P. *Carcinogenesis* (1983) **4**: 997-1000.

Chetsanga, C.J. and Grigorian, C. *Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med.* (1983) **44**: 321-331.

Chetsanga, C.J. and Grigorian, C. *Proc. Natl. Acad. Sci. U. S. A.* (1985) **82**: 633-637.

Chomczynski, P. and Sacchi, N. *Anal. Biochem.* (1987) **162**: 156-159.

Chung, M.H., Kasai, H., Jones, D.S., Inoue, H., Ishikawa, H. and Ohtsuka, E. *Mutat. Res.* (1991a) **254**: 1-12.

Chung, M.H., Kim, H.S., Ohtsuka, E., Kasai, H., Yamamoto, F. and Nishimura, S. *Biochem. Biophys. Res. Commun.* (1991b) **178**: 1472-1478. [published erratum appears in *Biochem. Biophys. Res. Commun.* (1991) **180**: 462]

Clark, G. (1989) *PhD. Thesis*, pp.166-171. University of Manchester.

Claycamp, H.G. and Luo, D. *Radiat. Res.* (1994) **137**: 114-117.

Cohen, S.M. and Ellwein, L.B. *Cancer Res.* (1991) **51**: 6493-6505.

Coleman, S.H. and Wild, D.G. *Nucl. Acids Res.* (1991) **19**: 3999.

Conlay, J.J. *Nature* (1963) **197**: 555-557.

Constantoulakis, P., Josephson, B., Mangahas, L., Papayannopoulou, T., Enver, T., Costantini, F.D. and Stamatoyannopoulos, G. *Blood* (1991) **77**: 1326-1333.

Cook, J.W., Hewitt, C.L. and Heiger, I. *J. Chem. Soc.* (1933) 395-405.

Coquerelle, T., Dosch, J. and Kaina, B. *Mutat. Res.* (1995) **336**: 9-17.

Costantini, F.D. and Lacy, E. *Nature* (1981) **294**: 92-94.

Costello, J.F., Futscher, B.W., Tano, K., Graunke, D.M. and Pieper, R.O. *J. Biol. Chem.* (1994) **269**: 17228-17237.

Cowled, P.A. and Forbes, I.J. *Br. J. Cancer* (1989) **59**: 904-909.

Cripps, W.H. *Trans. Pathol. Soc. London* (1882) **33**: 165-168.

Croy, R.G. and Wogan, G.N. *Cancer Res.* (1981) **41**: 197-203.

Darwin, C. Chapter 1: Variation under domestication, p.32. In: *On the Origin of Species by means of Natural Selection,* 4th edn. (1866) John Murray, London.

Daudel, P. and Daudel, R.C. (1966) *Chemical Carcinogenesis and Molecular Biology,* London. pp.1-5.

De Boer, H.A., Comstock, L.J. and Vasser, M. *Proc. Natl. Acad. Sci. U. S. A.* (1983) **80**: 21-25.

de Oliveira, R., van der Kemp, P.A., Thomas, D., Geiger, A. and Nehls, P. *Nucl. Acids Res.* (1994) **22**: 3760-3764.

De Wet, J.R., Wood, K.V., DeLuca, M., Helinski, D.R. and Subramani, S. *Mol. Cell Biol.* (1987) **7**: 725-737.

Demple, B., Jacobsson, A., Olsson, M., Robins, P. and Lindahl, T. *J. Biol. Chem.* (1982) **257**: 13776-13780.

Detke, S. *Biochem. Biophys. Res. Commun.* (1984) **125**: 812-816.

Di Mascio, P., Wefers, H., Do Thi, H.P., Lafleur, M.V. and Sies, H. *Biochim. Biophys. Acta* (1989) **1007**: 151-157.

DiPaolo, J.A., DeMarinis, A.J., Evans, C.H. and Doniger, J. *Cancer Lett.* (1981) **14**: 243-249.

Dizdaroglu, M. *Biochemistry* (1985) **24**: 4476-4481.

Doetsch, P.W. and Cunningham, R.P. *Mutat. Res.* (1990) **236**: 173-201.

Doll, R. and Peto, R. In: *The Causes of Cancer* (1981) Oxford University Press, Oxford.

Douillard, J.Y. and Hoffman, T. *Methods Enzymol.* (1983) **92**: 168-174.

Duesberg, P.H. *Proc. Natl. Acad. Sci. U. S. A.* (1987) **84**: 2117-2124.

Duwat, P., Ehrlich, S.D. and Gruss, A. *Appl. Environ. Microbiol.* (1992) **58**: 2674-2678.

Duwat, P., de Oliveira, R., Ehrlich, S.D. and Boiteux, S. *Microbiology.* (1995) **141**: 411-417.

Dybkaer, R. and Jorgensen, K. *Quantities and units in Clinical Chemistry* (1967), including Recommendations (1966) of the Commission on Clinical Chemistry of the International Union of Pure and Applied Chemistry and of the International Federation for Clinical Chemistry. Munksgaard, Copenhagen.

Engelward, B.P., Boosalis, M.S., Chen, B.J., Deng, Z. and Siciliano, M.J. *Carcinogenesis* (1993) **14**: 175-181.

Enterline, H.T. *Curr. Top. Pathol.* (1976) **63**: 97-112.

Fajen, J.M., Carson, G.A., Rounbehler, D.P., Fan, T.Y., Vitar, R., Goff, V.E., Wolf, M.H., Edwards, G.S., Fine, D.H. and Reinhold, V. *Science* (1979) **205**: 1262-1264.

Fan, C.Y., Potter, P.M., Rafferty, J., Watson, A.J., Cawkwell, L., Searle, P.F., O'Connor, P.J. and Margison, G.P. *Nucl. Acids Res.* (1990) **18**: 5723-5727.

Fan, C.Y. (1991) *PhD. Thesis.* University of Manchester.

Fan, T.Y., Morrison, J., Rounbehler, D.P., Ross, R. and Fine, D.H. *Science* (1977) **196**: 70-71.

Faux, S.P., Francis, J.E., Smith, A.G. and Chipman, J.K. *Carcinogenesis* (1992a) **13**: 247-250.

Faux, S.P., Gao, M., Chipman, J.K. and Levy, L.S. *Carcinogenesis* (1992b) **13**: 1667-1669.

Fearon, E.R. and Vogelstein, B. *Cell* (1990) **61**: 759-767.

Felgner, P.L., Gadek, T.R., Holm, M., Roman, R., Chan, H.W., Wenz, M., Northrop, J.P., Ringold, G.M. and Danielsen, M. *Proc. Natl. Acad. Sci. U. S. A.* (1987) **84**: 7413-7417.

Ferris, S.D., Sage, R.D. and Wilson, A.C. *Nature* (1982) **295**: 163-165.

Field, J.K. and Spandidos, D.A. *Anticancer Res.* (1990) **10**: 1-22.

Fine, D.H. *Oncology* (1980) **37**: 199-202.

Fishel, R., Lescoe, M.K., Rao, M.R., Copeland, N.G., Jenkins, N.A., Garber, J., Kane, M. and Kolodner, R. *Cell* (1993) **75**: 1027-1038.

Floyd, R.A., Watson, J.J., Wong, P.K., Altmiller, D.H. and Rickard, R.C. *Free Radic. Res. Commun.* (1986) **1**: 163-172.

Floyd, R.A., West, M.S., Eneff, K.L. and Schneider, J.E. *Arch. Biochem. Biophys.* (1989) **273**: 106-111.

Floyd, R.A. *Carcinogenesis* (1990) **11**: 1447-1450.

Floyd, R.A., West, M.S., Eneff, K.L. and Schneider, J.E. *Free Radic. Biol. Med.* (1990) **8**: 327-330.

Fowler, G.J.S., Rees, R.C. and Devonshire, R. *Photochem. Photobiol.* (1990) **52**: 489-494.

Fox, M., Brennand, J. and Margison, G.P. *Mutagenesis* (1987) **2**: 491-496.

Fraga, C.G., Shigenaga, M.K., Park, J.W., Degan, P. and Ames, B.N. *Proc. Natl. Acad. Sci. U. S. A.* (1990) **87**: 4533-4537.

Freeman, R.G., Murtishaw, W. and Knox, J.M. *J. Invest. Dermatol.* (1970) **54**: 164-169.

Freshney, R.I. *Culture of animal cells,* 2nd edn. (1987) Alan R. Liss, Inc., New York, NY.

Fridovich, I. *Adv. Enzymol. Relat. Areas. Mol. Biol.* (1974) **41**: 35-97.

Friedberg, E.C. *DNA Repair* (1984) W.H. Freeman, New York, NY.

Friedberg, E.C. *J. Cell Sci. Suppl.* (1987) **6**: 1-23.

Friedman, S. *J. Biol. Chem.* (1985) **260**: 5698-5705.

Friedmann, T. and Brown, D.M. *Nucl. Acids Res.* (1978) **5**: 615-622.

Fritz, G. and Kaina, B. *Biochim. Biophys. Acta* (1992) **1171**: 35-40.

Frosina, G. and Laval, F. *Carcinogenesis* (1987) **8**: 91-95.

Fucharoen, S., Fucharoen, G., Fucharoen, P. and Fukimaki, Y. *J. Biol. Chem.* (1989) **264**: 7780-7783.

Fuciarelli, A.F., Wegher, B.J., Gajewski, E. and Dizdaroglu, M. *Radiat. Res.* (1989) **119**: 219-231.

Fuciarelli, A.F., Wegher, B.J., Blakely, W.F. and Dizdaroglu, M. *Int. J. Radiat. Biol.* (1990) **58**: 397-415.

Furuichi, M., Yoshida, M.C., Oda, H., Tajiri, T., Nakabeppu, Y., Tsuzuki, T. and Sekiguchi, M. *Genomics* (1994) **24**: 485-490.

Gates, A.R. Maximising yield and developmental uniformity of eggs, pp.64-76. In: *Methods in mammalian embryology.* (1971) Ed: Daniel, J.C. W.H. Freeman, San Francisco, CA.

Gautier, C., Mehtali, M. and Lathe, R. *Nucl. Acids Res.* (1989) **17**: 8389.

Geiser, A.G. and Stanbridge, E.J. *Crit. Rev. Oncog.* (1989) **1**: 261-276.

Gelfand, D.H. and White, T.J. Chapter 16: Thermostable DNA polymerases, pp.13-20. In: *PCR Protocols,* 1st edn. (1990) Eds: Innis, M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J. Academic Press, Inc., San Diego, CA.

Gendron Maguire, M. and Gridley, T. *Methods Enzymol.* (1993) **225**: 794-799.

Gill, W.B., Taja, A., Chadbourne, D.M., Roma, M. and Vermeulen, C.W. *J. Urol.* (1987) **138**: 1318-1320.

Glauert, H.P., Schwarz, M. and Pitot, H.C. *Carcinogenesis* (1986) **7**: 117-121.

Goodenow, M., Huet, T., Saurin, W., Kwok, S., Sninsky, J. and Wain-Hobson, S. *J. AIDS* (1989) **2**: 344-352.

Gordon, J.W. and Ruddle, F.H. *Science* (1981) **214**: 1244-1246.

Gordon, J.W. *Methods Enzymol.* (1993) **225**: 747-771.

Graves, R.J., Felzenszwalb, I., Laval, J. and O'Connor, T.R. *J. Biol. Chem.* (1992b) **267**: 14429-14435.

Graves, R.J., Laval, J. and Pegg, A.E. *Carcinogenesis* (1992a) **13**: 1455-1459.

Groden, J., Thliveris, A., Samovitz, W., Carlson, M., Gelbert, L., Albertsen, H., Joslyn, G., Stevens, J., Spirio, L. and Robertson, M. *Cell* (1991) **66**: 589-600.

Guschlbauer, W., Duplaa, A.M., Guy, A., Teoule, R. and Fazakerley, G.V. *Nucl. Acids Res.* (1991) **19**: 1753-1758.

Gutteridge, J.M. and Halliwell, B. *Baillieres. Clin. Haematol.* (1989) **2**: 195-256.

Hall, J., Bresil, H., Serres, M., Martel Planche, G. and Wild, C.P. *Cancer Res.* (1990) **50**: 5426-5430.

Hall, J., Bresil, H., Donato, F., Wild, C.P., Loktionova, N.A., Komyakov, I.P., Lemekhov, V.G., Likhachev, A.J. and Montesano, R. *Int. J. Cancer* (1993) **54**: 728-733.

Halliwell, B. and Gutteridge, J.M. *Arch. Biochem. Biophys.* (1986) **246**: 501-514.

Halliwell, B. and Gutteridge, J.M. *Free Radicals in Biology and Medicine,* 2nd edn. (1989) Clarendon, Oxford.

Hammond, A.W. and D'Alessio, J.M. *Focus* (1986) **8**: 4-6.

Hampson, I.N., Pope, L., Cowling, G.J. and Dexter, T.M. *Nucl. Acids Res.* (1992) **20**: 2899.

Hanahan, D. *J. Mol. Biol.* (1983) **166**: 557-580.

Harris, L.C. (1990) *PhD. Thesis.* University of Manchester.

Harris, L.C., Potter, P.M. and Margison, G.P. *Biochem. Biophys. Res. Commun.* (1992) **187**: 425-431.

Hecht, S.S., Adams, J.D., Numoto, S. and Hoffmann, D. *Carcinogenesis* (1983) **4**: 1287-1290.

Hennings, H., Michael, D. and Patterson, E. *Proc. Soc. Exp. Biol. Med.* (1978) **158**: 1-4.

Hennings, H. and Yuspa, S.H. *J. Natl. Cancer Inst.* (1985) **74**: 735-740.

Henwood, J.M. and Heel, R.C. *Drugs* (1988) **36**: 429-454.

Higgins, D.M. and Anderson, R.N. *Arch. Pathol.* (1931) **12**: 186-202.

Hill, M.J., Hawksworth, G.M. and Tattersall, G. *Br. J. Cancer* (1973) **28**: 562-567.

Hill, R.B., Bensch, K.G. and King, D.W. *Exp. Cell Res.* (1960) **21**: 106-117.

Hirayama, T. (1979) *Gastric Cancer,* Ed.: Pfeiffer, C.J. New York, NY. pp.60-82.

Hodgson, S.V. and Maher, E.R. Chapter 1: General principles of cancer genetics, pp.1-19. In: *A Practical Guide to Human Cancer Genetics,* 1st edn. (1993) Cambridge University Press, Cambridge.

Hoffbrand, A.V. and Pettit, J.E. Chapter 10: Myeloproliferative Disorders, pp.182-193. In: *Essential Haematology,* 2nd edn. (1984) Blackwell Scientific Publications, Oxford.

Hogan, B., Costantini, F.D., and Lacy, E. *Manipulating the Mouse Embryo* (1986) Cold Spring Harbor Laboratory Press, New York, NY.

Hsie, A.W., Recio, L., Katz, D.S., Lee, C.Q., Wagner, M. and Schenley, R.L. *Proc. Natl. Acad. Sci. U. S. A.* (1986) **83**: 9616-9620.

Hsuan, J.J. *Anticancer Res.* (1993) **13**: 2521-2532.

Hueper, W.C., Wiley, F.H. and Wolfe, H.D. *J. Ind. Hyg. Toxicol.* (1938) **20**: 46-84.

Hunter, J.C., Burk, D. and Woods, M.W. *J. Natl. Cancer Inst.* (1967) **39**: 587-593.

Hutchison, F. *Prog. Nucleic Acid Res. Mol. Biol.* (1985) **32**: 115-154.

Ibeanu, G., Hartenstein, B., Dunn, W.C., Chang, L.Y., Hofmann, E., Mitra, S. and Kaina, B. *Carcinogenesis* (1992) **13**: 1989-1995.

Innis, M.A. and Gelfand, D.H. Chapter 1: Optimization of PCRs, pp.3-12. In: *PCR Protocols,* 1st edn. (1990) Eds: Innis, M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J. Academic Press, Inc., San Diego, CA.

Ishikawa, T., Takayama, S. and Kitagawa, T. *Cancer Res.* (1980) **40**: 4261-4264.

Ito, T. *Photochem. Photobiol.* (1978) **28**: 493-508.

Ivey, D.M. *Nucl. Acids Res.* (1990) **18**: 5882.

Jaenisch, R. *Proc. Natl. Acad. Sci. U. S. A.* (1976) **73**: 1260-1264.

Jagelman, D.G., DeCosse, J.J. and Bussey, H.J.R. *Lancet* (1988) **i**: 1149-1151.

Joshi, P.C. *Toxicol. Lett.* (1985) **26**: 211-217.

Kaina, B., Van Zeeland, A.A., Backendorf, C. and Thielmann, H.W. *Mol. Cell Biol.* (1987) **7**: 2024-2030.

Kaina, B., Fritz, G. and Coquerelle, T. *Environ. Mol. Mutagen.* (1993) **22**: 283-292.

Kalnik, M.W., Kouchakdjian, M., Li, B.F.L., Swann, P.F. and Patel, D.J. *Biochemistry* (1988) **27**: 100-115.

Kalnik, M.W., Li, B.F.L., Swann, P.F. and Patel, D.J. *Biochemistry* (1989) **28**: 6170-6192.

Karran, P., Lindahl, T., Ofsteng, I., Evensen, G.B. and Seeberg, E. *J. Mol. Biol.* (1980) **140**: 101-127.

Karran, P., Hjelmgren, T. and Lindahl, T. *Nature* (1982) **296**: 770-773.

Karran, P. and Hall, J. *Ann. Ist. Super. Sanita.* (1989) **25**: 21-26.

Kasai, H., Tanooka, H. and Nishimura, S. *Gann.* (1984) **75**: 1037-1039.

Kasai, H. and Nishimura, S. *Nucl. Acids Res.* (1984) **12**: 2137-2145.

Kasai, H., Crain, P.F., Kuchino, Y., Nishimura, S. and Ootsuyama, A. *Carcinogenesis* (1986) **7**: 1849-1851.

Kasai, H. and Nishimura, S. *Environ. Health Perspect.* (1986) **67**: 111-116.

Kasai, H., Nishimura, S., Kurokawa, Y. and Hayashi, Y. *Carcinogenesis* (1987) **8**: 1959-1961.

Kasai, H. and Nishimura, S. Chapter 4: Formation of 8-Hydroxydeoxyguanosine in DNA by Oxygen Radicals and its Biological Significance, pp.99-116. In: *Oxidative Stress: Oxidants and Antioxidants.* (1991) Academic Press Ltd, New York, NY.

Kasprzak, K.S., Diwan, B.A., Rice, J.M., Misra, M., Riggs, C.W., Olinski, R. and Dizdaroglu, M. *Chem. Res. Toxicol.* (1992) **5**: 809-815.

Kawabata, T., Matsui, M., Ishibashi, T. and Hamano, M. *IARC Sci. Publ.* (1983) **57**: 25-31.

Kawasaki, E.S. Chapter 19: Sample preparation from blood, cells, and other fluids, pp.146-152. In: *PCR Protocols,* 1st edn. (1990) Eds: Innis, M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J. Academic Press, Inc., San Diego, CA.

Kennaway, E.L. and Heiger, I. *Br. Med. J.* (1930) **1**: 1044-1046.

Kessel, D. *Photochem. Photobiol.* (1984) **39**: 851-859.

Kita, T., Brown, M.S. and Goldstein, J.L. *J. Clin. Invest.* (1980) **66**: 1094-1100.

Kleemann, D. *Laryngo-Rhino-Otol.* (1990) **69**: 437-439.

Klungland, A., Bjoras, M., Hoff, E. and Seeberg, E. *Nucl. Acids Res.* (1994) **22**: 1670-1674.

Knudsen, A.G. *Proc. Natl. Acad. Sci. U. S. A.* (1971) **68**: 820-823.

Koga, T., Shimada, Y., Kuroda, M., Tsujita, Y., Hasegawa, K. and Yamazaki, M. *Biochim. Biophys. Acta* (1990) **1045**: 115-120.

Koga, T., Fukuda, K., Shimada, Y., Fukami, M., Koike, H. and Tsujita, Y. *Eur. J. Biochem.* (1992) **209**: 315-319.

Kohda, K., Tada, M., Kasai, H., Nishimura, S. and Kawazoe, Y. *Biochem. Biophys. Res. Commun.* (1986) **139**: 626-632.

Kohda, K., Tada, M., Hakura, A., Kasai, H. and Kawazoe, Y. *Biochem. Biophys. Res. Commun.* (1987) **149**: 1141-1148.

Konig, K., Bockhorn, V., Dietel, W. and Schubert, H. *J. Cancer Res. Clin. Oncol.* (1987) **113**: 301-303.

Kouchakdjian, M., Bodepudi, V., Shibutani, S., Eisenberg, M., Johnson, F., Grollman, A.P. and Patel, D.J. *Biochemistry* (1991) **30**: 1403-1412.

Kriek, E. and Westra, J.G. *Carcinogenesis* (1980) **1**: 459-468.

Kroger, H., Gratz, R. and Grahn, H. *Int. J. Biochem.* (1984) **16**: 1387-1390.

Krulewski, T.F., Neumann, P.E. and Gordon, J.W. *Proc. Natl. Acad. Sci. U. S. A.* (1989) **86**: 3709-3712.

Kuchino, Y., Mori, F., Kasai, H., Inoue, H., Iwai, S., Miura, K., Ohtsuka, E. and Nishimura, S. *Nature* (1987) **327**: 77-79.

Kurokawa, Y., Hayashi, Y., Maekawa, A., Takahashi, M. and Kokubo, T. *J. Natl. Cancer Inst.* (1983) **71**: 965-971.

Lacy, E., Roberts, S., Evans, E.P., Burtenshaw, M.D. and Costantini, F.D. *Cell* (1983) **34**: 343-358.

Laemmli, U.K. *Nature* (1970) **227**: 680-685.

Lamberts, J.J. and Neckers, D.C. *J. Am. Chem. Soc.* (1983) **105**: 7465-7469.

Laval, F. and Laval, J. *Proc. Natl. Acad. Sci. U. S. A.* (1984) **81**: 1062-1066.

Laval, F. *Nucl. Acids Res.* (1994) **22**: 4943-4946.

Laval, J., Lopes, F., Madelmont, J.C., Godeneche, D., Meyniel, G., O'Connor, T.R. and Boiteux, S. *IARC. Sci. Publ.* (1991), pp.412-416.

Lawley, P.D. and Brookes, P.D. *Nature* (1961) **192**: 1081-1082.

Lawley, P.D. *Chemical Carcinogens* (1976), pp.83-244. Ed.: Searle, C.E. American Chemical Society, Washington, DC.

Lee, W.H., Bookstein, R., Hong, F., Young, L.J., Shew, J.Y. and Lee, E.Y. *Science* (1987) **235**: 1394-1399.

Lefebvre, P., Zak, P. and Laval, F. *DNA Cell Biol.* (1993) **12**: 233-241.

Leonard, G.A., Thomson, J., Watson, W.P. and Brown, T. *Proc. Natl. Acad. Sci. U. S. A.* (1990) **87**: 9573-9576.

Leonard, G.A., Guy, A., Brown, T., Teoule, R. and Hunter, W.N. *Biochemistry* (1992) **31**: 8415-8420.

Leppert, M., Dobbs, M., Scambler, P., O'Connell, P., Nakamura, Y., Stauffer, D., Woodward, S., Burt, R., Hughes, J. and Gardner, E. *Science* (1987) **238**: 1411-1413.

Lesko, S.A., Lorentzen, R.J. and Ts'o, P.O. *Biochemistry* (1980) **19**: 3023-3028.

Lewis, M.R. *Cancer* (1935) **25**: 305-310.

Lewis, M.R. *Anat. Rec.* (1945) **91**: 199-212.

Li, F.P. and Fraumeni, J.F. *Ann. Int. Med.* (1969) **71**: 747-752.

Li, F.P., Correa, P. and Fraumeni, J.F. *Cancer Epidemiol.* (1991) **1**: 91-94.

Lijinsky, W. and Kovatch, R.M. *Carcinogenesis* (1985) **6**: 1679-1681.

Lindahl, T. and Nyberg, B. *Biochemistry* (1972) **11**: 3610-3618.

Lindahl, T. and Karlstrom, O. *Biochemistry* (1973) **12**: 5151-5154.

Liscum, L., Luskey, K.L., Chin, D.J., Ho, Y.K., Goldstein, J.L. and Brown, M.S. *J. Biol. Chem.* (1983) **258**: 8450-8455.

Liu, L., Allay, E., Dumenco, L.L. and Gerson, S.L. *Cancer Res.* (1994) **54**: 4648-4652.

Loveless, A. *Nature* (1969) **223**: 206-207.

Maki, H. and Sekiguchi, M. *Nature* (1992) **355**: 273-275.

Margison, G.P. and O'Connor, P.J. *Chemical Carcinogens and DNA. Vol. I*, pp.111-159. (1979) Ed.: Grover, P.L. CRC Press, Boca Raton.

Margison, G.P. and Pegg, A.E. *Proc. Natl. Acad. Sci. U. S. A.* (1981) **78**: 861-865.

Margison, G.P., Cooper, D.P. and Brennand, J. *Nucl. Acids Res.* (1985) **13**: 1939-1952.

Margison, G.P. and Brennand, J. *J. Cell Sci. Suppl.* (1987) **6**: 83-96.

Margison, G.P., Brennand, J., Ockey, C.H. and O'Connor, P.J. *Bioessays* (1987) **6**: 151-156.

Material Safety Data Sheet. (1992) *Aldrich Chemical Co., UK.*

Matsukuma, S., Nakatsuru, Y., Nakagawa, K., Utakoji, T., Sugano, H., Sekiguchi, M. and Ishikawa, T. *Mutat. Res.* (1989) **218**: 197-206.

McCarthy, J.G., Edington, B.V. and Schendel, P.F. *Proc. Natl. Acad. Sci. U. S. A.* (1983) **80**: 7380-7384.

Meisler, M.H. *Trends. Genet.* (1992) **8**: 341-344.

Michaels, M.L., Pham, L., Cruz, C. and Miller, J.H. *Nucl. Acids Res.* (1991) **19**: 3629-3632.

Michaels, M.L. and Miller, J.H. *J. Bacteriol.* (1992) **174**: 6321-6325.

Michaels, M.L., Cruz, C., Grollman, A.P. and Miller, J.H. *Proc. Natl. Acad. Sci. U. S. A.* (1992) **89**: 7022-7025.

Michaels, M.L., Tchou, J., Grollman, A.P. and Miller, J.H. *Biochemistry* (1992) **31**: 10964-10968.

Middlebrook, M. Chapter 3: The Somme and the Germans, pp.60-62. In: *The First Day on the Somme - 1st July 1916.* (1984) Penguin Books Ltd., London.

Miller, E.C. *Cancer Res.* (1978) **38**: 1479-1496.

Miller, J.A. and Miller, E.C. *Prog. Exp. Tumour Res.* (1969) **11**: 273-301.

Miller, J.A. *Cancer Res.* (1970) **30**: 559-576.

Miyaki, M., Konishi, M., Kikuchi Yanoshita, R., Enomoto, M., Igari, T., Tanaka, K., Muraoka, M., Takahashi, H., Amada, Y. *Cancer Res.* (1994) **54**: 3011-3020.

Montesano, R. *J. Supramol. Struct. Cell Biochem.* (1981) **17**: 259-273.

Morgan, S.E., Kelley, M.R. and Pieper, R.O. *J. Biol. Chem.* (1993) **268**: 19802-19809.

Moriya, M. *Proc. Natl. Acad. Sci. U. S. A.* (1993) **90**: 1122-1126.

Morse, H.C. Chapter 1: The laboratory mouse - A historical perspective, pp.1-16. In: *The mouse in biomedical research. History, genetics and wild mice. (Vol. 1)* (1981) Ed: Foster, H.L. Academic Press, New York, NY.

Morson, B.C. Polyps in cancer of the large bowel, p.101. In: *The Gastrointestinal Tract.* (1977) Eds: Yardley, J.H., Morson, B.C., and Abell, M.R. Williams and Wilkins Company, Baltimore.

Morten, J.E., Bayley, L., Watson, A.J., Ward, T.H., Potter, P.M., Rafferty, J.A. and Margison, G.P. *Carcinogenesis* (1992) **13**: 483-487.

Morten, J.E. and Margison, G.P. *Carcinogenesis* (1988) **9**: 45-49.

Muller, E., Boiteux, S., Cunningham, R.P. and Epe, B. *Nucl. Acids Res.* (1990) **18**: 5969-5973.

Muller, W. and Crothers, D.M. *Eur. J. Biochem.* (1975) **54**: 267-291.

Mulligan, R.C. *Experimental Manipulation of Gene Expression* (1983), pp.155-173. Ed.: Inouye, M. Academic Press Inc., New York, NY.

Nackerdien, Z., Kasprzak, K.S., Rao, G., Halliwell, B. and Dizdaroglu, M. *Cancer Res.* (1991) **51**: 5837-5842.

Nghiem, Y., Cabrera, M., Cupples, C.G. and Miller, J.H. *Proc. Natl. Acad. Sci. U. S. A.* (1988) **85**: 2709-2713.

O'Connell, J.F., Klein-Szanto, A.J.P., DiGiovanni, D.M., Fries, J.W. and Slaga, T.J. *Cancer Res.* (1986) **46**: 2863-2865.

O'Connor, T.R., Boiteux, S. and Laval, J. *Nucl. Acids Res.* (1988) **16**: 5879-5894.

O'Connor, T.R., Boiteux, S. and Laval, J. *Ann. Ist. Super. Sanita.* (1989) **25**: 27-31.

O'Connor, T.R. and Laval, J. *Biochem. Biophys. Res. Commun.* (1991) **176**: 1170-1177.

O'Connor, T.R., Graves, R.J., de Murcia, G., Castaing, B. and Laval, J. *J. Biol. Chem.* (1993) **268**: 9063-9070.

Oster, G., Bellin, J.S., Kimball, R.W. and Schrader, M.E. *J. Am. Chem. Soc.* (1959) **81**: 5095-5098.

Palmiter, R.D., Brinster, R.L., Hammer, R.E., Trumbauer, M.E., Rosenfeld, M.G., Birnberg, N.C. and Evans, R.M. *Nature* (1982) **300**: 611-615.

Palmiter, R.D. and Brinster, R.L. *Cell* (1985) **41**: 343-345.

Palmiter, R.D. and Brinster, R.L. *Annu. Rev. Genet.* (1986) **20**: 465-499.

Paris, J.A. (1820) *Pharmacologia,* London.

Patel, S.T. and Lott, J.A. *Clin. Chem.* (1985) **31**: 993.

Pegg, A.E. *Adv. Cancer Res.* (1977) **25**: 195-269.

Pegg, A.E. *Rev. Biochem. Toxicol.* (1983) **5**: 83-133.

Pegg, A.E. and Byers, T.L. *FASEB J.* (1992) **6**: 2302-2310.

Pitot, H.C., Barsness, L., Goldsworthy, T. and Kitagawa, T. *Nature* (1978) **271**: 456-459.

Pitot, H.C. (1986) *Fundamentals of Oncology,* 3rd edn. Dekker, New York, NY.

Pott, P. In: *Chirurgical observations relative to the cataract, polypus of the nose, the cancer of the scrotum, the different kinds of ruptures and the mortification of the toes and feet.* (1775) London.

Potter, P.M., Wilkinson, M.C., Fitton, J., Carr, F.J., Brennand, J. and Margison, G.P. *Nucl. Acids Res.* (1987) **15**: 9177-9193.

Potter, V.R. *Environ. Health Perspect.* (1983) **50**: 139-148.

Preussmann, R. and Stewart, B.W. *Chemical Carcinogens,* 2nd edn. (1984), pp.643-828. Ed.: Searle, C.E. American Chemical Society, Washington, DC.

Pribnow, D. *Proc. Natl. Acad. Sci. U. S. A.* (1975) **72**: 784-788.

Price, A. and Lindahl, T. *Biochemistry* (1991) **30**: 8631-8637.

Raab, O. *Z. Biol.* (1900) **39**: 524-531.

Rafferty, J.A., Elder, R.H., Watson, A.J., Cawkwell, L., Potter, P.M. and Margison, G.P. *Nucleic. Acids Res.* (1992) **20**: 1891-1895.

Rafferty, J.A., Tumelty, J., Skorvaga, M., Elder, R.H., Margison, G.P. and Douglas, K.T. *Biochem. Biophys. Res. Commun.* (1994) **199**: 285-291.

Rainen, L. and Stollar, B.D. *Nucleic. Acids Res.* (1978) **5**: 3877-3889.

Rebeck, G.W., Coons, S., Carroll, P. and Samson, L. *Proc. Natl. Acad. Sci. U. S. A.* (1988) **85**: 3039-3043.

Reddy, M.V., Irvin, T.R. and Randerath, K. *Mutat. Res.* (1985) **152**: 85-96.

Refolo, L.M., Conley, M.P., Sambamurti, K., Jacobsen, J.S. and Humayun, M.Z. *Proc. Natl. Acad. Sci. U. S. A.* (1985) **82**: 3096-3100.

Rhiel, E. and Bryant, D.A. *Plant Physiol.* (1993) **101**: 701-702.

Ribeiro, D.T., Bourre, F., Sarasin, A., DiMascio, P. and Menck, C.F.M. *Nucl. Acids Res.* (1992) **20**: 2465-2469.

Rinchik, E.M., Bell, J.A., Hunsicker, P.R., Friedman, J.M., Jackson, I.J. and Russell, L.B. *Genetics* (1994) **137**: 845-854.

Roberts, J.J. *Adv. Rad. Biol.* (1978) **77**: 211-236.

Robson, C.N. and Hickson, I.D. *Nucleic. Acids Res.* (1991) **19**: 5519-5523.

Rowley, D.A. and Halliwell, B. *Biochim. Biophys. Acta* (1983a) **761**: 86-93.

Rowley, D.A. and Halliwell, B. *Clin. Sci.* (1983b) **64**: 649-653.

Saffhill, R., Margison, G.P. and O'Connor, P.J. *Biochim. Biophys. Acta* (1985) **823**: 111-145.

Sagher, D. and Strauss, B. *Biochemistry* (1983) **22**: 4518-4526.

Sai, K., Uchiyama, S., Ohno, Y., Hasegawa, R. and Kurokawa, Y. *Carcinogenesis* (1992) **13**: 333-339.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Ehrlich, H.A. and Arnheim, N. *Science* (1985) **230**: 1350-1354.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. *Science* (1988) **239**: 487-491.

Saiki, R.K. Chapter 2: Amplification of genomic DNA, pp.13-20. In: *PCR Protocols,* 1st edn. (1990) Eds: Innis, M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J. Academic Press, Inc., San Diego, CA.

Sakumi, K., Furuichi, M., Tsuzuki, T., Kakuma, T., Kawabata, S. and Sekiguchi, M. *J. Biol. Chem.* (1993) **268**: 23524-23530.

Sambamurti, K., Callahan, J., Luo, X., Perkins, C.P. and Jacobsen, J.S. *Genetics* (1988) **120**: 863-873.

Sambrook, K.J., Fritsch, E.F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual,* 2nd edn. (1989) Cold Spring Harbor Laboratory Press, New York, NY.

Sanger, F., Nicklen, S. and Coulson, A.R. *Proc. Natl. Acad. Sci. U. S. A.* (1977) **74**: 5463-5467.

Santibanez Koref, M., Elder, R.H., Fan, C.Y., Cawkwell, L., McKie, J.H., Douglas, K.T., Margison, G.P. and Rafferty, J.A. *Mol. Carcinog.* (1992) **5**: 161-169.

Schmerold, I. and Wiestler, O.D. *Cancer Res.* (1986) **46**: 245-249.

Schneider, J.E., Price, S., Maidt, L., Gutteridge, J.M. and Floyd, R.A. *Nucl. Acids Res.* (1990) **18**: 631-635.

Schreiber, J.H., Schisa, J.A. and Wilson, J.M. *Biotechniques* (1993) **14**: 818-823.

Schulte-Hermann, R. *Arch. Toxicol.* (1985) **57**: 147-158.

Sedgwick, B. *Mol. Gen. Genet.* (1983) **191**: 466-472.

Shibutani, S., Takeshita, M. and Grollman, A.P. *Nature* (1991) **349**: 431-434.

Shigenaga, M.K., Gimeno, C.J. and Ames, B.N. *Proc. Natl. Acad. Sci. U. S. A.* (1989) **86**: 9697-9701.

Shih, C.A. and Lin, J.K. *Carcinogenesis* (1993) **14**: 709-712.

Shine, J. and Dalgarno, L. *Eur. J. Biochem.* (1975) **57**: 221-230.

Shinkarenko, N.V. and Aleskovskii, V.B. *Russian Chem. Rev.* (1981) **50**: 220-228.

Sies, H. and Mehlhorn, R. *Arch. Biochem. Biophys.* (1986) **251**: 393-396.

Singer, B. and Kusmierek, J.T. *Annu. Rev. Biochem.* (1982) **51** 655-693.

Singer, B. *Cancer Invest.* (1984) **2**: 233-238.

Sloane, D.L., Goodman, M.F. and Echols, H. *Nucleic. Acids Res.* (1988) **16**: 6465-6475.

Snell, G.D. Congenic resistant strains of mice. In: *Origins of inbred mice.* (1978), pp.119-155. Ed: Morse, H.C. Academic Press, New York, NY.

Soriano, P., Gridley, T. and Jaenisch, R. *Genes Dev.* (1987) **1**: 366-375.

Stair, R.K., Nelson, C.J. and Mellors, J.W. *J. Virol.* (1991) **65**: 6339-6342.

Stein, R., Sciaky Gallili, N., Razin, A. and Cedar, H. *Proc. Natl. Acad. Sci. U. S. A.* (1983) **80**: 2422-2426.

Stuhlmann, H., Jaenisch, R. and Mulligan, R.C. *J. Virol.* (1989) **63**: 4857-4865.

Sturt, E. and Smith, C.A. *Cytogenet. Cell Genet.* (1976) **17**: 212-220.

Sugimoto, K., Hagishita, Y. and Himeno, M. *J. Biol. Chem.* (1994) **269**: 24271-24276.

Swann, P.F. *Mutat. Res.* (1990) **233**: 81-94.

Tatematsu, M., Nagamine, Y. and Farber, E. *Cancer Res.* (1983) **43**: 5049-5058.

Tchou, J., Kasai, H., Shibutani, S., Chung, M.H., Laval, J. and Nishimura, S. *Proc. Natl. Acad. Sci. U. S. A.* (1991) **88**: 4690-4694.

Tchou, J., Michaels, M.L., Miller, J.H. and Grollman, A.P. *J. Biol. Chem.* (1993) **268**: 26738-26744.

Teoule, R. and Cadet, J. *Mol. Biol. Biochem. Biophys.* (1978) **27**: 171-203.

Terwilliger, J.D. and Ott, J. *Handbook of human genetic linkage.* (1994) The Johns Hopkins University Press, Baltimore, MD.

Thomas, G. *Curr. Opin. Oncol.* (1994) **6**: 406-412.

Thomas, M. and Davis, R.W. *J. Mol. Biol.* (1975) **91**: 315-328.

Tindall, K.R. and Kunkel, T.A. *Biochemistry* (1988) **27**: 6008-6013.

Todd, P.A. and Goa, K.L. *Drugs* (1990) **40**: 583-607.

Tomlinson, I.P.M. *Trends Ecol. Evol.* (1993) **8**: 107-110.

Treffers, H.P., Spinelli, V. and Belser, O.N. *Proc. Natl. Acad. Sci. U. S. A.* (1954) **40**: 1064-1068.

Tudek, B., Boiteux, S. and Laval, J. *Nucl. Acids Res.* (1992) **20**: 3079-3084.

Uhlenhopp, E.L. and Krasna, A.I. *Biochemistry* (1971) **10**: 3290-3295.

U.S. Dept. of Health and Human Services. *The Health Consequences of Smoking* (1982) Cancer, Washington, DC.

Valerie, K., Henderson, E.E. and deRiel, J.K. *Nucleic. Acids Res.* (1984) **12**: 8085-8096.

Varley, H., Gowenlock, A.H., and Bell, M. Chapter 11: Quality control, pp.344-347. In: *Practical Clinical Biochemistry,* 5th edn. (1980) William Heinemann Medical Books Ltd, London.

Vickers, S., Duncan, C.A., Chen, I.-W., Rosegay, A. and Duggan, D.E. *Drug Metabolism and Disposition* (1990) **18**: 138-145.

Vidmar, J.J. and Cupples, C.G. *Can. J. Microbiol.* (1993) **39**: 892-894.

Vieira, J. and Messing, J. *Gene* (1982) **19**: 259-268.

Voigt, J.M. and Topal, M.D. *Carcinogenesis* (1995) **16**: 1775-1782.

von Hofe, E. and Kennedy, A.R. *Radiat. Res.* (1991) **127**: 220-225.

von Hofe, E., Fairbairn, L. and Margison, G.P. *Proc. Natl. Acad. Sci. U. S. A.* (1992) **89**: 11199-11203.

von Sonntag, C. *Free Radic. Res. Commun.* (1987) **2**: 217-224.

Wagner, E.F., Stewart, T.A. and Mintz, B. *Proc. Natl. Acad. Sci. U. S. A.* (1981) **78**: 5016-5020.

Wagner, G., Pott, U., Bruckschen, M. and Sies, H. *Biochem. J.* (1988) **251**: 825-829.

Wagner, T.E., Hoppe, P.C., Jollick, J.D., Scholl, R.L., Hodinka, R.L. and Gault, J.B. *Proc. Natl. Acad. Sci. U. S. A.* (1981) **78**: 6376-6380.

Wallace, S.S. *Environ. Mol. Mutagen.* (1988) **12**: 431-477.

Walter, C.A., Lu, J., Bhakta, M., Mitra, S., Dunn, W., Herbert, D.C., Hoog, T., Garza, P., Adrian, G.S. and et al. *Carcinogenesis* (1993) **14**: 1537-1543.

Wang, J.Y., Knudsen, E.S. and Welch, P.J. *Adv. Cancer Res.* (1994) **64**: 25-85.

Ward, J.F. and Kuo, I. *Radiat. Res.* (1976) **66**: 485-498.

Weiher, H., Noda, T., Gray, D.A., Sharpe, A.H. and Jaenisch, R. *Cell* (1990) **62**: 425-434.

Weitzman, S.A., Weitberg, A.B., Clark, E.P. and Stossel, T.P. *Science* (1985) **227**: 1231-1233.

White, G.R., Ockey, C.H., Brennand, J. and Margison, G.P. *Carcinogenesis* (1986) **7**: 2077-2080.

Wilkinson, M.C., Cooper, D.P., Southan, C., Potter, P.M. and Margison, G.P. *Nucl. Acids Res.* (1990) **18**: 13-16.

Willis, R.A. *The Spread of Tumours in the Human Body* (1952) Butterworth and Company, London.

Wilson, R.E., Hoey, B. and Margison, G.P. *Carcinogenesis* (1993) **14**: 679-683.

Wood, M.L., Dizdaroglu, M., Gajewski, E. and Essigmann, J.M. *Biochemistry* (1990) **29**: 7024-7032.

Wright, N. Chapter 25: The liver, pp.896-932. In: *The biology of epithelial cell populations (Vol. 2)* (1984) Eds: Wright, N. and Alison, M. Oxford University Press, Oxford.

Wynder, E.L. and Gori, G.B. *J. Natl. Cancer Inst.* (1977) **58**: 825-832.

Xiao, W., Derfler, B., Chen, J. and Samson, L. *EMBO J.* (1991) **10**: 2179-2186.

Yamagawa, K. and Ichikawa, K. *Mitteilungen Med. Facultat. Kaiserl. Univ. Tokyo* (1915) **15** (**2**): 295-344.

Yanisch-Perron, C., Vieira, J. and Messing, J. *Gene* (1985) **33**: 103-110.

Yanofsky, C., Cox, E.C. and Horn, V. *Proc. Natl. Acad. Sci. U. S. A.* (1966) **55**: 274-281.

Ying, T.S., Enomoto, K., Sarma, D.S.R. and Farber, E. *Cancer Res.* (1982) **42**: 876-880.

Yisraeli, J. and Syzf, M. *DNA methylation and its biological significance* (1984), p.353. Eds: Razin, A., Cedar, H., and Riggs, A.D. Springer-Verlag, New York, NY.

Yu, D.S., Chang, S.Y. and Ma, C.P. *J. Urol.* (1990) **144**: 164-168.

Zubay, G. *Annu. Rev. Genet.* (1973) **7**: 267-287.

Zuccotti, M., Grant, M. and Monk, M. *Methods Enzymol.* (1993) **225**: 557-567.