Repair of oligodeoxyribonucleotides containing $O^6$-alkylguanine by MGMT variant proteins

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Medical and Human Sciences

2013

Pattama Senthong

School of Medicine
Institute of Population Health
Centre for Occupational and Environmental Health
Chapter 1: Introduction

1.1 Classification of alkylating agents

1.2 Biological effects of alkylating agents

1.3 DNA damage caused by alkylating agents

1.4 Biological effects of specific DNA alkylation products
   1.4.1 7-methylguanine
   1.4.2 3-methyladenine
   1.4.3 O\textsuperscript{6}-alkylguanine

1.5 Human exposure to alkylating agents
   1.5.1 Endogenous sources
   1.5.2 Exogenous sources
      1.5.2.1 Tobacco smoking
      1.5.2.2 Occupational exposures
      1.5.2.3 Diet
      1.5.2.4 Cancer chemotherapy

1.6 Repair of DNA damage caused by alkylating agents

1.7 O\textsuperscript{6}-alkylguanine-DNA alkyltransferase
   1.7.1 The human MGMT protein
   1.7.2 Repair of O\textsuperscript{6}-alkGs by MGMT
   1.7.3 MGMT substrate specificity
   1.7.4 MGMT expression levels and MGMT activity
   1.7.5 Effect of zinc ions on MGMT activity
1.7.6 Effect of 5-methylcytosine on the rate of repair of $O^6$-alkGs by MGMT

1.7.7 Effect of DNA sequence context on the rate of repair of $O^6$-alkGs by MGMT

1.7.8 Human variants of $O^6$-alkylguanine-DNA alkyltransferase

1.7.9 Repair of $O^6$-alkGs by human MGMT variants

1.7.10 Pseudosubstrate inactivation of human MGMT variants

1.7.11 Association between MGMT polymorphisms and cancer risk

1.8 Aims and Objectives

Chapter 2: Materials and Methods

2.1 Expression of extended and normal MGMT in human lymphocytes

2.1.1 DNA sequencing

2.2 Generation, expression and purification of extended and normal MGMT

2.2.1 Generation of MGMT by PCR

2.2.2 DNA ligation

2.2.3 Transformation

2.2.4 PCR colony amplification

2.2.5 Plasmid purification

2.2.6 Expression of recombinant MBP fusion proteins

2.2.7 Purification of MBP fusion proteins

2.2.8 Western blotting

2.2.9 Cleavage of MBP-MGMT fusions proteins

2.2.10 Purification of MGMT

2.3 Generation, expression and purification of MGMT variant proteins

2.3.1 Site directed mutagenesis

2.4 Generation, expression and purification of Ogt protein

2.5 Functional activity of MGMT and Ogt proteins

2.5.1 Determination of activity
2.5.2 Kinetics of methyl group transfer to MGMT 82
2.5.3 Thermal stability 82
2.5.4 MGMT and Ogt inactivation by modified ODNs 82
2.5.5 The effects on MGMT activity of adding zinc to E. coli growth media 84
2.5.6 Surface plasmon resonance (SPR) assay of the interaction between MGMT Wt and variant proteins with \( O^6 \)-alkG containing ODNs 84
2.5.7 Enzyme-linked immunosorbent assay (ELISA) analysis of MGMT binding to double-stranded ODNs 85
2.5.8 Effect of cytosine methylation on the inactivation of MGMT by \( O^6 \)-MeG 86
2.5.9 Mass spectroscopic analysis of MGMT proteins 86
2.5.10 Alkyl group removal by MGMT assessed by restriction endonuclease site deprotection 87
2.5.11 MGMT activity and \( O^6 \)-CMG levels in azaserine- treated human melanoma cells 88

2.6 Statistical analysis 89

Chapter 3: Characterisation of MGMT variant proteins 91

3.1 Introduction 91
3.2 Generation, expression and purification of MGMT variant proteins 91
  3.2.1 Isolation of variant MGMTs by PCR 91
  3.2.2 Ligation, transformation, colony PCR and sequencing 91
  3.2.3 Expression and purification of MBP-MGMT variant fusion proteins 96
  3.2.4 Western analysis of purified proteins 97
  3.2.5 Cleavage of purified MBP-MGMT fusion proteins 98
  3.2.6 Purification of MGMT 98
3.3 Functional activity of MGMT variant proteins 101
  3.3.1 Determination of MGMT activity of purified proteins 102
  3.3.2 Kinetics of methyl group transfer to MGMT 103
  3.3.3 Thermal stability 103
3.3.4 MGMT inactivation by $O^6$-alkG containing ODNs

3.3.5 Mass spectroscopic analysis of alkyl group transfer to normal MGMT protein

3.3.6 SPR analysis of the interaction between Wt MGMT and variant proteins with $O^6$-alkG containing ODNs

3.4 Discussion

Chapter 4: Characterisation of extended and normal MGMT

4.1 Introduction

4.2 Expression of extended and normal MGMT in human lymphocytes

4.3 Generation, expression and purification of extended and normal MGMT
   4.3.1 Isolation of extended and normal MGMT cDNAs by PCR
   4.3.2 Ligation, transformation, colony PCR and sequencing
   4.3.3 Expression and purification of MBP-MGMT fusion proteins
   4.3.4 Western analysis of MBP-MGMT fusion proteins
   4.3.5 Cleavage of purified MBP-MGMT fusion proteins
   4.3.6 Purification of MGMT proteins

4.4 Characterisation of purified extended and normal MGMT proteins
   4.4.1 Determination of MGMT activity
   4.4.2 Kinetics of methyl group transfer to MGMT
   4.4.3 Thermal stability
   4.4.4 MGMT inhibition by calf thymus DNA
   4.4.5 MGMT inactivation by $O^6$-alkG containing ODNs
   4.4.6 The effects on MGMT activity of adding ZnCl$_2$ to E.coli growth media
   4.4.7 Effect of ZnCl$_2$ supplementation on MGMT inactivation by $O^6$-alkG containing single and double stranded ODNs
   4.4.8 Enzyme-linked immunosorbent assay (ELISA) analysis of MGMT binding to double-stranded ODNs
   4.4.9 Effect of 5-methylcytosine in $O^6$-MeG containing ODNs on the inactivation of extended and normal MGMT

4.5 Discussion
Chapter 5: \(O^6\)-carboxymethylguanine as a substrate for MGMT and the Ogt protein

5.1 Introduction 160
5.2 Alkyl group removal assessed by ELISA 160
5.3 Alkyl group removal assessed by restriction endonuclease site deprotection 163
5.4 Inactivation of MGMT and DNA damage in cultured human cells treated with azaserine 166
5.5 Ogt protein production, purification and characterisation 169
  5.5.1 Cloning, expression and purification Ogt 169
  5.5.2 Determination of Ogt methyl transfer activity 175
  5.5.3 Kinetics of methyl group transfer to Ogt protein 175
  5.5.4 Thermal stability 176
  5.5.5 Ogt inactivation by modified ODNs 177
5.6 Discussion 178

Chapter 6: Overall discussion 181

6.1 Further research 185
6.2 \(O^6\)-alkG adducts associated with occupational nitrosamine exposure and MGMT polymorphisms in Thailand’s rubber industry 187

References 189

Appendix 1: Thermal stability of variant MGMT proteins 211
Appendix 2: Site-directed mutagenesis, variant MGMT protein expression and purification 213
Appendix 3: Inactivation of MGMT variant proteins by ODNs 224
Appendix 4: Published paper 228

Word Count: 58,167
List of Figures

Figure 1.1: Structures of alkylating agents 27
Figure 1.2: Alkylation sites in DNA base-pairs 29
Figure 1.3: Metabolic activation of tobacco-specific nitrosamines and the formation of O\textsuperscript{6}-MeG and O\textsuperscript{6}-PobG 32
Figure 1.4: Mechanisms of the biological effects of O\textsuperscript{6}-MeG and their prevention by MGMT 34
Figure 1.5: Structure of tobacco alkaloids and the tobacco-specific nitrosamines 37
Figure 1.6: Human MGMT X-ray crystallographic structure, zinc binding site, and DNA binding mechanism 45
Figure 1.7: Primary amino acid sequence of the Wt MGMT, location of the amino acid changes in the polymorphic variant proteins and the codon changes responsible 51
Figure 1.8: Nucleotide sequence, exon structure and the location of polymorphisms in the MGMT cDNA 52
Figure 1.9: Linkage disequilibrium map of MGMT single nucleotide polymorphisms (SNPs) using Haploview version 4.2 55
Figure 1.10: The structure of MGMT showing the locations of the amino acids changed in the polymorphic variants 66
Figure 1.11: The comparison between Kozak consensus sequence and MGMT mRNA sequence around the initiation codon 68
Figure 2.1: Nucleotide sequence of part of the MGMT cDNA indicating the location (underlined) of the forward and reverse primers used in RT-PCR to assess expression of the extended and normal mRNAs in human peripheral blood mononuclear cells (PBMCs) 72
Figure 2.2: Schematic diagram of MGMT purification using amylose, DEAE-sepharose and a second amylose column after factor Xa digest 78
Figure 2.3: Schematic diagram of the directed mutagenesis procedure 79
Figure 2.4: Nucleotide sequence of the *E.coli* genome encompassing *ogt*, indicating the location (underlined) of the forward and reverse primers

Figure 2.5: Structures of the modified guanine residues in the ODN used in these studies

Figure 3.1: Agarose gel electrophoresis of PCR amplification products of pMAL-c2 Wt MGMT.

Figure 3.2: Agarose gel electrophoresis of PCR amplification products of 12 different *E.coli* colonies produced by transformation with the ligated pMAL-c2 F84 variant PCR products

Figure 3.3: Sequencing results and nucleotide sequence alignments of the MGMT variants

Figure 3.4: Expression of the F84 variant MBP-MGMT fusion protein

Figure 3.5: Purification of the F84 variant MBP-MGMT fusion protein by amylose affinity chromatography

Figure 3.6: Western blot of the purification fractions (indicated) for the F84 variant MBP-MGMT fusion proteins

Figure 3.7: Factor Xa cleavage of MBP-Wt MGMT fusion protein during incubation for various times

Figure 3.8: Factor Xa cleavage of MBP-MGMT fusion proteins

Figure 3.9: Purification of the Factor Xa cleaved F84 variant MBP-MGMT fusion protein by DEAE-sepharose

Figure 3.10: Re-purification of the F84 variant protein using an amylose column

Figure 3.11: Western blot of the amylose column re-purified fractions (indicated) for the Factor Xa-cleaved, DEAE-purified F84 variant MGMT protein

Figure 3.12: Transfer of [*3*H]-methyl groups from [*3*H]-methylated substrate DNA to MGMT with increasing amounts of Wt and variant MGMT proteins during incubation at 37°C for 1 hour.

Figure 3.13: The kinetics of methyl group transfer from substrate DNA by the Wt and variant MGMT proteins
Figure 3.14: Thermal stability of Wt and variant MGMT proteins during pre-incubation at A. 51°C, B. 44°C and C. 37°C

Figure 3.15: Half lives of Wt and variant MGMT proteins during pre-incubation at 44°C and 37°C

Figure 3.16: Effect of pre-incubation with a wide range concentration of twelve $O^6$-alkG containing ODNs on the remaining activity of the MGMT F84 variant

Figure 3.17: Effect of pre-incubation with specific concentration of six $O^6$-alkG containing ODNs on the remaining activity of the MGMT F84 variant

Figure 3.18: $IC_{50}$ values (nM) for the inactivation of the Wt and variant MGMT proteins by $O^6$-alkG containing ODNs

Figure 3.19: Inactivation of normal MGMT and variant proteins by five $O^6$-alkG containing ODNs

Figure 3.20: MALDI-TOF mass spectral analysis of tryptic digests of MGMT indicating the locations and masses of the fragments expected if alkyl group transfer occurred from ODNs containing: A. G, B. $O^6$-MeG, C. $O^6$-EtG, D. $O^6$-PrG, E. $O^6$-HOEtG, F. $O^6$-CMG, G. $O^6$-BzG, and H. $O^6$-PobG.

Figure 3.21: MALDI-TOF mass spectral analysis of tryptic digests of MGMT indicating the locations and masses of the fragments expected if alkyl group transfer occurred from ODNs containing: A. G, B. DAP, C. $O^6$-AEG, D. HOPr-DAP and E. $O^6$-MAG

Figure 3.22: SPR analysis of the binding to the immobilised $O^6$-alkG containing ODNs of A. MGMT protein followed by B. Atl1.

Figure 3.23: SPR analysis of the binding of Wt MGMT protein to the immobilised $O^6$-alkG containing ODNs

Figure 3.24: SPR analysis of the binding of F84 variant MGMT protein to the immobilised $O^6$-alkG containing ODNs

Figure 3.25: SPR analysis of the binding of Q128 variant MGMT protein to the immobilised $O^6$-alkG containing ODNs
Figure 3.26: SPR analysis of the binding of R160 variant MGMT protein to the immobilised $O^6$-alkG containing ODNs

Figure 3.27: SPR analysis of the binding to the immobilised $O^6$-alkG containing ODNs of A. V143/R178 and B. F84/R160 variant MGMT proteins

Figure 3.28: SPR analysis of the binding of the F84/V143/R178 variant MGMT protein to the immobilised $O^6$-alkG containing ODNs

Figure 3.29: SPR analysis of the binding of A. the V143/R178 MGMT variant to $O^6$-HOEtG containing ODN and B. the Q128 MGMT variant to $O^6$-PobG containing ODN

Figure 3.30: SPR analysis of the binding to $O^6$-BzG containing ODNs indicated of A. Wt MGMT showing complete removal of the benzyl group followed by no binding of Atl1. B. F84/V143/R178 MGMT showing most removal of the benzyl group followed by little binding of Atl1 and C. Q128 MGMT showing some removal of the benzyl group followed by the highest binding of Atl1

Figure 4.1: Amino acid sequence of MGMT

Figure 4.2: Schematic diagram of: A. Nucleotide sequence of the full length and putative truncated MGMT and B. amino acid sequence of the extended and normal MGMT

Figure 4.3: Scheme for investigating the expression of full length and truncated MGMT cDNAs

Figure 4.4: Agarose gel electrophoresis of PCR products obtained after amplification using the full length and truncated forward primers, and the common reverse primer, of cDNAs prepared from five human PBMC samples

Figure 4.5: Agarose gel electrophoresis of fragments of PCR-amplified full length and truncated MGMT after $SapI$ and $PstI$ digestion

Figure 4.6: Agarose gel electrophoresis of PCR products obtained after amplification of increasing amounts of pCMV6-AC-GFP vector using the truncated forward and common reverse primers
Figure 4.7: Agarose gel electrophoresis of the pCMV6-AC-GFP PCR amplification products for the full length and truncated MGMT cDNA

Figure 4.8: Agarose gel electrophoresis of PCR products of different *E.coli* clones produced by transformation using material generated by PCR using (A) full length and (B) truncated forward and common reverse primers

Figure 4.9: Nucleotide sequence alignments of the full length and truncated MGMT

Figure 4.10: Expression of MBP fusion proteins in *E.coli* extracts prepared at various times after addition of IPTG

Figure 4.11: Purification of the (A) extended and (B) normal MBP-MGMT fusions proteins by amylose affinity chromatography

Figure 4.12: Western blot of fractions obtained from the purification of MBP-MGMT fusion proteins

Figure 4.13: Factor Xa cleavage of extended and normal MBP-MGMT fusion proteins

Figure 4.14: DEAE-sepharose purification of MBP-MGMT factor Xa cleavage reaction products

Figure 4.15: Amylose column purification of the pooled fractions containing normal MGMT (purified as in Figure 4.14)

Figure 4.16: Western blot of the amylose column re-purification fractions for the factor Xa-cleaved, DEAE- purified normal MGMT protein.

Figure 4.17: Transfer of $[^3]H$-methyl groups from $[^3]H$-methylated substrate DNA to MGMT with increasing amounts of extended and normal MGMT proteins during incubation at 37°C for 1 hour.

Figure 4.18: The kinetics of methyl group transfer from substrate DNA to extended and normal MGMT

Figure 4.19: Thermal stability of MGMT proteins during pre-incubation at 44°C and 51°C
Figure 4.20: Inhibition of MGMT proteins by increasing concentrations of calf thymus DNA

Figure 4.21: Effect of pre-incubation with wide range concentration of $O^6$-alkG containing ODNs on the activity of extended and normal MGMT

Figure 4.22: Effect of pre-incubation with specific concentrations of $O^6$-alkG containing ODNs on the activity of extended and normal MGMT

Figure 4.23: IC$_{50}$ values (nM) for the inactivation of extended and normal MGMT proteins by ODNs in order of increasing potency

Figure 4.24: The effect of ZnCl$_2$ added to the *E. coli* growth media on the activity of purified MGMT proteins transfer [3H]-methyl group during incubation at 37ºC for 1 hour. Vertical lines indicate the protein limiting amounts of the individual proteins that were chosen for further studies.

Figure 4.25: Effect of pre-incubation with the $O^6$-MeG-containing single (ss) and double-stranded (ds) ODNs on the activity of extended and normal MGMT induced with and without ZnCl$_2$ in the *E. coli* growth media

Figure 4.26: Effect of ZnCl$_2$ added to the *E. coli* growth media on the activity of the purified MGMT proteins with single (ss) and double-stranded (ds) ODNs containing $O^6$-alkGs

Figure 4.27: Binding of MGMT (0-60 fmoles) to double-stranded (ds) ODNs assessed using an ELISA assay

Figure 4.28: Binding of MGMT (0-6 fmoles) to double-stranded (ds) ODNs assessed using an ELISA assay

Figure 4.29: MGMT binding to short and long double-stranded $O^6$-alkG containing ODNs (13 and 23 mer)

Figure 4.30: Effect of pre-incubation with the ODNs indicated on the activity of extended and normal MGMT

Figure 4.31: Effect of cytosine methylation on the IC$_{50}$ values of extended and normal MGMT proteins with modified ss and ds ODNs
Figure 5.1: Binding of ds $O^6$-alkG containing ODNs to SA-coated microtitre plates as determined by ELISA using A. anti-Atl1 antibody and B. anti-CMG antibody.

Figure 5.2: Binding of MGMT on ds $O^6$-alkG containing ODNs as determined by ELISA using A. anti-Atl1 antibody and B. anti-CMG antibody.

Figure 5.3: Location of the restriction endonuclease sites in the A.13-mer ODN and B. 23-mer ODN are indicated. The locations of the fluorescence tags, SIMA and HEX are shown in red.

Figure 5.4: Fluorescence image of non-denaturing PAGE gel of ds $O^6$-alkG containing ODNs, before (-) or after (+) digestion with A. NlaIII, B. CviKI, C. StyI and D. BsaJI.

Figure 5.5: Fluorescence images of non-denaturing PAGE of ds $O^6$-alkG containing ODNs.

Figure 5.6: MGMT activity in A375 cells extracts following exposure to azaserine.

Figure 5.7: MGMT specific activities (fmoles/µgDNA) calculated from the activity determined under protein-limiting conditions.

Figure 5.8: Effect of Atl1 pre-incubation with cell-free extracts of azaserine-treated A375 cells on their respective ability to bind to $O^6$-MeG-containing ODNs.

Figure 5.9: Effect of pre-incubation of anti-CMG antiserum with cell-free extracts of azaserine-treated A375 cells on their respective ability to bind to $O^6$-CMG-containing ODNs.

Figure 5.10: Effect of inactivation of MGMT by pre-incubation of A375 cells with Lomeguatrib on the levels of $O^6$-MeG and $O^6$-CMG after exposure to asazerine.

Figure 5.11: Agarose gel electrophoresis of the PCR amplification products of E.coli K12 genomic DNA.

Figure 5.12: Agarose gel electrophoresis of PCR amplification products of E.coli colonies produced by transformation with the ogt PCR products ligated into pMAL-c2.
Figure 5.13: Expression of MBP-Ogt fusion protein in extracts of *E.coli* prepared at various times after addition of IPTG.

Figure 5.14: Purification of the MBP-Ogt fusion protein analysed by 12% SDS-PAGE with Coomassie staining.

Figure 5.15: Western blot of the purification fractions of the MBP-Ogt fusion protein.

Figure 5.16: Factor Xa cleavage of MBP-Ogt fusion protein analysed by 12% SDS-PAGE with Coomassie staining.

Figure 5.17: Purification of the factor Xa cleaved MBP-Ogt fusion protein by DEAE-sepharose analysed by 12% SDS-PAGE with Coomassie staining.

Figure 5.18: Re-purification of Ogt using an amylose column.

Figure 5.19: Western blot of the amylose column re-purification fractions for the factor Xa-cleaved, DEAE-purified MBP-Ogt fusion protein.

Figure 5.20: Transfer of $[^3]$H-methyl groups from $[^3]$H-methylated substrate DNA to Ogt with increasing amounts of Ogt protein during incubation at 37°C for 1 hour. Vertical lines indicate the protein limiting amounts of the Ogt protein that was chosen for further studies.

Figure 5.21: The kinetics of methyl group transfer from substrate DNA.

Figure 5.22: Thermal stability of normal MGMT and Ogt protein during incubation at 44°C.

Figure 5.23: IC$_{50}$ values (nM) for the inactivation of normal MGMT and Ogt proteins by the five ss $O^6$-alkG containing ODNs.

Figure 5.24: Effect of pre-incubation of Ogt with ss $O^6$-alkG containing ODNs.
Appendix Figure 1A: Thermal stability of A. Wt, B. F84, C. Q128, D. R160, E. V143/R178, F. F84/R160 and G. F84/V143/R178 proteins during pre-incubation at 44°C.

Appendix figure 1B: Thermal stability of A. Wt, B. F84, C. Q128, D. R160, E. V143/R178, F. F84/R160 and G. F84/V143/R178 proteins during pre-incubation at 37°C.


Appendix figure 2B: Expression of the MBP-MGMT variant fusion protein in extracts of *E. coli* prepared at the times indicated after addition of IPTG. Extracts were analysed by 12% SDS-PAGE with Coomassie staining. A. Q128, B. R160 and F84/R160, C. V143/R178 and F84/V143/R178.

Appendix figure 2C: Purification of the MBP-MGMT variant fusion protein by amylose affinity chromatography. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining. A. Q128, B. R160, C. V143/R178, D. F84/R160 and E. F84/V143/R178.

Appendix figure 2D: Western blot of the purification fractions (indicated) for the MBP-MGMT variant fusion proteins. The membrane was probed with anti-MBP antibody and binding assessed by ECL. A. Q128, B. R160, C. V143/R178, D. F84/R160 and E. F84/V143/R178.

Appendix figure 2E: Purification of the Factor Xa cleaved Q128 variant MBP-MGMT fusion protein by DEAE-sepharose. Fractions indicated were analysed by 12% SDS-
PAGE with Coomassie staining. A. Elution with 25-75 mM NaCl, B. Elution with 100-500 mM NaCl.

Appendix figure 2F: Purification of the Factor Xa cleaved R160 variant MBP-MGMT fusion protein by DEAE-sepharose. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining. A. Elution with 25-75 mM NaCl, B. Elution with 100-500 mM NaCl.

Appendix figure 2G: Purification of the Factor Xa cleaved V143/R178 variant MBP-MGMT fusion protein by DEAE-sepharose. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining. A. Elution with 25-75 mM NaCl, B. Elution with 100-500 mM NaCl.

Appendix figure 2H: Purification of the Factor Xa cleaved F84/R160 variant MBP-MGMT fusion protein by DEAE-sepharose. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining. A. Elution with 25-75 mM NaCl, B. Elution with 100-500 mM NaCl.

Appendix figure 2I: Purification of the Factor Xa cleaved F84/V143/R178 variant MBP-MGMT fusion protein by DEAE-sepharose. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining. A. Elution with 25-75 mM NaCl, B. Elution with 100-500 mM NaCl.

Appendix figure 2J: Re-purification of the variant protein using an amylose column. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining. A. Q128, B. R160, C. V143/R178, D. F84/R160 and E. F84/V143/R178.

Appendix figure 2K: Western blot of the amylose column re-purified fractions (indicated) for the Factor Xa-cleaved, DEAE-purified MGMT variant protein. The
membrane was probed with anti-MGMT antibody.


Appendix Figure 3A: Effect of pre-incubation with 2.4 and 24.4 nM concentrations of twelve O\textsuperscript{6}-alkG containing ODNs on the remaining activity of MGMT variants.


Appendix Figure 3B: Effect of pre-incubation with specific concentration of six O\textsuperscript{6}-alkG containing ODNs on the remaining activity of MGMT variants. A. Q128, B. R160, C. V143/R178, D. F84/R160 and E. F84/ V143/R178.
List of Tables

Table 1.1: Alkylation products of DNA following reaction with alkylation agents
Table 1.2: Occupational exposure to nitrosamines
Table 1.3: N-nitroso compounds found in food and beverages
Table 1.4: Classification of chemotherapeutic alkylation agents and their mechanism
Table 1.5: The rate of repair by human MGMT of ODNs containing various $O^6$-alkGs
Table 1.6: Population prevalence of MGMT polymorphisms in exons 2, 3, 4 and 5
Table 1.7: Analysis of MGMT single nucleotide polymorphisms (SNPs) using Haploview version 4.2.
Table 1.8: Repair of $O^6$-MeG relative to other $O^6$-alkGs
Table 1.9: ODNs used to test the effect of sequence context on the rate of repair $O^6$-MeG/$O^6$-PobG
Table 1.10: Relative repair of $O^6$-MeG and $O^6$-PobG in different ODNs
Table 1.11: Inactivation of Wt and variants MGMT by 5 inactivators
Table 1.12: MGMT polymorphisms and risk of lung cancer
Table 1.13: MGMT polymorphisms and risk of endometrial, cervical, oral and pancreatic cancer
Table 1.14: MGMT polymorphisms and risk of head and neck, breast and colorectal cancer
Table 1.15: MGMT polymorphisms and risk of esophageal, bladder, prostate cancer and glioma
Table 2.1: Primary and secondary antibodies used in these studies
Table 2.2: Primers used to generate and sequence the MGMT variants
Table 2.3: Control (G) and $O^6$-alkG (X) containing ODNs and complementary ODNs used in ELISA assays
Table 2.4: ODNs used to assess the effect of 5-methylcytosine on the inactivation of MGMT by $O^6$-MeG
Table 3.1: Specific activities of Wt and variant MGMT proteins
Table 3.2: The ranking of the IC50 values (nM) for the inactivation of MGMT variant proteins by six O6-alkG containing ODNs

Table 3.3: Expected sizes of the tryptic fragments of MGMT containing the active site C145 with or without alkyl group covalent modification

Table 4.1: The difference in IC50 values between extended and normal MGMT by O6-alkG containing ODNs

Table 4.2: The specific activities of the extended and normal MGMT proteins with and without Zn and the protein limiting amounts

Table 4.3: p53 codon 245-spanning ODNs used to assess the effect of 5-methylcytosine on inactivation of MGMT by O6-MeG

Table 6.1: Summary of data for functional activity of MGMT variants compared with Wt protein

Table 6.2: Summary of data for functional activity comparison between extended and normal MGMT proteins

Table 6.3: Summary of data comparing functional activity of normal MGMT and Ogt proteins
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP site</td>
<td>Apurinic/apyrimidinic sites</td>
</tr>
<tr>
<td>APS</td>
<td>Ammonium persulphate</td>
</tr>
<tr>
<td>Atl</td>
<td>Alkyltransferase-like</td>
</tr>
<tr>
<td>Az</td>
<td>Azaserine</td>
</tr>
<tr>
<td>BB</td>
<td>Binding buffer</td>
</tr>
<tr>
<td>BCNU</td>
<td>1, 3-bis (2-chloroethyl)-1-nitrosourea</td>
</tr>
<tr>
<td>BER</td>
<td>Base Excision repair</td>
</tr>
<tr>
<td>BF</td>
<td>$O^4$-benzylfolic acid</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CCNU</td>
<td>1-(2-chloroethyl)-3-cyclohexyl-1-nitrosourea</td>
</tr>
<tr>
<td>CRC</td>
<td>Colorectal cancer</td>
</tr>
<tr>
<td>DAP</td>
<td>2, 6-diaminopurine</td>
</tr>
<tr>
<td>DEN</td>
<td>DiethylNitrosamine</td>
</tr>
<tr>
<td>DES</td>
<td>Diethyl sulphate</td>
</tr>
<tr>
<td>DMH</td>
<td>Dimethylhydrazine</td>
</tr>
<tr>
<td>DMN</td>
<td>DimethylNitrosamine</td>
</tr>
<tr>
<td>DMPT</td>
<td>Dimethylphenyltriazene</td>
</tr>
<tr>
<td>DMS</td>
<td>Dimethyl sulphate</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DSB</td>
<td>Double-strand breaks</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>EMS</td>
<td>Ethylmethanesulphonate</td>
</tr>
<tr>
<td>ENNG</td>
<td>$N$-methyl-$N'$-nitro-$N$-nitrosoguanidine</td>
</tr>
<tr>
<td>ENU</td>
<td>$N$-ethyl-$N$-nitrosourea</td>
</tr>
<tr>
<td>FHMBG</td>
<td>$O^6$-[4-[(γ-folyl)-oxymethyl]benzyl]guanine</td>
</tr>
<tr>
<td>HOPr-DAP</td>
<td>$N^6$-hydroxypropyl-2,6-diaminopurine</td>
</tr>
<tr>
<td>HR</td>
<td>Homologous recombination</td>
</tr>
<tr>
<td>HTH</td>
<td>Helix-turn-helix</td>
</tr>
<tr>
<td>IARC</td>
<td>International Agency for Research on Cancer</td>
</tr>
</tbody>
</table>
IPTG: Isopropylthiogalactoside
iso-NNAC: 4-(methylnitrosamino)-4-(3-pyridyl) butanoic acid
iso-NNAL: 4-(methylnitrosamino)-4-(3-pyridyl)-1-butanol
LB: Luria broth
MBP: Maltose binding protein
MEI: Methyl iodide
MGMT: Human $O^6$-alkylguanine- DNA alkyltransferase
MMR: Mismatch repair
MMS: Methyl methanesulphonate
MNNG: $N$-methyl-$N'$-nitro-$N$-nitroso guanidine
MNU: $N$-methyl-$N$-nitrosourea
NAB: $N'$-nitrosoanabasine
NAT: $N'$-nitrosoanatabine
NDBA: $N$-nitrosodibutylamine
NDEA: $N$-nitrosodiethylamine
NDELA: $N$-nitrosodiethanolamine
NDMA: $N$-nitrosodimethylamine
NDPhA: $N$-nitrosodiphenylamine
NER: Nucleotide Excision repair
NHEJ: Non-homologous end joining
NMOR: $N$-nitrosomorpholine
NNA: 4-(methylnitrosamino)-4-(3-pyridyl) butanal
NNAL: 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol
NNK: 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone
NNN: $N'$-nitrosonornicotine
NPIP: $N$-nitrosopiperidine
NPRO: $N$-nitrosoproline
NPYR: $N$-nitrosopyrrolidine
NTCA: $N$-nitrosothiazolidine-4-carboxylic acid
NTHZ: $N$-nitrosothiazolidine
$O^6$-AEG: $O^6$-aminoethylG
$O^6$-alkG: $O^6$-alkylguanine
$O^6$-BThG: $O^6$-(4-bromothenyl) guanine (lomeguatrib; PaTrin-2)
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>O&lt;sup&gt;6&lt;/sup&gt;-BzG</td>
<td>O&lt;sup&gt;6&lt;/sup&gt;-benzylguanine</td>
</tr>
<tr>
<td>O&lt;sup&gt;6&lt;/sup&gt;-BzoxoG</td>
<td>O&lt;sup&gt;6&lt;/sup&gt;-benzyl-8-oxo-guanine</td>
</tr>
<tr>
<td>O&lt;sup&gt;6&lt;/sup&gt;-CMG</td>
<td>O&lt;sup&gt;6&lt;/sup&gt;-carboxymethylguanine</td>
</tr>
<tr>
<td>O&lt;sup&gt;6&lt;/sup&gt;-EtG</td>
<td>O&lt;sup&gt;6&lt;/sup&gt;-ethylguanine</td>
</tr>
<tr>
<td>O&lt;sup&gt;6&lt;/sup&gt;-HOEtG</td>
<td>O&lt;sup&gt;6&lt;/sup&gt;-hydroxyethylguanine</td>
</tr>
<tr>
<td>O&lt;sup&gt;6&lt;/sup&gt;-MAG</td>
<td>O&lt;sup&gt;6&lt;/sup&gt;-methyladamantylguanine</td>
</tr>
<tr>
<td>O&lt;sup&gt;6&lt;/sup&gt;-MeG</td>
<td>O&lt;sup&gt;6&lt;/sup&gt;-methylguanine</td>
</tr>
<tr>
<td>O&lt;sup&gt;6&lt;/sup&gt;-PobG</td>
<td>O&lt;sup&gt;6&lt;/sup&gt;-[4-oxo-4-(3-pyridyl) butyl] guanine</td>
</tr>
<tr>
<td>O&lt;sup&gt;6&lt;/sup&gt;-PrG</td>
<td>O&lt;sup&gt;6&lt;/sup&gt;-propylguanine</td>
</tr>
<tr>
<td>ODNs</td>
<td>Oligodeoxyribonucleotides</td>
</tr>
<tr>
<td>PBMCs</td>
<td>Peripheral blood mononuclear cells</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-buffered saline</td>
</tr>
<tr>
<td>PBST</td>
<td>Phosphate-buffered saline Tween20</td>
</tr>
<tr>
<td>PCA</td>
<td>Perchloric acid</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethylsulphonylfluoride</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene difluoride</td>
</tr>
<tr>
<td>RE</td>
<td>Restriction endonuclease</td>
</tr>
<tr>
<td>RU</td>
<td>Relative units</td>
</tr>
<tr>
<td>SA</td>
<td>Streptavidin</td>
</tr>
<tr>
<td>SAM</td>
<td>S-adenosylmethionine</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulphate-polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SPR</td>
<td>Surface plasmon resonance</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris-Borate-EDTA</td>
</tr>
<tr>
<td>TBST</td>
<td>Tris-Buffered Saline Tween-20</td>
</tr>
<tr>
<td>TEMED</td>
<td>Tetramethylethylenediamine</td>
</tr>
<tr>
<td>TSNs</td>
<td>Tobacco specific nitrosamines</td>
</tr>
<tr>
<td>TMZ</td>
<td>Temozolomide</td>
</tr>
<tr>
<td>3FBDG</td>
<td>O&lt;sup&gt;6&lt;/sup&gt;-benzyl-3'-O-(γ-folyl)-2'-deoxyguanosine</td>
</tr>
<tr>
<td>5FBDG</td>
<td>O&lt;sup&gt;6&lt;/sup&gt;-benzyl-5'-O-(γ-folyl)-2'-deoxyguanosine</td>
</tr>
</tbody>
</table>
Abstract

Alkylating agents are a diverse family of compounds whose toxic, mutagenic and carcinogenic effects in living organisms are due to their ability to damage DNA. Humans are exposed to these agents through lifestyle, diet, occupation and some forms of chemotherapy, but they are also formed endogenously. To protect against these adverse effects, a variety of DNA repair processes have evolved. Among these is $O^6$-methylguanine-DNA methyltransferase (MGMT), a damage reversal protein that repairs $O^6$-alkylguanine ($O^6$-alkG) adducts by capturing the alkyl group onto a cysteine residue within the protein in an autoinactivating mechanism.

Human MGMT is known to be present in at least 12 polymorphic variant forms and numerous studies have partially characterised one or a few of these in relation to cancer susceptibility. The main objective in this thesis was to compare the substrate specificity of seven MGMT variant proteins, and also a putative size-variant of the normal MGMT, using short ODNs containing 12 different $O^6$-alkylguanines ($O^6$-alkGs). A variety of analytical methods was used in these studies, including radioisotope-based assays, enzyme-linked immunosorbent assays (ELISA), mass spectrometry (MS) and surface plasmon resonance (SPR).

The most potent inactivators of all MGMTs were ODNs containing $O^6$-BzG followed by $O^6$-MeG and $O^6$-CMG. For the other ODNs, the MGMT proteins could be classed in three groups: firstly Wt, and the F84, V143/R178 and F84/V143/R178 variants for which the potencies were $O^6$-PrG > $O^6$-PobG > $O^6$-EtG; secondly R160 and F84/R160 for which $O^6$-PrG > $O^6$-EtG > $O^6$-PobG and finally Q128 for which $O^6$-PobG > $O^6$-PrG > $O^6$-EtG. Alkyl group transfer from six of the ODNs to the active site cysteine of Wt MGMT was demonstrated by MS. The Q128 variant was the most resistant to inactivation by all modified ODNs. The R160 and F84/R160 variants were more resistant than the other variants to inactivation by the $O^6$-CMG and $O^6$-PobG containing ODNs. It was also shown that ZnCl$_2$ supplementation during Wt MGMT expression in *E. coli* significantly increased MGMT specific activity and decreased sensitivity to inactivation by the ODNs. In addition both 5-methylcytosine adjacent and opposite $O^6$-MeG significantly affected the inactivation of Wt MGMT. Finally, it was confirmed that $O^6$-CMG is, contrary to previous reports, a substrate for MGMT, and the inefficient repair of $O^6$-CMG by the *E. coli* Ogt alkyltransferase protein was shown to be the basis of the error. Studies with human cells suggest that MGMT also repairs $O^6$-CMG in *vivo*.

The results suggest that MGMT variants repair different $O^6$-alkGs at different rates and hence the consequence of alkylating agents exposure will depend not only on the nature of the exposure but also an individual’s own specific MGMT variant.
Declaration

No portion of the work referred to in the 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.

Copyright Statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, the University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses.
Acknowledgements

The work carried out in this thesis would not have been possible without the invaluable support and assistance of numerous individuals. I would like to thank my supervisors Dr Geoff Margison and Dr Andrew Povey for their help, support and guidance over the past four years. I thank them also for the opportunity to attend a conference and visit a lab in the USA, which greatly expanded my scientific horizons.

I sincerely thank the members of the Cancer Research UK Carcinogenesis Group at the Paterson Institute for Cancer Research, especially Mrs Gail McGown, Mrs Mary Thorncroft and Mrs Amanda Watson, who have guided me through so many aspects of the work I undertook there over three years. I also express my appreciation to Dr David Williams and Dr Oliver Wilkinson from the University of Sheffield who synthesised and purified all of the oligodeoxyribonucleotides (ODNs) containing O6-alkylguanines and other modified purines used in my studies. I would also like to thank Dr Claire Eyers for allowing me to use the mass spectrometry at Manchester Interdisciplinary Biocentre and also Miss Onrapak Reamthong for her help with these analyses. In addition, I express my gratitude to Dr Michael Fried and Dr Manana Melikishvili for allowing me to visit their laboratory at the Center for Structural Biology, University of Kentucky.

I am grateful to the Thai government for their scholarship and I would like to thank my workplace, the Prince of Songkla University, Surat Thani Campus (Thailand), for allowing me to study for a PhD at the University of Manchester for 4 years.
Chapter 1: Introduction

Alkylating agents are a structurally diverse family of compounds that cause a wide range of biological effects in living organisms, including cell death, chromosomal aberrations, mutagenesis and carcinogenesis (reviewed in Pegg, 2000). These effects are generally attributed to their ability to damage DNA giving rise to the formation of DNA adducts (reviewed in Drablos et al., 2004). Humans are exposed to these agents through lifestyle, diet, occupation, some forms of chemotherapy, and the environment and they are also formed endogenously (reviewed in Kaina et al., 2007).

The following sections deal firstly with the classification, structure, and basic reaction mechanisms of these agents. The biological effects are then considered, focussing on carcinogenesis in animal model systems. This is followed by a description of the nature of the DNA damage that might give rise to the biological effects and the cellular responses to that damage in terms of DNA repair processes, in particular, the damage reversal protein, MGMT. Aspects of human cancer are then considered, including examples of the routes of human exposure to alkylating agents. Finally the polymorphic variants in the human MGMT gene that might constitute cancer risk factors are described.

1.1 Classification of alkylating agents

The alkylating agents include alkyl halides, alkyl sulphates, alkyl alkanesulfonates, hydrazines, triazines, and N-nitroso compounds. The latter can be subdivided into 3 groups: nitrosoureas, nitrosoguanidines and nitrosamines. The nitrosamines and some triazenes are inert compounds that require metabolic activation to give rise to alkylating species. Examples of these and their structures are shown in Figure 1.1.
A. $S_N 1$ agents

\[
\text{MNU} \quad \begin{array}{c}
\text{H}_2\text{N} \quad \begin{array}{c}
\text{O} \\
\text{N} - \text{CH}_3 \\
\text{NO}
\end{array} \\
\text{NDMA} \quad \begin{array}{c}
\text{H}_3\text{C} \quad \begin{array}{c}
\text{N} - \text{N} \\
\text{O}
\end{array}
\end{array}
\end{array}
\]

\[
\text{MNNG} \quad \begin{array}{c}
\text{ON} - \text{N} - \text{C} - \text{NHNO}_2 \\
\text{BCNU} \quad \begin{array}{c}
\text{Cl} \\
\text{N} \\
\text{O} \quad \begin{array}{c}
\text{N} - \text{N} \\
\text{Cl}
\end{array}
\end{array}
\end{array}
\]

\[
\text{DMPT} \quad \begin{array}{c}
\text{N} = \text{N} - \text{N} - \text{CH}_3
\end{array} \\
\text{DMH} \quad \begin{array}{c}
\text{H}_3\text{C} \quad \begin{array}{c}
\text{N} - \text{N} \\
\text{CH}_3
\end{array}
\end{array}
\]

\[
\text{TMZ} \quad \begin{array}{c}
\text{N} - \text{CH}_3
\end{array}
\]

B. $S_N 2$ agents

\[
\text{DMS} \quad \begin{array}{c}
\text{H}_3\text{C} \quad \begin{array}{c}
\text{O} - \text{SO}_2 - \text{CH}_3
\end{array}
\end{array} \\
\text{MMS} \quad \begin{array}{c}
\text{H}_3\text{C} \quad \begin{array}{c}
\text{O} - \text{SO}_2 - \text{CH}_3
\end{array}
\end{array} \\
\text{MEI} \quad \begin{array}{c}
\text{H} \quad \begin{array}{c}
\text{C} - \text{H}
\end{array}
\end{array}
\]

\[
\text{SAM}
\]

\[
\text{adenine}
\]

\[
\text{H}_2\text{C} \quad \begin{array}{c}
\text{S} \\
\text{CH}_2 \\
\text{CH}_2
\end{array} \\
\text{H} - \text{C} - \text{NH}_3
\]

\[
\text{OH} \\
\text{OH} \\
\text{COO}^-
\]

Figure 1.1: Structures of alkylating agents. A. $S_N 1$ agents; MNU: $N$-methyl-$N$-nitrosoourea, NDMA: $N$-nitrosodimethylamine, MNNG: $N$-methyl-$N'$-nitro-$N$-nitroso guanidine, BCNU: 1,3-bis (2-chloroethyl)-1-nitrosourea. DMPT: Dimethylphenyl triazene, DMH: Dimethylhydrazine, TMZ: Temozolomide. B. $S_N 2$ agents; DMS: dimethyl sulphate, MMS: methylmethanesulphonate, MEI: Methyl iodide and SAM: $S$-adenosylmethionine.
The chemical reactions of alkylating agents take place by $S_N1$ (unimolecular nucleophilic substitution) or $S_N2$ (bimolecular nucleophilic substitution) type mechanisms (Warwick, 1963). The $S_N1$ reaction follows first order kinetics and the rate is dependent on a carbocation intermediate. In the $S_N2$ reactions, a transition complex is formed between the nucleophile and carbon centre which then releases a leaving group, leaving behind an alkyl group. The structures of some $S_N1$ and $S_N2$ alkylating agents are shown in Figure 1.1

1.2 Biological effects of alkylating agents

Alkylating agents are able to damage DNA through the formation of DNA adducts in a diversity of organisms including viruses, bacteria, plants and animals. Their mutagenic effects in bacteria have been known for decades, and mutagenicity in plants has been exploited in the improvement of crop plants and to develop new plant varieties (reviewed in Mba, 2013).

A wide range of animal species have been used to study biological responses to alkylating agents. While alkylating agents are toxic, mutagenic and can cause teratogenic effects, of particular interest is their carcinogenicity. One extensive study in several BD rat strains using 65 different $N$-Nitroso compounds demonstrated that many compounds (some of them shown in Figure 1.1) showed marked organ specific tumour induction (Druckrey et al., 1967). Similar organ specificities have been reported in other animal species including fish, Chinese and Syrian hamsters, and various strains of mice, rabbits and monkeys (Stanton, 1965; Ashley & Halver, 1968; Le Page & Christie, 1969; Mohr et al., 1974; Ishikawa et al., 1975; Hendricks et al., 1980; Brittelli et al., 1985; Ohgaki et al., 1986; Reznik & Padberg, 1991; Park et al., 1993; Lapis et al., 1997; Norton & Gardner, 2005; Hobbie et al., 2011; Kushida et al., 2011; Chen et al., 2012; Youssef et al., 2012).

The mutagenic effects of alkylating and other agents in rodents include bone marrow micronucleus induction, preneoplastic liver cell focus formation, mutations in gastrointestinal crypts, and the induction of dominant lethal mutations. Such effects have been exploited in the development of in vivo assays for the testing of agents that are being developed for use in humans (reviewed in Thurnherr et al., 1973; Wahnschaffe et al., 2005; Heindryckx et al., 2009).
The toxic effects of some alkylating agents, in particular the nitrosamides and triazenes have been exploited in their use in the treatment of various types of cancers. In general, the effectiveness of these agents is restricted by life-threatening toxicities in normal tissues, in particular the bone marrow, which limits the maximum doses of the agents that can be given to patients (Fairbairn et al., 1995; Clemons et al., 2000). All of the biological effects of the alkylating agents are probably mediated by their ability to introduce chemical damage into DNA including cancer development, and this is considered next.

1.3 DNA damage caused by alkylating agents

All alkylating agents react with the exocyclic oxygen and the ring nitrogen atoms of the DNA bases generating 4 oxygen adducts; $O^2$-methylcytosine ($O^2$-MeC), $O^6$-methylguanine ($O^6$-MeG), $O^2$-methylthymine ($O^2$-MeT) and $O^4$-methylthymine ($O^4$-MeT) and 7 nitrogen adducts (1-methyladenine (1-MeA), 3-methyladenine (3-MeA), 7-methyladenine (7-MeA), 3-methylcytosine (3-MeC), 3-methylguanine (3-MeG), 7-methylguanine (7-MeG) and 3-methylthymine (3-MeT). The sites of this damage are shown in Figure 1.2. In addition, two phosphotriester stereoisomers are formed by reaction with oxygen atoms in the phosphodiester backbone.

![Alkylation sites in DNA base-pairs. The numbers indicate the ring positions within the bases. The oxygen and nitrogen sites of attack are circled in black. dR; deoxyribose residue.](image)

**Figure 1.2:** Alkylation sites in DNA base-pairs. The numbers indicate the ring positions within the bases. The oxygen and nitrogen sites of attack are circled in black. dR; deoxyribose residue.
The relative amounts of these lesions are determined by the size of the alkyl group and the mechanism of alkylation. The majority of information relates to methylating agents, with far fewer reports on relative amounts from the higher alkylating species. In general the SN1 alkylating agents, such as MNU, alkylate both nitrogen and oxygen atoms while SN2 agents, such as MMS and SAM, mostly alkylate nitrogen atoms (Beranek, 1990). In addition, higher SN1 alkylating agents such as ethyl, generate relatively more oxygen alkylation than the SN1 methylaing agents as indicated in Table 1.1. Where this has been investigated, increasing the length of the carbon chain seems not to affect the relative levels of oxygen alkylation, but branched chain agents such as isopropyl generate yet higher relative levels (reviewed in Saffhill et al., 1985). The alkylation products of some classes of agents are considered below.

### Table 1.1: Alkylation products in DNA following reaction with alkylating agents.

<table>
<thead>
<tr>
<th>Agent</th>
<th>Source</th>
<th>Base alkylation products (% of total products)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Guanine</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N7</td>
</tr>
<tr>
<td>DMS</td>
<td>Salmon sperm (1)</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>Calf thymus (2)</td>
<td>75</td>
</tr>
<tr>
<td>DES</td>
<td>HeLa cells (3)</td>
<td>67</td>
</tr>
<tr>
<td>MNU</td>
<td>Salmon sperm (4)</td>
<td>75</td>
</tr>
<tr>
<td>ENU</td>
<td>HeLa cells (3)</td>
<td>11</td>
</tr>
<tr>
<td>MMS</td>
<td>Salmon sperm (5)</td>
<td>83</td>
</tr>
<tr>
<td>SAM</td>
<td><em>Micrococcus luteus</em> (6)</td>
<td>86</td>
</tr>
<tr>
<td>NOGC</td>
<td>Calf thymus (7)</td>
<td>76</td>
</tr>
<tr>
<td>NNK</td>
<td>Hamster lung (8)</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>Mouse lung (9)</td>
<td>87</td>
</tr>
</tbody>
</table>

ND indicates that the product was not detected.


Data taken from: 1 (Singer, 1975); 2 (Margison & O’ Connor, 1973); 3 (Sun & Singer, 1974); 4 (Lawley & Thatcher, 1973); 5 (Singer & Grunberger, 1983); 6 (Rydberg & Lindahl, 1982); 7 (Shuker & Margison, 1997); 8 (Rossignol et al., 1989); 9 (Peterson & Hecht, 1991).

Alkylating agents can directly or indirectly alkylate DNA. Two examples of alkylating agents that require metabolic activation to generate alkylating species are
the tobacco-specific nitrosamines (TSN), 4-(methylnitrosamino)-1-(3-pyridyl)-1-butane (NNK), and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) (also mentioned later in the context of human exposures). These are metabolized via both α-methylene and α-methyl hydroxylation. The former produces α-hydroxymethylene-NNK which spontaneously generates methanediazohydroxide and a keto aldehyde. The methanediazohydroxide is then converted to methanediazonium ion which is an SN1-type methylating agent. α-methyl hydroxylation yields α-hydroxymethyl which spontaneously loses formaldehyde generating a pyridyloxobutyl diazohydroxide that reacts with DNA and produces O^6-[4-oxo-4-(3-pyridyl) butyl] guanine (O^6-PobG; (Hecht, 1998). Another TSN, N-nitrosonornicotine (NNN) forms only the pyridyloxobutylating agent via a 2-α-hydroxylation pathway (Figure 1.3).

1.4 Biological effects of specific DNA alkylation products.

Little is known about the biological effects of most of the alkylation products in DNA mainly because most of them are generated in very small amounts (Table 1.1). The most abundant base lesions, 7-MeG, 3-MeA and O^6-MeG, which are generated by the more biologically active SN1 methylating agents, are better characterised, as are the O^6-alkGs generated by higher alkylating agents: these are considered below.

1.4.1 7-methylguanine

Although 7-MeG is the most abundant methylation product it is reported to be relatively harmless: it does not alter the coding properties of guanine, or interfere with transcription (Ludlum, 1970) or DNA replication (Prakash & Strauss, 1970). However, methylation at the N7 position destabilises the glycosylic bond resulting in chemical depurination and the formation of apurinic sites (AP site), which are highly cytotoxic and weakly mutagenic (Fronza & Gold, 2004). 7-MeG is also reported to undergo imidazole ring opening under physiological conditions, resulting in 2-6-diamino-4-hydroxy-5-N-methyl-formamidopyrimidine (Fapy; Chetsanga & Lindahl, 1979). Fapy blocks DNA synthesis and is therefore likely to be toxic, (O’ Connor et al., 1988) but it does not cause misparing in vitro and thus it is unlikely to be mutagenic (Boiteux & Laval, 1983). As the most abundant lesion, 7-MeG is
frequently used as a maker of exposure to methylating agents in epidemiological studies (Bianchini et al., 1993; Bianchini, 1994; Lees et al., 2007).

**Figure 1.3:** Metabolic activation of tobacco-specific nitrosamines and the formation of $O^6$-MeG and $O^6$-PobG (Drablos et al., 2004). NNK: 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, NNAL: 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol, NNN: $N^\prime$-nitrosonornicotine.
1.4.2 3-methyladenine

3-MeA blocks RNA polymerase and most DNA polymerases and is considered to be one of the more toxic lesions generated in DNA by methylating agents (Sedgwick et al., 2007). In addition, 3-MeA can cause sister chromatid exchanges and chromosome aberrations (Engelward et al., 1998). The glycosylic bond between 3-MeA and deoxyribose is more chemically labile that that of 7-MeG and it more rapidly forms AP sites in DNA (Margison & O’Connor, 1973). These can be repaired by AP endonucleases, but if not repaired, and they are replicated, this can give rise to double strand breaks, which, if not repaired, can be highly toxic (Fronza & Gold, 2004).

1.4.3 $O^6$-alkylguanine

The original suggestion that $O^6$-alkG may be mutagenic came from chemical experiments involving $S_N1$ and $S_N2$ alkylating agents and nucleosides (Loveless, 1969). In support of this, in vitro experiments indicated that replication of template DNA containing $O^6$-MeG frequently resulted in the incorporation of thymidine rather than cytosine in the product (Abbott & Saffhill, 1979). The same effect was seen when a plasmid containing $O^6$-MeG was replicated in bacteria (Topal et al., 1986). These miscoding properties of $O^6$-alkG explained how particularly the $S_N1$ alkylating agents induced G:C to A:T transition mutations and this provided a possible mechanism for their potent carcinogenicity, based on the assumption that carcinogenesis was initiated by mutations in critical genes including oncogenes such as Ras (The three human ras genes encode extremely similar proteins made up of chains of 188 to 189 amino acids, designated H-Ras, N-Ras, K-Ras4A and K-Ras4B (http://en.wikipedia.org/wiki/Ras_subfamily), or tumour suppressor genes such as p53 (see below, and reviewed in Esteller et al., 2000; Esteller et al., 2001; Nakamura, 2001).

It was later found that, as well as being mutagenic, $O^6$-MeG was toxic. The mechanism of cell killing involves one of the DNA repair pathways described below, but is presented here for completion (see Figure 1.4). During the first round of DNA synthesis following DNA alkylation, $O^6$-MeG pairs with thymine instead of cytosine resulting in an $O^6$-MeG:T mispair. The $O^6$-MeG:T mispair is recognised by
the DNA post replication mismatch repair (MMR) system and the binding of MMR components results in cleavage of the newly synthesized strand and removal of the misincorporated T residue. However, DNA repair synthesis restores the original $O^6$-MeG:T mispair and this re-initiates MMR in a “futile” DNA repair cycle. A second round of DNA replication of the gapped DNA template then generates a double strand break (DSB). Although DSB can be repaired (see below) they are potentially highly toxic lesions in DNA. This mechanism of cell killing has been reviewed several times (Olive, 1998; Jackson, 2002; Margison et al., 2002a; Goodsell, 2005; Helleday et al., 2007).

![Mechanisms of the biological effects of $O^6$-MeG and their prevention by MGMT (Margison et al., 2002).](image)

One study suggests that, of the $O^6$-MeG lesions generated in DNA in human tumor cell line, 80% cause G:C→A:T transition mutations ~20% cause
recombinations, in the form of sister chromatid exchanges and chromosome aberrations and less than 1% cause cell death (Rasouli-Nia et al., 1994).

The mutagenic, recombinogenic and toxic effects of the $O^6$-alkGs that are generated in DNA by the higher alkylating agents are less well investigated. There are reports that the $O^6$-ethylguanine: thymine mispairs are not a substrate for MMR and this may explain why these agents are less toxic than methylating agents (Pauly et al., 1998). Evidence for the mutagenicity of $O^6$-CMG was reported by Gottschalg et al., (2007) who showed that the carboxymethylating agent, potassium diazoacetate caused GC to AT transitions in $p53$. In addition, Zhang et al., (2013) have shown that $O^6$-CMG can form a stable base pair with T, implying that it will miscode to form transition mutations during DNA replication.

1.5 Human exposure to alkylating agents

A wide range of alkylation products are frequently found in human DNA (De Bont & van Larebeke, 2004). There is thus no question that exposure to alkylating compounds occurs. It is generally accepted that this is from both endogenous and exogenous sources, although the relative contributions of these sources to overall DNA alkylation damage have not been extensively investigated and hence their relative contribution to human cancer incidence has yet to be established. Since alkylation products can be present in DNA from people not knowingly exposed to exogenous agents (for example, Povey et al., 2000), endogenously formed agents can apparently be the exclusive cause. Endogenous and exogenous sources are discussed below.

1.5.1 Endogenous sources

Some of the suspected sources of endogenous alkylating agents are ethylene oxide and nitrosated bile and amino acids. The latter two types are known to be able to generate the DNA adducts $O^6$-CMG, $O^6$-MeG, 7-CMG and 3-CMA in vitro and in vivo (Zurio et al., 1982; Shuker & Margison, 1997; Sedgwick, 1997; O' Driscoll et al., 1999b; Burcham, 1999).

A clear case of an endogenous alkylating agent is $S$-adenosylmethionine (SAM; Figure 1.1). This is well known for its properties as a biological $S_N2$-type
methylating agent acting as a donor of methyl groups due to the positive charge of the sulphur activating the methyl group and causing it to be reactive (Cantoni, 1975). As well as acting as a biological methyl donor, SAM is also able to alkylate DNA in vitro (Table 1.1) and it seems likely that it will have the same effect in vivo. Different levels of SAM are present in different rodent tissues ranging in concentration from 25-50 µM (Eloranta, 1977) and the concentration of SAM in human plasma is 156 nM (Melnyk et al., 2000). At a concentration of 40 µM, SAM causes DNA alkylation at the same level as 20 nM MMS, indicating that the activity of SAM is 2000 fold weaker than MMS (Ryderberg & Lindahl, 1982).

1.5.2 Exogenous sources

Exogenous sources of alkylating agents are well known and include tobacco smoke, occupation and diet (Bugni et al., 2007). Included in this group are some of the N-nitroso compounds that are not actually alkylating agents unless they are acted upon by enzymes, commonly in the cytochrome P450 family. Laboratory studies have revealed that various human tissues can metabolize nitrosamines into alkylating compounds (Umbenhauer et al., 1985; Patterson et al., 2012).

1.5.2.1 Tobacco smoking

Tobacco smoking is a major risk factor for many types of human cancer (Hecht & Hoffmann, 1989). Tobacco smoke contains a very wide variety of toxic and mutagenic compounds including heavy metals, hydrocarbons, tar and the tobacco specific nitrosamines (TSNs) (Hammond & O'Connor, 2008) which are generated from the tobacco alkaloids; nicotine, nornicotine, anabasine and anatabine (Figure 1.5) during the curing and processing of tobacco (Irigaray et al., 2007). Seven TSNs have been reported and their structures are shown in Figure 1.5.

NNN, NNK and NAT are found in higher quantities in tobacco products compared to NNAL, NAB, iso-NNAL and iso-NNAC (Hecht, 1999). NNK, NNN and NNAL are the most carcinogenic of these compounds in laboratory animals, inducing tumors in mice, rats and hamsters (Hoffmann & Hecht, 1985); for example, NNN induces oesophageal tumours in rats (Hecht et al., 1983). Tobacco smoking is thought to be the principal causative agent in the development in humans of various
cancers, mainly lung but also cancers in the pancreas and oral cavity (reviewed in (Hecht & Hoffmann, 1989; Hoffmann et al., 1994; Martin-Moreno et al., 2008). Tobacco smoking is also causally related to oesophageal cancer (International Agency for Research on Cancer, 1978; Doecke et al., 2008).

In human tissues, the levels of alkyl adducts due to tobacco smoking have been investigated in several studies. The levels of 7-MeG in human bronchial cells and lymphocyte are higher in smokers (17.3 and 11.5 adducts/10^6 nucleotides, respectively) than non-smokers (4.7 and 2.3 adducts/10^6 nucleotides, respectively; Mustonen et al., 1993).

**Figure 1.5:** Structure of tobacco alkaloids and the tobacco-specific nitrosamines (Hecht, 1998).

The uppermost set of arrows represent nitrosation reactions that generate NNK (4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone), NNA (4-(methylnitrosamino)-4-(3-pyridyl) butanal), NNN (N′-nitrosornornicotine), NAB (N′-nitrosoanabasine), NAT (N′-nitrosoanatabine). The lower set of arrow represent reduction reactions that produce NNAL (4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol), iso-NNAL (4-(methylnitrosamino)-4-(3-pyridyl)-1-butanol) and iso-NNAC (4-(methylnitro samino)-4-(3-pyridyl) butanoic acid). The activation pathways for NNK, NNN and NNAL are shown in Figure 1.3.
1.5.2.2 Occupational exposures

Workers in certain industries have a significantly increased risk of several types of cancer such as leukaemia and cancers of the bladder, lung and stomach. High risk occupations include the rubber industry, leather tanning, foundry and metal working, chemical and fish processing industries. In these industries, relatively high concentrations of nitrosamines have been found in the ambient air and/or in grinding fluids and personal exposures have been estimated (Table 1.2, reviewed in Bartsch & Montesano, 1984). Of the compounds found, the International Agency for Research on Cancer (IARC) classify \( N\)-nitrosodimethylamine (NDMA) and \( N\)-nitrosodiethylamine (NDEA) to be probable human carcinogens, and \( N\)-nitrosopyrrolidine (NPYR) and \( N\)-nitrosomorpholine (NMOR) to be possible human carcinogens (International Agency for Research on Cancer, 1978).

Table 1.2: Occupational exposure to nitrosamines

<table>
<thead>
<tr>
<th>Industry</th>
<th>Major nitrosamines</th>
<th>Concentration range</th>
<th>Estimated exposure (µg/day/person)*</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leather tanneries</td>
<td>1,2</td>
<td>0.05-47 µg/m³</td>
<td>20-180</td>
<td>Rounbehler et al., 1979</td>
</tr>
<tr>
<td>Rubber</td>
<td>1, 2, 3, 4, 5</td>
<td>0.01-1230 µg/m³</td>
<td>15-150</td>
<td>Spiegelhalder &amp; Preussmann, 1983</td>
</tr>
<tr>
<td>Metal working</td>
<td>6</td>
<td>1.3-6 µg/m³</td>
<td>&gt;50</td>
<td>Fadlallah et al., 1990</td>
</tr>
<tr>
<td>Chemical</td>
<td>Rocket fuel</td>
<td>-</td>
<td>10-50</td>
<td>Fine &amp; Rounbehler, 1981</td>
</tr>
<tr>
<td></td>
<td>- Dye manufacture</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Surfactant</td>
<td>0.03-0.1 µg/m³</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>production</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foundries</td>
<td>1, 4</td>
<td>0.024-1.4 µg/m³</td>
<td>&gt;5</td>
<td>Ducos et al., 1988</td>
</tr>
<tr>
<td>Fish processing</td>
<td>1</td>
<td>0.01-0.06 µg/m³</td>
<td>&lt;5</td>
<td>Fajen et al., 1980</td>
</tr>
</tbody>
</table>

\( ^1N\)-nitrosodimethylamine; \( ^2N\)-nitrosomorpholine; \( ^3N\)-nitrosopyrrolidine; \( ^4N\)-nitrosodiethylamine; \( ^5N\)-nitrosodiphenylamine; \( ^6N\)-nitrosodiethanolamine.


The contribution of occupational exposures to alkyl DNA adducts is, however, not clear. Occupational nitrosamine exposure in a vehicle rubber seal manufacturing company had no significant association with 7-MeG levels in peripheral blood lymphocyte DNA (Rey et al., 2000). The concentrations of 7-MeG in workers exposed to nitrosamines ranged from 0.1-133.2 adducts/10⁷ nucleotides.
and in controls (no exposure) from 0.1-128.2 adducts/10^7 nucleotides. This suggests that, in this case at least, exogenous factors may not make a significant contribution to 7-MeG adduct levels.

### 1.5.2.3 Diet

A large number of N-nitroso compounds have been found at various levels in a wide variety of food and beverages which include cured and processed meat products, smoked preserved foods, foods subjected to drying by additives such as malt in the production of beer, pickled and salty preserved foods (Table 1.3). In addition to these pre-formed agents, diets can contain agents that may enhance endogenous processes that generate alkylation agents or their precursors. NDMA was the most potent adduct-forming compound in nitrite-containing foods, mostly being found in smoked/salted fish, cured meats, sausages and beer (Knekt et al., 1999). Thus nitrite, added as a preservative, can, under the appropriate conditions, form N-nitroso compounds by interaction with secondary amines (Lijinsky, 1999). Gut bacteria are also reported to mediate such nitrosation reactions (reviewed in Bingham et al., 2002).

Diets high in red and processed meat, which are well-established risk factors for colorectal cancer (CRC) (Norat et al., 2005) increase N-nitrosation within the colon, generating increased levels of N-nitroso compounds and also S-nitrosated compounds (Bos, 1989; Silvester & Cummings, 1995). In particular, increased levels of the nitrosated bile acid conjugate, N-nitrosoglycocholic acid (NOGC) are found in red meat-rich diets, causing increased levels of DNA alkylation products. Thus, exfoliated colon cell DNA contains O^6^-CMG, probably arising from N-nitrosation of the bile acid conjugate, glycocholic acid which also generates O^6^-MeG in DNA (Shuker et al., 1981; Song et al., 1982; Busby et al., 1985; Lewin et al., 2006). Indeed, human colorectal DNA has been reported to contain levels of O^6^-MeG that vary between <0.01-0.94 µmol O^6^-MeG/mol deoxyguanosine, with the highest occurring in the sigmoid colon and rectum, where most sporadic tumors occur (Hall et al., 1991; Povey et al., 2000; Povey et al., 2002). If these levels were present in animal tissues, they would be sufficiently high to cause adverse biological effects, particularly in cells that are deficient in DNA repair (see later).
Table 1.3: *N*-nitroso compounds found in food and beverages.

<table>
<thead>
<tr>
<th>Food and beverages</th>
<th>Major nitrosamines</th>
<th>Concentration range (µg/100g)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meat products:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacon</td>
<td>1,2,3,4,5,6,7</td>
<td>0.049-142.7</td>
<td>Stuff <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>Ham</td>
<td>1,2,3,4,5,6,7</td>
<td>0.004-46.1</td>
<td>Stuff <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>Sausage</td>
<td>1,2,3,4,6,7,8</td>
<td>0.001-11.9</td>
<td>Stuff <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>Cured meats</td>
<td>1</td>
<td>0.12</td>
<td>Palli <em>et al.</em>, 2001</td>
</tr>
<tr>
<td>Salami</td>
<td>1</td>
<td>0.05</td>
<td>Palli <em>et al.</em>, 2001</td>
</tr>
<tr>
<td>Fish products:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fried fish</td>
<td>1,2,4,5,6,7,9</td>
<td>0.003-4.6</td>
<td>Stuff <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>Smoked fish</td>
<td>1</td>
<td>2.6</td>
<td>Pobel <em>et al.</em>, 1995</td>
</tr>
<tr>
<td>Salted fish</td>
<td>1,2,3,4</td>
<td>2.8-454</td>
<td>Zou <em>et al.</em>, 1994</td>
</tr>
<tr>
<td>Dairy products:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole milk</td>
<td>1,2,3</td>
<td>0.003-0.014</td>
<td>Stuff <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>Margarine</td>
<td>1,8</td>
<td>0.026-0.049</td>
<td>Stuff <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>Butter</td>
<td>1,8</td>
<td>0.026-0.049</td>
<td>Stuff <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>Cheese</td>
<td>1,2,3,4,7</td>
<td>0.004-1.5</td>
<td>Stuff <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>Yoghurt</td>
<td>1</td>
<td>0.02</td>
<td>Palli <em>et al.</em>, 2001</td>
</tr>
<tr>
<td>Vegetables:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peas</td>
<td>2,6</td>
<td>0.025-0.3</td>
<td>Stuff <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>French fries</td>
<td>1,2</td>
<td>0.024-0.041</td>
<td>Stuff <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>Refries beans</td>
<td>1</td>
<td>0.033</td>
<td>Stuff <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>Pickles</td>
<td>1,2,3,4,6,7</td>
<td>0.079-2.095</td>
<td>Stuff <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>Beverages:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beer</td>
<td>1</td>
<td>0.04-0.202</td>
<td>Palli <em>et al.</em>, 2001; Stuff <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>White wine</td>
<td>1,2</td>
<td>0.025-0.109</td>
<td></td>
</tr>
</tbody>
</table>

*1*-*N*-nitrosodimethylamine; *2*-*N*-nitrosopyrrolidine; *3*-*N*-nitrosopiperidine; *4*-*N*-nitrosodiethylamine; *5*-*N*-nitrosodibutylamine; *6*-*N*-nitrosoproline; *7*-*N*-nitrosothiazolidine-4-carboxylic acid; *8*-*N*-nitrosomorpholine; *9*-*N*-nitrosothiazolidine.

The levels of *O*-CMG in colorectal DNA increase with a red meat diet (see above) and this strongly implies a role in CRC (Cupid *et al.*, 2004; Lewin *et al.*, 2006). However, nitrosation of the amino acid, serine, would also give rise to a carboxymethylating agent, so the origin of even these lesions is not clear.
Furthermore, the role of these agents in human cancer is obscured by the fact that cooked red meat also contains carcinogens such as heterocyclic amines and polycyclic aromatic hydrocarbons (Layton et al., 1995; Cross & Sinha, 2004; Jakszyn et al., 2011).

1.5.2.4. Cancer chemotherapy

Chemotherapeutic alkylating agents that attack the O6-position of guanine can be categorized into three major groups (Table 1.4): the methylating agents, such as procabazine, dacarbazine (DTIC), streptozotocin and temozolomide (structure shown in Figure 1.1), the chloroethylyating agents, such as BCNU (carmustine; structure shown in Figure 1.1) and CCNU (lomustine) and the bifunctional bis-chloroethylamines (Table 1.4). The chloroethylyating agents initially react with guanine in DNA and the product, O6-chloroethylguanine, forms the cyclic intermediate N1-O6-theganoguanine, which reacts with N3 of the opposing cytosine to form a crosslink between N1 of guanine and N3 of cytosine (Erickson et al., 1980). These DNA interstrand crosslinks are very toxic (Roos & Kaina, 2006). The bis-chloroethylamines predominantly react with DNA at the N7-positions of two guanine residues on opposite DNA strands generating interstrand crosslinks (reviewed in Brent, 1985).

Therapeutic methylating agents are used to treat melanoma, Hodgkin’s disease, glioblastoma and brain tumours (Souliotis et al., 1991; Spiro et al., 2000; Hegi et al., 2005). The chloroethylyating agents are commonly used in a variety of cancers such as glioblastoma (Parney & Chang, 2003), lymphoma (Bessell et al., 2004), malignant melanoma (Li & McClay, 2002) and Hodgkin disease (Stuart et al., 2001). In one study, BCNU treatment generated levels of 7-hydroxyethylguanine (7-HOEtG) in human brain tumor samples that were more than 800 times higher than in control samples (Bodell et al., 2001).
Table 1.4: Classification of chemotherapeutic alkylating agents and their mechanism

<table>
<thead>
<tr>
<th>Alkylating agent</th>
<th>Examples</th>
<th>Lethal DNA damage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methylating agents</td>
<td>Dacarbazine</td>
<td>O₆-MeG (3-meA)</td>
</tr>
<tr>
<td></td>
<td>Streptozotocin</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Procarbazine</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Temozolomide</td>
<td></td>
</tr>
<tr>
<td>Chloroethynitrosoureas</td>
<td>BCNU, CCNU</td>
<td>N1G:N3C interstrand crosslinks</td>
</tr>
<tr>
<td>Bischloroethylamines</td>
<td>Cyclophosphamide</td>
<td>N7G:N7G interstrand crosslinks</td>
</tr>
<tr>
<td></td>
<td>Melphalan</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chlorambucil</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ifosphamide</td>
<td></td>
</tr>
</tbody>
</table>

1.6 Repair of DNA damage caused by alkylating agents

In eukaryotic cells, different DNA alkylation products can be processed by six different repair mechanisms, and there is also some overlap between the pathways with some lesions being substrates for more than one repair pathway. The repair pathways are Base Excision Repair (BER), Nucleotide Excision Repair (NER), Recombination Repair (RR) which consists of Homologous Recombination (HR) and Non-Homologous End Joining (NHEJ) and Mismatch Repair (MMR) which has already been mentioned in the context of the toxic effects of O₆-MeG; section 1.4.3. There are also two damage reversal processes; AlkB-homologues (ABH2 and ABH3) and O₆-alkylguanine-DNA alkyltransferase (MGMT). As the latter is the main theme of this thesis, it is dealt with in a separate section.

In mammalian cells, the BER pathway removes 7-alkG, 3-alkG, 7-alkA and 3-alkA via the action of a DNA glycosylase, 3-alkyladenine-DNA glycosylase (AAG, also called APNG and MPG). This carries out hydrolytic cleavage of the N-glycosylic bond leaving an AP site which is processed by apurinic endonuclease, DNA polymerase and DNA ligase (Doetsch & Cunningham, 1990).

NER is reported to act on O₄-alkT and O₆-alkG lesions that involve longer chain alkyl groups such as butyl (Boyle et al., 1986; Boyle et al., 1987; Bol et al., 1999). It is initiated by damage recognition proteins or the stalling of transcription,
both of which recruit various multiprotein complexes that result in the elimination of a short stretch of DNA (around 24 nucleotides) containing the lesion. The section of ss DNA that remains acts a template for DNA polymerase to synthesize a short complementary sequence to fill the gap. The resulting SSB is ligated by DNA ligase (reviewed in Wood, 1997; de Laat et al., 1999).

DSBs are known as the most toxic of all DNA lesions and can be induced by different mechanisms such as exposure to ionizing radiation (IR). DSBs are repaired by two major pathways; homologous recombination (HR) and non-homologous end joining (NHEJ; reviewed in Valerie & Povirk, 2003). In HR, which occurs after DNA replication, the undamaged homologous chromatid acts as a template for the repair of a broken sister chromatid. NHEJ, which can occur at any time in the cell cycle, processes damaged DNA termini and re-ligates them, a process that is considered to be highly error-prone (reviewed in Valerie & Povirk, 2003; Lees-Miller & Meek, 2003).

Finally, ABH2 and ABH3 are dioxygenases that remove the methyl groups from 1-MeA and 3-MeC in single stranded (ss) and double stranded (ds) DNA using oxygen, 2-oxoglutarate and Fe (II) to oxidize the methyl group resulting in direct reversal of the damage upon spontaneous release of formaldehyde. ABH2 acts more rapidly on these lesions in ds DNA and ABH3 prefers ss DNA (reviewed in Aas et al., 2003; Falnes et al., 2004).

1.7 $O^6$-alkylguanine-DNA alkyltransferases

The earliest evidence for the active removal of $O^6$-alkG lesions from DNA was that, following the administration of NDMA to rats, rapid loss of $O^6$-MeG from hepatic DNA was seen, and this was attributed to an enzymatic repair process (O’Connor et al., 1973). The mechanism of this repair was elucidated following later observations in E.coli that exposure to low concentrations of MNNG increased resistance to the toxic effects of subsequent higher doses, a process termed the adaptive response. The gene controlling this response (called the ada gene) was cloned and later shown to encode a protein that removed methyl groups from $O^6$-MeG in DNA. The methyl groups became covalently attached to a cysteine residue in the active site of the protein so that the repair process was both autoinactivating
and stoichiometric. The \textit{ada} gene encodes a 39 kDa protein comprising alkytransferase domains of 18 kDa and 20 kDa the former acting on $O^6$-MeG (and $O^4$-MeT) and the latter acting on one of the phosphotriester stereoisomers, repair of which triggers the adaptive response mechanism. A second alkytransferase gene in \textit{E.coli}, the \textit{ogt} gene, constitutively expresses a 19 kDa protein with a repair function very similar to that of the 18 kDa subfragment of Ada (reviewed in Potter et al., 1989; Margison et al., 1990; Harris et al., 1992).

Many pro and eukaryotic alkytransferases have since been cloned and extensively studied. Figure 1.4 shows the basic damage reversal mechanism by which these proteins protect cells against the mutagenic (and hence carcinogenic) recombinogenic and toxic effects of alkylation agents. Below are discussed some of the established characteristics of the human alkytransferase, which is named MGMT, although it has also been referred to as ATase, AGT and AGAT.

\subsection{1.7.1 The human MGMT protein}

MGMT is encoded by a gene that is located on chromosome 10 at 10q26; it is approximately 300 kb in size and it has 5 exons, the first of which is reported to be non-coding. Exons 2-5 encode a protein that is usually considered to be 21.7 kDa in size and consist of 207 amino acids. It has an N terminal domain (residues 1-85) which has a bound zinc atom, and a C terminal domain (residues 86-207). The DNA-binding region consists of two helix-turn-helix (HTH) motifs at residues 114-121 and 127-136 and the active site motif spans residues 144-147 and contains the alkyacceptor cysteine residue at position 145 (Figure 1.6). The primary amino acid sequence is presented later, in the context of polymorphic variants of MGMT (Daniels et al., 2000; Tubbs et al., 2007).

\subsection{1.7.2 Repair of $O^6$-alkGs by MGMT}

Based on X-ray crystallographic studies of MGMT, the active site alky acceptor (Cys145) and second HTH DNA binding motifs are linked by Asn137 and this makes hydrogen bonds with Val139 and Ile143 and the thiol of Cys145. Binding to the minor groove of DNA is via the HTH motifs, and then the Arg128 ‘finger’ flips the guanine out of the DNA double helix and into the active site pocket (Figure 1.6). DNA binding causes the minor groove to widen more than 0.1 nm and causes
the DNA to bend approximately 15° away from the protein (Daniels et al., 2004). This binding is not considered to be $O^6$-alkG specific and the protein is thought to scan the DNA for substrate lesions by an on-off process. If a substrate lesion has been flipped into the active site pocket, the lone electron pair on the sulphur atom in Cys145 then attacks the alkyl group and forms a covalent bond with it, breaking the oxygen (6)-carbon bond on the substrate. The contribution of the zinc atom to the repair process is discussed later.

It has been reported that in cultured human tumour cells, the inactivated alkylated MGMT is then degraded by the ubiquitin-proteasomal pathway (Srivenugopal et al., 1996; Liu et al., 2001; Xu-Welliver & Pegg, 2002).

Figure 1.6: Human MGMT X-ray crystallographic structure, zinc binding site, and DNA binding mechanism. (A) Free MGMT (B) DNA-bound MGMT. The N-terminal domain (N) is shown in green, the C-terminal domain (C) is shown in yellow, and the HTH motifs are shown in purple (Tubbs et al., 2007).

1.7.3 MGMT Substrate specificity

The ability of MGMT to repair $O^6$-alkyl adducts was initially thought to decrease with increasing size of the alkyl group. For example, MGMT removes ethyl groups faster than isopropyl groups (Liem et al., 1994). However, it is reported that $O^6$-BzG (Pegg et al., 1993) and $O^6$-bromothenylG ($O^6$-BThG) (Clemons et al., 2005;
Shibata et al., 2006) in DNA are more effective substrates for MGMT than $O^6$-MeG, so the size of the alkyl group seems not to be the sole factor in the rate of alkyl transfer. Examples of this are $O^6$-hydroxyethylG ($O^6$-HOEtG; Shibata et al., 2006) and $O^6$-CMG (Shuker & Margison, 1997) which have previously been reported to be very poor substrates for MGMT.

Several studies of very large numbers of low molecular weight, mainly free base or mononucleoside, inactivators of MGMT show that the substrate range is very large (Dolan et al., 1985; Pegg et al., 1997; Pegg et al., 2001; Shibata et al., 2006b). Some of these pseudosubstrates have been tested in clinical trials in combination with alkylating agents in attempts to improve the effectiveness of therapy (reviewed in Middleton & Margison, 2003; Kaina et al., 2010).

The substrate specificity of MGMT is further discussed below in the context of MGMT polymorphic variants.

1.7.4 MGMT expression levels and MGMT activity

Different human tissues express different levels of activity, the rank order being liver > T-lymphocytes > small intestine > colon > myeloid precursors > lung > brain > bone marrow (Gerson et al., 1986). MGMT activity decreases further in bone marrow cells after in vitro treatment with the nitrosourea, BCNU, and remains low during DNA synthesis and cell proliferation, resulting in a ratio of DNA repair capacity to DNA synthesis which is one to two orders of magnitude lower in the bone marrow than in other tissues such as kidney and liver. The low MGMT expression in bone marrow has been implicated as the basis of the myelosuppressive action of nitrosoureas and related agents and also a risk factor for leukemia (Gerson et al., 1996).

The MGMT promoter region contains several transcription regulatory factor binding sites including simian virus 40 promoter factor 1 (Sp1), two glucocorticoid response elements and AP-1 and AP-2 sites. In addition, a 59 base pair element is present at the first intron-exon boundary and this has been shown to be required for the efficient expression of reporter-gene constructs. All of these elements may potentially be involved in regulating expression of MGMT, but the basis of the different levels of
MGMT in different tissues and between individuals has not been established (reviewed in Margison et al. 2003). In addition, one post-translational modification of MGMT in the form of phosphorylation has been reported to increase its activity in vitro (Lim et al., 2000) but this has also not been established to occur in human tissues.

It was originally established in cultured human cells that methylation of cytosine residues in CpG dinucleotides in the promoter region of the MGMT gene can silence its transcription and hence protein expression (Danam et al., 1999). It has also been shown that promoter methylation is present in a significant number of human cancers (Esteller et al., 1999) and, on the assumption that it correlates with low or non-expression of MGMT in the tumour, this parameter is currently used as a basis for selecting glioma patients for treatment with TMZ (Hegi et al., 2004). MGMT promoter methylation is also suspected to be a factor in the susceptibility of colorectal tissue to carcinogenesis by endogenous or exogenous alkylating agents (see above; Esteller et al., 2001).

Human tumour cells in vitro and human tumour xenografts expressing high levels of MGMT are resistant to killing by methylating and chloroethylating chemotherapeutic agents (reviewed in Verbeek et al., 2008). MGMT inhibitors or inactivators have been used to reduce the levels of active MGMT in order to increase sensitivity of such tumours to these agents (Gerson, 2002). One such agent is the MGMT pseudosubstrate, O6-BzG (Dolan et al., 1990; Moschel et al., 1992) which inactivates MGMT in vitro and in vivo and enhances tumour cells and xenografts to killing by BCNU or TMZ (Wedge & Newlands, 1996; Kreklau et al., 1999). However, the results from clinical trials of O6-BzG and the more potent agent, O6-(4-bromothenyl)guanine (lomeguatrib; PaTrin-2), show no clinical benefit in melanoma patients treated with DTIC or TMZ (Middleton et al., 2002; Clemons et al., 2005). This indicated the possibility that melanomas may be inherently resistant to anticancer drugs.

1.7.5 Effect of zinc ions on MGMT activity

In MGMT, the bound zinc atom is coordinated with Cys5, Cys24, His29 and His85 in the N-terminal domain of the protein (Daniels & Tainer, 2000). Occupancy of the N-terminal binding site by Zn$^{2+}$ has been reported to generate a more stable
conformation and enhanced the correct folding of the protein (Rasimas et al., 2003). Thus, when MGMT expression from an MGMT cDNA containing plasmid in *E.coli* was induced in growth in media supplemented with 0.1 mM ZnCl$_2$ there was a 5 fold increase in repair rate in comparison with that grown in media without additional ZnCl$_2$ (Rasimas et al., 2003), suggesting that Zn$^{2+}$ can be rate-limiting.

Later, the N and C terminal domains of MGMT were separately expressed and purified and their ability to transfer radioactivity from a $[^3$H]-methylated DNA substrate to the protein examined (Fang et al., 2005). Both domains alone were inactive and after addition of Zn$^{2+}$, very weak activity was seen only in the N terminal domain. However, when the N and C terminal domains were mixed together and Zn$^{2+}$ added, there was a 12 fold increase in activity (Fang et al., 2005).

The effect of nutritional zinc deficiency on MGMT activities has been examined in rats in various tissues including the liver, lung, kidney, spleen, brain, esophagus, forestomach, gastric-stomach and small intestine (Fong et al., 1988). Compared with zinc sufficiency, MGMT activities were significantly reduced in the oesophagus, spleen and lung. This result might explain an earlier report that dietary zinc deficiency enhances nitrosamine induced oesophageal carcinogenesis in rats (Fong et al., 1978).

1.7.6 Effect of 5-methylcytosine in the rate of repair of $O^6$-alkGs by MGMT

The presence of 5-MeC in DNA leads to a minor change in structure (Heinemann & Hahn, 1992): sugars and bases rise, twist and roll and the local curvature and glycosidic angles increase (Marcourt et al., 1999). In addition, the major groove decreases in depth and the minor groove narrows (Lefebvre et al., 1995), and this increases melting temperatures (Gill et al., 1974). These structural alterations enhance the three dimensional stability of DNA (Norberg & Vihinen, 2001) and increase water solubility (Sowers et al., 1987) which may have an influence on the rate of MGMT-mediated nucleotide flipping (Bentivegna & Bresnick, 1994).

The effect of the presence of 5-MeC in double stranded DNA on the repair of $O^6$-MeG within CG dinucleotides of the p53 codons 158, 245 and 248 has been
investigated (Guza et al., 2009). 5-MeC in codon 158 increased the rate of $O^6$-MeG repair compared with unmethylated DNA in the rank order: $O^6$-MeG base-paired with 5-MeC > 5-MeC located 5' to $O^6$-MeG > 5-MeC both base paired and 5' to $O^6$-MeG > no 5-MeC. In contrast, 5-MeC within p53 codon 248 decreased the rate of $O^6$-MeG repair by MGMT. Furthermore, repair of $O^6$-MeG in p53 codon 245 was slower in ODNs that had 5-MeC present in both DNA strands than for the ODNs containing no 5-MeC. In addition, MGMT-mediated repair of $O^6$-MeG was faster in ODNs that contained 5'-neighboring $MeC$ and base paired $MeC$ than without any 5-MeC. It was concluded that cytosine methylation had a small effect on the rate of MGMT repair of $O^6$-MeG within the sequence context of p53 codon 158, 245 and 248 (Guza et al., 2009) suggesting that 5-MeC might influence other kinetic steps related to repair such as the rate of base flipping.

### 1.7.7 Effect of DNA sequence context on the rate of repair of $O^6$-alkGs by MGMT

Previous studies have demonstrated that the rate of MGMT-mediated repair of $O^6$-alkG is affected by sequence context. H-ras codon 12 is a major cancer mutational hotspot induced by alkylating agents involving the second, but not the first G of this codon in this protooncogene (reviewed in Meyer et al., 2003). 15-mers spanning H-ras codon 12 (GGC GCT GGA $G^G$ GC GTG) have been synthesized but with the first or second guanine of codon 12 replaced by $O^6$-MeG (Georgiadis et al., 1991). The E.coli ada AGT repaired $O^6$-MeG at the first position 18 times faster than at the second position. Furthermore, substitution of the first guanine before $O^6$-MeG in the -TT$G^G$AG sequence with a thymine (-TT$G^G$AG-) increased the rate of repair 25 fold. Substitution of the second guanine of the -TT$G^G$AG- by adenine (i.e.-TT$G^G$AGA) reduced the repair rate 2.6 fold. This result indicated that the rate of repair by MGMT is dependent on the base sequence of DNA.

MGMT-mediated repair of $O^6$-MeG has been studied at 7 different positions within a 30-mer spanning K-ras codon 12 (5’-G$_1$TAG$_2$TTG$_3$G$_4$A G$_5$CTG$_6$G$_7$TG$_8$G$_9$C G$_{10}$TAG$_{11}$G$_{12}$CAAG$_{13}$AG$_{14}$T-3’). Under one set of conditions, the percent repair of $O^6$-MeG located at G$_5$-G$_{11}$ was 39, 61, 57, 61, 54, 42 and 57% respectively (Guza et
This observation supports the above conclusion that the rate of repair by MGMT varies with the local sequence context of O\textsuperscript{6} - MeG within ODNs.

MGMT repair of ODNs containing O\textsuperscript{6} - BzG, O\textsuperscript{6} - MeG, O\textsuperscript{6} - EtG, O\textsuperscript{6} - HOEtG and O\textsuperscript{6} - PobG has been investigated in three different sequence contexts (Table 1.5). MGMT repaired O\textsuperscript{6} - BzG, O\textsuperscript{6} - MeG and O\textsuperscript{6} - PobG at the first G position (sequence 1) faster than at the second position (sequence 2), again confirming that MGMT repairs O\textsuperscript{6} - alkG at different rates depending on the alkyl group and the sequence context (Coulter et al., 2007).

**Table 1.5:** The rate of repair by human MGMT of ODNs containing various O\textsuperscript{6} - alkG

<table>
<thead>
<tr>
<th>Alkyl group</th>
<th>Rate of repair (s\textsuperscript{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sequence 1</td>
</tr>
<tr>
<td>O\textsuperscript{6} - BzG</td>
<td>35±4</td>
</tr>
<tr>
<td>O\textsuperscript{6} - MeG</td>
<td>0.32±0.03</td>
</tr>
<tr>
<td>O\textsuperscript{6} - EtG</td>
<td>0.023±0.001</td>
</tr>
<tr>
<td>O\textsuperscript{6} - HOEtG</td>
<td>0.00093±0.7</td>
</tr>
<tr>
<td>O\textsuperscript{6} - PobG</td>
<td>0.000095±0.07</td>
</tr>
</tbody>
</table>

ND indicates not determined.
Sequences 1, 2, and 3 are 5’-GGCGCTGGAGGCGTG, 5’-GGCGCTGGAGGCGTG and 5’-AACA GCCATATGGCCC, respectively. Taken from Coulter et al., (2007).

In contrast, the MGMT repair rate constant (k) for 5 different sequence contexts namely 5’-TGT-3’, 5’-CGG-3’, 5’-GGT-3’, 5’-GGG-3’, 5’-GGA-3’ was 10.6, 9.7, 6.5, 10.3 and 11.5 (x 10\textsuperscript{-6} M\textsuperscript{-1}s\textsuperscript{-1}) in ds 29 mers (5’-ACTGACTGATG TTTGTXGXGACTGACTG, where G was O\textsuperscript{6} - EtG and was flanked 5’ and 3’ by different bases. This indicated that the MGMT repair rate for O\textsuperscript{6} - EtG did not significantly dependent on sequence context (Bender et al., 1996).

In conclusion, the rate of repair of O\textsuperscript{6} - alkG by MGMT depends on the local DNA sequence context and the nature of the alkyl group.
1.7.8 Human variants of $O^6$-alkylguanine-DNA alkyltransferase

The primary amino acid sequence of the MGMT protein is shown in Figure 1.7 which includes the location of the HTH motifs, the active site domain, and the $\text{Zn}^{2+}$ binding residues. These are also shown in Figure 1.8 which includes the nucleotide sequence and exon structure of MGMT, showing the location of the more abundant polymorphic variants.

For convenience, the most frequent MGMT variant will be called the wild type (Wt). The frequency of polymorphisms in different populations is listed in Table 1.6. At present, only 5 variants, namely W65C, L84F, I143V, K178R and G160R have been reported to influence the primary structure of MGMT (Pegg et al., 2007). Most studies have focused on the three common coding sequence polymorphisms including L84F, I143V and K178R.

\[
\text{Wt} \quad 1 \quad \text{mdkd} \quad \text{emkrt} \quad \text{tldsp1gkle} \quad \text{lsg} \quad \text{egglhe} \quad \text{iklngkgtsa} \quad \text{adavevpapa} \\
\text{E30K} \\
\text{Wt} \quad 51 \quad \text{avlggpeplm} \quad \text{qctawlnayf} \quad \text{pqpealeefp} \quad \text{vpallhpvfq} \quad \text{qesftrqv1w} \quad 100 \\
\text{P58S} \quad \text{W65C} \quad \text{L84F} \\
\text{Wt101} \quad \text{kllkvvkfge} \quad \text{vqgplaal} \quad \text{gnpkaarav} \quad \text{ggamrgnpvp} \quad \text{iliechr} \quad \text{150} \\
\text{R128Q} \quad \text{I143V} \\
\text{Wt151} \quad \text{sgavgnysg} \quad \text{glavkewlla} \quad \text{hghrlgkpg} \quad \text{lggsgllaga} \quad \text{wlkgagatsg} \quad 200 \\
\text{G160R} \quad \text{E166D} \quad \text{K178R} \\
\text{Wt201} \quad \text{sppagrn} \quad 207
\]

**Figure 1.7:** Primary amino acid sequence of the Wt MGMT, location of the amino acid changes in the polymorphic variant proteins and the codon changes responsible. The active site motif (pchr) is boxed. The DNA-binding regions of the helix-turn-helix motif are highlighted in grey and the amino acid residues involved in coordination with $\text{Zn}^{2+}$ are highlighted in blue. Silent polymorphisms are not shown. Genbank accession number NM_002412.3.
**Figure 1.8:** Nucleotide sequence, exon structure and the location of polymorphisms in MGMT cDNA. Exons 1, 2, 3, 4 and 5 are black, green, turquoise, red and purple respectively. The codon changes from wild-type are shown in magenta. The initiation codon is shown in capital letters. Genbank accession number: NM_002412.3.
### Table 1.6: Population prevalence of MGMT polymorphisms in exons 2, 3, 4 and 5

<table>
<thead>
<tr>
<th>Exon</th>
<th>Polymorphism (ID)</th>
<th>Frequency of heterozygotes and minor allele (%) by population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Europe</td>
</tr>
<tr>
<td>2</td>
<td>E30K (rs2020893)</td>
<td>&lt;1; ND; ND (1)</td>
</tr>
<tr>
<td></td>
<td>L53L (rs1803965)</td>
<td>21; 13; Sweden (2)</td>
</tr>
<tr>
<td></td>
<td>P58S (rs2308322)</td>
<td>&lt;1; ND; ND (1)</td>
</tr>
<tr>
<td></td>
<td>W65C (rs2282164)</td>
<td>ND; ND; ND</td>
</tr>
<tr>
<td>3</td>
<td>L84F (rs12917)</td>
<td>36 ; 28; Caucasian (6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25; 16; Sweden (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23.8; 17; Long Island (7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12 ; 12; UK (8)</td>
</tr>
<tr>
<td>4</td>
<td>R128Q (rs3750824)</td>
<td>&lt;1; ND; ND (1)</td>
</tr>
<tr>
<td></td>
<td>I143V (rs2308321)</td>
<td>21 ; 13; Caucasian (6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15 ; 15; Caucasian (9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22 ; 14; Sweden (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22.7; 15; Long Island (7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15 ; 12; UK (8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15 ; 7.3; Caucasian (10)</td>
</tr>
<tr>
<td>5</td>
<td>G160R (rs2308318)</td>
<td>2; 0.9;Caucasian (10)</td>
</tr>
<tr>
<td></td>
<td>A163A (rs2308319)</td>
<td>&lt;1; ND; ND (1)</td>
</tr>
<tr>
<td></td>
<td>E166D (rs2308320)</td>
<td>&lt;1; ND; ND (1)</td>
</tr>
<tr>
<td></td>
<td>K178R (rs2308327)</td>
<td>21 ; 13; Caucasian (6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22 ; 14; Sweden (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19.7; 13; Long Island (7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15 ; 15; UK (4)</td>
</tr>
<tr>
<td></td>
<td>A197A (rs2308324)</td>
<td>0.6; 0.6; Sweden (2)</td>
</tr>
</tbody>
</table>

ND indicates no data.

Reviewed in 1 (Margison et al., 2006); 2 (Ma et al., 2003); 3 (Chae et al., 2006); 4 (Margison et al., 2005); 5 (Li et al., 2005); 6 (Pegg, 2000); 7 (Shen et al., 2005); 8 (Jiao et al., 2006); 9 (Bugni et al., 2007); 10 (Kaur et al., 2000); 11 (Gerson et al., 1999); 12 (Yang et al., 2004).
The CEU (also known as CEPH or Utah residents with ancestors from northern and western Europe) HapMap population Phase III (http://hapmap.ncbi.nlm.nih.gov/downloads) was selected to study SNPs within the MGMT gene on chromosome 10 using Haplovew version 4.2. Human MGMT has at least 12 polymorphic sites but only seven of them were shown in HapMap data including rs2020893 (E30K), rs1803965 (L53L), rs2282164 (W65C), rs12917 (L84F), rs3750824 (R128Q), rs2308321 (I143V) and rs2308327 (K178R) (Table 1.7). In this population, the minor allele frequency varied between <0.001 and 0.157

Table 1.7: Analysis of MGMT single nucleotide polymorphisms (SNPs) using Haplovew version 4.2.

<table>
<thead>
<tr>
<th>No.</th>
<th>ID</th>
<th>Name</th>
<th>ObsHET</th>
<th>PredHET</th>
<th>HWpval</th>
<th>%Geno</th>
<th>MAF</th>
<th>Alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td>359</td>
<td>rs2020893</td>
<td>E30K</td>
<td>0.0</td>
<td>0.0</td>
<td>1.0</td>
<td>100.0</td>
<td>0</td>
<td>G:G</td>
</tr>
<tr>
<td>626</td>
<td>rs1803965</td>
<td>L53L</td>
<td>0.136</td>
<td>0.155</td>
<td>0.6797</td>
<td>98.9</td>
<td>0.085</td>
<td>C:T</td>
</tr>
<tr>
<td>627</td>
<td>rs2282164</td>
<td>W65C</td>
<td>0.0</td>
<td>0.0</td>
<td>1.0</td>
<td>97.8</td>
<td>0</td>
<td>G:G</td>
</tr>
<tr>
<td>628</td>
<td>rs12917</td>
<td>L84F</td>
<td>0.208</td>
<td>0.215</td>
<td>1</td>
<td>84.4</td>
<td>0.123</td>
<td>C:T</td>
</tr>
<tr>
<td>706</td>
<td>rs3750824</td>
<td>R128Q</td>
<td>0.0</td>
<td>0.0</td>
<td>1.0</td>
<td>87.8</td>
<td>0</td>
<td>G:G</td>
</tr>
<tr>
<td>723</td>
<td>rs2308321</td>
<td>I143V</td>
<td>0.339</td>
<td>0.282</td>
<td>0.3173</td>
<td>90</td>
<td>0.17</td>
<td>A:G</td>
</tr>
<tr>
<td>724</td>
<td>rs2308327</td>
<td>K178R</td>
<td>0.315</td>
<td>0.265</td>
<td>0.4487</td>
<td>87.8</td>
<td>0.157</td>
<td>A:G</td>
</tr>
</tbody>
</table>

ObsHET (observed heterozygosity), PredHET (predicted heterozygosity), HWpval (Hardy-Weinberg equilibrium p value: cutoff 0.05), %Geno (percentage of non-missing genotypes), MAF (minor allele frequency: minimum 0.001) and Alleles (major and minor alleles).

The haplovew plot is shown in Figure 1.9 with the criteria of pairwise linkage disequilibrium (D’) and LOD (log of the likelihood odds ratio). Figure 1.9A shows two haplotype blocks. Block 2 (rs2308321 and rs2308327) were in strong linkage disequilibrium (D’=1, r^2=1), as well as block 1 including rs1803965 and rs12917 (D’=1, r^2=0.702). Four white blocks with less linkage disequilibrium included rs12917- rs2308321 (D’=0.19, r^2=0.001), rs12917- rs2308327 (D’=0.06, r^2=0.00), rs1803965- rs2308321 (D’=0.15, r^2=0.01) and rs1803965- rs2308327 (D’=0.17, r^2=0.01).

Figure 1.9B shown the level of recombination in block 1 (CGC, haplotype frequency is 0.867; TGT, frequency 0.090; and CGT, frequency 0.043) and block 2

54
(AA, frequency 0.824; and GG, frequency 0.176). This figure also shows the level of recombination between adjacent blocks (multi allelic $D' = 0.13$). The frequency of recombination between (i) CGC (Block 1) and AA (Block 2) and (ii) CGC (block 1) and GG (Block 2) was greater than 10%. The frequency of recombination between the blocks with thin lines in Figure 1.9B (TGT: AA, TGT:GG and CGT:AA) was greater than 1%.

**Figure 1.9:** A. Linkage disequilibrium map of MGMT single nucleotide polymorphisms (SNPs) using Haploview version 4.2.
A. Five SNPs were identified as being in two separate haplotype blocks (in bold). Pairwise linkage disequilibrium ($D'$) is given for each SNP combination. Red empty squares indicate $D'$ values of 1.0 with LOD (log of the likelihood odds ratio) score $≥$ 2. White squares indicate $D'$ values $< 1.0$ with LOD scores $< 2$.
B. Two haplotype blocks were identified across the MGMT SNPs. The haplotype frequencies are shown to the right of each haplotype. Only haplotypes having a frequency greater than 1% are shown. The SNP numbers across the top of the haplotypes correspond to those in the Haploview plot. A multi-allelic $D'$ statistic, which indicates the level of recombination between two blocks, is shown in the crossing area. Connections from one block to the next are shown for haplotypes of greater than 10% frequency with thick lines and greater than 1% frequency with thin lines.

I143V and K178R are changes that are in almost perfect linkage disequilibrium (a nonrandom association of two genes on the same chromosome; Margison et al., 2005). The double variant, I143V/K178R, is more common in Caucasians where the frequency is 21% as compared to only 2% in Chinese or Japanese (Otsuka et al., 1996; Deng et al., 1999; Ma et al., 2003). The frequency of triple variant including L84F/I143V/K178R is 1.4% in Americans, 1.3% in Chinese and 1.3% in Caucasians (Shen et al., 2005; Huang et al., 2007; Liu et al., 2009).
1.7.9 Repair of $O^6$-alkGs by human MGMT variants

A number of groups have investigated the effect of MGMT polymorphisms on the function of MGMT in vitro and the effect of expressing the protein in E.coli and cultured human cells. Expression of the L84F variant protected ada/ogt deficient E.coli and HeLa cells against the toxic effects of MNNG as effectively as the Wt protein, but protection by W65C was reduced (Inoue et al., 2000). The W65C polymorphism reduced the level of MGMT protein expression even though the mRNA was produced at the same level as for Wt and L84F. It was concluded that the W65C variant was unstable in vivo possibly because the side chain of W65C is located in the first helix in the N-terminal domain and that without this interaction the protein is rapidly degraded.

Cultures of lymphocytes from 114 individuals have been established and NNK-induced chromosomal aberrations (CA) determined (Hill et al., 2005). The L84F polymorphism was associated with increased CA (P < 0.02) and suggested that this was due to its reduced capacity to repair the CA-inducing lesions. This effect was higher in subgroups of individuals, such as smokers, adolescents (< 43 years old) and males (P < 0.03). This result argues that younger individuals (13 ± 7) tended to smoke significantly less cigarettes per day than adults (21 ± 7) (P< 0.004) but younger smokers had increased NNK-induced CAs. This increase in DNA damage in younger individuals was not related to smoking intensity.

Short ODNs (5'-AACAGCCATATXGCCC where X is $O^6$-alkG) were used to examine the repair of the bulky $O^6$-alkGs, $O^6$-PobG, $O^6$-BuG and $O^6$-BzG relative to $O^6$-MeG by the Wt, I143V/K178R and G160R proteins (Mijal et al., 2004). I143V/K178R did not differ from Wt in the repair of these bulky adducts. Wt and I143V/K178R repaired $O^6$-MeG approximately twice as fast as $O^6$-PobG. G160R was consistently slower in the repair of bulky adducts and when presented with ODN containing both $O^6$-MeG and $O^6$-PobG, did not repair $O^6$-PobG until all of the $O^6$-MeG was removed. Wt, I143V/K178R and G160R displayed an approximately 2-3 fold preference for the removal of $O^6$-MeG over $O^6$-BuG. All of MGMT proteins preferentially repaired $O^6$-BzG relative to $O^6$-MeG (Table 1.8).
Table 1.8: Repair of $O^6$-MeG relative to other $O^6$-alkGs (Mijal et al., 2004).

<table>
<thead>
<tr>
<th>Protein</th>
<th>$O^6$-MeG/ $O^6$-PobG</th>
<th>$O^6$-MeG/ $O^6$-BuG</th>
<th>$O^6$-MeG/ $O^6$-BzG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>2.6 ± 0.5</td>
<td>2.4 ± 0.2</td>
<td>0.04 ± 0.04</td>
</tr>
<tr>
<td>I143V/K178R</td>
<td>2.6 ± 0.5</td>
<td>3.0 ± 0.3</td>
<td>0.09 ± 0.04</td>
</tr>
<tr>
<td>G160R</td>
<td>~7.5</td>
<td>3.5 ± 0.8</td>
<td>0.27 ± 0.04</td>
</tr>
</tbody>
</table>

The repair by Wt, L84F, I143V/K178R and G160R of $O^6$-PobG and $O^6$-MeG when present in ODN was examined within 5 different sequence contexts (Table 1.9) (Mijal et al., 2006).

Table 1.9: ODN used to test the effect of sequence context on the rate of repair $O^6$-MeG/$O^6$-PobG (Mijal et al., 2006)

<table>
<thead>
<tr>
<th>ODN</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5'$-$AACAGCCCATATG;GCCC</td>
<td>Sequence used in Mijal et al. (2004)</td>
</tr>
<tr>
<td>Sequence 1</td>
<td>5'$-$GGCGCTG$^*$,G$_2$AGGCCTG</td>
<td>H-ras; codon 12, adduct in G$_1$ of TGGA</td>
</tr>
<tr>
<td>Sequence 2</td>
<td>5'$-$GGCGCTG$_1$G$^*$,AGGCCTG</td>
<td>H-ras; codon 12, adduct in G$_2$ of TGGA</td>
</tr>
<tr>
<td>Sequence 3</td>
<td>5'$-$AACAGCTG$^*$,G$_2$AGGCCTG</td>
<td>H-ras like; codon 12, adduct in G$_1$ of TGGA</td>
</tr>
<tr>
<td>Sequence 4</td>
<td>5'$-$AACAGCTG$_1$G$^*$,AGGCCTG</td>
<td>H-ras like; codon 12, adduct in G$_2$ of TGGA</td>
</tr>
<tr>
<td>Sequence 5</td>
<td>5'$-$AACAGCTG$_1$G$^*$,AGGCCTG</td>
<td>H-ras like; codon 12, adduct in G$_2$ of TGGA</td>
</tr>
</tbody>
</table>

The relative repair of $O^6$-MeG/$O^6$-PobG was found to depend on the sequence context. Thus, repair of these adducts in sequences 1, 4 and 5 by L84F did not differ from Wt protein. Also, the I143V/K178R variant was most effective in the repair of $O^6$-PobG in sequence 5 and G160R was consistently the least effective in the repair of $O^6$-PobG although it did not differ from Wt in sequence 2, 3 and 5. In sequence 1 and 4, G160R was much less effective than Wt in the repair of these adducts (Table 1.10). This result indicated that human MGMT polymorphisms differ in their repair activity to these sequence contexts.
Table 1.10: Relative repair of $O^6$-MeG and $O^6$-PobG in different ODNs (Mijal et al., 2006)

<table>
<thead>
<tr>
<th>Protein</th>
<th>Repair ratio (amount of $O^6$-MeG repaired/amount of repaired $O^6$-PobG)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sequence 1</td>
</tr>
<tr>
<td>Wild type</td>
<td>2.0 ± 0.5</td>
</tr>
<tr>
<td>L84F</td>
<td>1.9 ± 0.1</td>
</tr>
<tr>
<td>I143V/K178R</td>
<td>1.8 ± 0.4</td>
</tr>
<tr>
<td>G160R</td>
<td>$^a\geq5$</td>
</tr>
</tbody>
</table>

Sequences 1 to 5 (Table 1.9) used: $^a >$ means no repair of $O^6$-PobG in the presence of $O^6$-MeG, $^b \geq$ means very little repair of $O^6$-PobG in the presence of $O^6$-MeG.

1.7.10 Pseudosubstrate inactivation of human MGMT variants

A number of studies have examined the ability of pseudosubstrates (any substance that mimics the substrate of an enzyme and thus inhibits its activity) to inactivate variant and Wt MGMT proteins. G160R has been reported to be 20 fold more resistant to inactivation by $O^6$-BzG than the Wt protein (Edara et al., 1996). It was suggested that this was because G160R was likely to form part of the active site pocket of MGMT and that a charged side chain interacts with other residues that might reduce the access of $O^6$-BzG to the catalytically active cysteine residue (Edara et al., 1996).

The inactivation of W65C, L84F and I143V/K178R by $O^6$-BzG, $O^6$-benzylfolic acid (BF), $O^6$-benzyl-3’-O-($\gamma$-folyl)-2’-deoxyguanosine (3FBDG), $O^6$-benzyl-5’-O-($\gamma$-folyl)-2’-deoxyguanosine (5FBDG) and $O^6$-[4-($\gamma$-folyl)-oxymethyl]benzyl]guanine (FHMBG) has been compared with Wt (Fang et al., 2008). W65C was slightly resistant to $O^6$-BzG, BF, 3FBDG and FHMBG and the I143V/K178R variant was resistant to all of these inactivators (Table 1.11). It was suggested that this was due to steric alterations in the active site pocket region where these inhibitors bind.
Table 1.11: Inactivation of Wt and variant MGMT by 5 inactivators (Fang et al., 2008)

<table>
<thead>
<tr>
<th>Protein</th>
<th>IC₅₀ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>O⁶-BzG</td>
</tr>
<tr>
<td>Wt</td>
<td>0.09 ± 0.01</td>
</tr>
<tr>
<td>W65C</td>
<td>0.27</td>
</tr>
<tr>
<td>L84F</td>
<td>0.09 ± 0.01</td>
</tr>
<tr>
<td>I143V/K178R</td>
<td>0.18 ± 0.01*</td>
</tr>
</tbody>
</table>

* Significantly different from Wt (P < 0.01). ** Significantly different from Wt (P < 0.001).

The L84F variant had little if any effect on the interaction with these inhibitors (Table 1.11) and this was consistent with a conservative change in the hydrophobic amino acid side-chain for another at a position that is a significant distance away from the active site.

The effect of expression of G160R on cell killing by BCNU has been examined in Chinese hamster ovary (CHO) cells following pretreatment of the cells with the MGMT pseudosubstrate, O⁶-BzG or its metabolite O⁶-benzyl-8-oxo-guanine (O⁶-BzoxoG) (Loktionova et al., 1999). G160R expressing cells were more resistant than Wt expressing cells to the combination of BCNU with O⁶-BzG or O⁶-BzoxoG. However, G160R was as effective as Wt in protecting cells against the combination of MNNG and BCNU. It was concluded that any G160R variant cancer patients would respond less well to therapy with O⁶-BzG in combination with an alkylating agent.

The L84F variant was reported to be degraded more rapidly than Wt in U87MG cells, but that this did not alter the sensitivity of these cells to TMZ or the combination of O⁶-BzG and TMZ (Remington et al., 2009).

1.7.11 Association between MGMT polymorphisms and cancer risk

Many studies have investigated the association between MGMT gene polymorphisms and the risk of a variety of human cancers (Tables 1.12, 1.13, 1.14, and 1.15, in which the homozygous genotypes are represented by CC, TT, GG, AA and heterozygous genotypes are presented by CT, AG, GA).
There are 3 reports of no significant associations between MGMT polymorphisms and lung cancer risk. The prevalence of L53L and L84F polymorphisms was not significantly different between lung cancer cases and controls in Koreans (Chae et al., 2006). The distribution of the K178R polymorphism was also not significantly different between lung cancer patients and controls in Caucasians (Yang et al., 2004). Finally, no evidence of an association between K178R and lung cancer risk has also been reported but it was suggested that the R alleles may protect against lung cancer, specifically in heavy smokers (Crosbie et al., 2008). In contrast, two papers reported that MGMT polymorphisms were associated with increased risk of lung cancer. Firstly, in a combined Caucasians and African-American population, the I143V polymorphism was associated with a significantly increased risk of lung cancer (OR = 2.1; 95% CI = 1.01 – 4.7) but this was of borderline significance. It is currently unclear why some, but not all, studies have reported significant associations with lung (and other) cancers. This may be related to sample size, differences in populations and control groups studied as well as potential differences in the type and magnitude of exposure to alkylating agents.

Polymorphisms in MGMT have been evaluated in relation to the risk of other types of cancers and these can be divided into 3 groups. The first group consists of those studies which report no significant effect between any polymorphisms and other types of cancer. L84F and I143V were not significantly associated with risk of colorectal cancer among cases and control in Americans (Tranah et al., 2006). L84F, I143V and K178R had no interaction with breast cancer risk, and in addition, there was no significant interaction between cigarette smoking and these polymorphisms (Shen et al., 2005). L84F and I143V were not significantly associated with pancreatic cancer (Jiao et al., 2006), and finally, W65C, L84F and I143V were not associated with the risk of oral cancer (Gal et al., 2005; Kietthubthew et al., 2006).

The second group contains five publications that report MGMT polymorphisms that were associated with increased risk of cancer. Firstly, two polymorphisms (I143V/K178R) were associated with an elevated risk for cervical carcinoma. In addition, women who carry a L84F/I143V/K178R haplotype were at a significantly increased risk for cervical carcinoma (Huang et al., 2007). Secondly, L53L and L84F when evaluated together led to a significantly increased risk of
bladder cancer (adjusted OR = 1.67; 95% CI = 1.01-2.77) (Li et al., 2005). Thirdly, the L84F genotype was associated with an increased risk of prostate cancer (OR = 1.99; 95% CI = 1.19-3.34) (Ritchey et al., 2005). Fourthly, R128Q (OR = 5.53; 95% CI = 2.58-7.16) and G160R (OR = 3.04; 95% CI = 1.48-6.31) were significantly associated with colorectal cancer (Khatami et al., 2008). Finally, L84F was associated with an increased risk of glioma in adults (≥ 18 years) who were exposed to ionizing radiation (adjusted OR = 5.95; 95% CI = 2.21-16.65) (Liu et al., 2009).

The third group reports that MGMT polymorphisms are associated with a decreased risk of cancer. L84F carriers had a significantly decreased risk of endometrial cancer (RR, 0.72; 95% CI = 0.53-0.96) compared with the wild-type genotype (Han et al., 2006). Also, L84F (OR = 0.71; 95% CI = 0.51-0.98) and I143V (OR = 0.66; 95% CI = 0.47-0.92) polymorphisms were associated with a decreased risk of head and neck cancer (Huang et al., 2005).

In addition to these studies on cancer risk, the associations between MGMT SNPs and alkylating agent treatment response have been investigated. In one study, a significant association between MGMT SNPs and cellular sensitivity to temozolomide was reported using a genome-wide association approach to assess temozolomide sensitivity in lymphoblastoid cell lines derived from white patients. rs477692 was associated with a differential response and rs531572 with response and MGMT transcript levels. (Brown et al., 2012).
<table>
<thead>
<tr>
<th>Type of cancer</th>
<th>MGMT polymorphisms</th>
<th>Genotype distribution</th>
<th>Population (source)</th>
<th>Population description</th>
<th>Odds Ratio (95% Confidence Interval)</th>
<th>Parameters Matched</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung cancer</td>
<td>L53L (CC/CT/TT)</td>
<td>Case: 79/20/1 Control: 79/19/2 Korean</td>
<td>Cases (432): newly diagnosed with primary lung cancer Controls (432): Hospital – randomly selected from healthy volunteers</td>
<td>1.03 (0.73-1.45) Matched Age, gender</td>
<td>Chae et al., 2006</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L84F (CC/CT/TT)</td>
<td>Case: 80/19/1 Control: 79/19/2 Korean</td>
<td>Cases (432): newly diagnosed with primary lung cancer Controls (432): Hospital – randomly selected from healthy volunteers</td>
<td>1.04 (0.74-1.47) Matched Age, gender</td>
<td>Chae et al., 2006</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Case: 74/23/3 Control: 75/23/2 Caucasians</td>
<td>Cases (1121): diagnosis with lung cancer Controls (1163): Clinics- cancer- free individuals</td>
<td>1.00 (0.82-1.21) Matched Age, gender, race</td>
<td>Wang et al., 2006</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I143V (AA/AG/GG)</td>
<td>Case: 77/23/0 Control: 85/15/0 Caucasians</td>
<td>Cases (53): Mixed histology, diagnosis with primary lung cancer Controls (55): Hospital – diagnosis unrelated to smoking</td>
<td>2.0 (0.78-5.7) Matched Age, gender, race</td>
<td>Kaur et al., 2000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Case: 75/24/1 Control: 78/21/1 Caucasians</td>
<td>Cases (1121): diagnosis with lung cancer Controls (1163): Clinics- cancer- free individuals</td>
<td>1.21 (0.99-1.47) Matched Age, gender, race</td>
<td>Wang et al., 2006</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Case: 86/14/0 Control: 94/6/0 African Americans</td>
<td>Cases (81): Mixed histology, diagnosis with primary lung cancer Controls (81): Hospital – diagnosis unrelated to smoking</td>
<td>2.3 (0.73-8.3) Matched Age, gender, race</td>
<td>Kaur et al., 2000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>K178R (AA/AG/GG)</td>
<td>Case: 78/22/0 Control: 81/19/0 Caucasians</td>
<td>Cases (92): Mixed histology, diagnosis with lung cancer Controls (85): Hospital – Urology and orthopaedic surgery clinics</td>
<td>1.19 (0.46-1.93) Matched Age, gender, race</td>
<td>Yang et al., 2004</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Case: 80/20/0 Control: 72/26/2 Caucasians</td>
<td>Cases (251): newly diagnosed with a tumor of the lung, trachea or bronchus Controls (358): North west lung cancer patients – free of benign and malignant tumors</td>
<td>0.67 (0.45-1.01) Matched Age, gender, smoking</td>
<td>Crosbie et al., 2008</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1.13: MGMT polymorphisms and risk of endometrial, cervical, oral, and pancreatic cancer

<table>
<thead>
<tr>
<th>Type of cancer</th>
<th>MGMT polymorphisms</th>
<th>Genotype distribution</th>
<th>Population (source)</th>
<th>Population description</th>
<th>Odds Ratio (95% Confidence Interval)</th>
<th>Parameters Matched</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endometrial cancer</td>
<td>L84F (CC/CT/TT)</td>
<td>Case: 79/19/2, Control: 76/22/2</td>
<td>Americans</td>
<td>Cases (434): female nurses with invasive endometrial cancer, Controls (1085): NHS- randomly selected from free of cancer</td>
<td>0.80 (0.61-1.05)</td>
<td>Year of birth, Menopausal status</td>
<td>Han et al., 2006</td>
</tr>
<tr>
<td></td>
<td>I143V (AA/AG/GG)</td>
<td>Case: 76/22/2, Control: 77/21/2</td>
<td>Americans</td>
<td>Cases (445): female nurses with invasive endometrial cancer, Controls (1089): NHS- randomly selected from free of cancer</td>
<td>1.06 (0.82-1.38)</td>
<td>Year of birth, Menopausal status</td>
<td>Han et al., 2006</td>
</tr>
<tr>
<td>Cervical cancer</td>
<td>L84F (CC/CT/TT)</td>
<td>Case: 69/29/2, Control: 74/25/1</td>
<td>Chinese</td>
<td>Cases (539): women with primary cervical cancer, Controls (800): Hospital - randomly selected from healthy women</td>
<td>1.25 (0.98-1.61)</td>
<td>Age</td>
<td>Ritchey et al., 2005</td>
</tr>
<tr>
<td></td>
<td>I143V (AA/AG/GG)</td>
<td>Case: 77/21/2, Control: 78/21/1</td>
<td>Chinese</td>
<td>Cases (539): women with primary cervical cancer, Controls (800): Hospital - randomly selected from healthy women</td>
<td>0.99 (0.76-1.30)</td>
<td>Age</td>
<td>Huang et al., 2007</td>
</tr>
<tr>
<td></td>
<td>K178R (AA/AG/GG)</td>
<td>Case: 78/20/2, Control: 81/18/1</td>
<td>Chinese</td>
<td>Cases (539): women with primary cervical cancer, Controls (800): Hospital - randomly selected from healthy women</td>
<td>1.07 (0.80-1.41)</td>
<td>Age</td>
<td>Huang et al., 2007</td>
</tr>
<tr>
<td>Oral cancer</td>
<td>W65C (GG/GC/CC)</td>
<td>Case: 100/0/0, Control: 100/0/0</td>
<td>Thai</td>
<td>Cases (106): diagnosis with oral cancer, Controls (164): Hospital – healthy participant</td>
<td>Not determined</td>
<td>Age, gender, smoking</td>
<td>Kietthubthew et al., 2006</td>
</tr>
<tr>
<td></td>
<td>L84F (CC/CT/TT)</td>
<td>Not determined</td>
<td>Americans</td>
<td>Cases (88): newly diagnosed with an oral cancer, Controls (55): Hospital – healthy participants</td>
<td>0.53 (0.28-1.02)</td>
<td>Age, gender, smoking</td>
<td>Gal et al., 2005</td>
</tr>
<tr>
<td></td>
<td>L84F (CC/CT/TT)</td>
<td>Case: 79/20/1, Control: 79/20/1</td>
<td>Thai</td>
<td>Cases (106): diagnosis with oral cancer, Controls (164): Hospital – healthy participants</td>
<td>1.11 (0.54-2.26)</td>
<td>Age, gender, smoking</td>
<td>Kietthubthew et al., 2006</td>
</tr>
<tr>
<td></td>
<td>I143V (AA/AG/GG)</td>
<td>Not determined</td>
<td>Americans</td>
<td>Cases (88): newly diagnosed with an oral cancer, Controls (66): Hospital – healthy participants</td>
<td>1.10 (0.66-1.41)</td>
<td>Age, gender, smoking</td>
<td>Gal et al., 2005</td>
</tr>
<tr>
<td>Pancreatic cancer</td>
<td>L84F (CC/CT/TT)</td>
<td>Case: 71/27/2, Control: 76/24/0</td>
<td>Americans</td>
<td>Cases (384): diagnosed with primary pancreatic cancer, Controls (357): Hospital – free of pancreatic cancer</td>
<td>1.22 (0.87-1.72)</td>
<td>Age, gender, race, smoking</td>
<td>Jiao et al., 2006</td>
</tr>
<tr>
<td></td>
<td>I143V (AA/AG/GG)</td>
<td>Case: 79/20/1, Control: 78/21/1</td>
<td>Americans</td>
<td>Cases (384): diagnosed with primary pancreatic cancer, Controls (357): Hospital – free of pancreatic cancer</td>
<td>0.98 (0.68-1.40)</td>
<td>Age, gender, race, smoking</td>
<td>Jiao et al., 2006</td>
</tr>
</tbody>
</table>
Table 1.14: MGMT polymorphisms and risk of head and neck, breast and colorectal cancer

<table>
<thead>
<tr>
<th>Type of cancer</th>
<th>MGMT polymorphism</th>
<th>Genotype distribution</th>
<th>Population (source)</th>
<th>Population description</th>
<th>Odds Ratio (95% Confidence Interval)</th>
<th>Parameters Matched</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head and neck cancer</td>
<td>L84F (CC/CT/TT)</td>
<td>Case: 75/23/2</td>
<td>Americans</td>
<td>Cases (514): diagnosed with cancer of the oral cavity and pharynx Controls (754): random-digit telephone dialing</td>
<td>0.75 (0.56-1.02)</td>
<td>Age, gender</td>
<td>Huang et al., 2005</td>
</tr>
<tr>
<td></td>
<td>I143V (AA/AG/GG)</td>
<td>Case: 81/18/1</td>
<td>Americans</td>
<td>Cases (536): diagnosed with cancer of the oral cavity and pharynx Controls (762): random-digit telephone dialing</td>
<td>0.72 (0.52-0.99)</td>
<td>Age, gender</td>
<td>Huang et al., 2005</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>L84F (CC/CT/TT)</td>
<td>Case: 73/25/2</td>
<td>Long Island</td>
<td>Cases (1064): adult female diagnosed with breast cancer Controls (1107): Hospital – random-digit dialing for women under age 65</td>
<td>1.10 (0.90-1.30)</td>
<td>Age</td>
<td>Shen et al., 2005</td>
</tr>
<tr>
<td></td>
<td>I143V (AA/AG/GG)</td>
<td>Case: 78/21/1</td>
<td>Long Island</td>
<td>Cases (1064): adult female diagnosed with breast cancer Controls (1103): Hospital – random-digit dialing for women under age 65</td>
<td>0.90 (0.70-1.10)</td>
<td>Age</td>
<td>Shen et al., 2005</td>
</tr>
<tr>
<td></td>
<td>K178R (AA/AG/GG)</td>
<td>Case: 80/19/1</td>
<td>Long Island</td>
<td>Cases (1067): adult female diagnosed with breast cancer Controls (1110): Hospital – random-digit dialing for women under age 65</td>
<td>0.90 (0.70-1.10)</td>
<td>Age</td>
<td>Shen et al., 2005</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>P58S (CC/CT/TT)</td>
<td>Case: 44/41/15</td>
<td>Iranian</td>
<td>Cases (201): adult female diagnosed with breast cancer Controls (200): Hospital – random-digit dialing for women under age 65 years</td>
<td>1.21 (0.77-1.89)</td>
<td>Age, gender</td>
<td>Khatami et al., 2008</td>
</tr>
<tr>
<td></td>
<td>L84F (CC/CT/TT)</td>
<td>Case: 79/18/3</td>
<td>Americans</td>
<td>Cases (200): diagnosed with colorectal cancer Controls (201): Hospital – free of colorectal cancer</td>
<td>0.81 (0.54-1.21)</td>
<td>Gender</td>
<td>Tranah et al., 2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Case: 20/80/0</td>
<td>Iranian</td>
<td>Cases (186): diagnosed with colorectal cancer Controls (2137): Hospital – free of colorectal cancer</td>
<td>1.23 (0.87-2.36)</td>
<td>Age, gender</td>
<td>Khatami et al., 2008</td>
</tr>
<tr>
<td></td>
<td>I143V (AA/AG/GG)</td>
<td>Case: 86/14/0</td>
<td>Americans</td>
<td>Cases (190): female newly diagnosed with colorectal cancer Controls (2151): blood donors</td>
<td>0.57 (0.37-0.89)</td>
<td>Gender</td>
<td>Tranah et al., 2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Case: 62/37/1</td>
<td>Iranian</td>
<td>Cases (200): diagnosed with colorectal cancer Controls (200): Hospital – free of colorectal cancer</td>
<td>1.68 (1.07-2.63)</td>
<td>Age, gender</td>
<td>Khatami et al., 2008</td>
</tr>
</tbody>
</table>
Table 1.15:  MGMT polymorphisms and risk of esophageal, bladder, prostate cancer and glioma

<table>
<thead>
<tr>
<th>Type of cancer</th>
<th>MGMT polymorphisms</th>
<th>Genotype distribution</th>
<th>Population (source)</th>
<th>Population description</th>
<th>Odds Ratio (95% Confidence Interval)</th>
<th>Parameters Matched</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Esophageal cancer</td>
<td>L84F (CC/CT/TT)</td>
<td>Case: 73/25/2; Control: 77/21/2</td>
<td>Australian</td>
<td>Cases (263): diagnosed with esophageal cancer Controls (1337): randomly selected from the Australian Electoral Roll</td>
<td>1.30 (1.00-1.80)</td>
<td>Age, gender, race, education, smoking</td>
<td>Doecke et al., 2008</td>
</tr>
<tr>
<td></td>
<td>I143V (AA/AG/GG)</td>
<td>Case: 66/30/4; Control: 75/23/2</td>
<td>Australian</td>
<td>Cases (263): diagnosed with esophageal cancer Controls (1337): randomly selected from the Australian Electoral Roll</td>
<td>1.40 (1.10-1.90)</td>
<td>Age, gender, race, education, smoking</td>
<td>Doecke et al., 2008</td>
</tr>
<tr>
<td>Bladder cancer</td>
<td>L53L (CC/CT/TT)</td>
<td>Case: 77/20/3; Control: 85/14/1</td>
<td>Chinese</td>
<td>Cases (167): newly diagnosed with bladder cancer Controls (204): Hospital – free of bladder cancer</td>
<td>1.57 (0.90-2.74)</td>
<td>Age, gender, smoking, alcohol</td>
<td>Li et al., 2005</td>
</tr>
<tr>
<td></td>
<td>L84F (CC/CT/TT)</td>
<td>Case: 79/20/1; Control: 85/14/1</td>
<td>Chinese</td>
<td>Cases (167): newly diagnosed with bladder cancer Controls (204): Hospital – free of bladder cancer</td>
<td>1.52 (0.88-3.66)</td>
<td>Age, gender, smoking, alcohol</td>
<td>Li et al., 2005</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>L84F (CC/CT/TT)</td>
<td>Case: 76/23/2; Control: 87/13/0</td>
<td>Chinese</td>
<td>Cases (161): newly diagnosed with prostate cancer Controls (240): random male over 18 years of age with no history of cancer</td>
<td>1.95 (1.15-3.30)</td>
<td>Age</td>
<td>Ritchey et al., 2005</td>
</tr>
<tr>
<td></td>
<td>I143V (AA/AG/GG)</td>
<td>Case: 96/3/1; Control: 98/2/0</td>
<td>Chinese</td>
<td>Cases (161): newly diagnosed with prostate cancer Controls (240): random male over 18 years of age with no history of cancer</td>
<td>1.55 (0.44-5.47)</td>
<td>Age</td>
<td>Ritchey et al., 2005</td>
</tr>
<tr>
<td>Glioma</td>
<td>L84F (CC/CT/TT)</td>
<td>CC/ CT+ TT Case: 81/19; Control: 74/26</td>
<td>Caucasians</td>
<td>Cases (373): newly diagnosed with glioma Controls (365): random-digit dialing through a contracting company</td>
<td>0.67 (0.45-0.95)</td>
<td>Age, gender, race</td>
<td>Liu et al., 2009</td>
</tr>
<tr>
<td></td>
<td>I143V (AA/AG/GG)</td>
<td>AA + AG/GG Case: 98/2; Control: 99/1</td>
<td>Caucasians</td>
<td>Cases (373): newly diagnosed with glioma Controls (365): random-digit dialing through a contracting company</td>
<td>1.95 (0.58-6.84)</td>
<td>Age, gender, race</td>
<td>Liu et al., 2009</td>
</tr>
<tr>
<td></td>
<td>K178R (AA/AG/GG)</td>
<td>AA + AG/GG Case: 98/2; Control: 99/1</td>
<td>Caucasians</td>
<td>Cases (373): newly diagnosed with glioma Controls (365): random-digit dialing through a contracting company</td>
<td>1.97 (0.57-6.67)</td>
<td>Age, gender, race</td>
<td>Liu et al., 2009</td>
</tr>
</tbody>
</table>
1.8 Aims and Objectives

The MGMT protein protects cells and tissues against the adverse effects of alkylating agents by repairing the $O^6$-alkG generated in DNA by these agents. There are many polymorphic variants of MGMT and our hypothesis was that different MGMT variants would carry different cancer risks, or affect the effectiveness of cancer treatments that employ alkylating agents. However, there are no studies in which the broad substrate specificity of a large number of MGMT variant proteins have been compared simultaneously.

The overall aim of this work was therefore to characterise biochemically different MGMT protein variants and in particular, their ability to repair different $O^6$-alkG substrates.

**Aim 1:** To generate and purify a number of MGMT variant proteins and to assess their relative ability to repair a wide range of $O^6$-alkGs using synthetic ODNs containing 12 different modified purine bases.

*Figure 1.10:* The structure of MGMT showing the locations of the amino acids changed in the polymorphic variants. Those in red are studied in the experimental sections of this thesis. From: [http://www.ncbi.nlm.nih.gov/structure](http://www.ncbi.nlm.nih.gov/structure) (MMDB (Molecular modeling database) ID: 13253).
The reasons for choosing these 6 variants were based upon their reported associations with cancer, their putative affect on MGMT function and to a lesser extent the minor allele frequency. The frequency of the minor F84 allele is ~ 20% and has been associated with an increased risk of prostate, bladder and breast cancer (Ritchey et al., 2005; Li et al., 2005; Shen et al., 2005). Furthermore, F84 is near His85, which is one of the residues that binds to Zinc and as Zinc binding enhances MGMT repair (Daniels & Tainer, 2000; Rasimas et al., 2003) this may alter activity. Q128 is very rare but this position is involved in the DNA binding property of the MGMT (Moore et al., 1994). R160 is also a rare variant but has been shown to confer resistance to O\textsuperscript{6}-BzG and hence may determine chemotherapy response using alkylating agents in combination with O\textsuperscript{6}-BzG (Edara et al., 1996). Q128 and R160 have been significantly associated with colorectal cancer (Khatami et al., 2008). V143/R178 is common in populations and it is very close to the active site Cys-145 which suggests that this amino acid change may alter repair rates (Bugni et al., 2007). This variant has been associated with an increased risk for cervical carcinoma (Huang et al., 2007). The F84 variant was used in combination with R160 and V143/R178 for double (F84/R160) and triple (F84/V143/R178) variants. In addition, women who carry the F84/V143/R178 variant have a significantly increased risk of cervical carcinoma (Huang et al., 2007).

Based on the literature (section 1.7.8), it was proposed to isolate six MGMT variants: F84, Q128, R160, V143/R178, F84/R160 and F84/V143/R178 (See Figure 1.10) and to determine their methyl transfer kinetics, thermal stability and substrate specificity using synthetic ODNs containing G, the O\textsuperscript{6}-position of which was modified with the following alkyl groups: methyl, ethyl, propyl, benzyl, carboxymethyl, hydroxyethyl, [4-oxo-4-(3-pyridyl)butyl] and methyladamantyl. In addition, ODNs containing the following modified purines were used: N\textsuperscript{6}-hydroxypropyl-2,6-diaminopurine, O\textsuperscript{6}-aminoethylG and 2,6-diaminopurine (Figure 2.5). The ODN sequences used are shown in Chapter 2 and the results are described in Chapter 3.
This was anticipated to provide data that would indicate whether any MGMT variant might provide a different level of protection against the exogenous or endogenous agents that might generate such damage in DNA.

**Aim 2:** Investigation of biochemical characteristics of a putative size variant of MGMT.

The MGMT mRNA contains an ATG codon (Figure 1.8) that was upstream of, and in frame with, the ATG that is generally considered to be the unique MGMT initiation codon which encodes the 21.7 kDa MGMT (in this thesis, referred to as “normal” MGMT). This upstream codon would give rise to an MGMT of 25 kDa (here referred to as “extended” MGMT). Literature reports of larger MGMT proteins seen in western blotting studies (Brent et al., 1990; von Wronski et al., 1991; Ostrowski et al., 1991; He et al., 1992) have been attributed to post translational modifications. However, the Kozak consensus sequence, g c c R c c A U G G, (upper case letters are highly conserved bases and AUG is the start codon) which is important in the initiation of translation (Kozak, 1981; Kozak, 1987a; Kozak, 1987b), indicated that the normal MGMT has G and A position +4 and -3 and the extended MGMT has an A at position -3. Therefore, while normal MGMT has a strong consensus, the extended MGMT has an “adequate” consensus (see Figure 1.11) and could give rise to a functionally active, higher molecular weight MGMT.

<table>
<thead>
<tr>
<th>Kozak</th>
<th>g c c A c c A U G G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extended</td>
<td>g g a A g g A U G C</td>
</tr>
<tr>
<td>Normal</td>
<td>g g a A a a A U G G</td>
</tr>
</tbody>
</table>

**Figure 1.11:** The comparison between Kozak consensus sequence and MGMT mRNA sequence around the initiation codon.

To test this hypothesis, it was proposed firstly to examine MGMT mRNA in human peripheral blood mononuclear cells to determine if there was evidence for transcription and splicing that could generate a truncated mRNA that would still encode the normal protein. Secondly, the putative extended protein would be
generated and purified and its characteristics compared with those of the normal protein.

Furthermore, given that previous studies had shown that Zinc supplementation of the growth media in which the *E.coli* clones that had been engineered to express recombinant MGMT proteins affected the activity of the purified protein (Rasimas *et al.*, 2003; Fang *et al.*, 2005), this was also investigated for the normal and extended proteins. In addition, as the presence of 5-MeC adjacent or opposite to $O^6$-MeG in ODN has been reported to affect MGMT activity (Bentivegna & Bresnick, 1994; Guza *et al.*, 2009), this was also investigated. This work is described in Chapter 4.

**Aim 3:** To examine to what extent $O^6$-CMG is a substrate for human MGMT.

The results generated in the first part of these studies indicated that $O^6$-CMG is a good substrate for human MGMT. However, this contradicted a previous report (Shuker & Margison, 1997; Kaina *et al.*, 2007) and it was considered that further confirmation of this finding was needed to ensure its validity. To investigate the basis of the original incorrect report, one of the *E.coli* alkyltransferase proteins, encoded by the *Ogt* gene was purified and characterised. These studies were extended to investigate the effects on MGMT activity and $O^6$-MeG and $O^6$-CMG DNA levels, of treatment of an MGMT-expressing cultured human cancer cell line with the carboxymethylating agent, azaserine. This work is described in Chapter 5.
Chapter 2: Materials and Methods

Methodology is presented in the context of the various projects undertaken, as described below. Except in those sections that mention MGMT variants, the sequence used is the common variant, referred to as Wild-type (Wt).

The polymerase chain reaction (PCR) is used extensively in this work. PCR is a biochemical technique used to amplify a specific region of a DNA strand, generating millions of copies of a particular DNA sequence. A basic PCR set-up requires several components including a DNA template, forward and reverse primers, *Taq* polymerase, PCR buffer and deoxyribonucleoside triphosphates (dNTPs) and a machine that heats and cools the tubes in a repeat cycle and in a programmable fashion.

The DNA template contains the target DNA region to be amplified and typically between 0.1 and 1 μg genomic DNA is used in a total reaction mixture of 100 μl. The forward and reverse primers are complementary to the 3' ends of both the sense and anti-sense strand of the DNA target. PCR primers should be 10-24 nucleotides in length with a GC content of between 40-60%. The primer should not be self-complementary or complementary to the other primer in the reaction mixture, to prevent primer-dimer or hairpin formation. *Taq* or other thermostable polymerases withstand the high temperatures needed for DNA-strand separation. The PCR buffer provides the optimum ionic and pH conditions for the activity and stability of the *Taq* polymerase and the dNTPs (dATP, dCTP, dGTP, dTTP), are required to synthesize new DNA strands.

During PCR amplification of target DNA, in principle, the DNA template is first denatured at 96°C, the reaction mix is cooled to the annealing temperature of the primers, and then the temperature is increased to the optimum for the polymerase activity. Then the reaction is heated to 96°C again. Each phase is usually completed in 30-60 seconds, and the denaturation-annealing-extension cycle is repeated usually 30 times. At each extension step, the amount of DNA product is doubled, leading to exponential amplification of the specific DNA template.
The PCR products are separated by gel electrophoresis depending on the size of the amplified fragment, and are visualized by staining with GelRed. The size of PCR products is compared with a DNA marker, which contains DNA fragments of known size, run on the gel alongside the PCR products (McPherson & Møller, 2000; Roux, 2004)

The melting and annealing temperatures of a primer are estimated as follows: if the primer is shorter than 25 nucleotides, the approximate melting temperature is calculated from the equation: \( T_m = 4(G + C) + 2(A + T) \). However, software is the more accurate way to assess secondary structures and \( T_m \). One example is Primer-Blast at www.ncbi.nlm.nih.gov/tools/primer-blast/ which provides two options: hetero-dimer and BLAST. The former assesses potential dimeric interactions of the primers and BLAST is the basic alignment search tool that takes the primer sequence and searches the entire database for matches. BLAST then shows a list of all matches found with an alignment score (Dieffenbach & Dveksler, 2003; Apte & Daniel, 2009). Using these tools can avoid primer self-interactions and non-specific amplification of non-target sequences.

qPCR was not used in my studies but is added here for information. qPCR can be used to measure the quantity of starting amounts of DNA, cDNA, or RNA templates in real-time PCR (RT-PCR) and for studying gene expression levels (Livak & Schmittgen, 2001). The method uses fluorescent dyes, for example Sybr Green and EvaGreen or fluorophore-containing DNA probes, such as TaqMan probes, to measure the increase in the amount of amplified product after each amplification cycle (Stahlberg et al, 2004). PCR efficiencies are usually calculated from standard curves based on either known DNA inputs or on dilution series of a reference cDNA sample. Alternatively, the target DNA concentration can be directly estimated from the fluorescence data, with no requirement for a calibration curve (Swillens et al., 2004).

2.1 Expression of extended and normal MGMT in human lymphocytes

Complementary DNAs from five human PBMC samples were generated by Mrs Gail McGown and Mrs Mary Thorncroft from the Carcinogenesis group at the
Paterson Institute for Cancer Research (PICR). These were amplified using a common reverse primer (5’- PO₄-TCAGTTTCGGCCAGCAGGCGGGGA-3’) and unique forward primers that were designed to anneal to the extended (5’-PO₄- ATGCTGGGACAGCCCGCAGCCCTCTA-3’) and normal (5’-PO₄-ACTGGACAAGGATTGTGAAATGAAACGCACGACCACACTGGACAGGCCCTTTGGGGAAGCTGGAGCTGTCTGGTTGTGAGCAGG GTCTGCAGACGATCCTGGGCAAGGGGACCAGTCTGACAGCCGGTGGGGATGCTGCAGCAGCGGAGCCGTGGGCAACTACT CCGGGAGGACTGGCCGTGAAGGAATGGCTTCTGGCCCATGAAGGCCACCG GTTGCTGGGAGGCTACCTCAGACTCAGGTGGTGCTGGGAGGGAGCTCAGGTCTGGCAGGGGCCTGG CTCAAGGGAGCGAGCTACCTCAGACTCAGGTGGTGCTGGGAGGGAGCTCAGGTCTGGCAGGGGCCTGG

CTCGGCCCCGCCCCGCAGGAGGTATGCTGGGACAGCCCGCAGCCCTCTA GAAACGGCTTTTGGCTCCGAGCCGCAGGAGGTCTCTGAGTGCTGCAGCAGGGTGGGGATTGTGAAATGAAACGAC CTGGGACAGCCCTTTGGGGAAGCTGGGAGCTGGTGTTGTGAGCAGG GTCTGCAGACGATCCTGGGCAAGGGGACCAGTCTGACAGCCGGTGGGGATGCTGCAGCAGCGGAGCCGTGGGCAACTACT CCGGGAGGACTGGCCGTGAAGGAATGGCTTCTGGCCCATGAAGGCCACCG GTTGCTGGGAGGCTACCTCAGACTCAGGTGGTGCTGGGAGGGAGCTCAGGTCTGGCAGGGGCCTGG CTCAAGGGAGCGAGCTACCTCAGACTCAGGTGGTGCTGGGAGGGAGCTCAGGTCTGGCAGGGGCCTGG

The sequence was from GenBank (accession number: NM_002412.3).

PCR amplifications were carried out using 0.5 µl of Taq polymerase, 10 µl of five times concentrated Taq buffer, 1 µl of 1 mM dNTPs (all from Promega), 20 pmol of each primer and 1 µl of DNA template, made up using double distilled water to a total volume of 50 µl. The cycling conditions were: 1 cycle of 30 sec at 96°C, followed by 25 cycles of 30 sec at 96°C, 30 sec at 55°C and 1 min at 72°C. An aliquot (8 µl) of the PCR products was mixed with 2 µl of non-denaturing gel loading buffer concentrate (10 mM Tris-HCl pH 7.5, 0.4 % w/v orange G, 0.03 % w/v bromophenol blue, 0.03 % w/v xylene cyanol FF, 15% v/v Ficoll® 400 and 50 mM EDTA) and then subjected to electrophoresis in a 1% agarose gel, precast with
GelRed (Biotium, used according to the manufacturer’s instructions) in Tris-Borate-EDTA (TBE; 89 mM Tris, 89 mM boric acid, 2 mM EDTA) at 50 volts for 45 min and visualized on an ultraviolet (UV) light transilluminator. Another aliquot (8 µl) of the PCR products was digested at 37°C for 1 h with 0.5 units of SapI and PstI in the buffer provided by the manufacturer (New England Biolabs: NEB) in a total volume of 10 µl.

2.1.1 DNA sequencing

The digestion products were separated and visualised as described above. DNA sequences after digestion (350-500 ng) were obtained using 10 pmol of the forward and reverse primers (see above) in a total volume of 12 µl. Sequencing was carried out using an ABI capillary sequencer by the Molecular Biology Core Facility at the PICR. The sequence obtained was aligned with that of the MGMT cDNA (see Figure 2.1), using Clustal W, available at [http://www.ebi.ac.uk/Tools/clustalw2/index.html](http://www.ebi.ac.uk/Tools/clustalw2/index.html).

2.2 Generation, expression and purification of extended and normal MGMT

2.2.1 Generation of MGMT by PCR

pCMV6-AC-MGMT-GFP (OriGene) was used as a source of both the extended and normal versions of MGMT. To establish a suitable amount of pCMV6-AC-MGMT-GFP DNA to use as template, serial dilutions from 0.001-100 pg/µl were PCR-amplified using the primers and conditions described in section 2.1.

2.2.2 DNA ligation

An aliquot (4 µl) of the PCR products obtained by amplification of 10 pg of pCMV6-AC-MGMT-GFP was ligated into XmnI (NEB) digested and phosphatase (NEB) treated pMAL-c2 (NEB) using 0.5 µl of T4 DNA ligase (Promega), 2 µl of ten times concentrated ligase buffer (660 mM Tris-HCl, 50 mM MgCl$_2$, 50 mM DTT and 10 mM ATP) in a total volume of 20 µl and incubated at room temperature for 1 h.

2.2.3 Transformation
The ligation product from section 2.2.2 was used to transform XL-1 blue *E.coli* (Promega). Competent cells (50 µl) were slowly thawed on ice and 2 µl of ligase reaction mix was added, mixed and placed on ice for 20 min. The reaction mix was heated at 42°C for 50 sec, then returned to ice for 2 min. Luria broth (LB; 1% w/v tryptone, 0.5% w/v yeast extract, 0.5% w/v NaCl; 950 µl) was added and the suspension incubated in a 37°C shaking incubator for 1 h. An aliquot (100 µl) was spread onto LB agar (1.5% agar in LB) plates containing ampicillin (100 µg/ml) and incubated overnight at 37°C.

2.2.4 PCR colony amplification

To confirm the presence of the extended and normal MGMT inserts, single colonies from the above plates were suspended in 100 µl LB containing ampicillin (100 µg/ml; LB/amp). PCR amplifications were carried out as described in 2.1 using 5 µl of these suspensions in a total volume of 50 µl. The PCR products were analysed as described in section 2.1.

2.2.5 Plasmid purification

Bacterial cultures (for example, 40 µl of the section 2.2.4) were added to universal tubes containing 10 ml of LB/amp. An aliquot (2 ml) was transferred to transformation tubes which were taken to the Molecular Biology Core Facility for robotic plasmid miniprep following overnight growth at 37°C with shaking. The miniprep DNA concentration was determined by absorption at 260 nm using an ND-1000 spectrophotometer (NanoDrop Technologies, hereafter called “Nanodrop”). The plasmid DNA sequence was obtained using a *mal*E primer (5’-GGTCGTCAG ACTGTCGATGAAGCC-3’) and the conditions as described in section 2.1.1.

2.2.6 Expression of recombinant MBP fusion proteins

To express recombinant proteins, a bacteria suspension from section 2.2.4 was inoculated into 10 ml of LB/amp containing 0.2% w/v glucose and incubated in a shaker at 37°C overnight. The overnight pre-culture (5 ml) was added to 100 ml of the same medium and further incubated until the optical density of the suspension at 595 nm reached 0.6. Protein expression was induced by the addition of
isopropylthiogalactoside (IPTG) to a final concentration of 0.4 mM. In some experiments, ZnCl₂ was added to a final concentration of 0.1 mM. After incubation for 1, 2 and 3 h, 10 ml of the culture were centrifuged at 2500g for 10 min at 4°C. The cell pellets were washed with 20 mM Tris-HCl (pH 8.3), resuspended in 1 ml of binding buffer (BB; 20 mM Tris-HCl pH 8.3, 200 mM NaCl, 1 mM EDTA) containing 5 µg/ml leupetin and transferred to a 1.5 ml eppendorf tube. Extracts were made by sonication (10 sec pulse, setting 4.5; Sonicator XL; Heat Systems) and phenylmethylsulphonylfluoride (PMSF; 10 µl of 8.7mg/ml solution in absolute ethanol) was immediately added. The extracts were then centrifuged at 17000g for 10 min at 4°C.

The supernatants were removed and the protein concentration was determined using the Bradford assay as follows. The standard curve was Bovine Serum Albumin (BSA) diluted in buffer I (50 mM Tris-HCl pH 8.3, 1mM EDTA, 3 mM DTT) to concentrations of 0, 0.02, 0.03, 0.04, 0.05, 0.06, 0.08 and 0.1 mg/ml. Bio-Rad protein reagent concentrate was diluted to a working solution and filtered through two sheets of 3MM filter paper prior to use. Forty µl of each standard and unknown (diluted as appropriate in buffer I) were added in duplicate to a microtitre plate and 200 µl of Bio-Rad reagent was added to each well and left at room temperature for 5 min. The absorbance was measured at 595 nm on a TECAN Genios plate reader. The results were accepted when the standard curve was linear (r > 0.99).

Aliquots of extract were analysed for expression by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) with Coomassie staining using a Mini Protein (BioRad) apparatus. A 12% gel was prepared by mixing 8 ml Protogel (37.5:1 acrylamide: bisacrylamide), 5ml 1.5 M Tris-HCl (pH 8.8), 200 µl 10% SDS, 6.6 ml ddH₂O, 200 µl 10% ammonium persulphate (APS) and 20 µl tetramethylethylenediamine (TEMED) and poured into the cassette until it was ¾ full, then water saturated butanol was layered onto the top. When the gel was set, the butanol was removed and 5% stacking gel (consisting of 3.4 ml Protogel, 2.5 ml 1 M Tris-HCl pH 6.8, 200 µl 10% SDS, 13.6 ml ddH₂O, 200 µl 10% APS and 20 µl TEMED) was added followed by the well-forming comb. When the gel was set, the wells were washed with ddH₂O, the cassette attached to the electrode block and then
running buffer (0.192 M glycine, 25 mM Tris-Base, 0.1% w/v SDS stored at 4°C) was added. Samples were mixed with loading buffer (50 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 1% β-mercaptoethanol, 12.5 mM EDTA, 0.02% bromophenol blue) and carefully layered into the wells. Twenty µl of the 1 ml extracts were analysed by 12%SDS-PAGE with Coomassie staining. Electrophoresis was carried out at 200V until the bromophenol blue tracking dye was approximately 5mm from the bottom of the gel. The gel was then removed and stained for 30 min with Coomassie Blue (0.2% Coomassie Brilliant Blue G250, 30% methanol, 0.5% acetic acid) followed by destaining in 20% methanol, 10% acetic acid. Images were captured using a Genesnap software (Syngene) and a Pulnix camera.

2.2.7 Purification of MBP fusion proteins

Amylose resin (1 ml) was added to an empty plastic column and washed with 8 ml of binding buffer (BB; 20 mM Tris-HCl pH 8.3, 200 mM NaCl, 1 mM EDTA). An aliquot (300 µl) of the total protein extract as prepared in section 2.2.6 was applied to the column, followed by washing with BB. Fractions (1 ml) were collected until trace amounts of protein (less than 0.1 mg/ml) were detected by measuring the absorption at 280 nm using a Nanodrop. The column was then eluted using BB containing 10 mM maltose and 1 ml fractions were collected. Fractions containing protein (again determined by Nanodrop) were analysed by 12%SDS-PAGE with Coomassie staining (see section 2.2.6).

2.2.8 Western blotting

Proteins were separated by 12%SDS-PAGE (see section 2.2.6) and transferred by electroblotting onto a polyvinylidene difluoride (PVDF) membrane in 25 mM Tris-Base, 192 mM glycine, 10% methanol for 1 h at 100V. The membrane was blocked with gentle rocking using 5% Marvel in TBST (50 mM Tris-HCl pH 7.5, 150 mM NaCl containing 0.1% v/v Tween-20) and then washed 3 times in TBST. According to the experiment, a primary antibody (Table 2.1) in TBST containing 0.5% Marvel was added at the dilution indicated and incubated for 1 h at room temperature. After three 5 min washes with TBST, the membrane was incubated with secondary antibody for 30 min at room temperature and again washed
3 times in TBST. The ECL detection solutions (Amersham, GE Healthcare) were mixed according to the manufacturer’s instructions and used to soak the membrane, which was then wrapped in Saran wrap and exposed to high performance autoradiography film (Amersham, GE Healthcare) for various times prior to development. Images were captured using a flatbed scanner (Epson).

**Table 2.1:** Primary and secondary antibodies used in these studies

<table>
<thead>
<tr>
<th>Antibody (Source)</th>
<th>Type</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary antibody</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Anti-MBP (NEB)</td>
<td>Mouse polyclonal</td>
<td>1/5000</td>
</tr>
<tr>
<td>2. Anti-MGMT R63 (in house)</td>
<td>Rabbit polyclonal</td>
<td>1/500</td>
</tr>
<tr>
<td>3. Anti-Atl1 (in house)</td>
<td>Rabbit polyclonal</td>
<td>1/500</td>
</tr>
<tr>
<td>4. Anti-MeG (Squarex)</td>
<td>Rabbit polyclonal</td>
<td>1/500</td>
</tr>
<tr>
<td>5. Anti-CMG (in house)</td>
<td>Rabbit polyclonal</td>
<td>1/500</td>
</tr>
<tr>
<td>6. Anti-Ogt (in house)</td>
<td>Rabbit polyclonal</td>
<td>1/25</td>
</tr>
<tr>
<td><strong>Secondary antibody</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Polyclonal Goat Anti-Rabbit Immunoglobulin/HRP (Dako)</td>
<td>Rabbit polyclonal</td>
<td>1/1000</td>
</tr>
<tr>
<td>2. Polyclonal Goat Anti-Mouse Immunoglobulin/HRP (Dako)</td>
<td>Mouse polyclonal</td>
<td>1/1000</td>
</tr>
</tbody>
</table>

### 2.2.9 Cleavage of MBP-MGMT fusions proteins

Aliquots (50 µg) of amylose-purified MBP-MGMT fusion proteins were digested with 1 µg of Factor Xa (NEB) in 50 µl of digest buffer (NEB) containing 1 mM DTT at room temperature for 1, 2, 4, 5 and 6 h. The efficiency of the digest was analysed by 12%SDS-PAGE with Coomassie staining (see section 2.2.6). Six-hour digests were found to be appropriate.

Figure 2.2 shows a schematic diagram of the procedure used for MGMT expression and purification.
Figure 2.2: Schematic diagram of MGMT purification using amylose, DEAE-sepharose and a second amylose column after factor Xa digest.
2.2.10 Purification of MGMT

The digestion mixture (40 µl) from section 2.2.9 was applied to a 1 ml column of DEAE-Sepharose (Sigma) pre-equilibrated with 10 mM Tris-HCl, 25 mM NaCl (pH 8.0). The column was washed with 5 ml of the same buffer, collecting 1 ml fractions, then was eluted stepwise with 1 ml fractions of 50 mM NaCl to 500 mM NaCl in 20 mM Tris-HCl (pH 8.0) in 50 mM NaCl increments. The protein concentration of the eluted fractions was determined by measuring the absorption at 280 nm by Nanodrop and those containing protein were analysed on 12%SDS-PAGE with Coomassie staining (see section 2.2.6). MGMT containing fractions were pooled and re-applied to amylose columns (see section 2.2.7) to remove MBP and any remaining uncleaved MBP fusion protein.

2.3 Generation, expression and purification of MGMT variant proteins

2.3.1 Site directed mutagenesis

To generate the 84Phe (F84), 128Gln (Q128), 160Arg (R160), 84Phe-160Arg (F84/R160), 143Val-178Arg (V143/R178) and 84Phe-143Val-178Arg (F84/V143/R178) variant cDNAs, PCR-mediated site-directed mutagenesis was used. pMAL-c2 containing the normal Wt MGMT sequence (Figure 2.3) was used as template and was amplified using the pairs of primers listed in Table 2.2.

![Figure 2.3: Schematic diagram of the directed mutagenesis procedure. The phosphorylated forward (mutagenic) and reverse primers were designed to anneal back to back on the plasmid.](image)

PCR amplifications were carried out using 0.5 µl of Phusion hot start DNA polymerase (NEB), 10 µl of five times concentrated buffer (NEB), 1 µl of 1 mM
dNTPs, 20 pmol of each primer and 10 pg of DNA template in a total volume of 50 µl. The cycling conditions were: 1 cycle of 30 sec at 96°C, followed by 25 cycles of 30 sec at 96°C, 30 sec at 45°C and 4 min at 72°C. An aliquot (8 µl) of the PCR products was electrophoresed on a 1% agarose gel containing GelRed in TBE at 50-volt for 45 min and visualized on a UV transilluminator.

**Table 2.2:** Primers used to generate and sequence the MGMT variants.

<table>
<thead>
<tr>
<th>MGMT variants</th>
<th>Primers used to generate and sequence variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal Wt</td>
<td>Forward: 5'- PO₄- ATG GAC AAG GAT TGT GAA ATG AAA C</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5'- PO₄- TCA GTT TGG CCA GCA GGC GGG GA</td>
</tr>
<tr>
<td>F84</td>
<td>Forward: 5'- PO₄- GGT CCG GCT TTT CAC CAT CCC GTT TTC CAG</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5'- PO₄- GGG GAA CTC TTC GAT AGC CTC GGG C</td>
</tr>
<tr>
<td>Q128</td>
<td>Forward: 5'- PO₄- AAG CCG CGC AAG CAG TGG GAG GAG CAA TGA</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5'- PO₄- GCC CAC GGC TCC GCT GCT GCA GAC C</td>
</tr>
<tr>
<td>R160</td>
<td>Forward: 5'- PO₄- AAC TAC TCC AGA GGA CTG GCC GTG AAG GAA</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5'- PO₄- GCC CAC GGC TCC GCT GCT GCA GAC C</td>
</tr>
<tr>
<td>V143</td>
<td>Forward: 5'- PO₄- CCC ATC CTC GTC CGG TGC CAC AGA GTG GTC</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5'- PO₄- GCC CAC GGC TCC GCT GCT GCA GAC C</td>
</tr>
<tr>
<td>R178</td>
<td>Forward: 5'- PO₄- CGG TTG GGG AGG CCA GGC TTG GGA GGG AGC</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5'- PO₄- GGT GGC CTT CAT GGG CCA GAA GCC ATT CCT</td>
</tr>
</tbody>
</table>

*Underlined bases indicate the codons changed.

The PCR products were circularized with T₄ DNA ligase (see section 2.2.2) and transformed into competent XL-1 blue *E.coli* (see section 2.2.3). After transformation, single colonies were suspended in LB/amp and used for PCR amplification with normal forward and reverse primers (Table 2.2). Positive colonies were subjected to the miniprep procedures (see section 2.2.5) and sequenced to ensure that the appropriate mutations had been generated (see section 2.1.1). Overexpression and purification of the MGMT variant-MBP fusion proteins was carried out as described in sections 2.2.6 to 2.2.10.

**2.4 Generation, expression and purification of Ogt protein**

The *ogt* gene was isolated by PCR amplification of *E.coli* K12 DNA using phosphorylated forward (5’-PO₄-ATGCTGAGATTACTTTGAAGAAAAAATTGC)
and reverse (5’-PO₄-TTACAGCAAAAGATAACCTTCATGGCG) primers as shown in Figure 2.4. The cycling conditions were: 1 cycle of 30 sec at 96°C, followed by 25 cycles of 30 sec at 96°C, 30 sec at 55°C and 1 min 30 sec at 72°C. The PCR product was ligated into XmnI digested and phosphatase treated pMAL-c2 and transformed into competent XL-1 blue E.coli as described in section 2.2.2. Following confirmatory DNA sequencing (see section 2.1.1), processing to obtain purified Ogt protein was as described in sections 2.2.7 to 2.2.10.

TCTGCCGATAGGTCCGGGTATTTTACCACGTGGTTGCTTTAGAGAGAAGG
ATGCTGAGATTACCTGAAAGAAAAATTGCAAGGACACTGCGGTCCACTGT
GGGTTAGATTTTGAGATGACATTTCTGCTGGGCGGGTTGGAATGGGAAGA
GTACAGCGCAAGCATGTCGCTGACCATCCATTATCGCAAAAGAA
GGCTATGAGGCATTCTGACCACCAATCCAGGCGGTTTTAAGCGCAAGCT
TCGTGAACTATTGCGGTAATCTTAGCATTATTGATACGCTTCCACTGC
TACGGGGGGGACGCGCATTACTATGCGCAAGTCTTGGAAAAACTACTGCAC
ATCCCGCTCGGCAGGTAATGCAATGCGCAATCCGCTGACGCAATGG
GCCGTCTGGTGCAGGTGCCGTGCTGGTCCGAACGAGATCGGACATTCC
CATCAGCATCGTCGCTACCTTGCCATCGGGTATTGCGGCAACGACGCACCA
TGACCAGATATGCGAGGCAGGTCCAGCGAAAAGATTGTGTATGCGCCA
TGAAGGTTATCTTTGTCTGTAAACATTAAAAATTTGTGCCAGCTTGTTC
ACACTTTTATGTAAGTTACCC

**Figure 2.4:** Nucleotide sequence of the E.coli genome encompassing ogt, indicating the location (underlined) of the forward and reverse primers. The sequence was from GenBank (accession number: NC_000913.2).

### 2.5 Functional activity of MGMT and Ogt proteins

#### 2.5.1 Determination of activity

To establish the amounts of the MGMT and Ogt proteins that were appropriate for accurate quantitation of activity and inactivation, a standard methyltransferase assay was used. Varying amounts of the proteins were incubated initially at 37°C for 1 h with 100 µl of N-[³H]-methyl-N-nitrosourea methylated calf thymus DNA in a total volume of 300 µl of buffer I containing 1mg/ml BSA (IBSA).
This substrate DNA contained approximately 800 cpm in O\textsuperscript{6}-MeG and was synthesised in house. After incubation, 100 µl of 10 mg/ml BSA, 100 µl of 4M perchloric acid (PCA) and 2 ml of 1M PCA were added and the DNA was hydrolysed at 77\degree C for 50 min. Samples were centrifuged at 2800 rpm for 10 min at 20\degree C, the supernatant aspirated and the pellet washed with 4 ml of 1M PCA. Water (300 µl) and scintillation fluid (3 ml; Ecoscint, Mensura Technology) were added and radioactivity, representing transfer of [\textsuperscript{3}H]-methyl groups to the proteins, was determined by scintillation counting. Results are presented as radioactivity (cpm) or fmoles of methyl groups transferred to protein, based on 0.133 cpm per fmole of methyl groups.

2.5.2 Kinetics of methyl group transfer to MGMT

Protein-limiting amounts of MGMTs or Ogt from the above assays (section 2.5.1) were incubated with [\textsuperscript{3}H]-methylated substrate DNA (100 µl) in a total volume of 200 µl of IBSA at 37\degree C for 0.5, 1, 2, 3, 4 and 5 h. The samples were processed as described in section 2.5.1 and the results are presented as fmoles of methyl groups transferred to protein versus incubation time.

2.5.3 Thermal stability

Protein-limiting amounts of MGMTs or Ogt were pre-incubated in a total volume of 100 µl of IBSA at 44\degree C or 51\degree C for 5, 10, 20, 30, 40, 50 and 60 min followed by the addition of [\textsuperscript{3}H]-methylated substrate DNA (100 µl: section 2.5.1) and further incubation for 1 h at 37\degree C. The samples were processed as described in section 2.5.1 and the results presented as % of control (\textit{i.e.} no pre-incubation) activity.

2.5.4 MGMT and Ogt inactivation by modified ODNs

ODNs containing modified \textit{O}\textsuperscript{6}-alkGs were synthesised using phosphoramidite chemistry on an Applied Biosystems 394 automated synthesiser by Dr David Williams and Dr Oliver Wilkinson, University of Sheffield.
The effect of preincubation of alkyltransferase proteins with various single-stranded ODN (5'-SIMA-GCCATGXCTAGTA) where SIMA (4,7-Dichloro-2',7'-diphenyl-fluorescein) is the fluorophore and X represents guanine (G) or a variety of modified purines (see Figure 2.5) was then determined. In some studies, the ODNs
were annealed to their complementary sequence to generate double stranded substrates, and in some cases the complementary strand contained biotin at the 5’-end.

A pilot experiment was first carried out using a protein-limiting amount of MGMT or Ogt (section 2.5.1) and 2.44 and 24.4 nM of each ODN. These were incubated for 1 h at 37°C followed by the addition of [3H]-methylated substrate DNA (100 µl) and further incubated for 1 h at 37°C. The samples were then processed as described in section 2.5.1. Results are presented as % MGMT activity remaining. To accurately determine the IC50 values (i.e. the amount of inactivator required to reduce the alkyltransferase activity by 50%) a narrower concentration range of each of the ODN was used in the assays.

2.5.5 The effects on MGMT activity of adding zinc to E.coli growth media

_E.coli_ harbouring the pMAL-c2 MGMT were grown in 10 ml of LB/amp containing glucose with shaking at 37°C overnight (see section 2.2.6) and an aliquot (5 ml) of this culture was added to 100 ml of the same medium containing 0.1 mM ZnCl2. Subsequent growth, IPTG induction, and processing to obtain purified protein was as described in section 2.2.6 to 2.2.10. The activities and ODN IC50 of these purified proteins were determined in parallel as described above (section 2.5.1).

2.5.6 Surface plasmon resonance (SPR) assay of the interaction between MGMT Wt and variant proteins with O6-alkG containing ODNs

SPR is a technique that is widely used to investigate molecular interactions, for example between proteins, between proteins and DNA and between proteins and lipids. SPR occurs when polarized light strikes an electrically conducting surface at the interface between two media. This generates electron charge density waves called plasmons, reducing the intensity of reflected light at a specific angle known as the resonance angle, in proportion to the mass on a sensor surface. To detect an interaction, the ligand (in this thesis, ds ODN) is initially immobilised onto the sensor surface. Its binding partner (in this thesis, MGMT protein) is injected under continuous flow through the flow cell. As the protein binds to the DNA, the accumulation of protein on the surface results in an increase in the refractive index.
This change in refractive index is measured in real time, and the result plotted as response or resonance units (RU) versus time generating a sensorgram.

ODN (5'-SIMA-GCCATGXCTAGTA) where X represents G, O6-MeG, O6-CMG, O6-BzG, O6-HOEtG and O6-PobG were annealed to the biotinylated complement (5'-TACTAGCCATGGC-Biotin) by mixing 1 nmole of each in a total volume of 100 μl of 0.05 M NaCl and heating in a dry-block at 80°C for 5 min, then slowly cooling to room temperature. The double-stranded ODNs were diluted to 10 nM in phosphate-buffered saline containing 0.005% Tween20 (PBST) and applied at room temperature to a ProtEon (BioRad) SPR machine containing a neutravidin-coated sensor chips (Bio-Rad) at a flow rate 30 μl/min until a signal of ~100 RU units was achieved. Each of the 6 lanes on the chip was thus coated with a different ODN. The machine then rotated the chip through 90° and aliquots of dilutions of the various MGMTs being assessed (0, 10, 20, 40, 80 and 100 nM) were flowed over the chip at 60 μl/min for 300 seconds (association time) followed by PBST for 900 seconds (dissociation time) at a flow rate of 100 μl/min, and signals recorded. Finally to qualitatively assess residual levels of O6-alkGs in the bound ODN, the lesion-binding S.pombe protein, Atl1 (20 nM, in house synthesis) was flowed over the chip at a flow rate 60 μl/min and signals recorded.

2.5.7 Enzyme-linked immunosorbent assay (ELISA) analysis of MGMT binding to double-stranded ODNs

The G and O6-alkG containing ODNs listed in Table 2.3 were annealed to 3'-biotinylated complements as described in section 2.5.6 and 50 and 500 fmoles/well in PBS/0.5mg/ml BSA were added to streptavidin (SA)-coated 96-well plates (Thermo scientific) that had been preblocked by addition of 400 μl per well of 3% (w/v) BSA in PBS and incubation at room temperature for 2 h. Increasing amounts of MGMT (0-60 fmoles) in 100 μl buffer I were added to the wells and incubated at 37°C for 1 h. The plates were washed three times with 100 μl PBS/0.5mg/ml BSA followed by sequential 1 h room temperature incubations with anti-MGMT (1/500 in PBS/0.5mg/ml BSA), and goat anti rabbit HRP (1/1000 in PBS/0.5mg/ml). Western Lightning reagent (Perkin Elmer) was added and binding quantified by measurement
of chemiluminescence on a TECAN GENios plate-reader. Results were the means (+/- SD) of triplicate determinations.

Table 2.3: Control (G) and $O^6$-alkG (X) containing ODNs and complementary ODNs used in ELISA assays.

<table>
<thead>
<tr>
<th>ODN*</th>
<th>$O^6$-alkG/G/complement</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’-SIMA-GCCATGGCTAGTA</td>
<td>G</td>
</tr>
<tr>
<td>5’-SIMA-GCCATGXCTAGTA</td>
<td>$O^6$-MeG</td>
</tr>
<tr>
<td>5’-SIMA-GCCATGXCTAGTA</td>
<td>$O^6$-BzG</td>
</tr>
<tr>
<td>5’-TACTAGCCATGGC-B</td>
<td>Complement</td>
</tr>
<tr>
<td>5’-GAACTGCAGCTCCGTGCTGGCCC</td>
<td>G</td>
</tr>
<tr>
<td>5’-GAACTXCAGCTCCGTGCTGGCCC</td>
<td>$O^6$-MeG</td>
</tr>
<tr>
<td>5’-GAACTXCAGCTCCGTGCTGGCCC</td>
<td>$O^6$-BzG</td>
</tr>
<tr>
<td>5’- B-GG GCCAGCAGGGAGCTGCAGTTC</td>
<td>Complement</td>
</tr>
</tbody>
</table>

*SIMA is the fluorophore, X indicates the lesion position and B indicates biotin.

2.5.8 Effect of cytosine methylation on the inactivation of MGMT by $O^6$-MeG

Inactivation of MGMT Wt protein by the ss and ds stranded $p53$ codon 245-spanning control, 5-methylcytosine (5-MeC) and/or $O^6$-MeG-containing ODNs (shown in Table 2.4) was then determined. MGMT proteins were incubated with ODNs for 1 h at 37ºC followed by the addition of excess [3H]-methylated substrate DNA (100 µl) and incubation for 1 h at 37ºC. The samples were processed as described in section 2.5.4 and the results are presented as % MGMT activity remaining.

2.5.9 Mass spectroscopic analysis of MGMT proteins

Single-stranded ODNs (5’-SIMA-GCCATGXCTAGTA) where X represents various $O^6$-alkGs, were incubated with MGMT proteins for 1 h at 37ºC. Negative controls contained no ODNs. Trypsin (1 µg; ratio of MGMT: trypsin, 50:1) was added and the reaction was incubated overnight at 37ºC. The digestion was terminated with the addition of 0.1% formic acid. MALDI-TOF analyses were performed on a Bruker UltraflexTM (Bruker Daltonics, Bremen, Germany). Full scans
of the peptide mixture from 800 to 4000 m/z and tandem mass spectral data of select ions were collected with α-cyano-4-hydroxycinnamic acid as the matrix. External calibration was performed with QCAL standard (Eyers et al., 2008), which is a standard for assessing instrument conditions for proteome analysis.

Table 2.4: ODNs used to assess the effect of 5-methylcytosine on the inactivation of MGMT by \(O^6\)-MeG.

<table>
<thead>
<tr>
<th>ODN abbreviations</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(\text{me}G) (ss)</td>
<td>5’-GCATGGGC (\text{me}G)GCATGAACCG-3’</td>
</tr>
<tr>
<td>(\text{me}C\text{me}G) (ss)</td>
<td>5’-GCATGGG (\text{me}C\text{me}G)GCATGAACCG-3’</td>
</tr>
<tr>
<td>CG (ss)</td>
<td>5’-GCATGGGC (\text{C}G) GCATGAACCG-3’</td>
</tr>
<tr>
<td>G: (\text{me}C) (ds)</td>
<td>5’-GCATGGGC GCATGAACCG-3’ (3’-\text{CGTACCCG (\text{C}CGTACTTGGC-5’})</td>
</tr>
<tr>
<td>(\text{me}G\text{C}(\text{ds})</td>
<td>5’-GCATGGGC (\text{CG}C)GCATGAACCG-3’ (3’-\text{CGTACCCG CCGTACTTGGC-5’})</td>
</tr>
<tr>
<td>(\text{me}C\text{me}G) (ds)</td>
<td>5’-GCATGGGG (\text{me}C\text{me}G)GCATGAACCG-3’ (3’-\text{CGTACCCG G CCGTACTTGGC-5’})</td>
</tr>
<tr>
<td>(\text{me}C\text{me}G:G\text{me}C) (ds)</td>
<td>5’-GCATGGGG (\text{me}C\text{me}G)GCATGAACCG-3’ (3’-\text{CGTACCCG G \text{me}C\text{me}GCGTACTTGGC-5’})</td>
</tr>
</tbody>
</table>

Bold font indicates modified bases or their complements.
A colon indicates that the bases are base-paired with the bases in the opposite strand.

2.5.10 Alkyl group removal by MGMT assessed by restriction endonuclease site deprotection

Control (G) and \(O^6\)-alkG-containing ODNs were annealed to the complementary oligonucleotide (5’-HEX-TAC TAG CCA TGG C-3’) where HEX (Hexachlorofluorescein) is the fluorophore as described in section 2.5.6. An aliquot (7 pmoles) of the resulting double-stranded ODNs was incubated with 21 pmoles MGMT in a total volume of 17 µl of buffer IBSA at room temperature for 3 h and then subjected to digestion with NlaIII, CviKI, StyI or BsaII in the buffer provided by the manufacturer (NEB) in a total volume of 10 µl at room temperature for 1 h. The samples were processed as described in section 2.2.6 except that a 15% non-denaturing PAGE gel was used.
2.5.11 MGMT activity and O\(^6\)-CMG levels in azaserine-treated human melanoma cells

MGMT-expressing human melanoma A375 cells were plated at a density of one million cells per 25ml flask in McCoy’s medium (Invitrogen) and incubated at 37°C until there were 5 x 10\(^6\) cells. For inactivation of MGMT, O\(^6\)-bromothenylguanine (O\(^6\)-BThG) was dissolved in dry DMSO at 20 mM and then added to a final concentration of 20 μM and incubated for 1 h. Azaserine (Sigma) was then added to final concentrations of 0, 29, 145, 290 and 580 μM and incubated overnight at 37°C. The cells were then harvested by trypsinisation by addition of 1 ml of trypsin preparation (Sigma) to the flask followed by incubation at room temperature for 5 min and centrifugation at 2500g for 10 min at 4°C. To determine MGMT activity, the cell pellets were washed with PBS and resuspended in 1 ml of buffer I containing 5 μg/ml leupetin then, transferred to 1.5 ml eppendorf tubes. Extracts were prepared as described in section 2.2.6 and assayed for MGMT activity as described in section 2.5.1.

To determine O\(^6\)-MeG and O\(^6\)-CMG levels, DNA was isolated from cell pellets using a QIAgen kit and quantified by picogreen assay. The standard curve was calf thymus (CT) DNA diluted in buffer I (50 mM Tris-HCl pH8.3, 1mM EDTA, 3 mM DTT) to concentrations of 0, 0.028, 0.055, 0.11, 0.2, 0.44, 0.88 and 1.76 μg/ml. Picogreen reagent buffer (Invitrogen). Ten μl of standard and unknown were added in duplicate to a microtitre plate and 100 μl of picogreen (Invitrogen) reagent (prepared by addition of 10 μl of stock concentrate to 10 ml of TE) was added to each well. The absorbance was measured at excitation 485 nm and emission 535 nm on a TECAN Genios plate reader. The results were accepted when the standard curve was linear (r > 0.99). Aliquots (1 μg) of DNA were digested for 1 h at 37°C with 10 units SauIIIa1 (Promega) in a total volume of 50 μl and the SauIIIa1 was then inactivated by heating at 65°C for 30 min. A biotinylated ODN (5’- PO\(_4\)-GATCAAAAAA-Biotin-3’) was then ligated using 1 unit T\(_4\) DNA ligase (Roche) and ligase buffer (Roche) by overnight incubation at 4°C. Serial dilutions (0, 25, 50 and 100 ng in 200 μl of PBS) of the above ligation products were then added to the wells of preblocked SA-coated microtitre plates (see section 2.5.7) and incubated for 1 h at room temperature. As a positive control for the ligation and subsequent steps,
serial dilutions (0, 25, 50 and 100 fmole) of O\textsuperscript{6}-MeG and O\textsuperscript{6}-CMG-containing ODN (5\textquotesingle-SIMA-GCCATGXCTAGTA-3\textquotesingle; where X is O\textsuperscript{6}-MeG or O\textsuperscript{6}-CMG) annealed to its biotinylated complement (5\textquotesingle-TACTAGCCATGGC-biotin-3\textquotesingle) was used. To confirm efficient binding of the ligated DNA, supernatants were removed and the DNA content determined by picogreen assay. The plates were washed three times with 100 
µl PBS containing 3 mg/ml BSA. For quantitation of both O\textsuperscript{6}meG and O\textsuperscript{6}CMG together, the Alkyltransferase-like (Atl1) protein was used as it has been shown to bind to all O\textsuperscript{6}-alkylguanines so far examined, including O\textsuperscript{6}meG and O\textsuperscript{6}CMG in single and double stranded DNA (Margison et al, 2007; Latypov et al, 2012). To the wells of the SA plates containing the DNA isolated from the Aza-treated cells were added 1 pmole of Atl1 protein (in house preparation) in 100 
µl PBS containing 3mg/ml BSA. The plates were incubated for 1 h at room temperature then washed three times with 100 
µl PBS containing 3 mg/ml BSA. Anti-Atl1 antibody (100uL: in house preparation diluted 1/500 in PBS containing 3mg/ml BSA) was added and incubated for 1 h at room temperature. The plates were processed and analysed as described above (section 2.5.7).

2.6 Statistical analysis

In the IC\textsubscript{50} experiments, each point on the MGMT inactivation curve can be used to determine the IC\textsubscript{50} since the majority of the lines show a linear relationship with ODN concentration in the assay. In these cases the values presented are the means of quintuplicate determinations. Statistical analysis was performed using one-way ANOVA when experiment comparing the IC\textsubscript{50} value more than two groups but when comparing the IC\textsubscript{50} between two groups, pairwise comparisons were used. Statistical significance was considered at p<0.05.

To compare the MGMT binding to ds ODNs in the ELISA experiments, triplicate determinations were undertaken and the means and SDs calculated from this. Statistical analysis was performed using pairwise comparisons with statistical significance considered at p<0.05.

For all other MGMT assays, no replicates were undertaken, but incrementally increasing amounts of MGMT proteins were used in all assays, and the linear
sections of the protein dependence curves usually gave correlation coefficients ($R^2$) of 0.99 or better indicating the accuracy of the determinations.
Chapter 3: Characterisation of MGMT variant proteins

3.1 Introduction

The human MGMT cDNA is present as the commonest version, referred to as wild type (Wt) and 12 rarer variants (see Figure 1.8). Based on literature information and relative frequency, a systematic study of six of the variants, namely F84, Q128, R160, V143/R178, F84/R160 and F84/V143/R178, was undertaken. The properties examined were the kinetics of methyl group transfer, thermal stability and ability to bind to and remove alkyl groups from ODNs containing a variety of $O^6$-alkG and other modified purines.

3.2 Generation, expression and purification of MGMT variant proteins

To compare the biochemical characteristics of the Wt and variant MGMTs, the encoding cDNAs were generated by PCR-based site directed mutagenesis, cloned into a fusion protein expression vector, and affinity purified (see Figure 2.2).

3.2.1 Isolation of variant MGMTs by PCR

High fidelity Phusion DNA polymerase was used to generate the MGMT cDNA variants. The Wt MGMT in the fusion protein vector pMAL-c2 was PCR-amplified using back to back 5’-phosphorylated mutagenic primers specific for each variant (see section 2.3.1). Figure 3.1 shows the products of PCR amplification separated by agarose gel electrophoresis. Based on the migration of the markers, the expected product (7270 bp) was obtained in all cases. The positive control for the PCR was 10 pg of pMAL-c2 MGMT (Wt) DNA and Wt forward and reverse primers, which produced the expected 624 bp fragment.

3.2.2 Ligation, transformation, colony PCR and sequencing

Following ligation to generate circular plasmids and transformation into competent XL-1 blue E.coli, individual ampicillin-resistant colonies were isolated and analysed by PCR using Wt forward and reverse primers as described in section 2.2.4. Figure 3.2 shows typical data for the F84 variant; based on marker migration, clone numbers 1-3, 5-6, 8-12 generated a PCR product of the expected size (624 bp).
Similar results were obtained with all of the other variants and these are shown in Appendix figure 2A.

**Figure 3.1:** Agarose gel electrophoresis of PCR amplification products of pMAL-c2 Wt MGMT. Products were obtained for all of the variants indicated. The positive control (Positive) was pMAL-c2 MGMT (Wt) DNA that was amplified using Wt forward and reverse primers.

**Figure 3.2:** Agarose gel electrophoresis of PCR amplification products of 12 different *E. coli* colonies produced by transformation with the ligated pMAL-c2 F84 variant PCR product. Positive indicates positive control.

Plasmids from the clones containing the MGMT variants were purified using a miniprep robot and the DNA sequenced and compared with the published MGMT cDNA sequence (GenBank; accession number: NM_002412.3). Figure 3.3A-C
shows the base changes on the electrophoretogram traces from the sequencer runs, alongside the determined and GenBank sequences for each of the variants.

A. Single variants

A.1 F84

A.2 Q128

A.3 R160
B. Double variants

B.1 V143/R178

B.2 F84/R160
C. Triple variant

C.1 F84/V143/R178

Figure 3.3: Sequencing results and nucleotide sequence alignments of the MGMT variants. **A.** single base variants, **B.** double base variants and **C.** triple base variant.
3.2.3 Expression and purification of MBP-MGMT variant fusion proteins

Clones containing the correct sequences of variants F84, Q128, R160, V143/R178, F84/R160 and F84/V143/R178 were induced with IPTG, harvested 0, 1, 2 and 3 h later and cell extracts were prepared and separated by 12%SDS-PAGE (see section 2.2.6). Figure 3.4 shows that, based on marker migration, the Coomassie-stained protein bands correspond to the predicted molecular weight for the MBP-MGMT fusion protein (63.7 kDa) and that this band was much more intense at 3 h. This F84 variant result is typical of the others variants, the results of which are shown in Appendix Figure 2B.

![Figure 3.4](image)

**Figure 3.4:** Expression of the F84 variant MBP-MGMT fusion protein in extracts of *E.coli* prepared at the times indicated after addition of IPTG. Extracts were analysed by 12% SDS-PAGE with Coomassie staining.

The MBP-MGMT fusion proteins were purified using amylose affinity columns as described in section 2.2.7 and fractions collected at the various stages were analysed by 12% SDS-PAGE. Figure 3.5 shows that the elution fractions contained bands that corresponded to the MBP-MGMT protein (63.7 kDa) and MBP protein (42 kDa). This F84 variant result is typical of the other variants, the results of which are shown in Appendix Figure 2C.
3.2.4 Western analysis of purified proteins

Proteins from duplicates of the gel shown in Figure 3.5 were electrotransferred onto PVDF membranes and probed with an anti-MBP antibody as described in section 2.2.8. Figure 3.6 shows that bands corresponding to 63.7 and 42 kDa contain the MBP-MGMT fusion protein, and MBP, respectively. This F84 variant result is typical of the variants, the results for which are shown in Appendix figure 2D.

Figure 3.5: Purification of the F84 variant MBP-MGMT fusion protein by amylose affinity chromatography. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining.

Figure 3.6: Western blot of the purification fractions (indicated) for the F84 variant MBP-MGMT fusion proteins. The membrane was probed with anti-MBP antibody and binding assessed by ECL.
3.2.5 Cleavage of purified MBP-MGMT fusion proteins

To establish the conditions for optimal cleavage of the MBP-MGMT fusion proteins, the MBP-Wt MGMT was digested with factor Xa at room temperature for 1 to 6 h as described in section 2.2.9. Figure 3.7 shows that complete digest of the MBP-Wt MGMT fusion protein to MBP and Wt MGMT was achieved in 6 h.

![Figure 3.7: Factor Xa cleavage of MBP-Wt MGMT fusion protein during incubation for various times and analysed by 12% SDS-PAGE with Coomassie staining.](image)

These conditions were used for MBP-MGMT variant fusion proteins. Figure 3.8 shows the products of the factor Xa digests of all of the variant MBP-fusion proteins corresponded to MBP and the MGMT proteins, based on the positions of the marker proteins.

3.2.6 Purification of MGMT

MGMT was purified after factor Xa cleavage using DEAE-Sepharose as described in section 2.2.10 and the elution fractions were analysed by 12% SDS-PAGE with Coomassie staining. Figure 3.9 shows that the 21.7 kDa band, corresponding to MGMT, eluted at 50-75 mM NaCl and that the 42 kDa band, corresponding to MBP eluted at 100-150 mM NaCl. This F84 variant result is typical
of the other variants the results for which are shown in Appendix figures 2E, 2F, 2G, 2H and 2I).

A. Single variants

![Figure 3.8: Factor Xa cleavage of MBP-MGMT fusion proteins. Ten µl of the 50µl digests indicated were analysed by 12% SDS-PAGE with Coomassie staining. For each variant - indicates before and + indicates after cleavage. A. Single variants (F84, Q128, R160) B. Double and triple variants (V143/R178, F84/R160, F84/V143/R178).]
Figure 3.9: Purification of the Factor Xa cleaved F84 variant MBP-MGMT fusion protein by DEAE-sepharose. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining. A. Elution with 25-75 mM NaCl, B. Elution with 100-500 mM NaCl.

The 50-75 mM NaCl elution fractions were pooled and applied to an amylose column to remove any trace amounts of MBP and uncleaved MBP fusion protein. Figure 3.10 shows that MGMT did not bind to the column and Figure 3.11 shows the result of the western blot of a duplicate gel, which confirms that the band at 21.7 kDa is the MGMT protein. Only the results for the F84 variant are shown: result for the other variants were similar and these are shown in Appendix Figures 2J and 2K.
MGMT containing fractions were pooled and the protein concentration was determined (see section 2.2.10). Aliquots were frozen and stored at -20°C.

Figure 3.10: Re-purification of the F84 variant protein using an amylose column. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining.

Figure 3.11: Western blot of the amylose column re-purified fractions (indicated) for the Factor Xa-cleaved, DEAE-purified F84 variant MGMT protein. The membrane was probed with anti-MGMT antibody.

3.3 Functional activity of MGMT variant proteins

The methyl transferase activity of the purified variant proteins was initially quantified in order to use closely similar amounts when comparing their
characteristics, which included methyl transfer kinetics, thermal stability and inactivation by and binding to modified ODNs.

3.3.1 Determination of MGMT activity of purified proteins

The activity of the purified MGMT variant proteins was determined using $[^3H]$-methylated substrate DNA as described in section 2.5.1. Figure 3.12 shows the results of titration curves, from which the specific activities of the proteins were determined (Table 3.1). The results in Figure 3.12 were used to calculate protein limiting amounts of MGMT using the Excel software set to linear fitting, which generally gave the equation $y = ax + b$ and showed correlation coefficients ($R^2$) of 0.99 or better. The protein limiting amounts of F84, Q128, R160, V143/R178, F84/R160, F84/V143/R178 and Wt proteins for subsequent experiments were chosen as 0.05, 0.27, 0.025, 0.018, 0.082, 0.05 and 0.01 µg respectively. The proteins had all been purified to the same apparent extent, so the large differences in specific activities (Table 3.1) were not expected. The possibility that this was a result of differences in the kinetics of methyl transfer was next examined.

![Figure 3.12: Transfer of $[^3H]$-methyl groups from $[^3H]$-methylated substrate DNA to MGMT with increasing amounts of Wt and variant MGMT proteins during incubation at 37°C for 1 hour. Vertical lines indicate the protein limiting amounts of the individual variants that were choosen for further studies.](image-url)
Table 3.1: Specific activities of Wt and variant MGMT proteins

<table>
<thead>
<tr>
<th>MGMT protein</th>
<th>Specific activity (fmoles/µg protein)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wt</td>
<td>4430</td>
</tr>
<tr>
<td>F84</td>
<td>1230</td>
</tr>
<tr>
<td>Q128</td>
<td>235</td>
</tr>
<tr>
<td>R160</td>
<td>1740</td>
</tr>
<tr>
<td>V143/R178</td>
<td>3060</td>
</tr>
<tr>
<td>F84/R160</td>
<td>579</td>
</tr>
<tr>
<td>F84/V143/R178</td>
<td>1170</td>
</tr>
</tbody>
</table>

3.3.2 Kinetics of methyl group transfer to MGMT

Most of the variant proteins Wt, F84, R160, V143/R178, F84/R160 and F84/ V143/R178 showed rapid transfer of [3H]-methyl groups, even within 30 min, and complete transfer at 2 h. In contrast, methyl transfer by the Q128 variant was much slower and was not complete until 3 h (Figure 3.13). Because slightly different amounts of variant MGMTs were used in this experiment, methyl transfer under these protein-limiting, excess substrate conditions does not plateau at the same level for each of the variants.

![Figure 3.13: The kinetics of methyl group transfer from substrate DNA by the Wt and variant MGMT proteins.](image)

3.3.3 Thermal stability

The Wt and variant MGMT proteins had similar patterns of thermal stability during pre-incubation at 51°C: activity dramatically decreased with 50% loss of
activity at around 4-10 min and complete loss of activity at 20 min (Figure 3.14A). This rapid loss of activity did not allow a very accurate comparison of relative half-life, so lower pre-incubation temperatures of 44°C and then 37°C, were employed and the results are shown in Figure 3.14B and 3.14C, respectively.

**Figure 3.14:** Thermal stability of Wt and variant MGMT proteins during pre-incubation at A. 51°C, B. 44°C and C. 37°C. Half-life values were determined from the curves and represent the time required to reduce the MGMT activity to 50%.
The results for the 44°C and 37°C studies are shown as individual graphs in Appendix Figure 1A and 1B, respectively. Trendlines were fitted by Excel software with intercepts fixed at 100%: all had correlation coefficients of 0.95 or better. The times corresponding to 50% loss of activity were determined from the slope of the lines and hence represented the mean of at least five data points.

Figure 3.15 summarises the data for 44°C and 37°C. There was no consistent pattern at the two temperatures: the F84 variant had the shortest half life at 44°C and the Q128 variant had the longest half life at 37°C. At physiological temperature, Wt, F84 and F84/R160 had closely similar half-lives and the others variants had longer half-lives.

![Figure 3.15: Half lives of Wt and variant MGMT proteins during pre-incubation at 44°C and 37°C.](image)

**3.3.4 MGMT inactivation by O⁶-alkG containing ODNs**

One of the objectives of this thesis was to investigate the relative ability of MGMT variants to repair a range of O⁶-alkGs in DNA. Previous reports indicated that WT MGMT acts at similar rates on ss and ds DNA, so it was considered that the initial studies would use only the ss version. To examine this, the effect of pre-incubation of Wt and variant MGMT for 1 h at 37°C with a single-stranded 13-mer
having the common sequence: 5’-SIMA-GCCATGXCTAGTA where X is G or the modified purines listed in Figure 2.5 was examined as described in section 2.5.4. Initial results using 2.4 and 24.4 nM ODN showed a very wide range of inactivation by these ODNs the most potent being the \( O^6 \)-BzG-containing ODN and the least potent were the G and DAP-containing ODNs (Figure 3.16). Data for all the other variants are shown in Appendix figure 3A.

**Figure 3.16:** Effect of pre-incubation with wide range concentration of twelve \( O^6 \)-alkG containing ODNs on the remaining activity of the MGMT F84 variant.

Based on these results, the six most effective ODNs (containing \( O^6 \)-MeG, \( O^6 \)-EtG, \( O^6 \)-PrG, \( O^6 \)-CMG, \( O^6 \)-BzG and \( O^6 \)-PobG) were assayed with all of the variants over concentration ranges that enabled determination of IC\(_{50}\) values. The detailed inactivation curves obtained with these ODNs and the F84 protein are shown in Figure 3.17. Data for all the other variants are shown in Appendix figure 3B.
Figure 3.17: Effect of pre-incubation with specific concentration of six $O^6$-alkG containing ODNs on the remaining activity of the MGMT F84 variant.

As the MGMT inactivation was in all cases essentially linear, each data point was used to extrapolate an IC$_{50}$ value and the mean value thus represents five data points. The data is presented in Figure 3.18 and summarised in Table 3.2.

When the variants were compared for each ODN individually, the Q128 variant was consistently the most resistant to inactivation (Figure 3.18A). Generally, all the other variants were broadly similar in sensitivity to inactivation by each ODN. The R160 and F84/R160 variants were more resistant than the other variants to inactivation by the $O^6$-CMG and $O^6$-PobG containing ODNs (Figure 3.18A).
Figure 3.18: IC$_{50}$ values (nM) for the inactivation of the Wt and variant MGMT proteins by $O^6$-alkG containing ODNs. A. Comparison between Wt and variant MGMT proteins for each ODN (* Significantly different from the Wt MGMT (P<0.05; one way ANOVA)) and B. Comparison between different $O^6$-alkG containing ODNs for each MGMT variant. (* Significantly different from the $O^6$-BzG (P<0.05; one way ANOVA)).
The ranking of the potency of the inactivators based on the IC₅₀ values is shown in Table 3.2. F84, V143/R178 and F84/V143/R178 variants have the same ranking order as Wt MGMT in which IC₅₀ values vary with O⁶-BzG > O⁶-MeG > O⁶-CMG > O⁶-PrG > O⁶-PobG > O⁶-EtG. There was a wide range of potency, with O⁶-PobG being approximately 100 times less potent than O⁶-BzG for the majority of the variants (Table 3.2). There appeared to be no relationship between IC₅₀ and the molecular weight of the alkyl group.

**Table 3.2:** The ranking of the IC₅₀ values (nM) for the inactivation of MGMT variant proteins by six O⁶-alkG containing ODNs.

<table>
<thead>
<tr>
<th>Ranking</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wt</td>
<td>O⁶-BzG</td>
<td>O⁶-MeG</td>
<td>O⁶-CMG</td>
<td>O⁶-PrG</td>
<td>O⁶-PobG</td>
<td>O⁶-EtG</td>
</tr>
<tr>
<td></td>
<td>(0.09 ± 0.02)</td>
<td>(0.93 ± 0.26)</td>
<td>(2.41 ± 0.63)</td>
<td>(9.14 ± 1.68)</td>
<td>(11.6 ± 1.44)</td>
<td>(16.02 ± 2.77)</td>
</tr>
<tr>
<td>F84</td>
<td>O⁶-BzG</td>
<td>O⁶-MeG</td>
<td>O⁶-CMG</td>
<td>O⁶-PrG</td>
<td>O⁶-PobG</td>
<td>O⁶-EtG</td>
</tr>
<tr>
<td></td>
<td>(0.26 ± 0.12)</td>
<td>(1.15 ± 0.33)</td>
<td>(2.84 ± 0.48)</td>
<td>(10.69 ± 0.49)</td>
<td>(13.51 ± 1.08)</td>
<td>(23.86 ± 3.59)</td>
</tr>
<tr>
<td>Q128</td>
<td>O⁶-BzG</td>
<td>O⁶-MeG</td>
<td>O⁶-CMG</td>
<td>O⁶-PobG</td>
<td>O⁶-EtG</td>
<td>O⁶-EtG</td>
</tr>
<tr>
<td></td>
<td>(1.34 ± 0.04)</td>
<td>(3.98 ± 0.02)</td>
<td>(16.42 ± 1.54)</td>
<td>(101.36 ± 9.70)</td>
<td>(104.50 ± 13.59)</td>
<td>(110.29 ± 9.38)</td>
</tr>
<tr>
<td>R160</td>
<td>O⁶-BzG</td>
<td>O⁶-MeG</td>
<td>O⁶-CMG</td>
<td>O⁶-PrG</td>
<td>O⁶-PobG</td>
<td>O⁶-EtG</td>
</tr>
<tr>
<td></td>
<td>(0.23 ± 0.15)</td>
<td>(1.79 ± 0.12)</td>
<td>(6.47 ± 0.81)</td>
<td>(11.94 ± 0.25)</td>
<td>(14.71 ± 1.52)</td>
<td>(92.43 ± 10.00)</td>
</tr>
<tr>
<td>V143/R178</td>
<td>O⁶-BzG</td>
<td>O⁶-MeG</td>
<td>O⁶-CMG</td>
<td>O⁶-PrG</td>
<td>O⁶-PobG</td>
<td>O⁶-EtG</td>
</tr>
<tr>
<td></td>
<td>(0.15 ± 0.06)</td>
<td>(0.91 ± 0.23)</td>
<td>(2.99 ± 0.77)</td>
<td>(10.68 ± 0.29)</td>
<td>(12.60 ± 0.97)</td>
<td>(23.01 ± 3.05)</td>
</tr>
<tr>
<td>F84/R160</td>
<td>O⁶-BzG</td>
<td>O⁶-MeG</td>
<td>O⁶-CMG</td>
<td>O⁶-PrG</td>
<td>O⁶-EtG</td>
<td>O⁶-PobG</td>
</tr>
<tr>
<td></td>
<td>(0.14 ± 0.01)</td>
<td>(1.02 ± 0.19)</td>
<td>(5.50 ± 1.07)</td>
<td>(8.91 ± 1.63)</td>
<td>(11.55 ± 2.29)</td>
<td>(69.14 ± 15.87)</td>
</tr>
<tr>
<td>F84/V143/ R178</td>
<td>O⁶-BzG</td>
<td>O⁶-MeG</td>
<td>O⁶-CMG</td>
<td>O⁶-PrG</td>
<td>O⁶-PobG</td>
<td>O⁶-EtG</td>
</tr>
<tr>
<td></td>
<td>(0.16 ± 0.09)</td>
<td>(0.77 ± 0.21)</td>
<td>(2.35 ± 0.58)</td>
<td>(6.82 ± 1.44)</td>
<td>(10.64 ± 0.79)</td>
<td>(16.03 ± 1.21)</td>
</tr>
</tbody>
</table>

The results of the initial screens using the ODNs containing O⁶-HOEtG, HOPr-DAP, O⁶-AEG, O⁶-MAG and DAP for each of the MGMTs are shown in Figure 3.19. Little or no inactivation was seen at the highest concentration of the ODNs used, so the IC₅₀s were higher than 24.4 nM.
Figure 3.19: Inactivation of normal MGMT and variant proteins by five $\gamma^6$-alkG containing ODNs. A. $\gamma^6$-HOEtG, B. HOPr-DAP, C. $\gamma^6$-AEG, D. $\gamma^6$-MAG and E. DAP.
3.3.5 Mass spectroscopic analysis of alkyl group transfer to normal MGMT protein

To determine if the inactivation seen above was a consequence of alkyl group transfer to the MGMT active site cysteine residue, MALDI-TOF analysis of tryptic digests of MGMT was undertaken (see section 2.5.9). It was considered that Wt protein would be representative of the variants so only this protein has so far been assessed. After digest of the unmodified Wt protein, the tryptic fragment containing the active site cysteine (GNPVPILIPCHR) produced an ion with mass/charge ratio \((m/z)\) of 1315.7. The expected mass changes in this tryptic fragment if an alkyl or other group had been transferred are shown in Table 3.3 and range from 14 (methyl) to 147 (pyridyloxobutyl) giving the predicted molecular weights listed.

**Table 3.3**: Expected sizes of the tryptic fragments of MGMT containing the active site C145 with or without alkyl group covalent modification

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Molecular weight</th>
<th>Additional group (Molecular weight)</th>
<th>Molecular weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNPVPILIPCHR</td>
<td>1315.7</td>
<td>Methyl (14)</td>
<td>1329.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Amino (15)</td>
<td>1330.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ethyl (28)</td>
<td>1343.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Propyl (42)</td>
<td>1357.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aminoethyl (43)</td>
<td>1358.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hydroxyethyl (44)</td>
<td>1359.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Carboxymethyl (57)</td>
<td>1372.7</td>
</tr>
<tr>
<td>Hydroxypropyl daminopurine (73)</td>
<td>1388.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benzyl (90)</td>
<td></td>
<td>1405.7</td>
<td></td>
</tr>
<tr>
<td>Methyladamantyl (135)</td>
<td>1450.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyridyloxobutyl (147)</td>
<td>1462.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.20 shows the relevant region of the mass spectra demonstrating that methyl, ethyl, propyl, benzyl, pyridyloxobutyl, carboxymethyl or hydroxyethyl groups are transferred from \(O^{6}\)-MeG, \(O^{6}\)-EtG, \(O^{6}\)-PrG, \(O^{6}\)-BzG, \(O^{6}\)-PobG, \(O^{6}\)-CMG and \(O^{6}\)-HOEtG to the active site of MGMT protein, under the incubation conditions used. Figure 3.21 shows a section of the mass chromatogram indicating the location of the tryptic fragment that would occur if groups had been transferred from HOPr-DAP, \(O^{6}\)-AEG, \(O^{6}\)-MAG and DAP to MGMT. In none of these cases was there any convincing evidence of alkyl or other group transfer.
Figure 3.20: MALDI-TOF mass spectral analysis of tryptic digests of MGMT indicating the locations and masses of the fragments expected if alkyl group transfer occurred from ODNs containing: A. G, B. $\text{O}^6$-MeG, C. $\text{O}^6$-EtG, D. $\text{O}^6$-PrG, E. $\text{O}^6$-HOEtG, F. $\text{O}^6$-CMG, G. $\text{O}^6$-BzG, and H. $\text{O}^6$-PobG.
**Figure 3.21:** MALDI-TOF mass spectral analysis of tryptic digests of MGMT indicating the locations and masses of the fragments expected if alkyl group transfer occurred from ODNs containing: A. G, B. DAP, C. $O^6$-AEG, D. HOPr-DAP and E. $O^6$-MAG.

**3.3.6 SPR analysis of the interaction between Wt MGMT and variant proteins with $O^6$-alkG containing ODNs**

To further examine protein-DNA interactions, SPR analyses were undertaken of the binding of the Wt and variant proteins to immobilised ds ODN as described in section 2.5.6. In each case, after the dissociation of the MGMT had been monitored, the sensor chips were analysed for residual alkyl groups by flowing the *S.pombe* $O^6$-alkG-binding protein, Atl1, over the chips. In Figures 3.23-3.28 the sensorgrams for
MGMT binding and dissociation are shown in the upper panels and the sensorgrams for Atl1 binding and dissociation are shown in the lower panels. In all cases the signal from the lane containing control (G) ODN has been subtracted from the others to show only $O^6$-alkG specific effects. Another control was provided by one lane over which MGMT was not flowed: the signal then generated by flowing Atl1 provided an indication of the amount of ODN that was bound to that lane.

The interpretation of this data was complicated by the fact that MGMT will not only bind to the ODNs immobilised on the neutravidin chips, but also remove alkyl groups from the $O^6$-alkG that are MGMT substrates (Figure 3.22A). It was not known if this would then result in the dissociation of the MGMT (Figure 3.22A), and if so, how rapidly this would occur. However, it was considered feasible to establish whether or not alkyl group removal had taken place by exploiting the ability of the *S.pombe* protein, Atl1, to bind strongly to all ODNs that contain $O^6$-alkG residues (Figure 3.22B), but to dissociate only slowly, particularly from the ODNs containing relatively complex residues (Latypov et al., 2012).

---

**Figure 3.22:** SPR analysis of the binding to the immobilised $O^6$-alkG containing ODNs of **A**, MGMT protein followed by **B**, Atl1. Red dots represent alkyl groups. Green dots represent MGMT and blue dots represent Atl1.
For the Wt MGMT, (Figure 3.23), there appeared to be very rapid binding to the ODN containing $O^6$-MeG and $O^6$-PobG and, somewhat slower binding to $O^6$-CMG but in all three cases, the signals decreased during the loading period. Binding to $O^6$-HOEtG was relatively slower and continued during the loading period and there was almost no signal for $O^6$-BzG. The sensorgrams for Atl1 showed no binding to the $O^6$-BzG-containing ODN, except in the lane that had not been exposed to MGMT. At the other extreme, binding to $O^6$-HOEtG-containing ODN was closely similar to the control no-MGMT lane. The other ODNs showed intermediate profiles, in particular the $O^6$-PobG profile where the maximum RU decreased as the amount of MGMT that had been applied increased.

![Sensorgrams for the binding of Wt MGMT protein to ODNs](image)

**Figure 3.23:** SPR analysis of the binding of Wt MGMT protein (upper panel: concentration range: 0, 10, 20, 40, 80 and 100 nM) to the immobilised $O^6$-alkG containing ODNs followed by Atl1 (lower panel: 4 nM). RU: Relative Units.

With the F84 variant, much lower intensity signals were generated with all of the ODNs (Figure 3.24). The Atl1-generated signals for the $O^6$-HOEtG and $O^6$-BzG-containing ODN were closely similar to those seen in Figure 3.23, but that for $O^6$-PobG and to a lesser extent, the $O^6$-MeG and $O^6$-CMG-ODNs were quite different.
Figure 3.24: SPR analysis of the binding of F84 variant MGMT protein (upper panel: concentration range as in Figure 3.23) to the immobilised $O^6$-alkG containing ODNs followed by Atl1 (lower panel: 4 nM). RU: Relative Units.

For the Q128 variant, no positive signals were seen, and there was even an indication that negative signals were generated by the $O^6$-MeG and $O^6$-BzG-containing ODNs (Figure 3.25). In contrast to the Wt and F84 variants, however, only the $O^6$-BzG-containing ODN showed a decrease in signal below the no-MGMT lane when Atl1 was flowed over the chip and this appeared to decrease in proportion to the amount of the Q128 variant that was used in the MGMT binding phase.

Figure 3.25: SPR analysis of the binding of Q128 variant MGMT protein (upper panel: concentration range as in Figure 3.23) to the immobilised $O^6$-alkG containing ODNs followed by Atl1 (lower panel: 4 nM). RU: Relative Units.
The R160 variant results are shown in Figure 3.26. In this case there appeared to be rapid and extensive binding to the $O^6$-HOEtG, $O^6$-CMG and $O^6$-PobG-containing ODNs while the profiles for the $O^6$-MeG and $O^6$-BzG-containing ODNs were similar to those for the Wt protein. However, the post Atl1-profiles were similar to those for the F84 variant.

**Figure 3.26**: SPR analysis of the binding of R160 variant MGMT protein (upper panel: concentration range as in Figure 3.23) to the immobilised $O^6$-alkG containing ODNs followed by Atl1 (lower panel: 4 nM). RU: Relative Units.

The MGMT profiles for the double variants, V143/R178 and F84/R160 were broadly similar (Figure 3.27) although the latter showed slightly more interaction for all but the $O^6$-BzG-containing ODN and the binding of Atl1 seemed to be correspondingly reduced. For both proteins, however, the binding of Atl1 to the $O^6$-BzG-containing ODN was essentially eliminated.
Finally, the triple variant F84/V143/R178 (Figure 3.28) was examined. In this case, there was signal decrease with all of the ODNs and only in the case of the \(O^6\)-HOEtG-containing ODN was there a positive signal. The post Atl1 signals were similar to those seen with the V143/R178 variant (Figure 3.27A).
**Figure 3.28:** SPR analysis of the binding of the F84/V143/R178 variant MGMT protein (upper panel: concentration range as in Figure 3.23) to the immobilised $O^6$-alkG containing ODNs followed by Atl1 (lower panels: 4 nM). RU: Relative Units.

### 3.4 Discussion

#### 3.4.1 Determination of MGMT activities of purified proteins

All of the intended variant MGMT cDNAs were successfully generated by site directed mutagenesis and the MBP-fusion proteins expressed, cleaved and purified to apparent homogeneity. Using standard assay conditions, however, the specific activities of the proteins varied from 4430 fmoles/µg protein for the Wt to 235 fmoles/µg protein for the Q128 variant. To examine the possible basis of this, the kinetics of methyl group transfer from methylated substrate DNA was determined. The reaction was generally rapid and $[^3]$H-methyl groups were extensively transferred from $[^3]$H-methylated substrate DNA to MGMT Wt and variant proteins within 60 min at 37°C except for Q128 which slowly transferred $[^3]$H-methyl groups over 3 h. The kinetics of methyl transfer by the latter probably explained the finding that it had the lowest specific activity. There are no other published reports of studies on the Q128 variant, the results from this study are supported by the demonstration by (Kanugula *et al.*, 1995) that the alteration of arginine-128 to alanine reduces the activity towards a methylated DNA substrate. (Moore *et al.*, 1994) have indicated that the MGMT region containing R128 is involved in DNA binding by MGMT. The amino acids from F94 to N137 form a
helix-turn-helix domain that is similar to some DNA binding motifs. Therefore, as the Q128 has a greatly reduced ability to transfer alkyl groups to DNA, this is consistent with this arginine residue playing a role in binding to the DNA substrate.

For the remainder of the variants, there appeared to be no correlation between methyl transfer rate and specific activity, but further experiments involving much shorter incubation times are needed to confirm this.

3.4.2 Thermal stability

The MGMT variant proteins had completely lost their activity after 20 min incubation at 51°C. This result is similar to that previously reported by (Margison et al., 2005) in which four variant proteins namely I143/K178, I143/R178, V143/K178 and V143/R178 rapidly lost activity after 10 min and had completely lost activity after 30 min at 51°C. During incubation at 44°C, the activity of MGMT variant proteins slowly decreased. The half-life of Wt MGMT was 27 min whilst it has been reported previously that MGMT from human lymphocytes rapidly lose activity at 45°C with a half-life of about 10 min (Brent, 1984). While most of the proteins lost activity in parallel, the F84 variant showed an unusual pattern of stability, losing activity relatively slowly at first, and then relatively rapidly at both 51°C and 44°C, but this was not evident at 37°C. The reasons for this have not been investigated.

Under physiological temperature conditions, the half lives of the Wt and variant proteins ranged from around 5 h to around 10 h. It is reasonable to suggest that the cellular feedback system that maintains steady-state levels of MGMT will compensate for such differences in stability in order to provide equivalent levels of repair capacity. However, if this is not the case, and the de novo rates of MGMT production are limiting, then people with different MGMT polymorphisms may have substantial differences in repair capacity. As a consequence such people might have much higher steady state levels of $O^6$-alkGs in their DNA. However, this needs to be established experimentally.

3.4.3 MGMT inactivation by modified oligodeoxyribonucleotides (ODNs)

In the preliminary screen, the $O^6$-MeG, $O^6$-EtG, $O^6$-PrG, $O^6$-BzG, $O^6$-PobG and $O^6$-CMG containing ODNs very effectively inactivated MGMT whereas those
containing guanine, $O^6$-HOEtG, HOPr-DAP, $O^6$-AEG, $O^6$-MAG and DAP were much less effective. With the latter group, it may have been possible to determine IC$_{50}$ values if sufficiently high concentrations of ODNs were used. However, for the other group of 6 ODNs the accurate assessment of the IC$_{50}$ values was possible, so that the Wt and variant proteins could be compared. The ODNs containing $O^6$-BzG and $O^6$-MeG were the most potent inactivators for Wt and all the variant proteins although there was substantially more variation in the IC$_{50}$ values of these proteins for the $O^6$-BzG than the $O^6$-MeG-containing ODN. The IC$_{50}$ values of 0.087 nM ($O^6$-BzG) and 0.93 nM ($O^6$-MeG) obtained using a 13 mer ODN (5'-SIMA-GCCATGXCTAGTA) with the Wt protein are different from those reported by (Shibata et al., 2006) IC$_{50}$ values were 0.2 nM ($O^6$-BzG) and 15.7 nM ($O^6$-MeG). Also and (Pegg, 2000) used a 11-mer ODN (5'-TGTGAXCTGTG) and reported IC$_{50}$ values of 1.3 nM ($O^6$-BzG) and; 10 nM ($O^6$-MeG). These differences may arise from differences in the pre-incubation and incubation times at 37ºC and the sequence context of $O^6$-alkG: Georgiadis, et al (1991); Delaney & Essigmann, (2001); Guza et al., (2006) and Coulter et al., (2007) reported that the repair of $O^6$-alkG is affected by sequence context. $O^6$-BzG is a better inactivator of MGMT activity than $O^6$-MeG. It seems likely that this is a consequence of the relative sizes and hydrophobicities of the benzyl and methyl groups. Thus the amino acid changes in the variants have less impact on the IC$_{50}$ values for methyl than for benzyl.

The Q128 variant was more resistant to inactivation than all of the other proteins for all of the ODNs. It is likely that this is a consequence of the lower rate of alkyl transfer by Q128 and if so, it raises the question of what are the most appropriate conditions to use for the comparison of IC$_{50}$ values. Thus longer incubation times for the Q128 protein may have resulted in lower IC50s, but this might also have been the case for the other variants. It may be that determining the rate of inactivation by the ODNs would provide a more physiologically significant value, especially if there is a limiting rate of de novo synthesis of MGMT protein (see above).
The results for the F84 variant are consistent with the earlier reports that this variant did not significantly differ from Wt MGMT in the inactivation by $O^6$-MeG (Inoue et al., 2000) and $O^6$-PobG (Mijal et al., 2006) but inactivation by $O^6$-BzG differs from that previously reported (Fang et al., 2008) possibly because this study used purified native MGMT but Fang used His$_6$-tagged MGMT. The results for the V143/R178 variant are in agreement with a previous report that the V143/R178 variant did not significantly differ from Wt in the repair of $O^6$-PobG (Mijal et al., 2004; Mijal et al., 2006). In contrast they disagree with that of (Fang et al., 2008) in the inactivation by $O^6$-BzG: these authors reported that this variant differed from Wt in the inactivation by $O^6$-BzG.

The R160 results obtained in this study are similar to those previously reported by (Mijal et al., 2004) and (Mijal et al., 2006) in which R160 was significantly different from Wt in the repair of $O^6$-PobG. In contrast, (Edara et al., 1996) reported that R160 was significantly resistant to inactivation by $O^6$-BzG. The results differed possibly because the pre-incubation time at 37°C differed between these studies (1 h) and the published work (30 min; Edara et al., 1996).

### 3.4.4 Mass spectrometry

The IC$_{50}$ values were assumed to reflect the inactivation of the MGMT proteins by transfer of the corresponding alkyl groups to the active site cysteine residue. To confirm that this was the case, the Wt protein was incubated with excess ODN and tryptic digests were analysed by mass spectrometry. Alkyl transfer to the active site peptide was confirmed for the $O^6$-MeG, $O^6$-EtG, $O^6$-PrG, $O^6$-BzG, $O^6$-PobG and $O^6$-CMG-containing ODNs. There was no evidence of this for the remaining 4 modified ODNs containing HOPr-DAP, $O^6$-AEG, $O^6$-MAG and DAP so that any inhibition of $[^3]$H-methyl transfer in the pilot assays was probably due to binding that did not result in alkyl group transfer.

### 3.4.5 Surface Plasmon Resonance

The interaction of Wt and variant MGMT proteins was finally investigated using SPR. The ProtEon software would not fit curves (data not shown) to any of the MGMT sensorgrams, so the data is qualitative.
The interpretation of these results might also be problematic, as binding of Atl1 would be reduced either if the alkyl groups had been removed, or if MGMT remained bound to the ODNs and prevented or inhibited the binding of Atl1. However, if using SPR, there was no or little apparent binding of MGMT and then complete apparent binding of Atl1 (as shown for \( O^6\)-HOEtG and \( O^6\)-PobG with the V143/R178 and Q128 variants in Figure 3.29 A and B, respectively), it would be safe to conclude that MGMT was unable to remove the alkyl group. In contrast, if there was no or little apparent binding of MGMT and then no binding of Atl1, this would indicate complete removal of the alkyl group as seen for the \( O^6\)-BzG-containing ODN and the Wt MGMT (Figure 3.30A). The same ODN with the F84/V143/R178 variant (Figure 3.30B) showed Atl1 binding only at the lowest MGMT concentration used whereas following the Q128 variant (Figure 3.30C) Atl1 binding decreased in proportion to the amount of the Q128 MGMT used. This suggests that these two variants might not protect cells against the adverse effect of benzylating agents as effectively as the Wt or the other variants. In contrast, for the \( O^6\)-PobG-containing ODN, the Wt protein would provide the most extensive protection against any agents that generate this lesion in DNA (for example TSN), with protection by the R160 and F84/R160 variants being much less and the other variants, probably negligible.

\[ \text{Figure 3.29: SPR analysis of the binding of A. the V143/R178 MGMT variant to } O^6\text{-HOEtG-containing ODN and B. the Q128 MGMT variant to } O^6\text{-PobG-containing ODN. The MGMT panels indicate no or little apparent binding of MGMT the Atl1 panels indicate A. complete or B. almost complete binding of Atl1.} \]
Figure 3.30: SPR analysis of the binding to $O^6$-BzG containing ODNs indicated of A. Wt MGMT showing complete removal of the benzyl group followed by no binding of Atl1. B. F84/V143/R178 MGMT showing most removal of the benzyl group followed by little binding of Atl1 and C. Q128 MGMT showing some removal of the benzyl group followed by the highest binding of Atl1.

In the case of the $O^6$-MeG-containing ODN, the Wt MGMT appears to be the most effective in methyl group removal, followed by the R160, then F84, then F84/R160 variants. Based on Atl1 binding, there was no indication that the Q128 variant removed methyl groups. Although the experimental systems are quite different, it may be that the IC$_{50}$ value obtained in the competition assays was a result of binding rather than alkyl transfer. This could be further examined by MS analyses of the variant proteins, or the use of MGMT mutant proteins in which the active site cysteine residue is replaced with another amino acid so that alkyl group transfer is not possible.
Chapter 4: Characterisation of extended and normal MGMT

4.1 Introduction

MGMT is widely accepted to be translated from an ATG initiation codon that generates a 21.7 kDa protein. However, previous studies have shown that purified human MGMT from human leukemic lymphoblasts (von Wronska et al., 1991) and human medulloblastoma cell lines (He et al., 1992) has a molecular weight of 25 kDa in 10% SDS-PAGE whereas human liver was reported to contain MGMT bands of 23.1 kDa and 21.7 kDa (Collier et al., 1996). Although these differences might be explained by the migration properties of MGMT in SDS-PAGE, examination of the MGMT mRNA sequence in Figure 1.8 indicates that there is an AUG in exon 1, upstream of, and in frame with, the accepted (“normal”) MGMT initiation codon. The encoded protein, which has been referred to here as “extended”, would have a predicted molecular weight of 25 kDa (Figure 4.1) and would be different from the “normal” protein which has a predicted molecular weight of 21.7 kDa.

![Amino acid sequence of MGMT. The amino-terminal methionines of the extended and normal proteins are indicated by M (Red). Sequence from GenBank (accession number: NM_002412.3).](image)

To establish that the normal transcript, and no alternatively spliced shorter version that would also encode the normal protein, is present in human cells, cDNA preparations from human peripheral blood mononuclear cells (PBMC) were amplified using a common reverse primer and forward primers that would amplify the expected and a truncated version of the MGMT transcripts, respectively, as described in section 2.1. Following this, to compare the functional activities of the extended and normal proteins, the encoding cDNAs were isolated by PCR, overexpressed in E. coli, purified and analysed by biotechnical techniques, including radioactivity based functional assays and ELISA, by methods described in section 125.
2.5. The approach was closely similar to that used to examine the MGMT variant proteins described in Chapter 3.

The full length and putative shorter version of the MGMT cDNAs have been referred to here as “full length” (716 bp) and “truncated” (624 bp) (Figure 4.2A). The normal cDNA could be translated to generate both the “extended” (25 kDa) and “normal” (21.7 kDa) amino acid sequences but the truncated cDNA could encode only the “normal” protein (Figure 4.2B).

Figure 4.2: Schematic diagram of:

A. Nucleotide sequence of the full length and putative truncated MGMT (the start codons (ATG) are shown in red) and

B. amino acid sequence of the extended and normal MGMT (The amino-terminal methionines (M) are shown in red).

4.2 Expression of extended and normal MGMT in human lymphocytes

The possibility that both full length and truncated MGMT cDNAs might be simultaneously expressed in human PBMC samples was investigated in the scheme shown in Figure 4.3. PCR amplifications were carried out using cDNA from human PBMC and the PCR products were digested at 37°C for 1 h with 0.5 units of SapI and PstI as described in section 2.1 and the scheme shown in Figure 4.3. It must be noted that this scheme cannot determine whether the truncated transcript is produced unless it is the sole transcript.
Figure 4.3: Scheme for investigating the expression of full length and truncated MGMT cDNAs.

A. The region to be amplified using different forward primers and a common reverse primer. B. The expected PCR products for full length and truncated MGMT and the position of SapI and PstI restriction endonuclease sites (shown in blue and red). C. Predicted products for full length and truncated MGMT cDNAs after digestion with SapI and PstI. The complete nucleotide sequence has been presented in Figure 2.1.
Figure 4.4 shows that all five of the cDNA samples analysed generated a PCR product that, based on the migration of the markers, closely corresponded to the predicted size for the full length cDNA (716 bp). Transcripts corresponding to a truncated cDNA (624 bp) were also generated, but may have arisen from either the extended or a shorter transcript. In no case was only a truncated transcript present. These results are discussed later.

![Figure 4.4: Agarose gel electrophoresis of PCR products obtained after amplification using the full length and truncated forward primers, and the common reverse primer, of cDNAs prepared from five human PBMC samples.](image-url)

To provide additional information for the identity of the bands prior to DNA sequencing, PCR product numbers 3 and 4 from Figure 4.4, were digested with SapI and PstI (Figure 4.3). Based on the migration of the markers, SapI produced 2 fragments of 367 and 350 bp for the full length MGMT and 369 and 255 bp for the truncated MGMT. In addition, PstI generated 3 fragments for both, with the shortest being unique to either the full length (173 bp) or truncated (120 bp) MGMT (Figure 4.5).

The original PCR products from Figure 4.4 were then purified and the nucleotide sequences were obtained as described in section 2.1.1. The sequences were found to be identical to the GenBank MGMT cDNA sequence.
4.3 Generation, expression and purification of extended and normal MGMT

To compare some of the biochemical characteristics of the normal and extended MGMTs, the encoding cDNAs were isolated by PCR, cloned into a fusion protein expression vector, and affinity purified.
4.3.1 Isolation of extended and normal MGMT cDNAs by PCR

Initially, to find a suitable amount of DNA template for the subsequent experiment that used PCR products for ligation without purification, serial dilutions of the OriGene vector, pCMV6-AC-GFP which contains the full length MGMT cDNA, were amplified using only the truncated forward and the reverse primer as described in section 2.2.1. Figure 4.6 shows that, based on the migration of the markers, the expected PCR product (624 bp) was obtained using 0.1 to 100 pg of DNA.

Figure 4.6: Agarose gel electrophoresis of PCR products obtained after amplification of increasing amounts of pCMV6-AC-GFP vector using the truncated forward and common reverse primers.

Figure 4.7: Agarose gel electrophoresis of the pCMV6-AC-GFP PCR amplification products for the full length and truncated MGMT cDNA.
Based on this result, the normal and truncated MGMT cDNAs were generated by PCR amplification of 10 pg of the vector using the primers for the two MGMTs. According to the migration of the markers, this produced the required 716 and 624 bp fragments (Figure 4.7).

4.3.2 Ligation, transformation, colony PCR and sequencing

The 716 and 624 bp PCR products were generated using 5’ phosphorylated primers. They could therefore be ligated directly into pMAL-c2 that had been digested with XmnI and treated with alkaline phosphatase (to prevent self-religation), transformed into competent XL-1 blue E.coli, and grown overnight at 37°C on LB/amp plates as described in section 2.2.3. To confirm the presence of the MGMT insert, individual colonies were picked and subjected to PCR (see section 2.2.4) using the same primers used to generate the products for cloning. Based on the positions of the marker bands, clone numbers 2, 4-6, 9-10, 12 (Figure 4.8A) and 1, 3-5, 7-9, 12 (Figure 4.8B) had PCR products of the expected size for full length (716 bp) and truncated (624 bp) MGMT cDNA respectively.

**Figure 4.8:** Agarose gel electrophoresis of PCR products of different E.coli clones produced by transformation using material generated by PCR using (A) full length and (B) truncated forward and common reverse primers.
The PCR products from Figure 4.8 were then purified and the nucleotide sequences were obtained as described in section 2.1.1. The sequences were found to be identical to the GenBank MGMT cDNA sequence (Figure 4.9).

<table>
<thead>
<tr>
<th>pMAL-c2 vector</th>
<th>Standard</th>
<th>Full length</th>
<th>Truncated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>GGATCGAGGGAAGATGGTGGAGACGCGCCCTAGAAGCTTTTGGTCGAGGCGCC 106</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full length</td>
<td>GGATCGAGGGAAGATGGTGGAGACGCGCCCTAGAAGCTTTTGGTCGAGGCGCC 120</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truncated</td>
<td>ATGGAGACGCTTGGTGGAGACGCGCCCTAGAAGCTTTTGGTCGAGGCGCC 83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>CGCAGTGCTCGTGCCGAGCTGCTTGGTGGAGACGCGCCCTAGAAGCTTTTGGTCGAGGCGCC 166</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full length</td>
<td>CGCAGTGCTCGTGCCGAGCTGCTTGGTGGAGACGCGCCCTAGAAGCTTTTGGTCGAGGCGCC 180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truncated</td>
<td>ATGGAGACGCTTGGTGGAGACGCGCCCTAGAAGCTTTTGGTCGAGGCGCC 143</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>GTGAAATGAAACGCACCACACTGGACAGCCCTTTGGGGAAGCTGGAGCTGTCTGGTTGTG 226</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full length</td>
<td>GTGAAATGAAACGCACCACACTGGACAGCCCTTTGGGGAAGCTGGAGCTGTCTGGTTGTG 240</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truncated</td>
<td>ATGGAGACGCTTGGTGGAGACGCGCCCTAGAAGCTTTTGGTCGAGGCGCC 180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>AGCAGGGTCTGCACGAAATAAAGCTCCTGGGCAAGGGGACGTCTGCAGCTGATGCCGTGG 286</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full length</td>
<td>AGCAGGGTCTGCACGAAATAAAGCTCCTGGGCAAGGGGACGTCTGCAGCTGATGCCGTGG 300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truncated</td>
<td>ATGGAGACGCTTGGTGGAGACGCGCCCTAGAAGCTTTTGGTCGAGGCGCC 203</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>TTGTGAAATTCGGAGAAGTGATTTCTTACCAGCAATTAGCAGCCCTGGCAGGCAACCCCA 526</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full length</td>
<td>TTGTGAAATTCGGAGAAGTGATTTCTTACCAGCAATTAGCAGCCCTGGCAGGCAACCCCA 540</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truncated</td>
<td>ATGGAGACGCTTGGTGGAGACGCGCCCTAGAAGCTTTTGGTCGAGGCGCC 443</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4.9:** Nucleotide sequence alignments of the full length and truncated MGMT. The start codons (ATG) and the stop codon (TGA) are shown in red and blue.
4.3.3 Expression and purification of MBP-MGMT fusion proteins

To express the MBP-MGMT fusion proteins, clones containing the correct sequences were grown overnight, scaled up and induced by IPTG as described in section 2.2.6. To establish a suitable time period of incubation following IPTG addition, samples were harvested 0, 1, 2 and 3 h later and cell extracts were prepared and analysed by 12% SDS-PAGE with Coomassie staining (see section 2.2.6). Figure 4.10 shows that, based on the migration of the protein markers, bands were obtained which corresponded to the predicted molecular weight for the extended and normal MBP-MGMT fusion proteins (67 kDa and 63.7 kDa), which represent the MBP protein (42 kDa) fused with the extended (25 kDa) or normal (21.7 kDa) MGMT proteins. They were not easily visible prior to the addition of IPTG and were more intense at 3 h. There was also a band at 42 kDa that corresponded to MBP alone, but this tended to slightly decrease in intensity after IPTG addition (Figure 4.10).

**Figure 4.10:** Expression of MBP fusion proteins in *E.coli* extracts prepared at various times after addition of IPTG, analysed by 12% SDS-PAGE with Coomassie staining.

Extracts of the *E.coli* growths harvested 3 h after addition by IPTG were purified using amylose affinity columns as described in section 2.2.7. The column load and fractions collected at the various washing and elution stages were analysed by 12% SDS-PAGE. Figure 4.11 shows that, based on the migration of the marker proteins, the load, some of the wash and all of the elution fractions contained bands that corresponded to the MBP-MGMT fusion proteins for the extended (Figure
4.11A) and the normal (Figure 4.11B) proteins. A band at 42 kDa again corresponded to the size of the MBP protein.

**A.**

![Western blot of MBP-MGMT fusion proteins](image)

**Figure 4.11:** Purification of the (A) extended and (B) normal MBP-MGMT fusions proteins by amylose affinity chromatography. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining.

**4.3.4 Western analysis of MBP-MGMT fusion proteins**

To confirm the identity of the major proteins, duplicates of the above gels were electrotransferred onto PVDF membranes and probed with an anti-MBP antibody as described in section 2.2.8. Figure 4.12 shows that the bands corresponding to 42, 67 and 63.7 kDa are all contain MBP, confirming that the lower MW band in Figure 4.11 is the MBP protein. The western blot also suggested that the second wash lane in Figure 4.11A probably contained some of the loaded sample, indicating column saturation, based on the intensities of the Coomassie stained and the ECL-visualised bands.
Figure 4.12: Western blot of fractions obtained from the purification of MBP-MGMT fusion proteins. The membranes were probed with anti-MBP antibody. A. Extended MBP-MGMT, B. Normal MBP-MGMT.

4.3.5 Cleavage of purified MBP-MGMT fusion proteins

The eluted fractions containing MBP-MGMT fusion proteins were pooled and incubated with factor Xa at room temperature for 6 h as described in section 3.2.5. Figure 4.13 shows that, for each of the MBP-MGMT variant proteins, factor Xa efficiently generated two major bands corresponding to the expected fragments of 42 kDa for MBP, 25 and 21.7 kDa for extended and normal MGMT, based on the positions of the maker proteins.

4.3.6 Purification of MGMT proteins

DEAE-Sepharose was used to purify the digestion mixtures containing MBP, factor Xa and MGMT as described in section 2.2.10. Digests (40 µl) were applied to 1 ml DEAE-Sepharose columns and the MGMT (21.7 kDa) and MBP (42 kDa) proteins were eluted at 50-75 mM NaCl and 100-150 mM NaCl respectively (Figure 4.14). Similar results were obtained for the extended MGMT protein (data not shown).
**Figure 4.13:** Factor Xa cleavage of extended and normal MBP-MGMT fusion proteins, as indicated; analysed by 12% SDS-PAGE with Coomassie staining.

**Figure 4.14:** DEAE-sepharose purification of MBP-MGMT factor Xa cleavage reaction products. **A.** Elution with 25-100 mM NaCl, **B.** Elution with 150-500 mM NaCl.
To remove any traces of MBP and any remaining uncleaved MBP fusion protein, fractions 1-3 from the 50 mM NaCl elution and fractions 1-2 from the 75 mM NaCl elution were pooled and applied to a 5 ml amylose column. Fractions were collected and analysed by 12% SDS-PAGE with Coomassie staining. Figure 4.15 shows that, as expected, MGMT did not bind to the column and was eluted mainly in binding buffer fraction 3. There was no evidence of any MBP or fusion protein in the elution fractions.

**Figure 4.15**: Amylose column purification of the pooled fractions containing normal MGMT (purified as in Figure 4.14).

To finally confirm that the band corresponding to 21.7 kDa is the normal MGMT protein a western blot of a duplicate gel was probed with anti-MGMT antibody as described in section 2.2.8. Figure 4.16 shows that the band corresponding to 21.7 kDa is indeed MGMT. Similar results were obtained for the extended MGMT protein (data not shown) and it was concluded that the MGMT prepared in this way is essentially pure. MGMT-containing fractions were pooled, the protein concentration determined (see section 2.2.6), and aliquots were frozen and stored at -20°C.
4.4 Characterisation of purified extended and normal MGMT proteins

To confirm that MGMT activity was retained after cleavage and purification and for subsequent characterisations, which included methyl transfer kinetics, thermal stability and inactivation by and binding to modified ODN (see below), the levels of methyl transferase activity were initially quantified.

4.4.1 Determination of MGMT activity

The activity of the purified MGMT proteins was determined using $[^{3}H]$-methylated substrate DNA as described in section 2.5.1. The results in Figure 4.17 were used to calculate protein limiting amounts of MGMT using the Excel software set to linear fitting, which generally gave the equation $y = ax + b$ and showed correlation coefficients ($R^2$) of 0.99 or better. The protein limiting amounts of extended and normal MGMT for subsequent assays were selected to be 0.114 and 0.020 µg, respectively. There was evidence that the maximum amounts of radioactivity transferred to protein was different (88 and 68.2 fmole) for extended and normal MGMT, but this was not pursued here, as the objective was to establish the protein-limiting amounts of the MGMTs for use in subsequent experiments.
Figure 4.17: Transfer of $[^3]H$-methyl groups from $[^3]H$-methylated substrate DNA to MGMT with increasing amounts of extended and normal MGMT proteins during incubation at 37°C for 1 hour. Vertical lines indicate the protein limiting amounts of the individual variants that were chosen for further studies.

The specific activities of the extended and normal proteins were 772 and 3410 fmoles/µg protein, respectively. One possible explanation for this difference was that the incubation time in this assay was 1 h, and the kinetics of methyl transfer by the two proteins might be quite different. This was therefore investigated.

4.4.2 Kinetics of methyl group transfer to MGMT

Protein-limiting amounts of the MGMT proteins were incubated with $[^3]H$-methylated substrate DNA at 37°C for up to 300 min to determine the kinetics of methyl group transfer and the results are shown in Figure 4.18.

Figure 4.18: The kinetics of methyl group transfer from substrate DNA to extended and normal MGMT.
The normal MGMT protein transferred [$^3$H]-methyl groups generally slightly more rapidly than the extended protein up to 180 min at which time both proteins had reached closely similar plateau-level activity. As the points were not replicates, it was not possible to indicate if the differences were statistically significant, but these results seem unlikely to explain the large difference in specific activity. Another possible explanation was that the two proteins differed in their stability during incubation at 37°C, and this was next investigated.

4.4.3 Thermal stability

It was previously mentioned (section 3.3.3) that the amino acid changes in the variant proteins may affect their thermal stabilities. To assess thermal stability, protein-limiting amounts of the extended and normal MGMT proteins were pre-incubated at 51°C or 44°C for 5, 10, 20, 30, 40 and 50 min, then [$^3$H]-methylated substrate DNA was added and incubation continued for 1 h at 37°C. Figure 4.19 shows that the extended and normal MGMT proteins had very similar thermal stabilities at both 51°C and 44°C with 50% loss of activity by 5 min and 27 min respectively. The slight differences, if any, were considered to be unlikely explain the large difference in specific activity or to have any biochemical impact.

![Figure 4.19: Thermal stability of MGMT proteins during pre-incubation at 44°C and 51°C.](image)
4.4.4 MGMT inhibition by calf thymus DNA

To examine if DNA binding might influence the specific activities of the extended and normal proteins, the effect on MGMT activity of pre-incubation of the MGMT proteins with calf thymus DNA was determined and the results are shown in Figure 4.20. The extended MGMT was somewhat more sensitive to competitive inhibition than normal MGMT, with IC$_{50}$ values of 75 µg and 91 µg respectively, suggesting slightly stronger binding of the extended protein to DNA.

![Figure 4.20: Inhibition of MGMT proteins by increasing concentrations of calf thymus DNA. IC$_{50}$ values are indicated by the vertical lines and are significantly different (P<0.05; pairwise comparisons).](image)

4.4.5 MGMT inactivation by $O^6$-alkG containing ODNs

One of the objectives of this thesis was to investigate the ability of MGMT to repair a range of $O^6$-alkGs in DNA. To examine this for the extended and normal MGMT proteins, their inactivation by ss 13 mer ODNs; 5'-SIMA-GCCATGXXCTA GTA-3' (where X is the modified purine base listed in Figure 2.5, Section 2.5.4) was determined by pre-incubation of the MGMT proteins with ODNs for 1 h at 37°C as described in section 2.5.4. In pilot experiments, concentrations of 2.4 and 24.4 nM were assessed for all ODNs with both proteins. The results showed that the potency of the ODNs varied widely from the most effective ODN containing $O^6$-BzG and $O^6$-...
CMG to the least effective containing DAP, the MGMT inhibition by which was indistinguishable from the G-containing ODNs; Figure 4.21 shows typical examples of these results.

Based on this initial screen, it was considered reasonable to focus on the eight most potent ODNs for the more detailed comparative studies described in this chapter.

**Figure 4.21:** Effect of pre-incubation with wide range concentrations of $O^6$-alkG containing ODNs on the activity of extended and normal MGMT.

The IC$_{50}$ values for the group of 8 potent ODNs were then determined using incrementally increasing amounts over concentration ranges that were appropriate for each ODN as indicated by the result of the initial screen and the results are presented in Figure 4.22.
Figure 4.22: Effect of pre-incubation with specific concentration of $\textit{O}^6$-alkG containing ODNs on the activity of extended and normal MGMT.
The results in Figure 4.22 were used to calculate the IC\textsubscript{50} values and the results are shown in Figure 4.23. O\textsuperscript{6}-BzG was the most potent inactivator for both the extended and normal MGMT and the others were in the order O\textsuperscript{6}-MeG > O\textsuperscript{6}-CMG > O\textsuperscript{6}-PrG > O\textsuperscript{6}-PobG > O\textsuperscript{6}-EtG > O\textsuperscript{6}-HOEtG > O\textsuperscript{6}-AEG. There was more than 1000-fold difference in the highest and the lowest potency ODN. The normal MGMT was statistically significantly more resistant than the extended MGMT to inactivation by all but the O\textsuperscript{6}-CMG-containing ODNs, which gave identical IC\textsubscript{50} values for the two proteins.

![Figure 4.23](image)

**Figure 4.23:** IC\textsubscript{50} values (nM) for the inactivation of extended and normal MGMT proteins by ODNs in order of increasing potency. Values are means +/- SD of 5 data points. *Significantly different from the extended MGMT (P<0.05; pairwise comparisons)

There was a general trend for the difference in IC\textsubscript{50} values between extended and normal MGMT to increase as the potency of the ODN for MGMT inactivation decreased (Table 4.1) but there was no such trend in the relative inactivation when the remaining activity for extended was calculated as a percentage of normal (Table 4.1).
**Table 4.1:** The difference in IC$_{50}$ values between extended and normal MGMT by $O^{6}$-alkG containing ODNs

<table>
<thead>
<tr>
<th>ODNs</th>
<th>difference in IC$_{50}$ (normal minus extended) (nM)</th>
<th>Extended as % of normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O^{6}$-BzG</td>
<td>0.023</td>
<td>70.5</td>
</tr>
<tr>
<td>$O^{6}$-MeG</td>
<td>0.43</td>
<td>49.4</td>
</tr>
<tr>
<td>$O^{6}$-CMG</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>$O^{6}$-PrG</td>
<td>6.02</td>
<td>23.6</td>
</tr>
<tr>
<td>$O^{6}$-PobG</td>
<td>7.35</td>
<td>21.9</td>
</tr>
<tr>
<td>$O^{6}$-EtG</td>
<td>12.71</td>
<td>16.7</td>
</tr>
<tr>
<td>$O^{6}$-HOEtG</td>
<td>44.58</td>
<td>18.9</td>
</tr>
<tr>
<td>$O^{6}$-AEG</td>
<td>71.78</td>
<td>30.4</td>
</tr>
</tbody>
</table>

**4.4.6 The effects on MGMT activity of adding ZnCl$_{2}$ to *E.coli* growth media**

The effect of ZnCl$_{2}$ added to the *E.coli* growth media during IPTG-induced expression of the fusion proteins on the activity of the extended and normal MGMT after factor Xa cleavage and amylose and DEAE-sepharose purification was then examined using the [$^3$H]-based assay as described in section 2.5.5. The results in Figure 4.24 were used to calculate protein limiting amounts of MGMT using the Excel software set to linear fitting, which generally gave the equation $y = ax + b$ and showed correlation coefficients ($R^2$) of 0.99 or better. The protein titration curves showed that ZnCl$_{2}$ supplementation increased the activity both proteins, but the extended protein more than the normal protein (Figure 4.24).

![Figure 4.24](image_url)

**Figure 4.24:** The effect of ZnCl$_{2}$ added to the *E.coli* growth media on the activity of purified MGMT proteins transfer [$^3$H]-methyl group during incubation at 37ºC for 1 hour. Vertical lines indicate the protein limiting amounts of the individual proteins that were chosen for further studies.
The specific activities of the proteins and the protein limiting amounts selected for subsequent experiments are shown in Table 4.2.

Table 4.2: The specific activities of the extended and normal MGMT proteins with and without Zn and the protein limiting amounts

<table>
<thead>
<tr>
<th>MGMT proteins</th>
<th>fmoles/µg protein</th>
<th>Protein-limiting amount (µg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extended</td>
<td>863</td>
<td>0.05</td>
</tr>
<tr>
<td>Extended + Zn</td>
<td>2730</td>
<td>0.015</td>
</tr>
<tr>
<td>Normal</td>
<td>5410</td>
<td>0.008</td>
</tr>
<tr>
<td>Normal + Zn</td>
<td>7530</td>
<td>0.004</td>
</tr>
</tbody>
</table>

4.4.7 Effect of ZnCl₂ supplementation on MGMT inactivation by O^6^-alkG containing single and double stranded ODNs

To examine the relative potency of ss and ds ODNs containing the O^6^-alkGs (see section 4.4.5 and Figure 2.5; section 2.5.4) as MGMT inactivators and the effect of ZnCl₂ added to the E.coli growth media, ds ODNs were generated by annealing to the 3'-biotinylated complement (see section 2.5.6). The ODNs were pre-incubated with the MGMT proteins for 1 h at 37°C as described in Section 2.5.4. Thus ss ODN was compared with ds ODN for the proteins induced without adding ZnCl₂, and the effect of adding ZnCl₂ was assessed using only ds ODNs. This would also provide the information needed to design later experiments in which it was intended to undertake comprehensive comparisons of all of the variant MGMTs using ss and ds ODN.

Figure 4.25 shows the inactivation curves obtained with the O^6^-MeG containing ODN presented as all possible pairwise comparisons. These and the data for the other ODNs are summarised in Figure 4.26 which shows the calculated IC₅₀ values. Ds ODNs containing O^6^-alkGs have lower IC₅₀ values except for extended MGMT repairing O^6^-BzG and O^6^-MeG containing ODNs and hence are statistically significantly more effective inactivators than the corresponding ss ODNs. The addition of ZnCl₂ to the growth media results in MGMT proteins that are significantly less sensitive to inactivation by ds ODNs for all ODNs except O^6^-PobG and O^6^-HOEtG.
Figure 4.25: Effect of pre-incubation with the $\delta$-MeG-containing single (ss) and double-stranded (ds) ODNs on the activity of extended and normal MGMT induced with and without ZnCl$_2$ in the *E.coli* growth media.
Figure 4.26: Effect of ZnCl$_2$ added to the *E. coli* growth media on the activity of the purified MGMT proteins with single (ss) and double-stranded (ds) ODNs containing $O^6$-alkGs. Values are means +/- SD of 5 data points * Significantly different from the -Zn (ds) (P<0.05; one way ANOVA). The scale on the Y axis has been adjusted to show differences for each ODN.

The ranking of the potency of the ODNs for both MGMT proteins was $O^6$-BzG > $O^6$-MeG > $O^6$-CMG > $O^6$-PobG > $O^6$-HOEtG consistent with the previous results for ss ODNs (section 4.4.5).
### 4.4.8 Enzyme-linked immunosorbent assay (ELISA) analysis of MGMT binding to double-stranded ODNs

The binding of the extended and normal MGMT proteins induced to express in the absence (-Zn) and presence (+Zn) of ZnCl$_2$ in the *E.coli* growth media, to $O^6$-MeG, $O^6$-BzG and G containing ODNs was investigated using ELISAs as described in section 2.5.7. Preliminary experiments used plates coated with 50 and 500 fmoles of the biotinylated 13-mer ds ODNs and a wide range of amounts of normal +Zn MGMT (0-60 fmoles). Figure 4.27 shows that there was no significant ECL signal with the control (G containing) ODNs and there was no apparent difference in the signals between 50 and 500 fmoles of the $O^6$-MeG and $O^6$-BzG containing ODNs. A marked increase in the signal was seen at the lower MGMT concentrations and a slower increase over the higher concentrations. These results indicated that the signals generated were determined by the amounts of MGMT used and not the amounts of ODN in the wells.

**Figure 4.27:** Binding of MGMT (0-60 fmoles) to double-stranded (ds) ODNs assessed using an ELISA assay. Each point represents the mean and standard deviation of triplicate determinations; where error bars are not visible, they are masked by the symbol.
Based on the above result, plates were coated with 50 fmoles of the ODNs and used in assays with 0-6 fmoles of normal+Zn MGMT. Figure 4.28 shows an almost linear relationship between MGMT concentration and ECL signal.

![Figure 4.28: Binding of MGMT (0-6 fmoles) to ds ODNs assessed using an ELISA assay. Each point represents the mean and standard deviation of triplicate determinations.](image)

It was therefore considered appropriate to use this assay to compare the MGMT proteins being investigated and so 50 fmoles of the ODN and 0-6 fmoles of MGMT proteins were used in subsequent experiments. In addition, the effect of ODN length was examined by comparing short (13 mer; 5'-SIMA-GCCATG XCTAGTA) and long (23 mer; 5'-GAACTXCAGCTCCGTGCTGGCCC) ds ODNs. Short and long ODNs were assessed in the same experiment and the results are therefore directly comparable. The results (Figure 4.29) show that again, binding of all proteins to G containing ODNs was only slightly above background levels. Binding to 13-mer O6-alkG-containing ODNs was lower than to the 23-mer ODNs, but only slightly. In contrast, in all cases and at all MGMT levels, binding of the normal protein was generally 50% of that of the extended protein. Furthermore, binding of the +Zn MGMT was only around 50% of that of the -Zn MGMT.
A. Extended MGMT; 13 mer  

B. Normal MGMT; 13 mer

C. Extended MGMT; 23 mer  

D. Normal MGMT; 23 mer

Figure 4.29: MGMT binding to short and long double-stranded $O^6$-alkG-containing ODNs (13 and 23 mer). Values shown are means +/- S.D. of triplicate determinations.

4.4.9 Effect of 5- methylcytosine in $O^6$-MeG-containing ODNs on the inactivation of extended and normal MGMT

The effect of the presence of 5-methylcytosine (5-MeC) in the same or opposite strands in G or $O^6$-MeG containing ODNs encompassing the p53 codon 245 (Table 4.3) on MGMT inactivation was determined as described in section 2.5.8.
**Table 4.3**: *p*53 codon 245-spanning ODNs used to assess the effect of 5-methyl cytosine on inactivation of MGMT by *O*6-MeG

<table>
<thead>
<tr>
<th>ODN abbreviations</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>C&lt;sup&gt;me&lt;/sup&gt;G (ss)</td>
<td>5'-GCATGGG&lt;sup&gt;me&lt;/sup&gt;G GCATGAACCG-3'</td>
</tr>
<tr>
<td>C&lt;sup&gt;me&lt;/sup&gt;C&lt;sup&gt;me&lt;/sup&gt;G (ss)</td>
<td>5'-GCATGGG&lt;sup&gt;me&lt;/sup&gt;C&lt;sup&gt;me&lt;/sup&gt;G GCATGAACCG-3'</td>
</tr>
<tr>
<td>CG (ss)</td>
<td>5'-GCATGGG C G GCATGAACCG-3'</td>
</tr>
<tr>
<td>G:&lt;sup&gt;me&lt;/sup&gt;C (ds)</td>
<td>5'-GCATGGGC GCATGAACCG-3' 3'-CGTACCCG&lt;sup&gt;me&lt;/sup&gt;CGTACTTGGC-5'</td>
</tr>
<tr>
<td>G:&lt;sup&gt;me&lt;/sup&gt;G:C (ds)</td>
<td>5'-GCATGGGC&lt;sup&gt;me&lt;/sup&gt;G GCATGAACCG-3' 3'-CGTACCCG&lt;sup&gt;me&lt;/sup&gt;CGTACTTGGC-5'</td>
</tr>
<tr>
<td>G:&lt;sup&gt;me&lt;/sup&gt;C:&lt;sup&gt;me&lt;/sup&gt;G (ds)</td>
<td>5'-GCATGGG&lt;sup&gt;me&lt;/sup&gt;G C&lt;sup&gt;me&lt;/sup&gt;G GCATGAACCG-3' 3'-CGTACCCG G CCATCGCCG-5'</td>
</tr>
<tr>
<td>G:&lt;sup&gt;me&lt;/sup&gt;G:&lt;sup&gt;me&lt;/sup&gt;C (ds)</td>
<td>5'-GCATGGG&lt;sup&gt;me&lt;/sup&gt;G&lt;sup&gt;me&lt;/sup&gt;G GCATGAACCG-3' 3'-CGTACCCG&lt;sup&gt;me&lt;/sup&gt;CGTACTTGGC-5'</td>
</tr>
<tr>
<td>C:&lt;sup&gt;me&lt;/sup&gt;G:&lt;sup&gt;me&lt;/sup&gt;G:C (ds)</td>
<td>5'-GCATGGG&lt;sup&gt;me&lt;/sup&gt;G&lt;sup&gt;me&lt;/sup&gt;G&lt;sup&gt;me&lt;/sup&gt;G GCATGAACCG-3' 3'-CGTACCCG&lt;sup&gt;me&lt;/sup&gt;CGTACTTGGC-5'</td>
</tr>
</tbody>
</table>

Bold font indicates modified bases or their complements.
A colon indicates that the bases are base-paired with the bases in the opposite strand.

Figure 4.30 shows the data for the inactivation of the extended and normal MGMT proteins (induced in the absence of supplementary zinc) with only the ds ODN. There was very slight and closely similar inactivation with the control G:<sup>me</sup>C ODN and a range of inactivation by the other ODNs. These data were used to determine the IC<sub>50</sub> values which are shown in Figure 4.31 along with and those from the other ODNs. Because the extent of the inactivation was so low, it was not possible to determine IC<sub>50</sub> values for inactivation by the ss G, C:<sup>me</sup>C:G and G:<sup>me</sup>C controls.
Figure 4.30: Effect of pre-incubation with the ODNs indicated on the activity of extended and normal MGMT. A. Extended MGMT; B. Normal MGMT

Figure 4.31A shows that for each of the ODN, the IC$_{50}$ was lower for the extended than the normal MGMT, as was seen for all other ODN used in these studies but only $C^{\text{me}}G$ (ss), $\text{me}C^{\text{me}}G$ (ss) and $\text{me}G: \text{me}C$ (ds) were significantly different between extended and normal MGMT. Figure 4.31B shows that 5-MeC in the ss ODN did not affect the inactivation either the extended or normal MGMT but 5-MeC in ds ODNs, did increase the inactivation of MGMT. For normal MGMT, IC$_{50}$ values
for $\text{meG:meC (ds)}$ and $\text{meC:meG:GmeC (ds)}$ were significantly lower than that of $\text{meG:C (ds)}$ but only that for $\text{meC:meG:GmeC (ds)}$ was significantly different for the extended MGMT.

**Figure 4.31:** Effect of cytosine methylation on the IC$_{50}$ values of extended and normal MGMT proteins with modified ss and ds ODNs. Bars represent the mean and standard deviation of quintuplicate data points. **A.** comparison between extended and normal Wt MGMT proteins (* Significantly different from extended Wt MGMT; P<0.05 using one way ANOVA) and **B.** comparison between 5-MeC containing ODNs (* Significantly different from the C$^{\text{meG}}$ (ss) and * Significantly different from the meG:C (ds) P<0.05 using one way ANOVA).

### 4.5 Discussion

#### 4.5.1 Expression of extended and normal MGMT in human lymphocytes

In this limited study, all five human PBMC cDNA samples generated a PCR product that was expected from the normal cDNA transcript, and none produced only
a signal that would have originated from a putative shorter MGMT splice variant. However, there appeared to be differences in the relative intensities of the bands for different PBMC preparations, particularly in sample 1 where the extended was more intense than the normal, and sample 3 where the normal was more intense than the extended (Figure 4.4). It is therefore possible that both full length and truncated transcripts are present in these samples. In further studies, northern analysis and quantitative real time PCR might be used to confirm this observation and generate quantitative data, and a substantial number of lymphocyte cDNAs should be examined.

Given that there are literature reports based on the results of western analyses that longer versions of the MGMT protein can be expressed (Brent et al., 1990; von Wronski et al., 1991; Ostrowski et al., 1991; He et al., 1992) it was considered reasonable to investigate if an extended version of MGMT would be functionally active and how it might compare with the normal protein.

4.5.2. Properties and functional activity of extended and normal MGMT

Both the extended and normal MGMT proteins were found to be functionally active. Considering that the extended and normal MGMT proteins were purified in the same way and appeared to be single bands on Coomassie-stained gels, the reasons for the differences in specific activities (Figure 4.17) are obscure. The kinetics of methyl transfer did appear to be slightly different (Figure 4.18), but not sufficient to explain the marked difference in specific activity. The thermal stabilities of both proteins at 44°C and 51°C were almost identical (Figure 4.19). Pre-incubation of the proteins with calf thymus DNA resulted in greater inhibition of the extended protein, suggesting that non-specific binding was more extensive: this might explain the apparently slower transfer rates for this extended protein (Figure 4.20). The effect of supplementary zinc in the E.coli growth media is discussed later.

The inactivation of the two proteins by ODN containing a variety of 6-alkGs was carried out using fixed amounts of active protein and identical pre-incubation and incubation conditions. It was considered that the results would not be influenced by any differences in specific activity, and so could be compared. For most of the
ODN, the extended protein was much more susceptible to inactivation than the normal protein. The exception to this was the \( O^6 \)-CMG-containing ODN which gave identical IC\(_{50}\) values for the two proteins (Figure 4.23). The most effective inactivator was the ODN containing \( O^6 \)-BzG and the least effective was \( O^6 \)-AEG which was \(~1000\) times less potent. The absolute differences between extended and normal proteins tended to increase as the effectiveness of the ODN increased, but there was no consistent pattern in the relative extent of the differences (Table 4.1).

### 4.5.3 The effects on MGMT activity of adding ZnCl\(_2\) to E.coli growth media

One possibility to explain the much lower specific activity of the extended protein is that it might have been expressed in much higher levels in E.coli, and if the trace amounts of Zinc that were presumably present in the growth media were limiting, then more of the extended protein may have been lacking Zn. To investigate this, ZnCl\(_2\) was added to the E.coli growth media during the induction of MGMT expression. The purified proteins were assayed (Figure 4.24) and the specific activities (Table 4.2) indicated that while there was a \( \sim3\) -fold increase in the specific activity of the extended protein following induction in Zn-supplemented media, there was also an increase with the normal protein, although to a lesser extent, and its specific activity was less than three times that of the extended protein. A previous study by (Rasimas \textit{et al.}, 2003) indicated that MGMT expressed in E.coli grown in media supplemented with 100 \( \mu\)M ZnCl\(_2\) had a 5 fold increase in activity compared with media without additional ZnCl\(_2\). The differences between those and the present results may reflect the media used in the two studies. This was not investigated further.

Five of the \( O^6 \)-alkG-containing ODNs were used in experiments to compare their effectiveness in ss and ds forms without Zn supplementation and in the ds form only comparing with and without supplementation. The results for all one way ANOVA show statistically significant differences (Figure 4.26) and the summary of the data show that for all \( O^6 \)-alkGs, ds ODNs were significantly more potent than ss ODNs (approximately 1.5-3 times) and that ZnCl\(_2\) supplementation results in MGMT proteins that are significantly less sensitive to inactivation by ds ODNs (approximately 2 times). If the binding levels seen are inversely proportional to the
repair of the alkyl group, this suggests that the normal protein is much more effective in repair than the extended protein and that supplementary ZnCl₂ approximately doubles the efficiency of repair.

ZnCl₂ supplementation in the *E.coli* growth media has been reported to enhance the stability of the interaction between N and C terminal domains (Daniels *et al.*, 2000; Rasimas *et al.*, 2003). It seems likely that a less flexible structure is generated and that this restricts the access of substrates into the active site region, thus making the protein less susceptible to inactivation. The results of the ELISA analyses support this hypothesis as binding of the MGMT proteins expressed in ZnCl₂-supplemented media was consistently lower than in media without additional Zinc. The physiological effects of Zinc binding in human cells may include protein folding and turnover, substrate binding and alkyl transfer, and proteasome degradation and merit further investigation.

### 4.5.4. Enzyme-linked immunosorbent assay (ELISA) analysis of MGMT binding to double-stranded ODNs

These radioisotope-based assays have been assumed to be an indication that the MGMT was inactivated by pre-incubation with the ODN. However, as already discussed in Chapter 3, they cannot distinguish between inactivation and binding that would not necessarily result in inactivation but competitively inhibit the action of the MGMT on the radiolabelled substrate. It was therefore considered worthwhile to attempt to establish an ELISA assay using rabbit polyclonal antibodies to MGMT that had been produced in house. Initially, it was considered unlikely that this would be successful, as the literature indicates that MGMT binds transiently and non-specifically to DNA (Bender *et al.*, 1996; Rasimas *et al.*, 2003; Melikishvili *et al.*, 2008) and the SPR results presented in Chapter 3 show that binding is transient, and repair, when it occurs, happens quite rapidly.

When MGMT was added to the wells of SA-coated microtitre plates that had been pre-incubated with G, O⁶-MeG or O⁶-BzG-containing ODN, there was a very marked MGMT binding to the O⁶-alkG ODN and close to background levels for the G ODN (Figure 4.28). As MGMT is reported to act rapidly on O⁶-MeG and even more rapidly on O⁶-BzG (see introduction), these results were not expected.
However, the signal from the $O^6$-BzG ODN was somewhat lower than that from the $O^6$-MeG ODN, particularly at higher levels of MGMT possibly indicating repair of the lesion and dissociation of the MGMT.

ELISA was then used to investigate the binding of extended and normal MGMT to these ODN, and the effect of Zn supplementation. Any impact that the length of the ODN might have on binding was also examined, but using only two ODNs: a 13-mer and a 23-mer. These results demonstrate that, under the experimental conditions used, the length of the ODN did not substantially affect the binding of MGMT (Figure 4.29). However, the extended protein binds to the ODNs on the plates to a significantly higher level than to the normal protein and the presence of ZnCl$_2$ in the *E.coli* growth media results in proteins that have considerably (~50%) reduced binding levels in this assay.

These results suggest that the amino-terminal region of the extended protein has an influence on the action of MGMT by allowing it to bind more extensively than the normal protein to the ODN, and in the radioactivity based functional assay, resulting in a reduced ability to act on the $O^6$-MeG in the radiolabelled substrate and hence more effective inactivation. In addition, with the radiolabelled substrate, the methyl group transfer to the extended protein is slower than that of the normal protein. These experiments do not exclude the possibility that the normal protein dissociates more readily than the extended protein from the $O^6$-alkG-containing ODN. While CT DNA binding seemed to result in somewhat greater inhibition of the extended protein, this might be the result of slower transfer rates, especially since there appeared to be little or no binding of either the extended or normal MGMT to the G-containing ODN in the ELISA.

These results appear to contradict the SPR-based observations in Chapter 3 that MGMT binding to $O^6$-alkG-containing ODN is transient. In fact, since there was little binding seen with the G-containing ODN, the ELISA results suggest that the binding is both strong and alkyl specific. Preliminary ELISA results (not shown) indicate that proteinase K digest is able to remove the bound MGMT, and that, based on a reduced level of binding by Atl1, the alkyl lesion is removed. One possible explanation is therefore that under the conditions used in ELISA, but not those in
SPR, the alkylated MGMT stays in contact with the ODN. This clearly requires further examination using ODN of substantially different lengths and possibly SA-coated magnetic beads to avoid the use of a plastic or gold-plated support matrix.

4.5.5 Effect of cytosine methylation on the inactivation of extended and normal MGMT by \(O_6\)-MeG

Some preliminary studies were carried out on short ODNs spanning the \(p53\) codon 245 containing a CpG known to undergo cytosine methylation. ODNs containing several combinations of methylated and non-methylated cytosines and \(O_6\)-MeGs were produced and examined in the standard competition assay. To simplify the task, only the -Zn normal and extended MGMT proteins were used. For all of these ODNs, the extended protein was more effectively inhibited than the normal protein, in agreement with previous results in this chapter. For both proteins, the most potent inhibitor was the ODN in which both of the Cs and the “bottom” strand G were methylated. However, as previously discussed, the more extensive inhibition may be due either to greater binding or greater inactivation and this would need to be explored in future studies. There is one other published report of studies on the inactivation of MGMT by \(O_6\)-MeG in the context of cytosine methylation: (Guza et al., 2009) used a ss ODN spanning \(p53\) codon 245 and found that the rate of repair of \(O_6\)-MeG by Wt MGMT was the same if the base 5’ to \(O_6\)-MeG was 5-MeC or C. It is worth noting that for other \(p53\) codons, relative to C located 5’ to the \(O_6\)-MeG, repair rates were either increased (codon 158) or decreased (codon 248) when 5-MeC was located 5’ to the \(O_6\)-MeG. Guza et al., (2009) also used codon 245 ds ODNs as in the present report, and measured repair rates. They found, relative to ds ODN containing C, when 5-MeC was located 5’ to, or opposite, \(O_6\)-MeG the rate was increased, but when located both 5’ and opposite \(O_6\)-MeG the rate decreased. The IC\(_{50}\) values shown in the present report indicated decreases for all three of these ODNs, suggesting an increase in repair rate for all three. The basis of these differences might be due to the use of a histidine-tagged MGMT in the work of Guza et al., (2009), or the use of IC\(_{50}\) values rather than kinetic measurements.
Chapter 5: $O^6$-carboxymethylguanine as a substrate for MGMT and the Ogt protein

5.1 Introduction

Previous reports suggested that $O^6$-CMG was not a substrate for MGMT (Shuker & Margison, 1997), but the observations in Chapter 3 and 4 contradicted this. To further confirm these findings and to investigate the reason for the earlier conclusion, ELISAs and an restriction endonuclease (RE) site deprotection assay were used in *in vitro* studies. These studies were then extended to examine the inactivation of MGMT and the formation of DNA damage in cultured human cells by treatment with azaserine, which both methylates and carboxymethylates DNA. Finally, the *E.coli ogt* gene was isolated and the Ogt protein expressed, purified and directly compared with MGMT for inactivation by $O^6$alkG-containing ODN. Some of this work has been published (see Appendix 4).

5.2 Alkyl group removal assessed by ELISA

To establish ELISA methods for the quantitation of $O^6$-MeG and $O^6$-CMG in DNA, SA-coated microtitre plates were coated with 0-80 fmoles of biotinylated $O^6$-MeG, $O^6$-CMG and G-containing ds ODNs (13mer; 5'-SIMA-GCCATGXC TAGTA where X was G or the modified purine base). These were probed with antibodies that recognise $O^6$-CMG (in house) commercial antibodies to $O^6$-MeG (Squarex) or with 1 pmole of purified Atl1 which was then detected using antibodies to Atl1 (in house). In all cases, peroxidase-tagged secondary antibodies were used to generate an ECL signal (see section 2.5.7). Figure 5.1A shows that 1 pmole of Atl1 bound to ODNs containing $O^6$-MeG and $O^6$-CMG but not to ODNs containing G. Figure 5.1B shows that the anti-CMG antibody detected $O^6$-CMG but not $O^6$-MeG or G in the ODNs. In both cases, the intensity of the ECL signals was reasonably proportional to the amounts of ODNs in the wells. Eighty fmoles of ODNs were easily detected by these methods and this amount was used in subsequent experiments. The lower limit of quantitation was around 10 fmoles.
Figure 5.1: Binding of ds $O^6$-alkG containing ODNs to SA-coated microtitre plates as determined by ELISA using A anti-Atl1 antibody and B, anti-CMG antibody. Each point represents the mean and standard deviation of the ECL signals obtained for triplicates wells in the plates. Error bars represent the standard deviation (SD); where error bars are not visible, they are masked by the symbol.

In the following experiments, only normal Wt MGMT that had been grown in media without supplementary Zn was used. Increasing amounts of MGMT (0-210 fmoles in 100 µl of buffer I) were incubated at 37°C for 1 h in triplicate wells coated with 80 fmoles of the ODNs. After probing the plates using the two detection methods, there was an almost linear decrease in the ECL signals generated by both
the Atl1-based (Figure 5.2A) and anti-\(O^6\)-CMG antibody-based (Figure 5.2B) methods. This supported the previous findings (Chapters 3 and 4) that the two alkyl groups are removed from the ODN and since the signals decreased almost in parallel, removal of both lesions had similar efficiencies. In addition, Figure 5.2B confirms that the anti- \(O^6\)-CMG antibody does not recognise \(O^6\)-MeG-containing ODN under these conditions.

**Figure 5.2:** Binding of MGMT on ds \(O^6\)-alkG containing ODNs as determined by ELISA using A, anti-Atl1 antibody and B, anti-CMG antibody. Each point represents the mean and standard deviation of the ECL signals obtained for triplicate wells in the plates. Error bars represent the standard deviation (SD); where error bars are not visible, they are masked by the symbol.
5.3 Alkyl group removal assessed by restriction endonuclease site deprotection

As shown using mass spectrometry (Chapter 3), the action of MGMT on the $O^6$-CMG involved removal of the CM group. Therefore, by analogy with the effect of MGMT on $O^6$-MeG, a blocked RE site should be restored, providing additional confirmation of the repair mechanism. The RE sites present in the 13-mer ODN used in most of these studies are shown in Figure 5.3A and the $PstI$ site location in the 23-mer is shown in Figure 5.3B.

![Diagram of restriction endonuclease sites](image)

**Figure 5.3:** Location of the restriction endonuclease recognition and cleavage sites in the **A**, 13-mer ODN and **B**, 23-mer ODN are indicated. The locations of the fluorescence tags, SIMA and HEX are shown in red.

To initially assess which RE would be suitable, 7 pmoles of the G, $O^6$-MeG and $O^6$-CMG-containing 13-mer were incubated with $Nla$III, CviKI, $Sty$I or $Bsa$II prior to analysis by non-denaturing PAGE with fluorescence detection of parent
ODN (see section 2.5.10). Figure 5.4A-5.4C shows that NlaIII, CviKI and StyI efficiently digested control (G), O\textsuperscript{6}-MeG and O\textsuperscript{6}-CMG-containing ODNs whereas BsaJI cleavage of both O\textsuperscript{6}-MeG and O\textsuperscript{6}-CMG containing ODN was substantially less that of the G-containing ODN (Figure 5.4D).

A. NlaIII  
B. CviKI  
C. StyI  
D. BsaJI

**Figure 5.4:** Fluorescence image of non-denaturing PAGE of ds O\textsuperscript{6}-alkG containing ODNs, before (-) or after (+) digestion with A, NlaIII; B, CviKI; C, StyI and D, BsaJI. In all cases, the upper band is the intact ODN and the lower band is the digestion product or some unannealed ss ODN.

The above result indicated that BsaJI was able to differentiate between G and O\textsuperscript{6}-alkG-containing ODNs. The effect of pre-incubation of each of these ODNs and the 23-mer ODN that contained a PstI site that was known to be blocked by O\textsuperscript{6}-MeG and was found to be blocked by O\textsuperscript{6}-CMG (see below) was then determined. Figure 5.5 shows that for each ODN, incubation with MGMT alone had no effect: the ds
band and some excess ss ODN migrated unchanged during the non-denaturing PAGE. *Bsa*II (Figure 5.5A) and *Pst*I (Figure 5.5B) were able to digest the G-containing ODNs almost to completion, and this was not affected by pre-incubation with MGMT. *Bsa*II and *Pst*I were unable to digest the *O*<sup>6</sup>-MeG or *O*<sup>6</sup>-CMG-containing ODNs, but for both RE, this was almost completely reversed by pre-incubation with MGMT. Thus the action of MGMT on DNA containing either *O*<sup>6</sup>-MeG or *O*<sup>6</sup>-CMG results in the removal of the alkyl group and the formation of guanine.

A.

![Fluorescence images of non-denaturing PAGE of ds *O*<sup>6</sup>-alkG-containing ODNs.](image)

B.

![Fluorescence images of non-denaturing PAGE of ds *O*<sup>6</sup>-alkG-containing ODNs.](image)

**Figure 5.5:** Fluorescence images of non-denaturing PAGE of ds *O*<sup>6</sup>-alkG-containing ODNs. In all cases, the upper band is the intact ODN, the intermediate band is the digestion product and the lower band is some unannealed ss ODN. **A**, *Bsa*II and **B**, *Pst*I.
5.4 Inactivation of MGMT and DNA damage in cultured human cells treated with azaserine

To investigate whether or not $O_6$-CMG is a substrate for MGMT in vivo, cultured human melanoma (A375) cells were incubated with increasing doses of azaserine (0-580 µM) and cell free extracts were prepared and initially assayed for MGMT activity (as described in section 2.5.11). Untreated control cells expressed a high level of MGMT activity but this was decreased in cells treated with 29 µM azaserine and all higher doses completely ablated the activity (Figure 5.6 and 5.7).

![Figure 5.6:](image)

**Figure 5.6:** MGMT activity in A375 cells extracts following exposure to azaserine. Incrementally increasing amounts of each extract (based on the DNA content of the extract, as indicated) were incubated with substrate DNA as described in section 2.5.11. Substrate limiting conditions were reached for extracts of the control and 29 µM azaserine-treated cells.

As azaserine both methylates and carboxymethylates DNA (Harrison et al., 1997; Shuker & Margison, 1997; Harrison et al., 1999) inactivation of MGMT might have been due only to the formation of $O_6$-MeG in cellular DNA. To establish that both $O_6$-MeG and $O_6$-CMG were generated, increasing amounts of crude cell-free extracts, which contained cellular DNA as well as protein, were pre-incubated with standardised amounts of either Atl1 or anti-CMG antiserum. The mixture was then added to microtitre plates coated with $O_6$-MeG or $O_6$-CMG-containing ODNs, respectively, and signals generated by probing with anti-Atl1 in the case of Atl1 and then the corresponding secondary antibodies.
Figure 5.7: MGMT specific activities (fmoles/µg DNA) calculated from the activity determined under protein-limiting conditions.

Figures 5.8 and 5.9 show that there was a decrease in the binding of Atl1 and anti-CMG to the $O^6$-MeG or $O^6$-CMG-containing ODNs, respectively. The decrease was apparently more extensive for Atl1 than anti-CMG but these data are only qualitative: to determine the absolute amounts of the lesions, it was necessary to generate calibration curves by preincubation of the Atl1 and anti-CMG with $O^6$-MeG and $O^6$-CMG-containing ODNs.

Figure 5.8: Effect of Atl1 pre-incubation with cell-free extracts of azaserine-treated A375 cells on their respective ability to bind to $O^6$-MeG-containing ODNs. Each point represents the mean and standard deviation of the ECL signals obtained for triplicate wells in the plates. Error bars represent the standard deviation (SD); where error bars are not visible, they are masked by the symbol.
Having demonstrated that both $O^6$-MeG and $O^6$-CMG were generated in cellular DNA by azaserine, it was reasonable to examine if MGMT affected the levels of both $O^6$-MeG and $O^6$-CMG in DNA following treatment of the cells with azaserine. To assess this, cells were pre-incubated with an MGMT inactivating dose of the agent, Lomeguatrib (LM), which would be expected to increase the amounts of lesions detected in extracted DNA if they were normally removed by MGMT. One hour later, the cells were treated with azaserine (29 $\mu$M) and 16 h later, the cells were harvested, DNA was extracted, and $O^6$-alkG levels determined as described in section 2.5.11. This single experiment was carried out in triplicate. Figure 5.10 shows that in comparison with azaserine only treated cells, the DNA extracted from LM and azaserine treated cells contained significantly higher levels of $O^6$-CMG (detected using anti-CMG antibody; P<0.05) and to a lesser extent $O^6$-MeG (detected with an antibody to $O^6$-MeG; P< 0.05). Atl1 can bind both $O^6$-MeG and $O^6$-CMG and the anti-Atl1 antibody generated the highest signal, approximately corresponding to the other two signals and also showing a difference in binding by Atl1 (P<0.05).
Figure 5.10: Effect of inactivation of MGMT by pre-incubation of A375 cells with Lomeguatrib on the levels of $O^6$-MeG and $O^6$-CMG after exposure to asazerine. Isolated DNA was analysed as indicated using antibodies to $O^6$-MeG and $O^6$-CMG and the binding At11 to both of these lesions. Values are means +/- SD of 3 data points. * Significantly different from without Lomeguatrib (P<0.05).

5.5 Ogt protein production, purification and characterisation

In order to investigate previous claims that alkyltransferases are unable to repair $O^6$-CMG in DNA, the *E.coli* Ogt alkyltransferase-encoding gene was isolated and the Ogt protein expressed and purified. Its activity was then compared with that of normal Wt MGMT. Radioactivity-based assays were used to investigate the kinetics of methyl group transfer to proteins, thermal stability and inactivation by modified ODNs.

5.5.1 Cloning, expression and purification Ogt

The *ogt* gene was amplified from 10 pg of *E.coli* K12 template DNA by PCR using *ogt* 5’-phosphorylated forward and reverse primers as described in section 2.4. Figure 5.11 shows the expected 516 bp product.
Figure 5.11: Agarose gel electrophoresis of the PCR amplification products of E. coli K12 genomic DNA using the primer pairs (5'-PO4-ATGCTGAGATTACTTG AAGAAAAATTGC) and (5'-PO4-TTACAGCAAAAGATAACCTTCATGGCG).

The 516 bp PCR product was isolated and ligated into XmnI digested and phosphatase treated pMAL-c2 and transformed into competent XL-1 blue E. coli as described in section 2.2.2 and 2.2.3. Ampicillin resistant clones were screened for the Ogt insert, by colony PCR using the ogt forward and reverse primers as described in section 2.2.4. Figure 5.12 shows that clones 3-10 and 12 contained a band representing the expected product (516 bp).

Figure 5.12: Agarose gel electrophoresis of PCR amplification products of E. coli colonies produced by transformation with the ogt PCR products ligated into pMAL-c2. The positive control was PCR amplification of E. coli K12 DNA.

Clone number 5 was subjected to plasmid miniprep and the insert was sequenced as described in section 2.1.1. ClustalW showed that the obtained sequence was identical to the published ogt sequence shown in Figure 2.4.
Clone 5 was induced with IPTG, harvested 0, 1, 2 and 3 h later and cell extracts were prepared and separated by 12% SDS-PAGE (see section 2.2.6). Figure 5.13 shows that Coomassie blue-stained protein bands corresponding to the predicted molecular weight for the MBP-Ogt fusion protein (61 kDa) and for MBP (42 kDa) were detected and that these bands were much more intense at 3 h.

![Figure 5.13](image1)

**Figure 5.13:** Expression of MBP-Ogt fusion protein in extracts of *E.coli* prepared at various times after addition of IPTG. Extracts were analysed by 12% SDS-PAGE with Coomassie staining.

The 3 h induction sample was purified using amylose affinity columns as described in section 2.2.7 and fractions collected at the various stages were analysed by 12% SDS-PAGE. Figure 5.14 shows that the elution fractions contained bands that corresponded to the MBP-Ogt fusion protein and MBP.

![Figure 5.14](image2)

**Figure 5.14:** Purification of the MBP-Ogt fusion protein analysed by 12% SDS-PAGE with Coomassie staining.
Proteins from a duplicate of the gel shown in Figure 5.14 were electrotransferred onto PVDF membranes and probed with an anti-MBP antibody as described in section 2.2.8. Figure 5.15 shows that both the 61 kDa and 42 kDa bands were detected by this antibody.

**Figure 5.15:** Western blot of the purification fractions of the MBP-Ogt fusion protein. The membrane was probed with anti-MBP antibody and binding assessed by ECL.

The MBP-Ogt protein was cleaved using factor Xa by incubation at room temperature for 6 h as described in section 2.2.9. Figure 5.16 shows complete cleavage to two bands corresponding to MBP (42 kDa) and the Ogt protein (19 kDa).

**Figure 5.16:** Factor Xa cleavage of MBP-Ogt fusion protein analysed by 12% SDS-PAGE with Coomassie staining.
The factor Xa cleavage products were purified using DEAE-sepharose as described in section 2.2.10. Figure 5.17 shows that a band corresponding to Ogt (19 kDa) eluted at 50-75 mM NaCl and MBP (42 kDa) eluted at 100-150 mM NaCl.

A. Figure 5.17: Purification of the factor Xa cleaved MBP-Ogt fusion protein by DEAE-sepharose analysed by 12% SDS-PAGE with Coomassie staining. A. Elution with 25-75 mM NaCl, B. Elution with 100-500 mM NaCl.

Amylose chromatography was used as a final step to purify the elution fractions containing Ogt. Figure 5.18 shows that, as expected, the cleaved Ogt did
not bind to the column and was eluted in the binding buffer washes. Figure 5.19 shows the result of the western blot of a duplicate gel, which confirms that band at 19 kDa is the Ogt protein.

**Figure 5.18:** Re-purification of Ogt using an amylose column. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining.

**Figure 5.19:** Western blot of the amylose column re-purification fractions for the factor Xa-cleaved, DEAE-purified MBP-Ogt fusion protein. The membrane was probed with anti-Ogt antibody (in house).
5.5.2 Determination of Ogt methyl transfer activity

The activity of the purified Ogt protein was determined using \([^3\text{H}]\)-methylated substrate DNA as described in section 2.5.1. The results in Figure 5.20 were used to calculate protein limiting amounts of MGMT using the Excel software set to linear fitting, which generally gave the equation \(y = ax + b\) and showed correlation coefficients \((R^2)\) of 0.99 or better. From Figure 5.20 the protein limiting amount of Ogt for subsequent assays was selected as 0.24 µg.

![Figure 5.20](image)

**Figure 5.20:** Transfer of \([^3\text{H}]\)-methyl groups from \([^3\text{H}]\)-methylated substrate DNA to Ogt with increasing amounts of Ogt protein during incubation at 37°C for 1 hour. Vertical lines indicate the protein limiting amounts of the Ogt protein that was chosen for further studies.

5.5.3 Kinetics of methyl group transfer to Ogt protein

The kinetics of \([^3\text{H}]\)-methyl group transfer from \([^3\text{H}]\)-methylated substrate DNA was determined for Ogt and the normal MGMT protein (Chapter 4) over a 5 h incubation period as described in section 2.5.2. Figure 5.21 shows that while the maximum amounts of radioactivity transferred were ultimately closely similar, the kinetics were slightly different: normal human MGMT displayed more rapid transfer than Ogt.
5.5.4 Thermal stability

The activity of Ogt decreased rapidly during pre-incubation at 44°C, with a 50% loss of activity by 4 min and a complete loss of activity at 20 min. In comparison, the normal MGMT protein was much more stable, losing 50% of activity by ~27 min (Figure 5.22; the MGMT data is also shown in Chapter 4; Figure 4.19).

![Figure 5.21: The kinetics of methyl group transfer from substrate DNA.](image1)

![Figure 5.22: Thermal stability of normal MGMT and Ogt protein during incubation at 44°C.](image2)
5.5.5 Ogt inactivation by modified ODNs

The effect of pre-incubation of Wt and variant MGMT for 1 h at 37ºC with ss ODNs having the common sequence: 5’-SIMA-GCCATGXXCTAGTA where X is G or the modified purines listed in Figure 2.5 was examined as described in section 2.5.4. The initial broad range of ODN concentrations (data not shown) identified that the five most potent inactivators of Ogt were the ODNs containing $O^6$-MeG, $O^6$-PrG, $O^6$-EtG, $O^6$-BzG or $O^6$-CMG. These were assayed with Ogt along with the normal MGMT over a narrower concentration range, the IC$_{50}$ values were determined (Figure 5.23). The ODN containing $O^6$-BzG caused only very weak inactivation of Ogt whereas it was, in confirmation of the results in Figure 4.23, the most potent inactivator for human MGMT. The ODN containing $O^6$-CMG was by far the weakest inactivator of Ogt. For the $O^6$-MeG, $O^6$-EtG and $O^6$-PrG containing ODNs, Ogt had closely similar IC$_{50}$ values to the normal MGMT.

![Figure 5.23](image)

**Figure 5.23:** IC$_{50}$ values (nM) for the inactivation of normal MGMT and Ogt proteins by the five ss $O^6$-alkG containing ODNs. Values are means +/- SD of 5 data points. * Significantly different from the normal MGMT (P<0.05; pairwise comparisons)

The other 6 modified ODNs containing $O^6$-HOEtG, HOPr-DAP, $O^6$-AEG, $O^6$-MAG, $O^6$-PobG and DAP were ineffective inactivators of Ogt. Pre-incubation resulted in IC$_{50}$ values that were considerably greater than 24.4 nM (Figure 5.24) and in most cases inhibition of Ogt activity was to the same or similar extent as the G-
containing control ODN, the IC50 for which was not calculated, but extrapolation indicated that it was in excess of 24.4 nM.

![Figure 5.24](image)

**Figure 5.24:** Effect of pre-incubation of Ogt with ss $O^6$-alkG containing ODNs.

5.6 Discussion

5.6.1 $O^6$-carboxymethylguanine as a substrate for MGMT

Earlier reports using cell-free extracts from *E.coli* and a human cell line suggested that alkyltransferases do not act on $O^6$-CMG in DNA (Shuker & Margison, 1997) and this was supported by studies in human cell lines treated with carboxymethylating agents (O’Driscoll *et al.*, 1999). The conclusions drawn by O’Driscoll *et al* (1999) might be due to the possibility that $O^6$-CMG can be repaired by both MGMT and NER or that other products of carboxymethylating agents that are processed by NER are more toxic than $O^6$-CMG in the cells used by those authors. Given the compelling evidence for a role for MGMT in protecting against CRC, this suggested that $O^6$-CMG might not be a major factor in CRC risk, despite the observation that its abundance increases in high risk diet situations (see Chapter 1). However, the results in Chapters 3 and 4 showed that, in contrast to the previous report, $O^6$-CMG is a potent inactivator of MGMT in *vitro* (see Figures 3.18 and 4.23) and mass spectral analysis indicated that carboxymethyl groups are transferred from $O^6$-CMG in ODN to the active site cysteine of MGMT (see Figure 3.20F).
The results in this chapter firstly demonstrate that the anti- $O^6$-CMG antibody specifically detects this lesion in ODN and that the Atl1 protein is able to detect both $O^6$-CMG and $O^6$-MeG in ODN. The sensitivity of these methods enabled studies of the processing of these lesions simultaneously. In confirmation of the previous MS analyses (Chapter 3), it was then shown that, in vitro, MGMT removes alkyl groups from $O^6$-CMG and $O^6$-MeG with similar efficiencies, and that, based on RE digests, this regenerates guanine.

To extend these observations to intact human cells, it was then shown treatment of melanoma cells with azaserine inactivated MGMT and generated damage in DNA that was detected using antibodies to $O^6$-MeG and $O^6$-CMG and also using the ability of the Atl1 protein to bind to both of these lesions. These results are supported by the report of (Harrison et al., 1999) that azaserine produces both $O^6$-CMG and $O^6$-MeG in DNA. However, these authors found 38 times more $O^6$-CMG than $O^6$-MeG in DNA whereas here the ratio was 4:1. This may be explained by the fact that (Harrison et al., 1999) reacted purified DNA in vitro and with much higher concentrations of azaserine (10 mM). In cultured cells treated with low doses of azaserine, the relative amounts of $O^6$-MeG and $O^6$-CMG may be affected by the action of the NER system as suggested by O'Driscoll et al., 1999 (see above).

MGMT inactivation in A375 cells might have been due to its action on either of these lesions, so cells were pre-incubated with the MGMT inactivating agent, LM, and then exposed to azaserine. The LM pretreatment resulted in a statistically significant increase in the levels of both $O^6$-MeG and $O^6$-CMG, as indicated by the use of specific antibodies, and confirmed by the use of Atl1, which detects both lesions. These results show that MGMT inactivation increases the levels of both lesions and hence that MGMT can act on both lesions.

5.6.2. Properties and functional activity of Ogt protein

The previous claim that MGMT does not act on $O^6$-CMG in DNA (Shuker & Margison, 1997) was based on the use of semi-purified E.coli alkyltransferases. In order to investigate this further, the E.coli Ogt protein was overexpressed and
purified and partially characterised, before its sensitivity to inactivation by \( O^6 \)-CMG-containing, and several other, ODNs was compared with that of MGMT.

Thermal stability measurements at 44°C showed that Ogt lost 50% of activity by 4 min and a complete loss of activity at 20 min at incubation 44°C. This is similar to a previous report by (Rebeck et al., 1989) in which at 55°C, Ogt rapidly lost activity with a half-life of about 2 min and had completely lost activity after 3 min.

The IC\(_{50}\) for the inactivation of Ogt by ODN containing \( O^6 \)-MeG was statistically significantly greater that that of MGMT. This contrasts with previous report by (Elder et al., 1994) in which Ogt was reported to be more sensitive than MGMT to inactivation by \( O^6 \)-MeG. The extrapolated IC\(_{50}\) values were of a similar order of magnitude in both of these studies, so there may be some methodological or ODN sequence-based explanation for the differences. However, in the case of the \( O^6 \)-BzG-containing ODN, (Elder et al., 1994) reported Ogt and MGMT to have similar sensitivity, and this is in marked contrast to the finding here that MGMT was ~300 times more sensitive than Ogt. The present result is consistent with that of (Pegg et al., 1993) showing that the Ogt was inactivated by \( O^6 \)-BzG but was much less sensitive than MGMT, so the results of (Elder et al., 1994) remain to be explained.

Ogt repaired \( O^6 \)-MeG more effectively than \( O^6 \)-EtG and \( O^6 \)-PrG and this is supported by data previously reported by Wilkinson et al., (1989) and Morimoto et al., (1985) respectively.

Finally, Ogt was found to be more than 20 times more resistant than human MGMT to inactivation by the \( O^6 \)-CMG containing ODN. It seems likely that this explains the conclusion of (Shuker & Margison, 1997) that alkyltransferases are unable to repair \( O^6 \)-CMG in DNA because these authors used only the \( E.coli \) protein and extrapolated their findings to MGMT.
Chapter 6: Overall discussion

Humans are exposed to alkylating compounds produced endogenously and in the environment (Margison et al., 2003). These agents produce DNA adducts and these adducts can be genotoxic, inducing mutagenic, carcinogenic, recombinogenic, genotoxicity and cytotoxic effects (Margison et al., 2002). O\textsuperscript{6}-alkylguanine-DNA alkyltransferase (MGMT) is a DNA repair protein that repairs O\textsuperscript{6}-alkG adducts by removing the alkyl group from DNA to a specific cysteine residue of the protein. Although several reports have investigated the association between MGMT polymorphisms and cancer risk in various patient populations, the evidence is unclear as epidemiological and laboratory studies are often contradictory (Ritchey et al., 2005; Han et al., 2006; Chae et al., 2006), possibly because it is difficult to ascertain exposure to alkylating agents in human studies (Reh et al., 2000). Our hypothesis was that MGMT variant proteins might provide a different level of protection against different alkylating agents as they might repair O\textsuperscript{6}-alkG to different extents. Hence this might alter cancer risk, or affect the effectiveness of cancer treatments that employ alkylating agents.

To examine this hypothesis, wild type MGMT (extended and normal), six MGMT variants and Ogt have been expressed and functionally characterised. Tables 6.1, 6.2 and 6.3 summarise the characteristics of the MGMT variants (Table 6.1), the extended MGMT protein (Table 6.2) and Ogt (Table 6.3) in comparison with normal Wt MGMT.

There was clear evidence that repair of O\textsuperscript{6}-alkG within DNA by MGMT depends upon both the nature of the alkyl group and the MGMT variant protein. In particular the Q128 MGMT variant was the most resistant to inactivation by all of the O\textsuperscript{6}-alkG containing ODNs (O\textsuperscript{6}-MeG, O\textsuperscript{6}-EtG, O\textsuperscript{6}-PrG, O\textsuperscript{6}-BzG, O\textsuperscript{6}-PobG and O\textsuperscript{6}-CMG) and this variant was the slowest in binding and transferring the alkyl group from the ODNs. These results are consistent with an important role for the length of the arginine side chain (Arg; R) at position 128 to the binding and flipping out of the alkyl group from the active site pocket (Daniels et al., 2004). Although the functional significance of these altered characteristics in vivo remains to be determined, it is possible that they may result in altered repair of O\textsuperscript{6}-alkG adducts and hence
potentially altered cancer risk. In addition, there was evidence that R160 and F84/R160 were significantly more resistant than the Wt to inactivation by ODNs containing $O^6$-CMG and $O^6$-PobG because arginine at position 160 is relatively large and is a highly charged side chain in the active site pocket of MGMT that might discriminate against a large adduct ($O^6$-PobG) or a charged adduct ($O^6$-CMG).

**Table 6.1:** Summary of data for functional activity of MGMT variants compared with Wt protein.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Wt</th>
<th>F84</th>
<th>Q128</th>
<th>R160</th>
<th>V143/ R178</th>
<th>F84/ R160</th>
<th>F84/V143/ R178</th>
</tr>
</thead>
</table>
| **Kinetics**
| time of $[^3]$H-methyl groups transfer to protein (hour) | 2        | =   | ↑    | =    | =          | =         | =             |
| **Thermal stability**
| (half life; mins) | 51ºC     | 5 ± 0.1 | ↑    | ↑    | ↓          | ↑         | ↓             |
|                      | 44ºC     | 26 ± 0.3 | =    | ↑    | ↑          | ↑         | ↑             |
|                      | 37ºC     | 270 ± 0.4 | ↓   | ↑    | ↑          | ↑         | ↑             |
| **3. Inactivation by $O^6$-alkG-containing ODNs (IC$_{50}$; nM)** |          |      |      |      |            |           |               |
| $O^6$-BzG             | 0.09 ± 0.02 | ↑    | ↑    | ↑    | ↑          | ↑         | ↑             |
| $O^6$-MeG             | 0.93 ± 0.26 | ↑    | ↑    | ↑    | ↑          | ↑         | ↑             |
| $O^6$-CMG             | 2.41 ± 0.63 | ↑    | ↑    | ↑    | ↑          | ↑         | ↑             |
| $O^6$-PrG             | 9.14 ± 1.68 | ↑    | ↑    | ↑    | ↓          | ↓         | ↓             |
| $O^6$-PobG            | 11.60 ± 1.44 | ↑    | ↑    | ↑    | ↑          | ↑         | ↑             |
| $O^6$-EtG             | 16.02 ± 2.7  | ↑    | ↑    | ↓    | ↓          | ↓         | =             |
| **Alkyl group repair from SPR data** |          |      |      |      |            |           |               |
| $O^6$-BzG             |          | =    | ↓    | =    | =          | =         | ↑             |
| $O^6$-MeG             |          | ↓    | ↓    | ↓    | ↓          | ↓         | ↓             |
| $O^6$-CMG             |          | =    | =    | ↑    | =          | =         | =             |
| $O^6$-HOEtG           |          | ↓    | ↓    | =    | ↓          | ↓         | ↓             |
| $O^6$-PobG            |          | ↓    | ↓    | ↓    | ↓          | ↓         | ↓             |

= Equal to Wt; ↑ significantly higher than the Wt (P<0.05); ↓ significantly lower than the Wt (P<0.05); ↑ higher than Wt 10% but not significant; ↓ 10% lower than Wt but not significant

For SPR data: ↑ and ↓ indicate results qualitatively different

182
Table 6.2: Summary of data for functional activity comparison between extended and normal MGMT proteins

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Extended vs Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kinetics</strong></td>
<td></td>
</tr>
<tr>
<td>Rate of $[^1H]$-methyl group transfer to protein</td>
<td>=</td>
</tr>
<tr>
<td><strong>Thermal stability</strong> (half life; mins)</td>
<td></td>
</tr>
<tr>
<td>51°C</td>
<td>=</td>
</tr>
<tr>
<td>44°C</td>
<td>=</td>
</tr>
<tr>
<td><strong>Inhibition by calf thymus DNA (µg)</strong></td>
<td>↑</td>
</tr>
<tr>
<td><strong>Inactivation by ss $O^6$-alkG-containing ODNs (IC$_{50}$; nM)</strong></td>
<td>↑</td>
</tr>
<tr>
<td>$O^6$-BzG</td>
<td></td>
</tr>
<tr>
<td>$O^6$-MeG</td>
<td></td>
</tr>
<tr>
<td>$O^6$-CMG</td>
<td></td>
</tr>
<tr>
<td>$O^6$-PrG</td>
<td></td>
</tr>
<tr>
<td>$O^6$-PobG</td>
<td></td>
</tr>
<tr>
<td>$O^6$-EtG</td>
<td></td>
</tr>
<tr>
<td>$O^6$-HOEtG</td>
<td></td>
</tr>
<tr>
<td>$O^6$-AEG</td>
<td></td>
</tr>
<tr>
<td><strong>5. Effect of ZnCl$_2$ supplementation on MGMT inactivation</strong></td>
<td></td>
</tr>
<tr>
<td>by ds $O^6$-alkG containing ODNs  (IC$_{50}$; nM)</td>
<td>↑</td>
</tr>
<tr>
<td>$O^6$-BzG</td>
<td></td>
</tr>
<tr>
<td>$O^6$-MeG</td>
<td></td>
</tr>
<tr>
<td>$O^6$-CMG</td>
<td></td>
</tr>
<tr>
<td>$O^6$-HOEtG</td>
<td></td>
</tr>
<tr>
<td>$O^6$-PobG</td>
<td></td>
</tr>
<tr>
<td><strong>ELISA analysis of MGMT binding to ds ODNs</strong></td>
<td></td>
</tr>
<tr>
<td>ODN 13 mer</td>
<td>↓</td>
</tr>
<tr>
<td>ODN 23 mer</td>
<td>↓</td>
</tr>
<tr>
<td><strong>Effect of cytosine methylation on MGMT inactivation by $O^6$-alkG</strong></td>
<td></td>
</tr>
<tr>
<td>containing ODNs  (IC$_{50}$; nM)</td>
<td>↑</td>
</tr>
<tr>
<td>$C^{me}$G</td>
<td></td>
</tr>
<tr>
<td>(ss)</td>
<td></td>
</tr>
<tr>
<td>$C^{me}$G</td>
<td></td>
</tr>
<tr>
<td>(ss)</td>
<td></td>
</tr>
<tr>
<td>(ds)</td>
<td></td>
</tr>
<tr>
<td>(ds)</td>
<td></td>
</tr>
<tr>
<td>C (ds)</td>
<td></td>
</tr>
<tr>
<td>$^{me}$C (ds)</td>
<td></td>
</tr>
</tbody>
</table>
= Equal to Wt; ⤵ significantly higher than the extended Wt (P<0.05); ⤵ higher than extended Wt 10% but not significant; ⤵ 10% lower than extended Wt but not significant

The putative extended and normal MGMT proteins do have some different functional activities (Table 6.2). While the methyl transfer rates and thermal stabilities were similar, the normal MGMT had significantly higher IC₅₀ values than extended MGMT for inactivation by O₆-alkG containing ODNs. However, since it has not yet been established that the extended and normal proteins are co-expressed in human tissue samples, the physiological significance of these observations is not clear.

There is clear evidence that MGMT alters CRC risk (Lewin et al., 2006). Furthermore, promoter hypermethylation of MGMT, which results in the downregulation of MGMT expression, is associated with the presence of G:C to A:T transition mutations in p53 in human colorectal tumorigenesis (Esteller et al., 2001) and hence considered an important risk factor in this disease. Another risk factor is dietary red meat (Cross et al., 2010), which has been shown to increase the formation of O₆-CMG in DNA (Lewin et al., 2006). Despite these observations, a previous study had reported that O₆-CMG was not a substrate for MGMT, implying that the repair of O₆-CMG by MGMT was unlikely to play an important role in colorectal carcinogenesis. However, it is important to note that bacterial alkyltransferase was used as a proxy for MGMT (Shuker & Margison, 1997). In contrast to this report, the results in this thesis clearly demonstrate that O₆-CMG is a very good substrate for MGMT and hence MGMT might well reduce cancer risk not only from methylating, but also carboxymethylating carcinogens.

The present results show that the E.coli Ogt protein is more than 20 times more resistant than MGMT to inactivation by O₆-CMG-containing ODNs but similarly susceptible to inactivation by O₆-MeG-containing ODNs. Furthermore normal MGMT had significantly lower IC₅₀ values for inactivation by ss O₆-alkG-containing ODNs and a longer half life at 44°C than Ogt protein (Table 6.3). However, these results clearly demonstrate that O₆-CMG is almost as effective a substrate for MGMT as is O₆-MeG. Thus, decreased levels of MGMT would be expected to increase the risk of malignant transformation initiated by O₆-CMG (Van
den Bussche et al., 2012). These observations therefore support the concept that, in addition to O6-MeG, O6-CMG may be an important factor in the aetiology of human CR cancer and hence that MGMT may play a key role in protection against this disease.

**Table 6.3:** Summary of data comparing functional activity of normal MGMT and Ogt proteins

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Normal MGMT vs Ogt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kinetics</td>
<td></td>
</tr>
<tr>
<td>time of [3H]-methyl groups transfer to protein (hour)</td>
<td>↓</td>
</tr>
<tr>
<td>Thermal stability at 44ºC (half life; mins)</td>
<td>↑</td>
</tr>
<tr>
<td>Inactivation by ss O6-alkG-containing ODNs (IC50; nM)</td>
<td></td>
</tr>
<tr>
<td>O6-BzG</td>
<td></td>
</tr>
<tr>
<td>O6-MeG</td>
<td></td>
</tr>
<tr>
<td>O6-CMG</td>
<td></td>
</tr>
<tr>
<td>O6-PrG</td>
<td></td>
</tr>
<tr>
<td>O6-EtG</td>
<td></td>
</tr>
</tbody>
</table>

↑ significantly higher than the normal MGMT (P<0.05); ↓ significantly lower than the normal MGMT (P<0.05); † higher than normal MGMT 10% but not significant; ‡10% lower than normal MGMT but not significant

**6.1 Further research**

The results of these studies provide strong evidence that MGMT variants differ in their functional activity and hence they are worthy of further study. Such work could include:

1. Examining the functional consequences of different MGMT variant proteins by transfecting different variant MGMTs into an MGMT deficient cell line and investigating to what extent the resulting cell lines differ in their sensitivity to different alkylating agents.

2. Further studies of normal and putative truncated transcripts in PBMC samples. Northern analysis and quantitative real time PCR might be used to generate quantitative data, and a substantial number of lymphocyte cDNAs should be
examined. In parallel, the presence of normal and extended MGMT proteins in PBMC extracts would be assessed by western blotting and MS.

3. Based on my results in Chapter 4, Wt MGMT inactivation by ss was significantly different from ds ODNs. In further studies, all of the MGMT variants should be reassessed for IC\textsubscript{50} using both ss and ds ODNs.

4. Better identifying populations exposed to alkylating agents using MGMT as a reagent to identify $O^\beta$-alkG lesions. DNA isolated from human tissues could be treated with purified MGMT and the resulting alkylated MGMT digested with trypsin and analysed by MS.

5. A study of the associations between MGMT polymorphisms and cancer risk in populations exposed to alkylating agents. In a case-control study of lung cancer in smokers and ex-smokers, the associations between MGMT haplotypes and lung cancer risk could be determined by isolating DNA from tissues and genotyping with RFLP-PCR analysis.

6. The normal and $O^\beta$-alkG-containing ODN binding characteristics and stoichiometry of WT and variant MGMT proteins could be further investigated by electrophoretic mobility shift assays (EMSA) using non-denaturing PAGE. For visualization, the ODNs would be labelled with a fluorescent dye which would demonstrate the the effect of adding MGMT on the migration of the ODN. Band intensities would be quantified using a phosphorimager. The gels would then be subjected to electroblotting onto a PVDF membrane and then probed with anti-MGMT antibody to demonstrate the location of the MGMT in relation to the ODN.

7. While the majority of alkyltransferase structures have been determined using X-ray crystallography, Nuclear Magnetic Resonance (NMR) has also be used to study the structures of alkyltransferases in solution, for example, that of the thermophilic alkyltransferase from \textit{Methanococcus Jannaschii} (Roberts \textit{et al.}, 2006). The advantage of NMR is that, unlike in crystal structures, flexibilities in the molecule, which may have implications in its functional activity, can be observed. This may be particularly releveant to the polymorphic variants of MGMT, and give additional insight into the possible effect of the amino acid changes on the binding
to, or alkyl transfer kinetics from, ODN containing a variety of $O^6$alkG lesions. The disadvantage of NMR is that the protein or ODN needs to be labelled with an isotope that is suitable for NMR measurements, usually $^{15}$N and/or $^{13}$C.

6.2 $O^6$-alkG adducts associated with occupational nitrosamine exposure and MGMT polymorphisms in Thailand’s rubber industry

Workers in the rubber industry have a significantly increased risk of several cancers such as leukemia and cancer of the bladder, lung and larynx (International Agency for Research on Cancer, 1982) and exposure to nitrosamines including NDMA, NDEA, NDPA, NPIP, NPYR and NMOR are a suspected cause of this increased risk. The highest concentrations of nitrosamines, measured in the rubber industry, occur during the curing and vulcanisation of rubber products (Reh & Fajen, 1996; Straif et al., 2000) which uses agents such as tetramethyl thiurame disulphide, zinc-diethylthiocarbamate and morpholinomercaptobenzothiazole. These agents generate NDMA (1060 µg/m$^3$) and NMOR (4700 µg/m$^3$), which are the most prevalent volatile nitrosamines found during vulcanisation. NDMA (90 µg/m$^3$) and NMOR (380 µg/m$^3$) were found during personal monitoring (Spiegelhalder & Preussmann, 1983). Other Nitrosamines have been detected in rubber industry including: NDEA, NDPA, NPIP, and NPYR (Spiegelhalder & Preussmann, 1983).

The Occupational Safety and Health Administration (OSHA), the National Institute for Occupational Safety and Health (NIOSH) and the American Conference of Governmental Industrial Hygienists (ACGIH) classify NDMA as an occupational carcinogen and other nitrosamines as weakly carcinogenic (NIOSH, 1997), but there are no occupational exposure limits value for nitrosamines. In Germany, there is a limit for total occupational exposure to nitrosamines: in general industries and rubber vulcanization, such exposure may not exceed 1 µg/m$^3$ and 2.5 µg/m$^3$, respectively (reviewed in Oury et al., 1997)

There are several rubber factories around my workplace in Thailand and I propose to investigate whether DNA adducts are associated with occupational nitrosamine exposure and MGMT polymorphisms. This research will measure the
concentration of \( O^6 \)-alkG adducts in DNA isolated from blood samples collected at the same time as personal air samples.

As controls, unexposed workers will be included in this research. To determine the MGMT variants in these workers, extracted DNA will be amplified by PCR and genotyped.

Prior to the study, local workers would complete a consent form and also questionnaires that assess factors that might influence nitrosamine metabolism including age, height, weight, tobacco use, alcohol intake, diet and medical history.

Air samples would be collected and analyzed in a NIOSH laboratory for NDMA, NDEA, NDPA, NPIP, NPYR and NMOR. Samples will be collected by attaching a Thermosorb/N\(^{TM}\) tube to the participant’s clothes in the breathing zone of exposed and control workers. The tube will be connected to a personal air sample pump run at a rate 1 l/min for approximately 6 h. The sample tube will be desorbed with 2 ml of a mix solution of 75% dichloromethane and 25% of methanol. Different nitrosamines will be separated using gas chromatography and nitrosamines specifically detected using a thermal energy analyser.

Blood samples (40 ml) will be taken from the workers at the end of the work shift, lymphocytes will be isolated from the whole blood and DNA will be extracted from the lymphocytes. Total \( O^6 \)-alkG adducts in DNA will be quantified using an Atl1-based ELISA.

In total, 5 MGMT single nucleotide polymorphisms (SNPs) including L84F, R128Q, I143V, K178R and G160R will be determined. The extracted DNA will be amplified by PCR using primers appropriate to the region of SNPs. The PCR products will be visualized on a 1% agarose gel and the size of the DNA will be ascertained by comparison with an authentic size standard.

Urine samples (25 ml) will be collected at the end of the work shift from both exposed and control workers. Nitrosamines will determine using a gas chromatography - thermal energy analyser according to a NIOSH standard method.
Results would then establish the extent to which these workers are exposed to alkylating agents, the important factors that determine such exposure and also the association between MGMT polymorphism and $O^6$-alkG adducts in DNA.

References


the first intron and codon 178 of the DNA repair, $O^6$-alkylguanine-DNA alkyltransferase. Int.J.Cancer, **122**, 791-795.


International Agency for Research on Cancer. IARC monographs on the evaluation of the carcinogenic risk of chemicals to humans (the rubber industry). 1982. 28


Loktionova, N., Welliver, M., Crone, T., Kanugula, S. & Pegg, A. (1999) Protection of CHO cells by mutant forms of O\textsuperscript{6}-alkylguanine-DNA alkyltransferase from killing by 1,3-bis-(2-chloroethyl)-1-nitrosourea (BCNU) plus O\textsuperscript{6}-benzylation or O\textsuperscript{6}-benzyl-8-oxoguanine. Biochem Pharmacol, 58, 237-244.


Rebeck, G., Smith, C., Goad, D. & Samson, L. (1989) Characterization of the major DNA repair methyltransferase activity in unadapted *Escherichia coli* and


Tranah, G., Bugni, J., Giovannucci, E., Ma, J., Fuchs, C., Hines, L., Samson, L. & Hunter, D. (2006) $O^6$-methylguanine-DNA methyltransferase Leu84Phe and Ile143Val polymorphisms and risk of colorectal cancer in the Nurses' Health...
Study and Physicians' Health Study (United States). *Cancer Causes Control*, 17, 721-731.


Appendix 1: Thermal stability of variant MGMT proteins

Appendix Figure 1A: Thermal stability of A. Wt, B. F84, C. Q128, D. R160, E. V143/R178, F. F84/R160 and G. F84/V143/R178 proteins during pre-incubation at 44°C.
Appendix figure 1B: Thermal stability of A. Wt, B. F84, C. Q128, D. R160, E. V143/R178, F. F84/R160 and G. F84/V143/R178 proteins during pre-incubation at 37°C.
Appendix 2: Site-directed mutagenesis, variant MGMT protein expression and purification

Appendix figure 2B: Expression of the MBP-MGMT variant fusion protein in extracts of *E.coli* prepared at the times indicated after addition of IPTG. Extracts were analysed by 12% SDS-PAGE with Coomassie staining. A. Q128, B. R160 and F84/R160, C. V143/R178 and F84/V143/R178.
Appendix figure 2C: Purification of the MBP-MGMT variant fusion protein by amylose affinity chromatography. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining. A. Q128, B. R160, C. V143/R178, D. F84/R160 and E. F84/ V143/R178
Appendix figure 2D: Western blot of the purification fractions (indicated) for the MBP-MGMT variant fusion proteins. The membrane was probed with anti-MBP antibody and binding assessed by ECL. A. Q128, B. R160, C. V143/R178, D. F84/R160 and E. F84/ V143/R178.
Appendix figure 2E: Purification of the Factor Xa cleaved Q128 variant MBP-MGMT fusion protein by DEAE-sepharose. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining. A. Elution with 25-75 mM NaCl, B. Elution with 100-500 mM NaCl.
Appendix figure 2F: Purification of the Factor Xa cleaved R160 variant MBP-MGMT fusion protein by DEAE-sepharose. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining. A. Elution with 25-75 mM NaCl, B. Elution with 100-500 mM NaCl.
Appendix figure 2G: Purification of the Factor Xa cleaved V143/R178 variant MBP-MGMT fusion protein by DEAE-sepharose. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining. **A.** Elution with 25-75 mM NaCl, **B.** Elution with 100-500 mM NaCl.
Appendix figure 2H: Purification of the Factor Xa cleaved F84/ R160 variant MBP-MGMT fusion protein by DEAE-sepharose. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining. A. Elution with 25-75 mM NaCl, B. Elution with 100-500 mM NaCl.
Appendix figure 2I: Purification of the Factor Xa cleaved F84/V143/R178 variant MBP-MGMT fusion protein by DEAE-sepharose. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining. A. Elution with 25-75 mM NaCl, B. Elution with 100-500 mM NaCl.
Appendix figure 2J: Re-purification of the variant protein using an amylose column. Fractions indicated were analysed by 12% SDS-PAGE with Coomassie staining. A. Q128, B. R160, C. V143/R178, D. F84/R160 and E. F84/ V143/R178.
Appendix figure 2K: Western blot of the amylose column re-purified fractions (indicated) for the Factor Xa-cleaved, DEAE-purified MGMT variant protein. The membrane was probed with anti-MGMT antibody. A. Q128, B. R160, C. V143/R178, D. F84/R160 and E. F84/V143/R178.
Appendix 3: Inactivation of MGMT variant proteins by ODNs
Appendix Figure 3A: Effect of pre-incubation with 2.4 and 24.4 nM concentrations of twelve O6-alkG containing ODNs on the remaining activity of MGMT variants. A. Q128, B. R160, C. V143/R178, D. F84/R160 and E. F84/ V143/R178.
Appendix Figure 3B: Effect of pre-incubation with specific concentration of six $O^6$-alkG containing ODNs on the remaining activity of MGMT variants. A. Q128, B. R160, C. V143/R178, D. F84/R160 and E. F84/ V143/R178.
Appendix 4: Published paper

Nucleic Acids Research, 2013, Vol. 41, No. 5 3047–3055

The nitrosated bile acid DNA lesion $O^6$-carboxymethylguanine is a substrate for the human DNA repair protein $O^6$-methylguanine-DNA methyltransferase.

Pattama Senthong,¹ Christopher L. Millington,² Oliver J. Wilkinson,² Andrew S. Marriott,³ Amanda J. Watson,³ Onrapak Reamtong,⁴ Claire E. Eyers,⁴ David M. Williams,² Geoffrey P. Margison,³,* and Andrew C. Povey¹

¹Centre for Occupational and Environmental Health, Faculty of Medical & Human Sciences, University of Manchester, Manchester M13 9PL, UK, ²Department of Chemistry, Centre for Chemical Biology, Krebs Institute, Richard Roberts Building, University of Sheffield, Sheffield, S3 7HF, UK, ³Cancer Research-UK Carcinogenesis Group, Paterson Institute for Cancer Research, University of Manchester, Manchester M20 9BX, UK and ⁴School of Chemistry, Michael Barber Centre for Mass Spectrometry, Manchester Institute of Biotechnology, University of Manchester, Manchester M13 9PL, UK

Abstract

The consumption of red meat is a risk factor in human colorectal cancer (CRC). One hypothesis is that red meat facilitates the nitrosation of bile acid conjugates and amino acids, which rapidly convert to DNA-damaging carcinogens. Indeed, the toxic and mutagenic DNA adduct $O^6$-carboxymethylguanine ($O^6$-CMG) is frequently present in human DNA, increases in abundance in people with high levels of dietary red meat and may therefore be a causative factor in CRC. Previous reports suggested that $O^6$-CMG is not a substrate for the human version of the DNA damage reversal protein $O^6$-methylguanine-DNA methyltransferase (MGMT), which protects against the genotoxic effects of other $O^6$-alkylguanine lesions by removing alkyl groups from the $O^6$-position. We now show that synthetic oligodeoxyribonucleotides containing the known MGMT substrate $O^6$-methylguanine ($O^6$-MeG) or $O^6$-CMG effectively inactivate MGMT in vitro (IC$_{50}$ 0.93 and 1.8 nM, respectively). Inactivation involves the removal of the $O^6$-alkyl group and its transfer to the active-site cysteine residue of MGMT. $O^6$-CMG is
therefore an MGMT substrate, and hence MGMT is likely to be a protective factor in CRC under conditions where $O^6$-CMG is a potential causative agent.

Introduction

Most sporadic colorectal cancers (CRCs) arise through an adenoma–carcinoma sequence, and the molecular pathways have been well-characterized (1,2). Known risk factors for CRC include red and processed meat (3-5), but the mechanisms by which these dietary factors modify CRC risk remain to be fully elucidated (6,7). Cooked red meat, for example, may contain carcinogens such as heterocyclic amines and polycyclic aromatic hydrocarbons (3). Furthermore, faeces contain potent mutagens and other DNA-reactive agents (8,9), and studies of DNA isolated from colorectal mucosa have shown damage arising from exposure to a variety of genotoxic agents, including alkylating agents (AAs) (10-13).

AAs are a diverse set of chemicals, which include N-nitroso compounds that can be potent mutagens, clastogens and carcinogens (14,15). Human exposure to such agents is unavoidable and may arise from ingestion of pre-formed AAs or through in situ formation, typically mediated by the bacterial or chemical nitrosation of compounds containing amino groups (15). Red and processed meat intake increases N-nitrosation within the colon, generating increased levels of N-nitroso compounds and also S-nitrosated compounds (16).

The toxicological effects of AAs can be attributed largely to their ability to alkylate DNA, and the biological properties of some of the alkyl DNA adducts formed, especially $O^6$-alkylguanine lesions, are well-characterized (17). $O^6$-methylguanine ($O^6$-MeG) is a known toxic, mutagenic and carcinogenic base modification in DNA that, in the absence of repair, can induce $G\rightarrow AT$ transition mutations and recombinations in the form of sister chromatid exchanges (18). Human CR DNA is known to contain $O^6$-MeG, the levels of which vary 100-fold, with the highest occurring in the sigmoid colon and rectum, where most sporadic tumors occur (13,19). These levels are sufficiently high to cause adverse biological effects, particularly in cells that are deficient in DNA repair (20). In addition, exfoliated colon cell DNA contains $O^6$-carboxymethylguanine ($O^6$-CMG), probably
arising from \(N\)-nitrosation of the bile component glycocholic acid (21), which also generates \(O^6\)-MeG in DNA (22). That the levels of \(O^6\)-CMG in DNA from exfoliated colonic cells increase with a red meat diet strongly implies a role in CRC (21).

\(O^6\)-MeG is eliminated from DNA by the DNA repair protein \(O^6\)-methylguanine DNA methyltransferase (MGMT) in a stoichiometric process that results in the transfer of the methyl group to Cys145 in the active site of the protein (18,23,24). MGMT overexpression in cell and animal models thus protects against the mutagenic, carcinogenic and toxic effects of AAs (25). In human colorectal mucosa, MGMT activity is highly variable, in part due to exposure to AAs and dietary factors but also as a consequence of MGMT polymorphisms, which have been found to modify CRC risk, depending on the diet (26-28). Furthermore, colorectal tumors can occur in gastrointestinal regions expressing low MGMT activity, and low activity in normal colon tissue has been associated with the presence of \(K\)-ras \(GC\rightarrow AT\) transition mutations in colorectal tumors (29,30). In addition, cytosine-methylation of CpG islands within the promoter region of the MGMT gene is associated both with reduced MGMT expression and with an increased frequency of \(GG\rightarrow AT\) transition mutations in \(K\)-ras in CRCs (31-33). Indeed, adenomas containing a \(K\)-ras \(GC\rightarrow AT\) mutation have lower MGMT levels (relative to adjacent normal tissue) than adenomas without this mutation (33). As MGMT removes \(O^6\)-alkylguanine lesions from DNA, these observations strongly support the hypothesis that AAs are involved in the aetiology of at least a proportion of CRC.

Although MGMT is known to have broad substrate specificity, an earlier report using cell-free extracts from \textit{Escherichia coli} overexpressing the \textit{E. coli} \(O^6\)-alkylguanine-DNA alkyltransferase-encoding genes \textit{ada} or \textit{ogt} and a human cell line expressing endogenous MGMT suggested that alkyltransferases do not act on \(O^6\)-CMG in DNA (22). Subsequent studies suggested that \(O^6\)-CMG was processed by the nucleotide excision repair pathway (34). Given the compelling evidence for a role for MGMT in protecting against CRC, this would suggest that \(O^6\)-CMG is unlikely to be a significant factor in CRC risk, despite the observation that its abundance increases in high-risk diet situations.
To address this apparent inconsistency, we have prepared short oligodeoxyribonucleotides (ODNs) containing $O^6$-alkylguanines (35) and used these as substrates for purified *E. coli* Ogt and MGMT proteins. We found that an $O^6$-MeG-containing ODN inactivated both alkyltransferases with similar efficiency. However, although the $O^6$-CMG-containing ODN was a poor substrate for Ogt, it was an effective inactivator of MGMT. To confirm this observation, we used a combination of methods and show that, in vitro, the CM group is removed from $O^6$-CMG, that this regenerates guanine and that the CM group is transferred to the active-site cysteine residue in MGMT to generate S-carboxymethylcysteine (CM-cysteine). Thus $O^6$-CMG is a substrate for MGMT. These results are discussed in relation to the hypothesis that $O^6$-CMG is a human colorectal carcinogen and that MGMT is a major protection factor.

**Materials and Methods**

**Expression and purification of MGMT, Ogt and Atl1**

Proteins were expressed as maltose-binding protein (MBP) fusion proteins from pMAL-2c expression vector constructs essentially as described by Pearson *et al.* (36), with minor modifications as follows. Single bacterial colonies were inoculated into 100 ml of rich medium with glucose [w/v: 1% tryptone, 0.5% yeast extract, 0.5% NaCl, 0.2% glucose, containing 100 µg/ml ampicillin (Sigma)] and incubated at 37°C overnight in an orbital shaking incubator. Twenty millilitres of this culture were then used to inoculate 4 l of the same medium. The culture was grown to an optical density of 0.6 (at 600 nM) and then isopropylthiogalactoside from a 1 M stock solution was added to a final concentration of 0.4 mM, to induce protein expression. Cells were incubated for a further 3 h at 37°C and then harvested by centrifugation at 2500×g for 10 min. Cell pellets were washed with 20 mM Tris-HCl (pH 8.3) and re-suspended in 20 ml of binding buffer [BB: 20 mM Tris-HCl (pH 8.3), 200 mM NaCl, 1 mM EDTA]. Extracts were prepared by sonication (four 20 s pulses), with cooling on ice for 1 min between pulses. The extracts were then centrifuged at 17 000×g for 20 min and the supernatants pooled. The protein concentration of the extract was determined by the Bradford assay.
Amylose resin (New England Biolabs (NEB), MBP binding capacity 3 mg/ml) was pre-equilibrated with BB. The bacterial extract was diluted to 2.5 mg/ml in a total volume of 50 ml of BB, applied, washed with BB and eluted in 1-ml fractions using BB containing 10 mM maltose. The protein concentration of the eluted fractions was then determined by Bradford assay.

MBP-Atl1 fusion protein (40 mg) was cleaved with 0.1% w/v of Factor Xa (1 mg/ml, NEB) at room temperature for 2 h. The efficiency of the reaction was assessed by resolving the cleavage products on a 15% sodium dodecyl sulphate-polyacrylamide gel (SDS-PAGE). The cleavage reaction was then applied to a Superdex 200 column (10 × 300 mm; GE Heathcare) that was pre-equilibrated with 50 mM Tris-HCl (pH 8.3), 100 mM NaCl. The column was eluted at a flow rate of 0.8 ml/min, and 1.6-ml fractions were collected. Protein elution was monitored by absorption at 215 nm using a flow cell. Pooled MBP-Atl1-containing fractions were further purified through amylose columns to remove remaining uncleaved MBP fusion protein.

The other MBP fusion proteins were cleaved using factor Xa (NEB, 1 µg) with 5 µl of 10× concentrated buffer (NEB), 50 µg of MBP-MGMT and MBP-Ogt proteins and 1 mM dithiothreitol (DTT) in a total volume of 50 µl for 6 h at room temperature. The efficiency of the digest was analysed on 12% SDS–PAGE with Coomassie staining. The digestion mixtures were applied to DEAE-Sepharose (Sigma) columns (2 ml bed volume) that were pre-equilibrated with 10 mM Tris-HCl containing 25 mM NaCl (pH 8.0). The column was then washed with the same buffer for five column volumes and then eluted with five 1-ml fractions of 10 mM Tris-HCl containing 50 mM NaCl and then 75 mM NaCl. MGMT- and Ogt-containing fractions were pooled and applied to amylase columns to remove MBP and any remaining uncleaved MBP fusion protein. The protein concentration of the eluted fractions was determined by measuring the absorption at 280 nm using a Nanodrop ND-100 spectrophotometer, and those containing protein were analysed on 12% SDS–PAGE with Coomassie staining.
Determination of specific activity of MGMT and Ogt

The activities of the MGMT and Ogt proteins were determined by quantifying the amount of radioactivity transferred to protein from $[^3H]$-methylated substrate DNA (100 µl containing 100 fmole $O^6$-MeG) at 37°C for 1 h. After incubation, 100 µl of 10 mg/ml bovine serum albumin (BSA) was added, and the DNA was hydrolysed in a total volume of 2.5 ml of 1 M perchloric acid (PCA) by heating at 77°C for 50 min. Samples were centrifuged at 2800 r.p.m. for 10 min at 20°C and the supernatant aspirated. The precipitate was washed with 4 ml of 1 M PCA and then 300 µl of double-distilled water and 2 ml of scintillation fluid (Ecoscint: Mensura Technology) were added. Radioactivity was quantified by scintillation counting.

Synthesis and purification of modified ODNs

ODNs of sequence 5′-GCCATGG*CTAGT, 5′-SIMA-GCCATGG*CTAGT or 5′-GAACTG*CAGCTCCGTGCTGGCC where G* was G, $O^6$-MeG or $O^6$-CMG were synthesized using commercially available phosphoramidites and sulphone phosphoramidite (35) on an Applied Biosystems 394 automated synthesizer. For incorporation of $O^6$-CMG, base-labile phosphoramidites were used, and ODNs containing 2-amino-6-methylsulphonylpurine at the G* position were initially synthesized and then converted to $O^6$-CMG-containing ODNs using methylglycolate, as described (35,37). The ODNs were purified by reverse phase-high performance liquid chromatography (RP-HPLC) and characterized by electrospray ionisation mass spectrometry. Complementary ODNs containing 3′-biotin or 3′-HEX were synthesized by Sigma.

Inactivation of MGMT by modified ODNs

Purified MGMT or Ogt (100 fmole) were incubated with varying concentrations of ODN (0–24.4 nM) in a total volume of 200 µl of buffer I (50 mM Tris-HCl pH 8.3 containing 1 mM EDTA and 3 mM DTT) for 1 h at 37°C, followed by the addition of excess $[^3H]$-methylated substrate DNA and processing, as described above. Results were used to calculate IC$_{50}$, i.e. the amount of ODN
required to reduce MGMT activity to 50%, and were the means (±SD) of triplicate measurements.

Mass spectral analysis of alkyl group transfer to MGMT

Control (G) and O\(^6\)-alkG-containing ODNs (5'-SIMA-G CCA TGGG* CTA GTA-3', where G* is G or O\(^6\)-MeG or O\(^6\)-CMG) were annealed to the complementary oligonucleotide (5'-TAC TAG CCA TGG C-3') in a volume of 20 µl of 50 µM NaCl by heating to 80°C for 5 min and cooling to room temperature (>2 h) and then transferred to ice. ODNs (130 pmole) were incubated with MGMT (130 pmole by activity) in 28 µl of buffer I for 1 h at 37°C. Trypsin (1 µg; ratio of MGMT:trypsin, 50:1) was added and the reaction was incubated overnight at 37°C. The digestion was terminated by the addition of 100% formic acid to 0.1% (v/v). MALDI-TOF analyses were performed on a Bruker UltraflexTM (Bruker Daltonics, Bremen, Germany). Full scans of the peptide mixture from 800 to 4000 m/z and tandem mass spectral data of selected ions were collected with α-cyano-4-hydroxycinnamic acid as the matrix. External calibration was performed with QCAL standard (38).

Alkyl group removal assessed by ELISA

ODNs were annealed to 3'-biotinylated complement, as described previously, and immobilized on streptavidin-coated 96-well plates (80 fmole/well in PBS/5 mg/ml BSA). Increasing amounts of MGMT (0-210 fmole) in 100 µl buffer I were added to the wells and incubated at 37°C for 1 h. The plates were washed three times with 100 µl PBS/3 mg/ml BSA. ODNs were detected using Atl1 as a reagent or antisera to O\(^6\)-CMG. In the former case, plates were incubated with 1 pmole of Atl1 protein in 100 µl PBS/3 mg/ml BSA for 1 h at room temperature, followed by sequential 1-h room temperature incubations with anti-Atl1 antisera (1/500 in PBS/3 mg/mL BSA, the generation and characterization of which will be reported elsewhere) and then goat anti-rabbit HRP (1/1000 in PBS/3 mg/ml). In the latter case, rabbit polyclonal antisera to O\(^6\)-CMG were obtained by immunization with O\(^6\)-CMG coupled to keyhole limpet haemocyanin (Biogenes). Western Lightning
reagent (Perkin Elmer) was added and binding quantified by measurement of chemiluminescence on a Tecan GENios plate-reader. Binding was expressed as percentage of control values obtained without added MGMT. Results were the means (±SD) of triplicate measurements.

**Alkyl group removal assessed by restriction endonuclease site deprotection**

Control (G) and O<sup>6</sup>-alkG-containing ODNs were annealed to the complementary ODN (terminating with 5'-HEX or biotin as appropriate), as previously described. The resulting double-stranded ODNs (7 pmoles) were incubated with 21 pmoles MGMT in a total volume of 17 µl of IBSA (Buffer I containing 0.1% BSA) at room temperature for 3 h and then subjected to digestion with *Nla*<sub>III</sub>, *Cvi*<sub>KI</sub>, *Sty*I, *Bsa*<sub>JI</sub> or *Pst*I in the buffer provided by the manufacturer (New England Biolabs) in a total volume of 10 µl at room temperature for 1 h. Non-denaturing PAGE gel-loading buffer (30% glycerol, 0.25% xylene cyanol, 0.25% bromophenol blue: 2 µl) was added to each sample, which was then applied to a 20% PAGE gel and subjected to electrophoresis at 100 volts for 1 h. The gels were scanned on a Pharos phosphorimager (BioRad) using the fluorescent gel scanner settings.

**Results**

**Inhibition of MGMT by O<sup>6</sup>-alkG-containing ODNs**

Duplex ODNs containing *O<sup>6</sup>*-MeG and *O<sup>6</sup>*-CMG, but not the control ODN (which contained G in place of the alkylated base), potently inhibited the action of MGMT on [³H]-methylated substrate DNA. Under the assay conditions used, the IC<sub>50</sub> (50% inhibition) values for *O<sup>6</sup>*-MeG- and *O<sup>6</sup>*-CMG-containing ODNs were 0.9 and 1.7 nM, respectively (Figure 1). In contrast, the *E. coli* Ogt protein that was assessed in previous studies (22), while having a similar IC<sub>50</sub> for the *O<sup>6</sup>*-MeG-containing ODN (IC<sub>50</sub> 1.7 nM), was substantially more resistant to inhibition by the *O<sup>6</sup>*-CMG-containing ODN (IC<sub>50</sub> 38 nM: Figure 1).
Mechanism of MGMT inhibition

The inhibition of MGMT activity by $O^6$-CMG-containing ODN may have been either competitive inhibition, as a consequence of the ODN binding tightly to MGMT, or inactivation, due to alkyl group transfer from the $O^6$-CMG to MGMT. In the former case, the lesion would still be present in the ODN after incubation with MGMT. We therefore used ELISA methodology to examine the effect of increasing amounts of MGMT on the levels of $O^6$-CMG remaining in the ODN after incubation and we compared this with $O^6$-MeG-containing ODN. We exploited the ability of the *Schizosaccharomyces pombe* Atl1 protein to bind to both of these lesions and detected this binding using an anti-Avl1 antiserum. We also used an anti-$O^6$-CMG antiserum that we generated and showed to be suitable for this purpose (Supplementary Figure S1). Increasing amounts of MGMT produced an almost linear decrease in the enhanced chemiluminescence signals generated by both the Atl1-
based (Figure 2A) and anti-$O^6$-CMG antibody-based (Figure 2B) detection methods, strongly suggesting that alkyl groups are removed from the ODN.

Figure 2. Incubation of duplex ODNs containing $O^6$-MeG (squares) or $O^6$-CMG (triangles) with MGMT decreases the binding of Atl1 (A) and anti-CMG antibody (B). No binding to control (G-containing) ODN (diamonds) was seen. To prepare the substrates, ODNs (5'-GCCATGG*-CTAGTA, where G* is G, $O^6$-MeG or $O^6$-CMG were annealed to complementary ODN containing 5'-termini biotin monitored, as described in Materials and Methods.
There is ample literature to support this conclusion for DNA containing $O^6$-MeG. However, for $O^6$-CMG, about which considerably less is known, we could not exclude the possibility that strong binding of MGMT to this lesion was blocking access of Atl1 or the anti-$O^6$-CMG antibody or in situ modification of the lesion had occurred so that it was no longer detected by these reagents.

**Alkyl group removal assessed by restriction endonuclease site restoration**

We reasoned that if MGMT repairs $O^6$-CMG by alkyl group transfer to the active-site cysteine residue, as it does for $O^6$-MeG, this would result in the generation of guanine in the ODN. The ODN sequence used in these studies (5′-GCCATGG*CTAGTA, where G* is G, $O^6$-MeG or $O^6$-CMG) contained several restriction endonuclease sites (Supplementary Figure S2). Of these, we firstly demonstrated that BsaJI was able to differentiate between G and $O^6$-alkylG (Supplementary Figure S3). We then showed that pre-incubation of the ODN with MGMT almost completely restored the blocked BsaJI sites in both the $O^6$-MeG- and $O^6$-CMG-containing ODNs (Figure 3A). The same effect was seen with a different ODN sequence′ (5′-GAACTG*CAGCTCCGTGCTGGCCC) in which the $O^6$-alkylG was part of the PstI recognition sequence (Figure 3B). Thus the action of MGMT on ODN containing either $O^6$-MeG or $O^6$-CMG in two different sequence contexts results in the removal of both the methyl and the carboxymethyl group, resulting in the formation of guanine within the ODN.

**Mass spectrometric analysis of alkyl group transfer to MGMT**

To demonstrate that the action of MGMT resulted in the transfer of the alkyl group to the active-site cysteine residue in the protein, we undertook MALDI-TOF mass spectrometry. Tryptic digests of MGMT generated multiple peptides, the tryptic fragment (Tp20) containing the active cysteine (GNPVPILIPCHR) producing an [M + H]$^+$ ion with $m/z$ 1315.7 (Figure 4A), the identity of which was confirmed by tandem mass spectrometry. The expected mass changes of this tryptic fragment if an alkyl group had been transferred are 14 (methyl) and 57 (carboxymethyl i.e. CH$_2$CO$_2$) producing [M + H]$^+$ ions at $m/z$ 1329.7 and 1372.7, respectively (Figure 4B and C). The results of these analyses indicate that, as with the methyl group of $O^6$-
MeG, the carboxymethyl group of $O^6$-CMG is indeed transferred to the active-site cysteine residue in MGMT, forming $S$-carboxymethylcysteine.

**Figure 3.** Pre-incubation with MGMT of duplex ODN containing $O^6$-MeG and $O^6$-CMG enables cleavage by (A) *BsaI* (in duplexed 5'-GCCATGG*CAGCTGG*CTAGTA) and (B) *PstI* (in duplexed 5'-GAACTGG*CAGCTGGCGTGGCCC). Lanes 1-4, 5-8 and 9-12 show G-, $O^6$-MeG- and $O^6$-CMG-containing duplex ODNs, respectively; lanes 1, 5 and 9 show untreated ODN (7 pmoles); lanes 2, 6 and 10 show ODN incubated with MGMT (21 pmoles) only, and then digested with *BsaI* or *PstI* only; and lanes 4, 8 and 12 show ODN incubated with MGMT and then digested with *BsaI* or *PstI*. Non-denaturing PAGE gels were imaged as described in Materials and Methods.
Figure 4. MALDI-TOF mass spectral analysis of tryptic (Tp) peptides of unmodified MGMT (A; control) or MGMT incubated with ODNs containing $O^6$-MeG (B) or $O^6$-CMG (C). Modified peptides (Tp20* MeG and Tp20** CMG) are observed at $m/z$ 1329.7 and 1372.7 in (B) and (C), respectively. MBP described tryptic peptides derived from the MBP affinity tag on MGMT.

Discussion

There is considerable evidence to implicate $O^6$-CMG in the aetiology of human CRC and to suggest that lower levels of MGMT activity in more susceptible regions of the gut are a predisposing factor in this process. However, the limited previous data suggest that $O^6$-CMG is not a substrate for alkyltransferases (22,24), so
it is not clear why decreased MGMT activity would result in predisposition to CRC if $O^{6}$-CMG is not a substrate for MGMT: thus there appears to be no correlation between these two factors. Our recent development of a generic method to generate ODNs containing specific $O^{6}$-alkG residues (35) has allowed us to readdress this question. MGMT substrates have mostly been assessed through the inactivation or inhibition of MGMT, and to a lesser extent the re-generation of G in the substrate or the transfer and covalent attachment of the alkyl group to the cysteine residue in the active site of MGMT. Such studies have shown that MGMT has a broad substrate range, acting on $O^{6}$-alkylguanines with a variety of alkyl substituents but at different rates. For the straight-chain alkyl series, the efficiency of alkyl transfer decreases with increasing chain length (23,39,40). However, this is not simply a question of alkyl group size, because $O^{6}$-benzylG and $O^{6}$-bromothenylG are better substrates than $O^{6}$-MeG when present in short ODNs (35,41).

Previously, DNA treated with the nitrosated bile acid $N$-nitrosoglycocholic acid (NOGC) was shown to inhibit the action of the two *E. coli* alkyltransferases encoded by the *ada* and *ogt* genes, and also MGMT (22). NOGC can both methylate and carboxymethylate DNA, and to determine whether repair of both lesions was taking place, HPLC was used to analyse enzymic hydrolysates of NOGC-treated DNA after incubation with alkyltransferases. It was thus found that although $O^{6}$-MeG was a substrate for the *E. coli* ada protein, it was unable to act on DNA containing $O^{6}$-CMG. From this observation it was concluded that $O^{6}$-CMG was not recognized by any alkyltransferases (22). Later studies in human cells treated with azaserine or potassium diazoacetate showed that MGMT did not provide measurable protection against the toxicity of these agents, whereas the nucleotide excision repair pathway did, again suggesting that $O^{6}$-CMG is not a substrate for MGMT (34).

In the present report, we have used a variety of methods to analyse ODN and MGMT protein and for the first time conclusively demonstrate that $O^{6}$-CMG is in fact a substrate for MGMT. Indeed ODNs containing $O^{6}$-CMG inactivated MGMT almost as effectively as those containing $O^{6}$-MeG; both of these lesions were converted to G; and both of the alkyl groups became covalently attached to the cysteine residue in the MGMT tryptic peptide encompassing the active site. It is
reasonable to conclude that the previous report implying that O\textsuperscript{6}-CMG is not a substrate for alkyltransferases in general was an extrapolation from bacterial to mammalian MGMT that, under the conditions used, was experimentally unsupported. Indeed the present results show that the E. coli Ogt protein is >20 times more resistant than MGMT to inactivation by O\textsuperscript{6}-CMG-containing ODNs but similarly susceptible to inactivation by O\textsuperscript{6}-MeG-containing ODN and this probably explains the previous observation.

Promoter hypermethylation of MGMT, which results in the down-regulation of MGMT expression, is associated with the presence of G:C to A:T transition mutations in p53 in human colorectal tumorigenesis and hence considered an important risk factor in this disease (42). Another risk factor is dietary red meat (4,6), which has been shown to increase the formation of O\textsuperscript{6}-CMG in DNA (21). These two risk factors seem at variance with the earlier reports that O\textsuperscript{6}-CMG is not a substrate for MGMT. However, our present results clearly demonstrate that O\textsuperscript{6}-CMG is almost as effective a substrate for MGMT as is O\textsuperscript{6}-MeG. Thus, decreased levels of MGMT would be expected to increase the risk of malignant transformation initiated by the presumed promutagenic lesion (43) O\textsuperscript{6}-CMG. These observations therefore support the concept that MGMT may play a key role in protection against human CRC by repairing not only O\textsuperscript{6}-MeG but also O\textsuperscript{6}-CMG.

Supplementary data

Supplementary Data are available at NAR Online: Supplementary Figures 1-3 and Supplementary references [44, 45].

Funding

Conflict of interest statement. None declared.

References


**Supplementary material**

**Supplementary Data:**

Supplementary Information: Senthong *et al.*

**Supplementary Results**

*Detection of $O^6$-MeG and $O^6$-CMG in DNA*

We previously showed that the *S. pombe* protein, Atl1, binds to ODNs containing $O^6$-MeG (44, 45). In the present report we have demonstrated, using a rabbit polyclonal antiserum raised to purified Atl1, similar and extensive binding to ODN containing $O^6$-MeG and $O^6$-CMG (Fig. S1A), but not to ODN containing G. We also generated a rabbit polyclonal antibody to $O^6$-CMG and showed by ELISA that it detected $O^6$-CMG but not $O^6$-MeG in ODN (Fig. S1B). In both cases, the intensity of the signals were essentially directly proportional to the amounts of ODN in the wells. Eighty femoles of ODN were easily detected by these methods and this amount was used in subsequent experiments. The lower limit of quantitation was around 10 femoles.
Figure S1. Binding of Atl1 (A) and anti-CMG antibody (B) to increasing amounts of ODNs containing G (diamonds) $O^6$-CMG (triangles) or $O^6$-MeG (squares). Atl1 binding was visualised using anti-Atl1 antisera and HRP-conjugated goat-anti-rabbit secondary antibody, which was also used to detect the binding of anti-CMG antibody as described in the methods section.

**Restoration of blocked restriction endonuclease sites**

The RE sites present in the ODN used in most of these studies are shown in Figure S2.

Figure S2. Location of the restriction endonuclease sites in the ODN.
If these sites were blocked by the presence of an alkyl group at the $O^6$-position of G, the action of MGMT would restore the site and allow cleavage. We therefore used PAGE with fluorescence detection of parent ODN and ODN following incubation with NlaIII, CviKI, StyI, and BsaJI to assess the ability of these REs to digest control (G) and $O^6$-alkG-containing ODN duplexes.

A. NlaIII

B. CviKI

C. StyI

D. BsaJI

**Figure S3.** Effect of the presence of G, $O^6$-MeG and $O^6$-CMG in ODN (as indicated) on their digestability by the restriction endonucleases, NlaIII (A) CviKI (B), StyI (C) and BsaJI (D). – and + indicate without or with RE. Pictures are fluorescence images of the 5'-HEX-labelled ds ODN after non-denaturing PAGE. Lanes 1-2, 3-4 and 5-6, ds G, $O^6$-MeG and $O^6$-CMG containing ODNs respectively; Lane 1, 3 and 5, control (not incubated with restriction enzymes); Lane 2, 4 and 6 incubated with NlaIII (A), CviKI (B), StyI (C) and BsaJI (D).
NlaIII, CviKI and StyI efficiently digested control (G), O\textsuperscript{6}-CMG and O\textsuperscript{6}-MeG-containing ODNs (Figure S3A-S3C). However, cleavage of both of the O\textsuperscript{6}-alkG-containing ODN by BsaJI was substantially less that that of the G-containing ODN (Figure S3D), allowing us to use this to monitor the de-alkylation of these lesions by MGMT.

Supplementary references
