Characterizing pro-mutagenic *O*^{6}-alkylguanine adducts in human colorectal tissue

A thesis submitted to the University of Manchester for the degree of

Doctor of Philosophy

in the Faculty of Biology, Medicine and Health Sciences.

2017

Rasha Abdelhady Mohamed Abdelhady

School of Health Sciences

Division of Population Health, Health Services Research and Primary Care

Centre for Epidemiology

Centre for Occupational and Environmental Health

Word count: 56,025 words

List of Figures

List of Tables

List of Appendices

List of Abbreviations

%	Percentage
˚C	Degrees centigrade
µg	Micro gram
µL	Microliter
µM	Micro molar
AA	Alkylating agent
ACVR1B	Activin A receptor type 1B
AFAP	Attenuated familial adenomatous polyposis
AICR	American Institute for Cancer Research
AMBIC	Ammonium Bi Carbonate
AMER1	APC membrane recruitment protein 1
AOM	Azoxymethane
APC	Adenomatous polyposis coli
APS	Ammonium persulfate
ASP	Active site peptide
Atl	Alkyl transferase like proteins
Atl1	Alkyl transferase like protein 1 from '*Schizosaccaromyces pombe'*
ATNOC	Apparent total NNOC
BB	Binding buffer
BER	Base excision repair
BRAF	V-raf murine sarcoma viral oncogene homolog B
BSA	Bovine serum albumin
CDH10	Cadherin 10
Cf.	Confer
CIN	Chromosomal instability
CpG	Cytosine phosphate guanine
CT-DNA	Calf thymus DNA
CTNNB1	Catenin beta 1
CYP450	Cytochrome P450
Da	Dalton
ddH_2O	Double distilled water
dG	deoxyguanosine
DMH	1,2 dimethylhydrazine
DNA	Deoxyribonucleic acid
DOCK2	Dedicator of cytokinesis 2
DS	Double stranded
DSBR	Double strand break repair
DTT	Dithiothreitol
EB	Elution buffer
ECL	Enhanced Chemiluminescence
EDNRB	Endothelin receptor type B
EDTA	Ethylenediaminetetra acetic acid
EMSA	Electrophoretic mobility shift assay
FAP	Familial adenomatous polyposis
FAT4	FAT atypical cadherin 4
FBXW7	F-box and WD repeat domain containing 7
FeNO	nitrosyl haem
fmole	Femtomoles
g	Gram
GG-NER	global genome nucleotide excision repair
GI	Gastrointestinal

$H_2NO_2^+$	Nitrous acidium ion
HCA	Heterocyclic amines
His-MGMT	his-tagged MGMT fusion protein
HNO_2	Nitrous acid
HNPCC	Hereditary nonpolyposis colon cancer
HPLC	High pressure liquid chromatography
HR	Homologous recombination
HTH	Helix-turn-helix
IARC	International Association for Research on Cancer
IBD	Inflammatory bowel disease
IPTG	Isopropylthiogalactoside
KDA	Potassium diazoacetate
K-ras	Kirsten rat sarcoma viral oncogene homolog
LB	Loading buffer
LOD	Limit of detection
LUB	Late unbound protein
M	Molar
m/z	Mass to charge ratio
MALDI-TOF	Matrix assisted laser desorption/ionization-time of flight
MAP	MYH gene associated polyposis
MAP3K21	Mitogen-activated protein kinase kinase kinase 21
MBP	Maltose binding protein
MeIQx	2-amino-3,8-dimethylimidazo[4,5-f]quinoxaline
mg	Milli gram
MGMT	O^6-alkylguanine DNA-alkyl transferase
mL	Milli liter
MLH1	MutL homolog 1
mM	Milli molar
MMR	Mismatch repair
MMS	Methyl methanesulphonate
MNU	N-methyl-N-nitrosourea
MS	Mass spectrometry
MSH2/6	MutS homolog 2/6
MSI	Microsatellite instability
MSS	Microsatellite stable
MWM	Molecular weight marker
MYH	MutY homologue
N^1-alkylA	N^1-alkyladenine
N_2O_3	Nitrous anhydride
N^3-alkylA	N^3-alkyladenine
N^3-alkylC	N^3-alkylcytosine
N^3-alkylG	N^3-alkylguanine
N^3-alkylT	N^3-alkylthymine
N^7-alkylA	N^7-alkyladenine
N^7-alkylG	N^7-alkylguanine
NDBA	N-nitrosodibutylamine
NDEA	N-nitrosodiethylamine
NDMA	N-nitrosodimethylamine
NEM	N-ethylmaleimide
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
nm	nanometer
nM	Nanomolar
NMEA	N-nitrosomethylethylamine

NMOR	N-nitrosomorpholine
NNAL	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol
NNK	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone
NNN	N′-nitrosonornicotine
NNOC	N-nitroso compounds
NO	Nitric oxide
NO_2	Nitrites
NO_3	Nitrates
NOGC	N-nitrosoglycocholic
NPIR	N-nitrosopiperidine
NPRO	N-nitrosoproline
NPYR	N-nitrosopyrolidine
N-ras	Neuroblastoma RAS viral oncogene homolog
O^2-alkyC	O^2-alkycytosine
O^2-alkylT	O^2-alkylthymine
O^4-alkylT	O^4-alkylthymine
O^6-alkylG	O^6-alkylguanine
O^6-benzylG	O^6-benzylguanine
O^6-butylG	O^6-butylguanine
O^6-CEG	O^6-carboxyethylguanine
O^6-CMG	O^6-carboxymethylguanine
O^6-ethylG	O^6-ethylguanine
O^6-HOEtG	O^6-hydroxyethylguanine
O^6-MeG	O^6-methylguanine
O^6-pobG	O^6-[4-oxo-4-(3-pyridyl)butyl]guanine
O^6-propylG	O^6-propylguanine
ODN	Oligodeoxyribonucleotides
One-way ANOVA	One-way analysis of variance
P	Probability
PA	Peak area
PAGE	Polyacrylamide gel electrophoresis
PAH	Polycyclic aromatic hydrocarbons
PBST	0.1% Tween-20 in Phosphate Buffered Saline
PBSTM	5% Marvel in PBST
PCA	Perchloric acid
PhIP	2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine
PIK3CA	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha
PMF	Peptide Mass Fingerprint
pmole	Picomoles
PMSF	Phenylmethylsulfonyl fluoride
PTEN	Phosphatase and tensin homolog
R^2	Correlation coefficient
$R_2N–NO$	N-nitrosamines
RE	Restriction enzyme
RIA	Radioimmunoassay
RN(NO)COR	N-nitrosamides
rpm	Revolutions per minute
RT	Toom temperature
S/N	Signal to noise
SAM	*S*-adenosylmethionine
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SMAD 2,3,4	SMAD family member 2, 3, 4

SNO	nitroso thiols
SOX9	SRY-box 9
SS	Single stranded
STK11	Serine/threonine kinase 11
TAE	Tris/Acetic Acid/EDTA
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TCF7L2	Transcription factor 7 like 2
TC-NER	Transcription-coupled nucleotide excision repair
TE buffer	Tris-EDTA buffer
TEMED	Tetramethylethylenediamine
TGFBR2	Transforming growth factor, beta receptor 2
TMZ	Temozolomide
TP	Tryptic peptides
TP53	Tumour protein p53
TTN	Titin
UB	Unbound protein fraction
UHPLC-MS/MS	Ultra-high pressure liquid chromatography–tandem mass spectrometry
UV	Ultra violet
v/v	Volume per volume
vs.	*Versus*
WCRF	World Cancer Research Fund
WHO	World Health Organization
wt/wt	Weight per weight

Amino acid abbreviations

Amino acid	Abbreviation	
	Three letters	One letter
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

List of Publications

Conference abstracts

National Cancer Research Institute (NCRI) Cancer Conference, Liverpool, November 2017.
Poster presentation: Initial characterisation of the alkyl DNA adductome in human colorectal DNA by mass spectroscopic analysis of MGMT tryptic peptides.

Division of Population Health, Health Services Research and Primary Care Away Day, Manchester UK, June 2017.
Oral presentation: Mass spectroscopic analysis of MGMT tryptic peptides allows detection of O^6-alkylguanine adducts in temozolomide modified calf thymus DNA and human colorectal cancer DNA.

Faculty of Biology, Medicine and Health Doctoral Academy PhD conference, Manchester UK, May 2017.
Poster presentation: Mass spectroscopic analysis of MGMT tryptic peptides allows detection of O^6-alkylguanine adducts in oligodeoxyribonucleotides and Calf Thymus DNA treated with temozolomide.

American Association of Cancer Research Annual Meeting, Washington DC, USA, April 2017.
Poster presentation: Mass spectroscopic analysis of MGMT tryptic peptides allows detection of O^6-alkylguanine adducts in oligodeoxyribonucleotides, temozolomide modified calf thymus DNA and human colorectal cancer DNA.

Centre of Occupational and Environmental Health Lane Lecture and Symposium, Manchester UK, October 2016.
Oral presentation: Alkylating agent exposure and human colorectal cancer risk.

Centre of Occupational and Environmental Health Lane Lecture and Symposium, Manchester UK, October 2016.
Poster presentation: Mass spectroscopic analysis of MGMT tryptic peptides allows detection of O^6-alkylguanine adducts in oligodeoxyribonucleotides and Calf Thymus DNA treated with temozolomide.

United Kingdom Environmental Mutagenesis Society Conference 2016 (UKEMS) Annual Meeting, London UK, June 2016.
Poster presentation: Mass spectroscopic analysis of MGMT tryptic peptides allows detection of O^6-alkylguanine adducts in oligodeoxyribonucleotides and Calf Thymus DNA treated with temozolomide.

10th UK and Ireland Occupational and Environmental Epidemiology Conference, Buxton UK, April 2016.
Poster presentation: Mass spectroscopic analysis of MGMT tryptic peptides allows detection of O^6-alkylguanine adducts in oligodeoxyribonucleotides.

Institute of Population Health Showcase, Manchester UK, March 2016.
Poster presentation: Mass spectroscopic analysis of MGMT tryptic peptides allows detection of O^6-alkylguanine adducts in oligodeoxyribonucleotides.

Centre of Occupational and Environmental Health Lane Lecture and Symposium, Manchester UK, November 2015.
Poster presentation: Detection of O^6-alkylguanine adducts by mass spectroscopic analysis of MGMT.

Abstract

Colorectal cancer is the third most commonly diagnosed cancer worldwide. The positive association between red and/or processed meat consumption and human colorectal cancer risk is well established potentially *via* heme-catalysed formation of carcinogenic N-nitrosocompounds (NNOC). These NNOC can form a range of O^6-alkylguanine (O^6-alkylG) adducts that may increase mutational and cancer risk. The overall aim of this study was to develop a novel approach to characterize the overall load of O^6-alkylG adducts in human colorectal DNA using the known action of the DNA repair protein, O^6-alkylguanine O^6-alkyltransferase (MGMT) to irreversibly transfer the alkyl group from O^6-alkylG to the MGMT active site peptide (MGMT-ASP). Incubation of chemically synthesised single (SS) and double stranded (DS) oligodeoxyribonucleotides (ODNs) containing O^6-methylguanine (O^6-MeG), O^6-carboxyethylguanine (O^6-CEG), but only DS O^6-carboxymethylguanine (O^6-CMG) -23 mer containing ODNs with MGMT resulted in MGMT inactivation. There was a preference for repair of DS ODNs. Recorded IC_{50} values were 0.17±0.02, 0.07±0.01, 15.7±2.1, 27.00±1.70 and 7.50±0.50 nanomolar (nM) for SS O^6-MeG, DS O^6-MeG, DS O^6-CMG, SS O^6-CEG, and DS O^6-CEG, respectively. This repair was confirmed using a restriction endonuclease site deprotection assay as O^6-MeG and O^6-CEG inhibition of PstI digestion of ODNs, whilst, O^6-CMG inhibited MboI. This inhibition was removed by pre-incubation with MGMT. Alkyl group transfer from these ODNs to a Maltose binding protein-MGMT (MBP-MGMT) fusion protein was demonstrated as alkylated active site peptides (ASPs) were detected using with Matrix assisted laser desorption/ionization-time of flight (MALDI-TOF). *S*-methylcysteine (*m/z* 1329.7), and *S*-carboxyethylcysteine (*m/z* 1387.7) modified ASPs were detected following incubation of MBP-MGMT incubated with SS or DS O^6-MeG, and O^6-CEG ODNs, respectively. *S*-carboxymethylcysteine (*m/z* 1373.7) was detected with DS but not SS O^6-CMG 23 mer ODN. Only the unmodified MGMT-ASP (*m/z* 1315.7) was found after incubating MGMT with unmodified ODN. Quantitative mass spectrometry (MS) findings highlighted that MGMT preferentially repaired specified adducts when presented in DS ODNs over SS ones. Temozolomide (TMZ) methylated calf thymus DNA (CT-DNA) and human colorectal cancer DNA samples were incubated with his tagged MGMT fusion protein (his-MGMT) and his-MGMT recovered using Ni-coated beads, digested *in situ* with trypsin and analysed by MALDI-TOF. MS analysis of tryptic digests of MGMT incubated with methylated CT-DNA revealed methylated MGMT-ASP, detecting as low as 50 fmole of O^6-MeG/mg CT-DNA. This approach was applied to 13 human colorectal DNA samples, 12 of them showed evidence of O^6 alkylation damage. Of 13 samples, 12 contained O^6-MeG, 8 contained O^6-CMG and 1 contained O^6-HOEtG. A further 7 peaks were detected with a signal/noise (S/N) ratio of >10, potentially, correspond to alkylated MGMT-ASP ions but identities of the alkyl groups are as yet unknown. These results demonstrate proof of principle and confirm that human colorectal cancer DNA contains a range of O^6-alkylG adducts that constitute a fraction of DNA alkyl adductome.

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

Dedication

To my Family.

To my Mother, I would have never achieved this without your unlimited love, support and sacrifices.

Acknowledgments

I would like to thank and express my overwhelming gratitude, first and foremost, to Al-Mighty God for everything.

I would also like to express my gratitude to my supervisors, Dr. Andrew Povey and Prof. Perdita Barran, for their immense support, academic and otherwise. Your help has been great and much appreciated.

Special thanks go to Prof. Geoffrey Margison for all of his contributions to this thesis either *via* training or providing materials. Many thanks go to Dr David Williams, University of Sheffield for providing a range of ODNs that were used for this study. I would also like to thank Mr. Liam Hanson for the synthesis and purification of LH1 ODN.

I gratefully acknowledge my colleagues, in the Biomarker lab at the University of Manchester for their kind help and support as well as for making the journey of research a pleasant one.

Last but not least, I would like to thank my family, my parents, my son Ahmed, my daughter Jannah, my sister and my brother for their endless support and encouragement through the years.

Finally, I am indebted to Egyptian Ministry of Higher Education for their financial funding.

1 Chapter 1: Introduction

Colorectal cancer is a major cause of mortality and morbidity (1). Colorectal cancer, also known as bowel cancer, is a term that refers to slowly developing cancers that occur in various anatomical sites of the colon (e.g. caecum, sigmoid colon, transverse colon, ascending colon or descending colon) as well as rectum (2,3). Figure 1.1 shows the distribution of diagnosed colorectal cancer cases in UK, by anatomical sites, demonstrating the pattern of cancer distribution within the large bowel highlighting that the highest malignancy incidence rates were reported in rectum then sigmoid colon followed by caecum (4). Colorectal cancer has been classified into two major categories: sporadic cases (representing approximately 65-75% of colorectal cancer cases) and familial or hereditary cases (25-35%) (5).

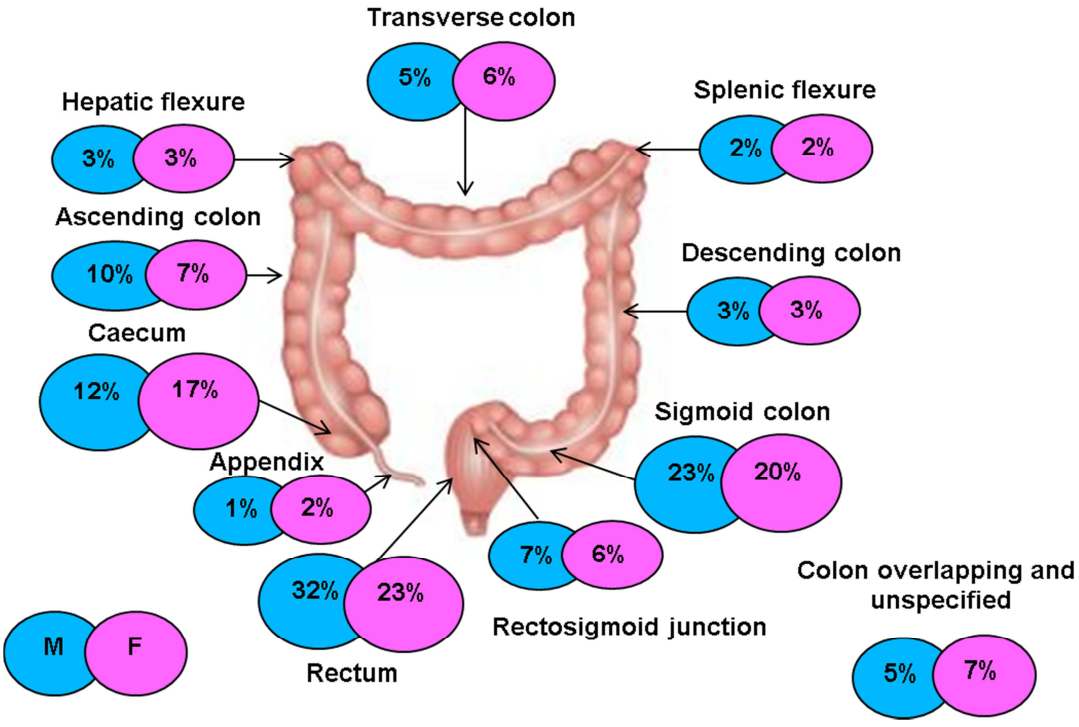


Figure 1.1: Distribution of colorectal cancer cases by anatomical sites in the UK.

Percents do not sum due to rounding to the nearest 10. Figure adapted with modifications from (6). M (blue) and F (pink) denote males and females, respectively.

1.1 Epidemiology of colorectal cancer

Globally, colorectal cancer is the third most common cancer type after lung and prostate cancer in males and second most common cancer after breast cancer in females (7). In addition, colorectal cancer is the fourth leading cause of cancer death and it accounts for approximately 8.5% of total deaths due to cancer (8,9). In 2012 in 184 countries worldwide, approximately 1.4 million people were diagnosed with colorectal cancer of which 746 000 (55%) were men and 614 000 (45%) in women. This represents 9.7% of all totally diagnosed cancer cases. Whilst, in 2012, 694 000 people died of colorectal cancer of both sexes (9).

In the UK, colorectal cancer is the third most common cancer type in both males and females as it accounts for 11% of the totally diagnosed new cancer cases in both males and females (10–13). Nearly, 41,265 new colorectal cancer cases were identified in 2014 in the UK: 22,844 (55%) in men and 18,421 (45%) in women with a male: female ratio is 1.2 (10–13). Moreover, colorectal cancer is the second most common cause of cancer death in the UK (in 2014) as it accounted for 10% of the total cancer deaths where 15 903 patients died of colorectal cancer of both sexes (14,15). In addition, colorectal cancer incidence and mortality rates in the UK were 63.9 per 100,000 and 24.6 per 100,000 persons, respectively, in 2014.

In addition, a recent study reported that, approximately, 135,000 new colorectal cancer cases are projected to occur in the United States in 2017 (58% in patients aged 65 or older), plus 50 000 patients are expected to die of the disease (16). In addition, colorectal cancer incidence and mortality rates in the United States were 40.7 per 100,000 and 14.8 per 100,000 persons, respectively, during 2009 through 2013.

Colorectal cancer incidence and mortality rates are 30% and 40% higher in men than in women, respectively (16,17). The root cause of sex disparities is not very well understood; however, it could be ascribed to sex hormones or differences in exposure to environmental carcinogens as well as the interactions between these elements (18,19).

The incidence of colorectal cancer varies globally with up to 10-fold difference in the incidence across the world (9) which reflects the impact of life style factors on the incidence of this disease. Notably, the highest colorectal cancer occurrence is reported in western or industrialized countries such as Australia, New Zealand and Western Europe where it accounts for 63% of the total global burden (20). The lowest incidence rates were reported in South-Central Asia and Africa except South Africa (21). In addition, studies on migrants from low risk to high risk countries indicated an increase in colorectal cancer incidence rate in both migrants and their off springs toward the typical rate of the host country (1). This phenomenon was demonstrated clearly in European migrants to Australia as well as Japanese migrants to Hawaii and the United States (22). However, the pattern of colorectal cancer mortality is quite different since 52% of recorded deaths were in less developing countries (9) that is mainly attributed to early detection and screening programs in developed countries that contributed significantly to lowering colorectal cancer mortality rates.

1.2 Risk factors of colorectal cancer

Colorectal cancer incidence rates are associated with both modifiable (e.g. nutritional practices and cigarette smoking) and non-modifiable factors (e.g. age and inherited genetic risk).

1.2.1 Age

Colorectal cancer risk is strongly related to age where 90% of diagnosed cases are reported in patients aged 50 or older (23). The median age of diagnosis of colon cancer is 68 in men and 72 in women; for rectal cancer the median age of diagnosis is 63 for both sexes (24). Figure 1.2 shows both average number of new cases per year as well as age-specific incidence rates per 100,000 population, in the UK (2012-2014), demonstrating that age-specific incidence rate increases sharply from age group 50-54 with the highest rates in the age group 85-89 for both sexes (25). In the United States, it was reported that incidence rate is 50-fold higher in patients aged 60-79 compared to that in patients younger than 40 years (1).

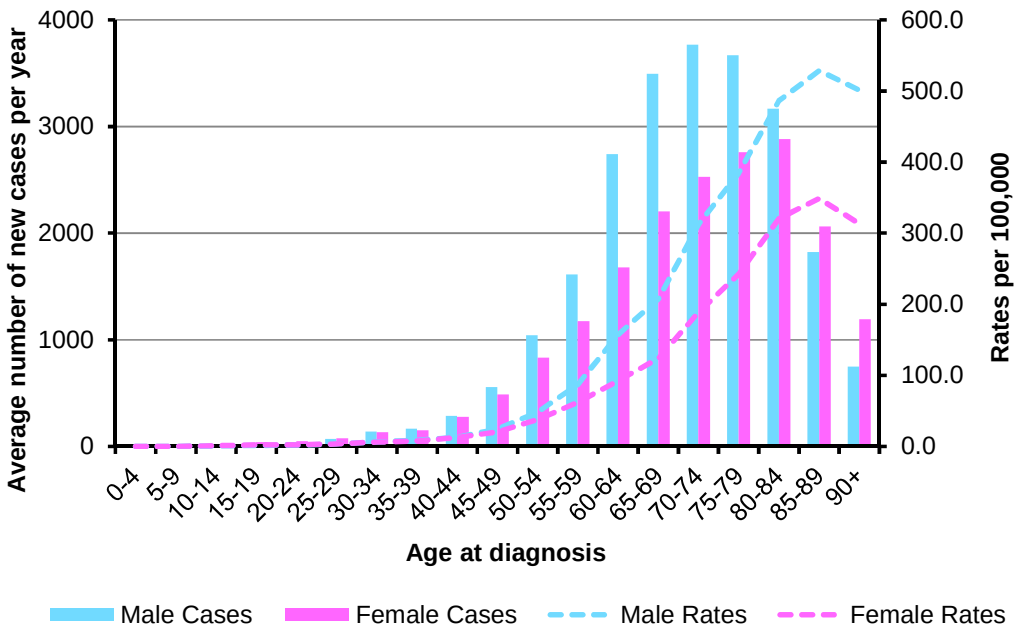


Figure 1.2: Average number of new cases per year and age-specific incidence rates per 100,000 population, in the UK (2012-2014).

Figure adapted with modification from (25).

1.2.2 Hereditary (genetic) factors

Hereditary factors are involved in the development of colorectal cancer cases, and approximately 3-5% of diagnosed cases arise in the setting of inherited syndromes (26,27) where the risk of colon cancer is very high. These include hereditary nonpolyposis colon cancer (HNPCC) or Lynch Syndrome, familial adenomatous polyposis (FAP) and several extremely rare hamartomatous polyposis syndromes. Inherited forms of colorectal cancer are usually classified into two main categories, namely, nonpolyposis syndrome characterized by the development of few or no polyps such as Lynch syndrome (28). Whilst, the second is polyposis syndrome characterized by the development of multiple colorectal polyps that could be sub-divided to adenomatous polyposis (e.g FAP), hamartomatous polyposis and hyperplastic polyposis. Colorectal cancer inherited syndromes are summarised in Table 1.1.

Table 1.1: Colorectal cancer inherited syndromes.

Syndrome	Inheritance	Involved genes	Colon presentation	Lifetime colon cancer risk	Extracolonic manifestations
Lynch syndrome	Autosomal dominant	Mismatch repair (MMR) genes MLH1, MSH2, MSH6 and PMS2	Few adenomas often present. This disorder is characterized by early onset of colorectal cancer; average age of colorectal cancer onset in Lynch syndrome is 45 years (29,30). The most common form of hereditary colorectal cancer (28). It accounts for 1-3% of all colorectal malignancies (31).	50- 80%	Endometrial and ovarian cancer
FAP	Autosomal dominant	Adenomatous polyposis coli (APC) tumour suppressor gene	Hundreds to thousands of colorectal adenomatous polyps. Average age of polyposis onset, 16 years. The second most common form of hereditary colorectal cancer; it accounts for 0.5-1% of colorectal cancer cases (31).	~100%	Duodenal cancer
Attenuated familial adenomatous polyposis (AFAP)	Autosomal dominant	APC	Variable number of colorectal adenomatous polyps (10-100) with a more proximal colonic distribution. Average age of polyposis onset is 26 (26,27).	70- 80%	Similar to FAP
MYH gene associated polyposis (MAP)	Autosomal recessive	MutY homologue (MUTYH, also called MYH)	Colonic phenotype similar to AFAP.	Undefined	Duodenal polyposis

Reproduced from (26,27).

Table 1.1: Colorectal cancer inherited syndromes (continued).

Syndrome	Inheritance	Involved genes	Colon presentation	Lifetime colon cancer risk	Extracolonic manifestations
Peutz-Jeghers syndrome	Autosomal dominant	Serine/threonine kinase 11 (STK11)	Moderate to large, but few, hamartomatous polyps (31).	39%	Breast, pancreatic, small intestine, gastric and oesophageal, carcinomas
Juvenile polyposis syndrome	Autosomal dominant	SMAD family member 4 (SMAD4)	Colorectal juvenile polyposis, often beginning in the first decade (26,27).	17–22% by age 35, ~ 68% by age 60	Gastric polyps, gastric adenocarcinoma, pancreatic and small intestine carcinoma
Hyperplastic polyposis	Unknown	Unknown	Hyperplastic polyps, sessile serrated adenomas, traditional serrated adenomas and mixed adenomas (26,27).	Undefined	None reported

1.2.3 Inflammatory bowel disease (IBD)

The risk of developing colorectal cancer is increased with both types of IBD, which are ulcerative colitis and Crohn's syndrome (32). IBD-associated colorectal cancer accounts for 1-2% of all colorectal cancer cases (33) in addition, it accounts for 10-15% of total deaths in patients with IBD (34,35). Mean age of developing IBD-associated colorectal cancer is lower than that of sporadic colorectal cancer which was reported to be approximately 43 years (36). The risk increases with the longer duration of colitis, the greater extent of inflammation, a family history of colorectal cancer and extensive mucosal involvement (37).

Moreover, the incidence of colorectal cancer in IBD patients was reported to be 60% higher than general population by an American study (38,39). However, a meta-analysis of 116 studies reported that the overall prevalence of colorectal cancer in any ulcerative colitis patient was estimated to be 3.7% (36). They also stated a non-significant association between duration of colitis and the increased colorectal cancer risk (2, 8 and 18% at 10, 20 and 30 years of the disease) (36). Notably, a more recent meta-analysis study in 2014 reported lower colorectal incidence rates than previously estimated in ulcerative colitis patients and they ascribed that to better control of inflammation, adherence to screening programs as well as higher prophylactic colectomy rates (40).

The increased colorectal cancer risk in IBD patients is thought to be due to both genetic and acquired factors. At a molecular level, the pathogenesis of colorectal cancer is related to inflammatory changes including mucosal inflammatory mediators (e.g. Cyclo-oxygenase-2, interleukin-6, interleukin-23, tumour necrosis factor-α, nuclear factor-κB, and chemokines) and oxidant stress.

1.2.4 Dietary factors

Colorectal cancer is more common in populations with a "westernized" diet. Therefore, diet is thought to be a key factor determining colorectal cancer risk, and changes in foods habits might reduce up to 70% of this cancer burden (41).

1.2.4.1 Increased red and processed meat consumption

Accumulating epidemiologic evidence over the past decades highlighted the strong association between red and/or processed meat consumption, but not white meat or fish, and colorectal cancer incidence (42–49). Both World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) and International Association for Research on Cancer (IARC) have classified processed meat as a cause of colon and rectal cancers (Group 1). While red meat have been classified as a probable cause of colorectal cancer (Group 2A) (50–52). The term 'red meat' refers to beef, lamb, pork and goat of domesticated animals; while 'processed meat' refers to meats preserved by smoking, curing, salting or by addition of chemical preservatives (e.g. salami, bacon and sausages).

Approximately, 21% of diagnosed colorectal cancer cases in UK in 2010 have been linked to eating red and/or processed meat (past exposure) (53). In addition, a recent meta-analysis of cohort studies stated that, the daily intake of 100-120 g red meat or 25-50 g of processed meat was associated with an increase in colorectal cancer risk by approximately 17-30% and 9-50%, respectively (54–56). This is supported by the findings of studies, with high number of participants (478,040 men and women from 10 European countries), reporting that daily intake of 160 g of red meat increases colorectal cancer risk by 35% compared to those who consume less than 20 g per day (41,57). With respect to processed meat consumption, the estimated increase in colorectal cancer risk reaches 49% with a consumption rate of only 25 g/day (57). Conclusively, the risk estimate is higher for processed meat compared to red meat.

For the analysis of this causal relationship, several pathways or plausible mechanisms of action have been proposed. Firstly, heme iron catalyzed endogenous formation of N-nitrosocompounds (NNOC) is a possible explanation because red and processed meat, but not white meat or fish, cause a dose-dependent increase in total NNOC (58) which is believed to be associated with subsequent increase in the formation of nitroso compound-specific DNA adducts O^6-alkylG adducts. The pathway of heme iron induced endogenous formation of nitrosating agents (59) will be explained in detail in a later section. In addition, heme iron can promote colorectal carcinogenesis through lipid peroxidation and/or cytotoxic and genotoxic effects of heme on epithelial cells (44,60,61). Apart from heme iron, nitrates and nitrites

preservatives in processed meat contribute significantly to body burden of NNOC. NNOC formation could either occur inside the meat *per se* or *via* endogenous metabolic processes in human intestine following reaction with amines or amides (42).

Another proposed explanation for the association between colorectal cancer risk and red or processed meat intake is *via* the potential mutagenic effect of heterocyclic amines (HCA) and polycyclic aromatic hydrocarbons (PAH) contained in meat. HCA are formed as pyrolysis products when meat is cooked at high temperatures or for extended times especially by grilling, frying or barbecuing (62).

Various HCA are formed according to the type of meat, the heating temperature and the chemical environment (e.g. water, oil, onion). Based upon the chemical structure, HCA could be classified into four major subgroups: amino-carbolines, imidazo-quinolines, imidazo-quinoxalines and imidazo-pyridines. The most abundant HCA in meat are 2-amino-3,8-dimethylimidazo[4,5-f]quinoxaline (MeIQx) and 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP) (Figure 1.3) (63). In general, HCA are known potent carcinogens and potent mutagens in laboratory animals and they could induce a variety of histologic types of tumours in multiple organs including colon and small intestine (64). Moreover, previous evidence demonstrated the association between colorectal cancer risk and very well cooked meat which are likely to contain high levels of HCA (65). It should be noted that the production of HCA is not specific to red and processed meat since HCA are formed in poultry as well.

In contrast with HCA, PAH are produced from the incomplete combustion of organic compounds. PAHs consist of 3 or more fused benzene rings without any acyclic groups and they form due to condensation of smaller units into stable PAHs and the condensation process mainly occurs at high temperatures such as when grilling of meat (66). PAH include benzo[a]pyrene, benzo[a]anthracene and dibenzo[a,h] anthracene which are the mostly commonly occurring PAHs in the average diet (Figure 1.3) (66). These three food PAHs are classified as probable carcinogens (Group 2A) to humans whilst other PAH have been classified as possible human carcinogens (Group 2B) (67). Conclusively, both HCA and PAH are food borne carcinogens and genotoxic agents that cause point mutation, deletions and insertions initiating colorectal carcinogenesis as shown in animal studies (68).

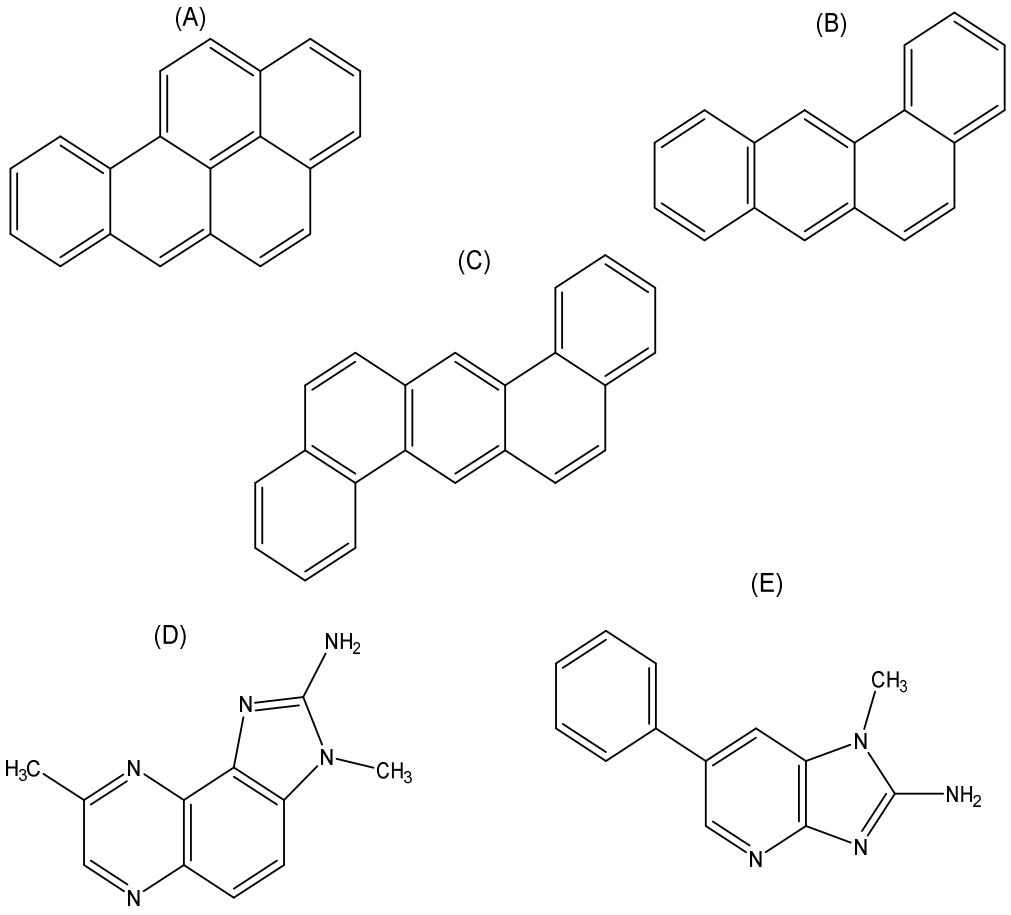


Figure 1.3: Carcinogenic PAHs (A-C) and HCAs (D-E) produced in cooked meat.

A: benzo[a]pyrene, B: benzo[a]anthracene, C: dibenzo[a,h] anthracene, D: MeIQx (2-amino-3,8-dimethylimidazo[4,5-f]quinoxaline), E: PhIP (methyl-6-phenylimidazo[4,5-b]pyridine). Figure adapted with modifications from (66).

1.2.4.2 Low dietary fibre intake

Conversely, diets high in vegetables, fresh fruit, fish and fibre intake have been found to decrease colorectal cancer risk (69). Approximately, 12% of colorectal cancer cases in UK in 2010 have been linked to too little fibre intake (70). In the early 1970s, Burkitt hypothesized that dietary fibre may protect against colorectal cancer, based on the observation that colorectal cancer was rare in rural Africans, and they ate a diet rich in fibre from unrefined grains and/or leafy vegetables (71). Since then, many case-control studies reported an inverse association between dietary fibre intake and the risk

of colorectal cancer (72). In addition, a recent international cohort study in Europe has reported an inverse association between fibre intake and colorectal cancer risk (73). Furthermore, many animal models showed inhibitory effects of various types of fibre on colon tumour development (74).

1.2.5 Alcohol consumption

Alcoholic beverages are classified by WCRF/AICR as well as IARC as a colorectal cancer carcinogen (Group1) (51). Approximately, 11% of colorectal cancer cases in the UK in 2010 were linked to alcohol consumption (53). The strong association between chronic alcohol consumption and colorectal cancer incidence is principally documented by meta-analysis studies (75,76) reporting 15% increase in colorectal cancer risk for an increase of 100 g of alcohol intake per week (77).

In addition, ethanol is classified as Group 1 carcinogen to human by IARC (78,79). Several mechanisms have been suggested to delineate alcohol consumption and cancer development that include increasing intestinal permeability enhancing environmental carcinogen uptake (80). In addition, increased formation of both reactive oxygen species and reactive nitrogen species may contribute to ethanol mediated colorectal cancer carcinogenesis *via* triggering membrane lipid peroxidation where lipid peroxidation products interact with DNA of colonic mucosal cells forming DNA adducts termed etheno DNA adducts (81). Finally, ethanol consumption enhances the carcinogenicity of some known colon carcinogenic agents, e.g. 1,1-dimethylhydrazine, that requires metabolic activation by liver metabolizing enzymes *via* induction of such enzymes (82).

1.2.6 Tobacco smoking

There is a strong association between smoking and colorectal cancer and the association appears to be dose related (83) and also evident in case of passive smoking (84) as reported by a meta-analyses and systematic reviews (85). The incidence rate increases by 17-21% in the current smokers compared to never-smokers (83). Approximately 8% of diagnosed colorectal cancer cases in UK in 2010 were linked to tobacco smoking (past exposures) (70). Consistently, tobacco smoking

has been classified as a colorectal cancer carcinogen (Group 1) by IARC (51). Furthermore, a positive correlation between colorectal cancer risk and the number of cigarettes smoked per day was reported by a previous research highlighting that smoking 10 cigarettes per day increases colorectal cancer risk by 7-11% (86). Smoking-associated colorectal cancer risk was suggested to be unmodified by common genetic polymorphisms involved in the metabolism of tobacco smoke carcinogens (87).

Notably, tobacco smoke contains more than 60 carcinogenic and genotoxic agents such as PAHs, HCAs, nitrosamines and aromatic amines (88). Mechanism by which tobacco smoking increases colorectal cancer risk is attributed to carcinogenic nitrosating AAs available in tobacco smoke that could reach colorectal mucosa either through the alimentary tract or the circulatory system (89).

1.2.7 Physical activity and obesity

Approximately, 13% and 3% of diagnosed colorectal cancer cases in UK in 2010 were attributable to obesity and physical inactivity, respectively (53). Furthermore, body fatness is classified as a cause of colon cancer by IARC and WCRF/AICR (90). Epidemiological studies suggest that obesity is associated with 30-70% increased colorectal cancer risk in men whilst the evidence is less consistent in women (91). The underlying mechanisms linking obesity to colorectal cancer are areas of much ongoing research. However, it could be ascribed to increased insulin resistance and decreased gut motility. On the other hand, physical activity is proposed to be inversely correlated to colon cancer risk in both sexes by a meta-analysis study with no effect on rectal cancer in either sex (92). In contrast to obesity, sustained moderate physical activity seems to increase metabolic rate and oxygen uptake by body cells. In addition, it contributes to improving gut motility as well as decreasing insulin resistance and blood pressure (93). However, biologic mechanisms correlating physical inactivity to colorectal cancer risk are still a matter of debate.

Other factors contribute to colorectal cancer risk including exposure to ionizing radiation (2%) e.g. X-radiation and gamma radiation as well as all forms of asbestos.

1.3 Molecular basis of colorectal cancer

Several molecular events contribute to colorectal cancer development and progression. Among these different events or pathways, the most important are the chromosomal instability (CIN) pathway and microsatellite instability (MSI) pathway and cytosine phosphate guanine (CpG) island methylation pathway (94).

1.3.1 Chromosomal instability (CIN)

CIN is the most common type of genomic instability driving colorectal cancer development. Nearly, 85% of sporadic colorectal cancer cases develop through the CIN pathway (95–97). CIN induces significant changes in both chromosomal copy number and chromosomal structure (98). CIN is considered an effective mechanism for mutational inactivation of many tumour suppressor genes whose normal function is to oppose the malignant phenotype such as APC , Tumour protein p53 (TP53) and SMAD4 (99).

The underlying mechanisms for CIN include rare inactivating mutations of genes responsible for maintaining chromosomal stability during cell replication. In addition, CIN may be driven by the occurrence of chromosome segregation abnormalities at mitosis. These abnormalities result from either alteration in kinase activities, excessive activation of telomerase enzyme, or aberration in checkpoints (100).

1.3.2 Microsatellite instability (MSI)

The MSI pathway is observed in approximately 15% of sporadic colorectal cancers and about 90% of Lynch syndrome cases (101). The MSI pathway mainly results from mutational inactivation of the genes responsible for DNA MMR system, namely, MLH1, MSH2 and MSH6, (102). The protein products of those genes have a variety of functions in the recognition and repair of DNA mismatches (103). The inactivation of MMR genes could be either inherited as in Lynch syndrome or acquired

via somatic silencing of the promoter region of MLH1 gene *via* enzymatic methylation leading to inactivation of MMR pathway (104–106).

Moreover, MSI could lead to inactivation of tumour suppressor genes whose functional regions contain mononucleotide or dinucleotide repeat sequences such as those encoding transforming growth factor β (TGF-β) receptor type II (TGFBR2) and BCL2-associated X protein (BAX) (107). MSI is characterized by a massive accumulation of frame-shift mutations in micro-satellite sequences. Microsatellites are repeated DNA sequences which occur approximately every 50–100 kb pairs throughout the human genome (108).

1.3.3 CpG island methylation pathway

Specifically, CpG island methylation pathway is an epigenetic abnormality that could contribute to 15-20% of colorectal malignancy (109). A CpG Island is a DNA sequence greater than 200–500 bases in length with more than 50% GC content and it overlaps the promoter region of 60–70% of all genes. Under normal physiological conditions, CpG islands promote gene transcription and tend to be protected from enzymatic methylation by DNA methylases; however, they can become aberrantly methylated in cancer by DNA methylase enzymes. When this island is enzymatically methylated, gene silencing and inhibition of gene transcription occur with subsequent silencing of downstream transcriptional units leading to abnormal cellular function which can potentially result in carcinogenesis (110). CpG islands are found in the promoter regions of tumour suppressor genes such as: MLH1 gene, APC and the DNA repair protein; MGMT gene. Where methylation of the CpG island of these genes results in inactivation of such genes and loss of their carcinogenesis-opposing function leading to colorectal cancer development (111).

1.4 Genetic aberrations in colorectal cancer

Colorectal cancer is driven by multiple and sequential genetic mutations that could lead to either activation of oncogenes or deactivation of tumour suppressor genes. A previous study identified a median of 76 and 9 genes altered by point mutation and copy number change in colorectal cancer, respectively (112).

Among those, the most important genes are tumour suppressor genes such as: APC and TP53 genes (113). In addition oncogenes such as K-ras and BRAF oncogenes, that have a high frequency of genetic aberrations in colorectal cancer, are known to play an essential role in colorectal cancer development (114). Table 1.2 summarises the most frequently mutated genes in colorectal cancer. In addition, recent publications aimed at defining genomic landscape of human colorectal cancer confirmed APC, TP53 and K-ras mutations as the most predominant genetic defects in colorectal cancer alongside the identification of other colorectal cancer associated genes, as shown in Figure 1.4 (115–117). Other frequently mutated genes in human colorectal cancer genomes include TTN, PIK3CA, FBXW7, SMAD4, TCF7L2, N-ras, AMER1, SMAD2, CTNNB1 and SOX9. K-ras and N-ras genes had oncogenic codon 12 and 13 or codon 61 mutations, whereas the remaining genes had inactivating mutations (115–117). Notably, a study in 2014 identified novel colorectal cancer associated genes (FAT4, CDH10 and DOCK2) (115). FAT4, CDH10 are negative mediators of Wnt/βcatenin pathway whilst DOCK2 is an oncogene.

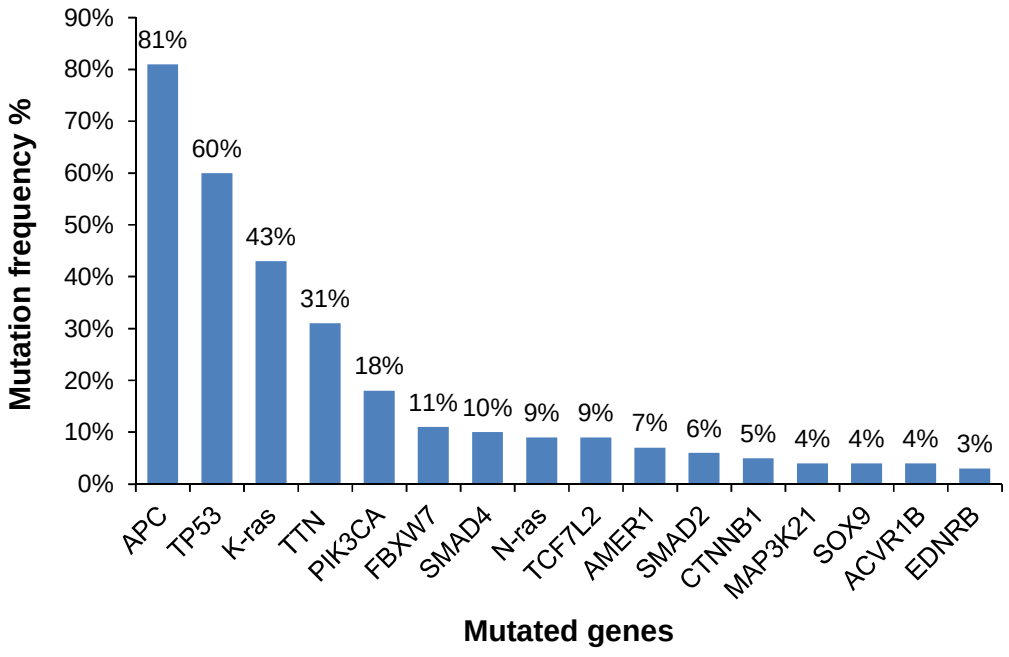


Figure 1.4: Significantly mutated genes in colorectal cancer genomes.
Figure adapted with modifications from (117).

Table 1.2: Selection of some of the most frequently mutated genes in colorectal cancer.

Affected Gene	Frequency	Normal function	Nature of Defect	Comments
APC	85%	Tumour suppressor gene encoding APC protein, a component of β-catenin degradation complex (118). APC protein regulates cell proliferation *via* binding to and degrading β-catenin oncoprotein that promote cellular proliferation (119).	Both germ-line and somatic mutations result in inactivation of APC gene with subsequent loss of the β-catenin oncoprotein degradation leading to inappropriate activation of the Wnt signaling pathway (120).	- The activation of the Wnt signaling pathway is regarded as the initiating event in colorectal cancer (121). - Germ-line mutation in FAP (31). - GC:AT transition mutations were reported in APC gene in colorectal cancer cases. - APC transition mutations were associated with processed meat consumption.
TP53	35-55%	Tumour suppressor gene (122); encoding wild-type p53 protein that mediates cell-cycle arrest and cell-death as well as enhances cellular autophagy (123).	Inactivating missense mutations and 17p chromosomal deletion contribute to cellular malignant transformation due to loss of the wild-type p53 protein (124) Point mutations occurring at codons 175, 245, 248, 273 and 282.	- Mutational inactivation of TP53 is the second key genetic step in colorectal cancer and coincides with the transition of large adenomas into invasive carcinomas (107). - GC:AT transition mutations have been reported in TP53 gene in colorectal cancer cases (125).
K-ras	35-45%	Oncogene, is a member of the RAS gene family, encoding K-ras G-protein implicated in G-protein signaling pathways that modifies or adjusts differentiation and/or proliferation of normal cells (126).	Somatic activating mutations causing activation of RAS/RAF/MAPK signaling pathways, thereby promoting cell survival and apoptosis suppression (127). Point mutations at codons 12 (84-87%), 13 (13-16%) and 61 (much lesser frequency) (128).	These point mutations are mostly G to A transitions or G to T transversions (128). Poehlmann *et al.* (129) documented that the frequency of G to A transitions in colorectal cancer was 63%, whilst the frequency of G to T transversions was 29%.

Table 1.2: Selection of some of the most frequently mutated genes in colorectal cancer (continued).

Affected Gene	Frequency	Normal function	Nature of Defect	Comments
TGFBR2	25–30%	Tumour suppressor gene encoding TGF-β2 peptide mediating cell growth arrest and apoptosis (113).	Frameshift inactivating mutations causes inactivation of TGF-β pathway (130).	Mutation present in >90% of tumours with MSI and 15% of microsatellite stable colon cancers (130).
SMAD4, SMAD2, SMAD3	10–35%	Tumour suppressor genes; encoding SMAD proteins, components and transcriptional mediators of TGF-β signaling pathway (113).	SMAD4 is located on chromosome 18q and inactivated by 18q chromosomal deletion or mutation (131).	Germ-line mutations in SMAD4 gene are reported in familial juvenile polyposis, (132).
MLH1, MSH2, MSH6	15–25%	DNA MMR genes (113).	Defects in DNA MMR pathway permitting the accumulation of oncogenic mutations and tumour-suppressor loss (133).	Inherited mutation in Lynch syndrome or epigenetic silencing of MLH1 gene (134).
BRAF	8–12%	Oncogene and a member of the RAF gene family that encodes a serine-threonine protein kinase, a downstream effector of activated RAS (113).	Activating mutation in the BRAF mimics the biologic consequences of K-ras mutation (135).	Associated with hyperplastic polyposis, with increased incidence in serrated adenomas (135).
PTEN	10-15%	Tumour suppressor gene encoding a protein which negatively regulates PI3K pathway (113).	Inactivating mutations results in cell survival signaling and apoptosis suppression (113).	

APC: Adenomatous Polyposis Coli, TP53: Tumour protein p53, K-ras: Kirsten rat sarcoma viral oncogene homolog, TGFBR2: Transforming growth factor, beta receptor 2, SMAD 2, 3, 4: SMAD family member 2, 3, 4, MLH1: mutL homolog 1, MSH2/6: MutS protein homolog 2/6, BRAF: v-raf murine sarcoma viral oncogene homolog B, PTEN: Phosphatase and tensin homolog.

1.5 Development of colorectal cancer

Approximately, 95% of colorectal cancers develop through a stepwise genetic model known as adenoma-carcinoma sequence as illustrated in Figure 1.5, presented by Vogelstein *et al* (136). In this model colorectal cells are transformed from normal to dysplastic epithelium then to malignant cells involving a series of histological, morphological, and genetic aberrations that accumulate over a period of time (127).

The sequence usually starts with the formation of benign growths termed polyps (137), which are benign and localized (and cannot penetrate the underlying basement membrane) lesions that result from abnormal proliferation of cells caused by the accumulation of both genetic and epigenetic mutations (113,138). Therefore, polyps of several histologic subtypes are considered as pre-neoplastic lesions for colorectal carcinoma. More than 90% of colorectal carcinomas are adenocarcinomas originating from epithelial cells of the colorectal mucosa (139). Other rare types of colorectal carcinomas include neuroendocrine, squamous cell, adenosquamous, spindle cell and undifferentiated carcinomas. Polyps are classified as either hyperplastic or adenomatous polyps, otherwise known as adenomas. The term ‘hyperplastic polyps’ refers to polyps that remain benign with a low likelihood of becoming cancerous. However adenomas have a higher chance of becoming malignant (135).

Invasive cancers are classified into stage I, II, III and IV cancers. Stage I and II cancers are characterized by being confined within the wall of the colon and rectum, and are curable by surgical removal of the tumour. When the disease has spread to regional lymph nodes it is classified as a stage III tumour which is curable by surgery combined with adjuvant chemotherapy with ~ 73% survival rate. However at stage IV, the malignant cells have metastasized to distant sites and other body organs and the disease is usually incurable (140). Figure 1.5 shows a classical progression model of colorectal carcinogenesis is illustrated with genes whose alterations are responsible for each of the progressive steps (141).

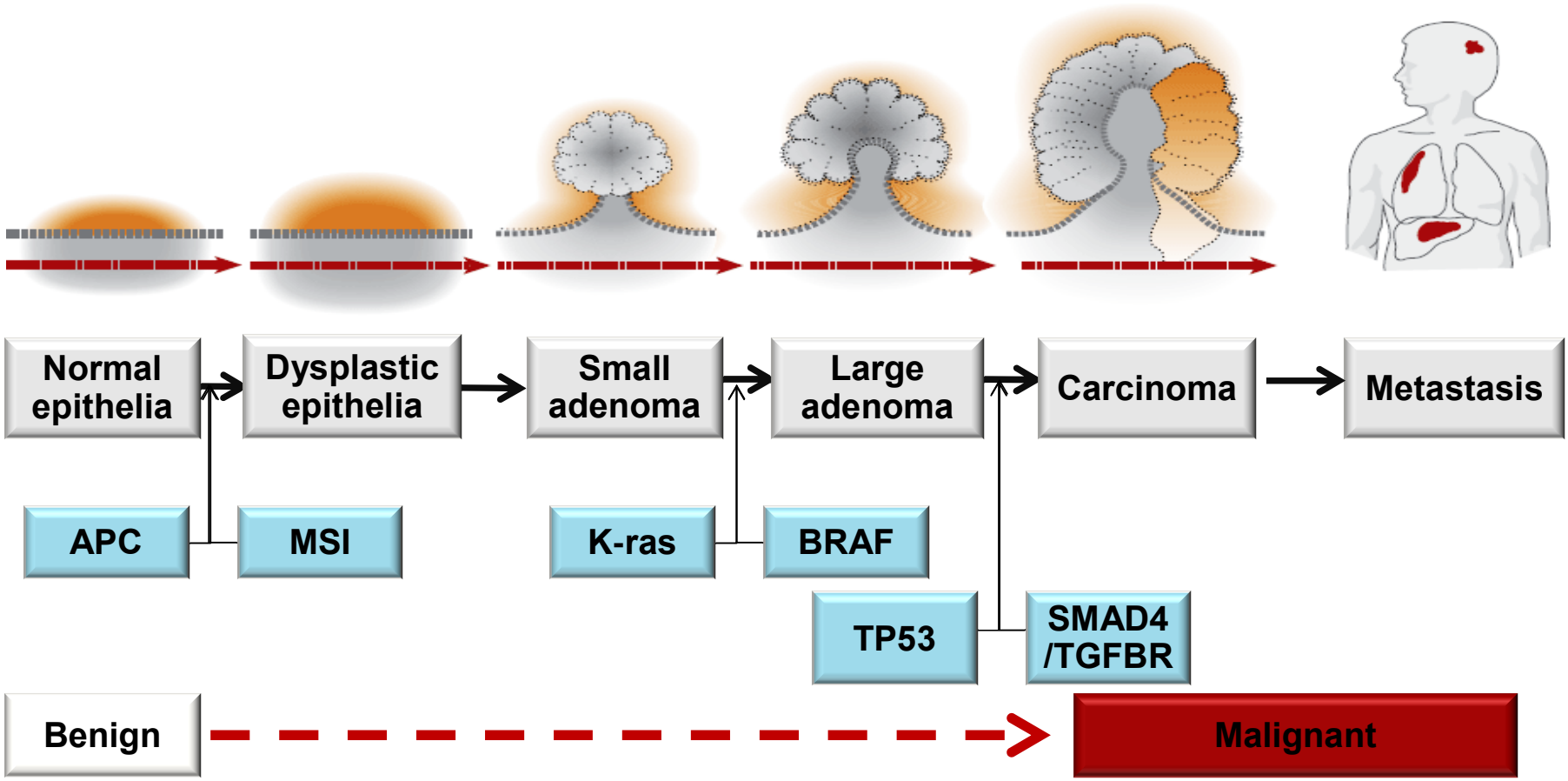


Figure 1.5: A classical progression model of colorectal carcinogenesis.

Figure adapted with modifications from (113).

1.6 Mutational landscape of colorectal cancer

Numerous studies have attempted to investigate colorectal cancer mutational landscape, among those, a study in 2007 reported that, 9.3% of the examined genes were mutated plus 93% of the detected mutations were single base substitutions, predominantly, C to T transitions at 5′-CpG-3′, whilst, the remaining 7% of the alterations was insertions or deletions (116). Notably, these results were supported by the findings of a study in 2014 confirmed that C:G→T:A transitions were the predominant nucleotide change in colorectal cancer (115). In addition, previous attempts to unravel mutational signatures of colorectal cancers indicated that 15% of colorectal cancer cases were associated with DNA MMR failure (142) known as microsatellite unstable cancers characterized by base substitutions, predominantly, C to T transitions and C to A transversions (143).

A study in 2013, investigated the landscape of mutational signatures in 30 types of human cancer including colorectal cancer. Mutational signatures were extracted using 2 parameters: firstly base substitutions, secondly, local DNA base sequence context of each mutation. Since there are six classes of base substitution C→A, C→G, C→T, T→A, T→C, T→G and since they incorporated information on the bases immediately 5′ and 3′ to each mutated base, there are 96 possible mutations in this classification. This 96 substitution classification is particularly useful for distinguishing mutational signatures which cause the same substitutions but in different sequence contexts as reported by Alexandrov *et al*. (144)

Three mutational signatures were identified in colorectal cancer out of 21 distinct signatures, namely, Signature1B, 6 and 10; and the percentage contribution of each signature is shown in Figure 1.6. Notably Signature 1B was characterized by the predomination of C→T base substitution (Figure 1.7A). Similar to Signature 1B, Signature 6 is characterized by higher prevalence of C→T mutations, however, it's distinct from mutational signature 1B in that it contains small indels (1 bp); this pattern of indels is termed as MSI which characterizes the cancers with defect in MMR pathway such as colorectal cancer (Figure 1.7B). With regard to Signature 10, it's characterized by the prominence of two substitution mutations namely, C→A and C→T where both exceeded 20% (Figure 1.7C)(144).

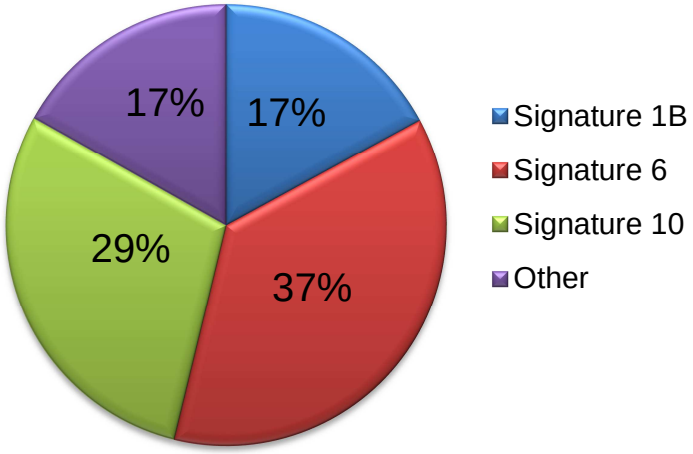


Figure 1.6: Contribution of the signature of the mutational processes operative in colorectal cancer.

Signature 1B and Signature 6 were characterized by the predomination of C→T base substitution, whilst signature 10 was characterized by the predomination of C→A and C→T substitutions. Produced based on data from (144).

Additionally, a publication in 2015 identified two mutational signatures in colorectal cancer genomes which were signature 1 (predominated by C>T transitions at CpG nucleotide) and signature 5 illustrated in Figure 1.8 (145). A recent publication in 2017 identified three distinct mutational signatures in colorectal cancer genome (146). The first two were signature 1 and 5 that were closely related to those previously described by the previous study in 2015. In addition, signature 18 was identified in colorectal cancer genome which was characterized by C>A transversions and it was associated with MUTHY gene mutation highlighting the potential role of the base excision repair (BER) system instability in the colorectal cancer tumourigenesis (Figure 1.9).

Previous data highlighted that C>T transition mutation is the most predominant nucleotide change in colorectal cancer mutated genes. This genetic aberration is potentially attributed to a variety of factors including deamination of the enzymatically methylated 5-methylcytosine at CpG islands or O^6-MeG resulting from human exposure to environmental alkylating agents (115,147).

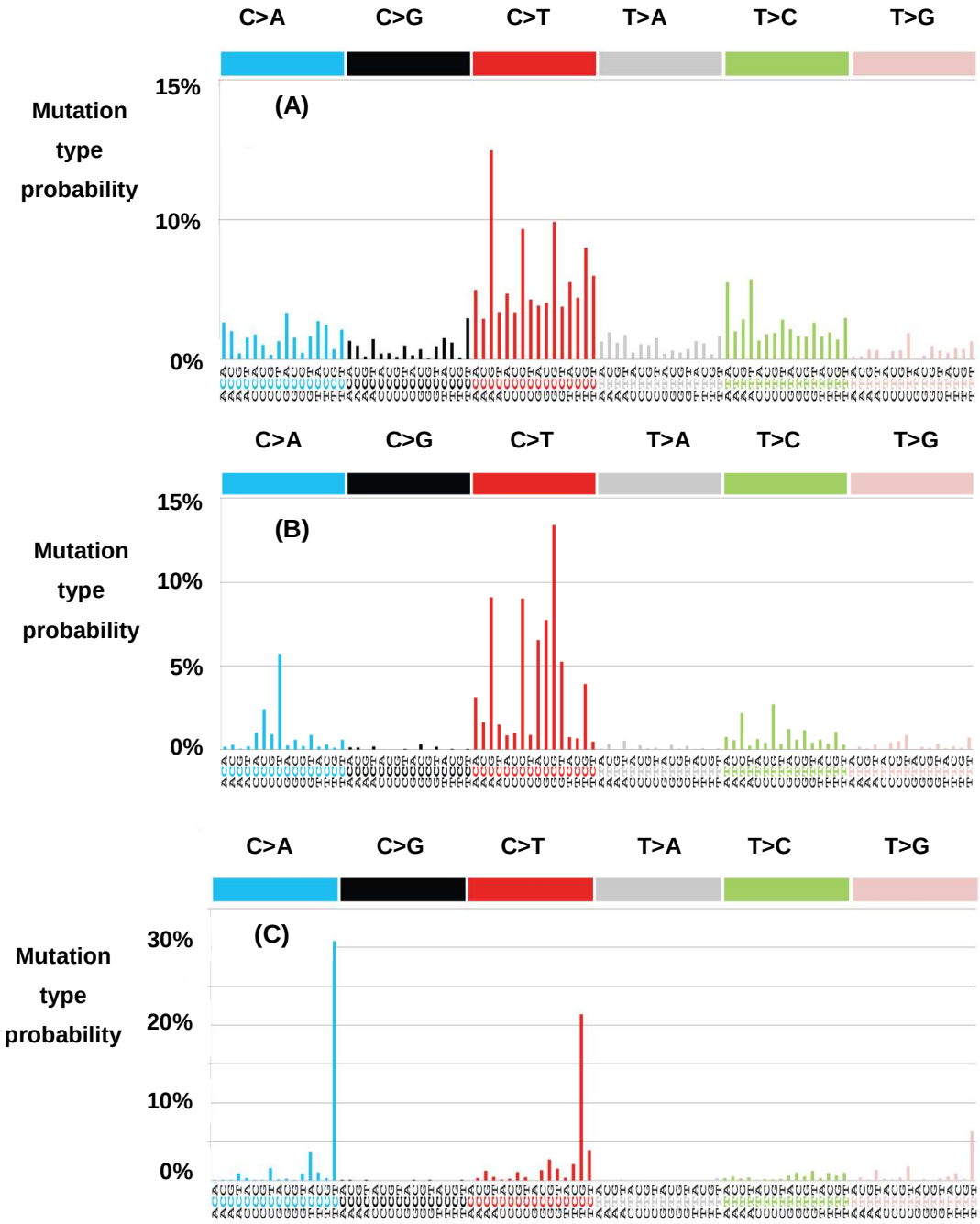


Figure 1.7: Mutational signatures found in human colorectal cancer (A) Signature 1B, (B) Signature 6 and (C) Signature 10.

Figure adapted with modifications from (144).

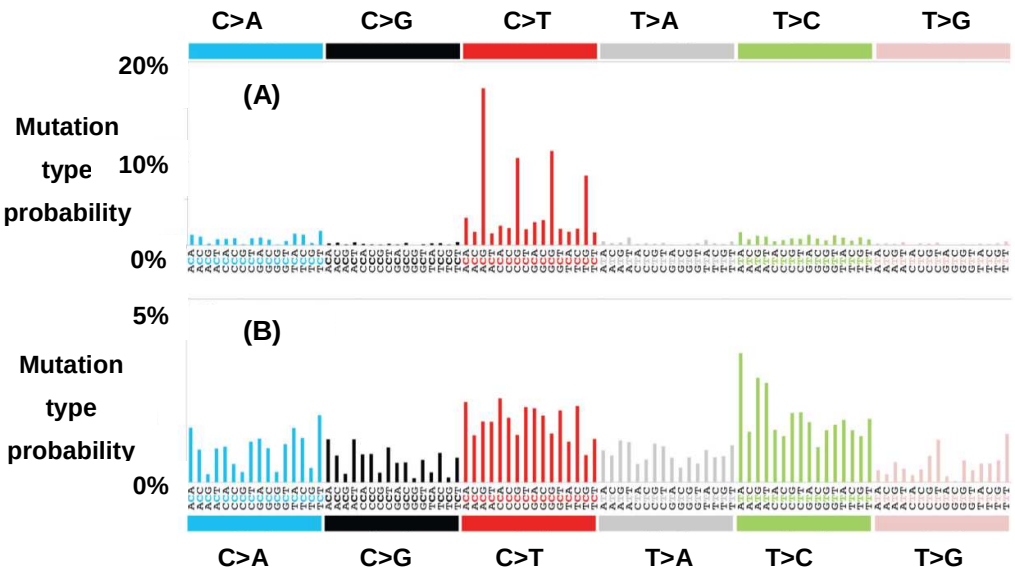


Figure 1.8: Patterns of mutational signatures detected in 559 colorectal cancer (A) Signature 1 and (B) signature 5.

Figure adapted from (145).

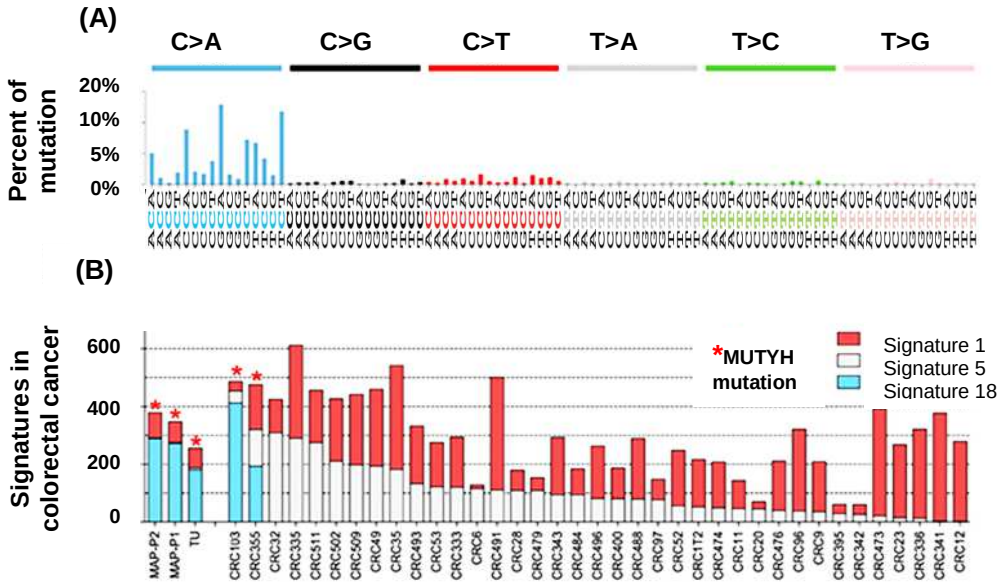


Figure 1.9: Mutational signatures identified in a cohort of 40 colorectal cancers.

(A) Mutational signature 18 is depicted using a 96-substitution classification defined by the substitution type and sequence context immediately 5' and 3' to the mutated base. (B) Contribution of the three signatures identified to each of examined tumours. Figure adapted from (146).

1.7 DNA Alkylating agents (AAs)

AAs are considered pro-mutagenic and genotoxic agents (148,149). In addition, AAs are known as potent human and animal carcinogens. Some AAs are intrinsically reactive however; most of them require enzyme-mediated bioactivation for the generation of biologically active intermediates (150). Generally, AAs are defined as chemical compounds characterized by their high ability to transfer an alkyl group to nucleophilic sites of cellular macromolecules either proteins, RNA or DNA molecules through an electrophilic attack (151).

DNA alkylators could act in either a mono- or bifunctional manner. Monofunctional alkylators transfer only one alkyl group to DNA with the subsequent formation of covalently modified DNA bases known as alkyl DNA adducts (152). On the other hand, bifunctional alkylators can react at two different sites in the DNA leading to the formation of either an interstrand or intrastrand DNA crosslink (153). The formation of DNA adducts cause disastrous consequences *via* structural damage of DNA that affects crucial biological processes contributing to cellular toxicity and mutation (150).

AAs are a large group of different chemical classes with a diverse portfolio of chemical structures (alkylhalides, tobacco specific N-nitosamines or NNOC). They can react with DNA *via* either a S_N1 mechanism that follows 1^{st} order kinetics where the formation of a carbocation intermediate is the rate limiting step (Figure 1.10A) or a S_N2 mechanism of action which is characterised by the formation of a transition complex, between the nucleophile and carbon centre, as well as following 2^{nd} order kinetics (Figure 1.10B) (151). S_N1 agents include N-methyl-N-nitrosourea (MNU), 1,2 dimethylhydrazine (DMH) and TMZ as shown in Figure (1.10C-E), whilst S_N2 agents include *S*-adenosylmethionine (SAM) (Figure 1.10F).

An important class of AAs is NNOC, including N-nitrosamines, N-nitrosamides and nitrosoguanidines. They are characterised by a nitroso group (-N=O) attached to a nitrogen atom (-N-N=O). Human exposure to NNOC is *via* either endogenous formation inside human body or in diet upon ingestion of a range of food products and beverages such as processed meat, smoked fish, beer, cheese and pickles (154,155). NNOC are known to be potent carcinogens and the alkylation of human DNA is likely to be the major mechanism of action (156,157).

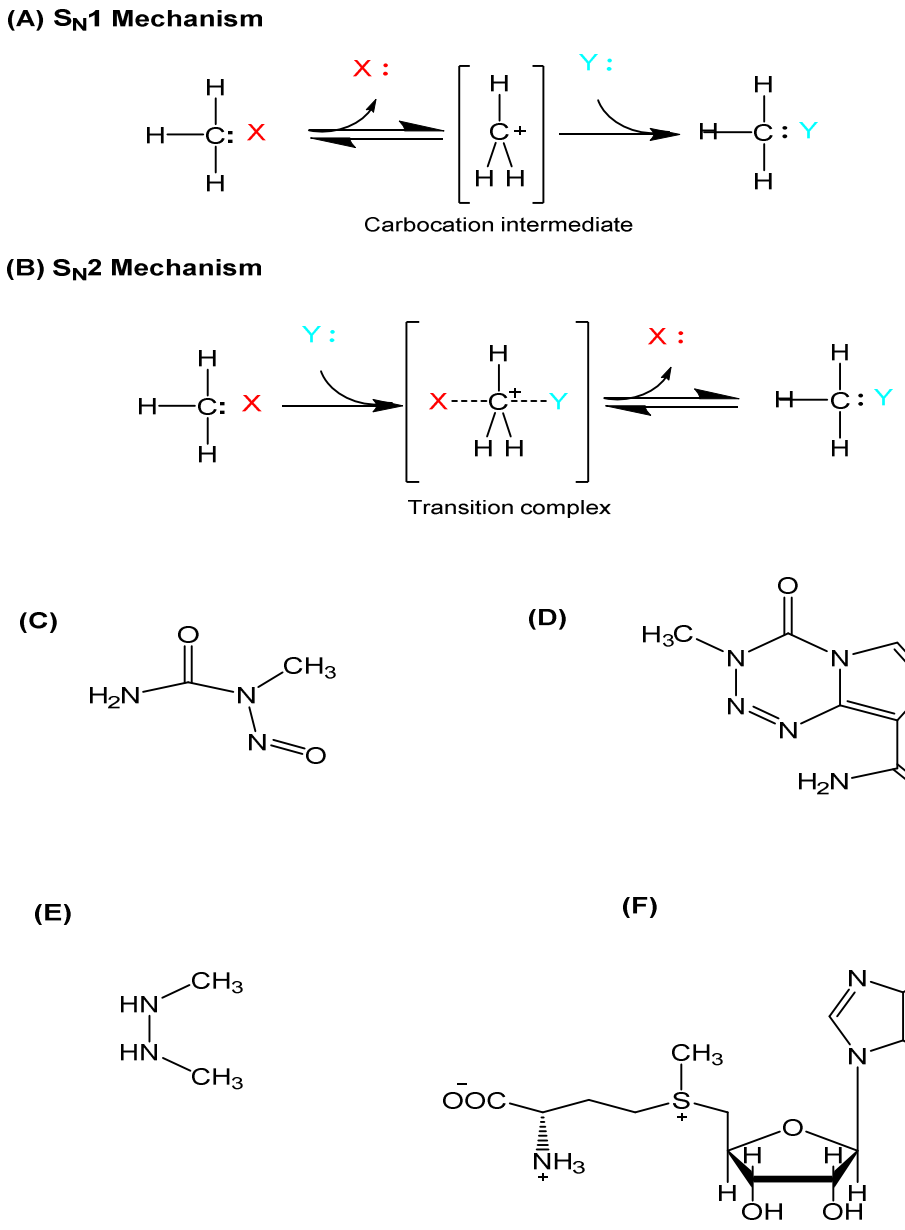


Figure 1.10: Alkylation mechanisms and selected AAs structures.

(A) S_N1 alkylation mechanism, (B) S_N2 alkylation mechanism, (C) N-methyl-N-nitrosourea, (D) Temozolomide, (E) 1,2-dimethylhydrazine, (F) S-adenosyl methionine. X (red) and Y (blue) denote leaving groups and nucleophiles, respectively, for S_N1 (A) and S_N2 (B) alkylation mechanisms.

1.8 Human exposure to AAs

AAs are ubiquitous and they arise from exogenous sources or they could be generated within the body during metabolism. Humans are exposed to AAs from occupational, environmental, lifestyle or dietary sources. In addition, AAs are used as cytostatic drugs in cancer chemotherapy in significantly high doses, such as TMZ (158).

1.8.1 Exogenous exposure to AAs

Total human exogenous exposure to N-nitrosamines is estimated to be 1.10 µmol/day; the major exposure sources are the diet (72%), occupational exposure (25%), cigarette smoking (2%), and miscellaneous minor sources, including pharmaceutical products, cosmetics, indoor and outdoor air (1%) as reported by (159).

1.8.1.1 Dietary ingestion of NNOC

Diet is the major source of exogenous exposure to NNOC for most individuals as they could be formed in food *per se* prior to food consumption (159). Preformed dietary NNOC are detected in nitrite/nitrate preserved or food that was exposed to nitric oxide (NO) during food processing. Most notably processed meat regarded as the principal source of dietary intake of NNOC as nitrosation of amines or amides in food is most likely to occur during the curing process of meat by nitrosating agents (160). Smoked fish, cheese and beer are potential sources of dietary NNOC. A range of N-nitrosamines are potentially detected in food such as N-nitrosodimethylamine (NDMA) and N-nitrosodiethylamine (NDEA), N-nitrosopyrolidine (NPYR), N-nitrosopiperidine (NPIR), N-nitrosodibutylamine (NDBA), N-nitrosomorpholine (NMOR), N-Nitrosomethylethylamine (NMEA) and N-nitrosoproline (NPRO), although the presence of the chemically more unstable nitrosamides is less likely (154,161). The chemical structures of these dietary NNOC are shown in Figure 1.11.

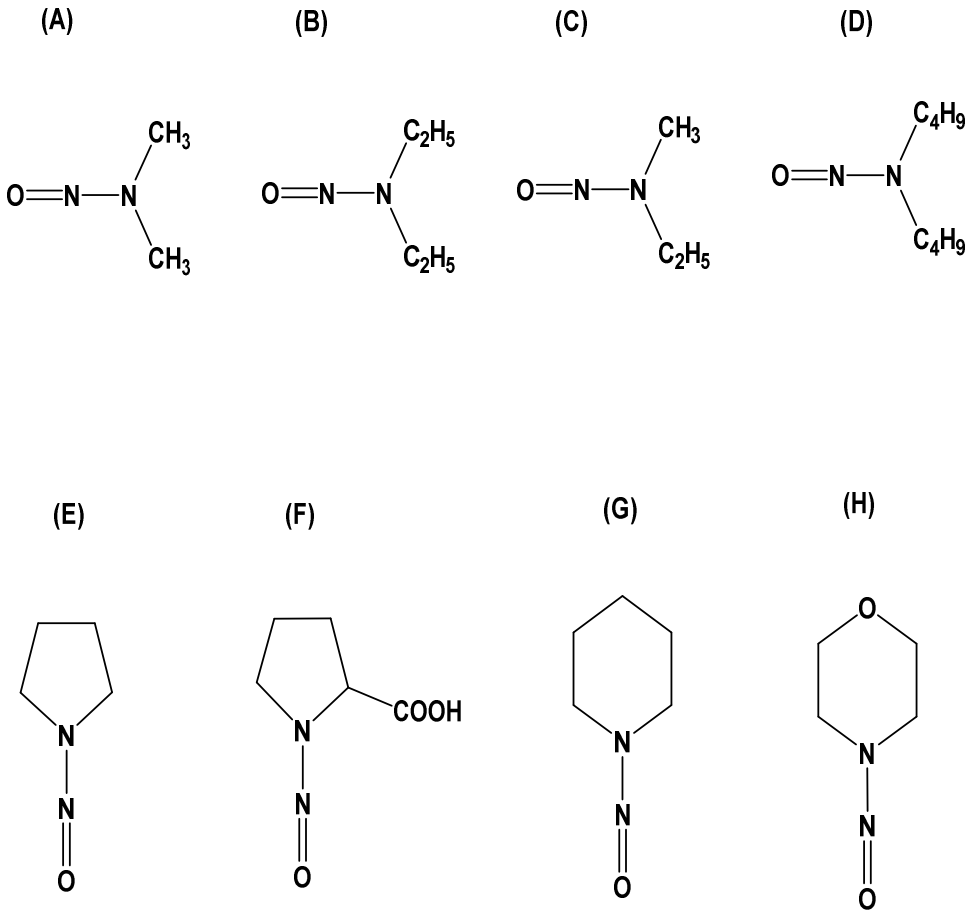


Figure 1.11: Structures of most common dietary NNOC.

(A) N-nitrosodimethylamine; (B) N-nitrosodiethylamine; (C) N-nitrosomethylethylamine; (D) N-nitrosodibutylamine; (E) N-nitrosopyrolidine; (F) N-nitrosoproline; (G) N-nitrosopiperidine; (H) N-nitrosomorpholine. Figure adapted with modifications from (161).

NDMA and NDEA were classified as 2A (probably carcinogenic to human) by IARC whereas, NMEA, NDBA, NPIP, NPYR, and NMOR were classified as 2B (possibly carcinogenic to human) (162,163). Levels of detected N-nitrosamines in food are listed in Table 1.3.

Table 1.3: Preformed N-nitrosamines in food and beverages.

Product	Detected nitrosamines	Concentration range (µg/kg)	Reference
Red meat and Poultry:			
Beef and beef products	ND	-	(163)
Pork meat flesh	ND	-	(163)
Pork meat, in salt	1,2	0.06-0.28	(164)
Pork liver	1,6	0.21-0.64	(163)
Smoked pork brisket	1,2,3,4,5,6	0.04-37	(164)
Chicken	ND	-	(163)
Processed meat:			
Ham	1,2,3,4,5	0.04-3.45	(163)
Smoked ham	1,2,3,5,6	0.04-12	(84)
Bacon	1,2,5	0.17-3.92	(83)
Sausages	1,2,3,5,6	0.02-7.9	(163)
German sausages	1,2,3,5,6	0.02-7.5	(164)
Salami	1,2,3,5	0.04-12	(84)
Fish products:			
Tuna	1	0.12	(163)
Salted fish	1,2	7.64-322.92	(84)
Smoked fish	1	0.26	(165)
Fish cake	1,2	0.2-6.44	(166)
Dairy Products:			
Dried milk powder	1,2,4,5,6	< 0.05- 0.3	(167)
Cheese	1,2,5	< 0.04-4	(163)
Yogurt	ND	-	(163)
Butter	6	0.2	(163)
Margarine	1,2,6	0.2-5.8	(163)
Vegetables:			
Pickled vegetables	1,2,4,5,	0.01-0.2	(168)
French fries	1,4	0.0024-0.0041	(168)
Fresh vegetables	1	ND-3.3	(164)
Frozen vegetables	2,5	1.7-3.6	(164)
Alcoholic drinks:			
Beer	1,2,4,6	0.2-1.87	(163)
Wine	1,4	0.03-0.11	(168)
Whisky	1,2,4,6	0.16-2.5	(163)
Malt beverage	2,6	0.2-0.73	(166)

[1]N-nitrosodimethylamine; [2]N-nitrosodiethylamine; [3]N-nitrosodibutylamine; [4]N-nitrosopyrolidine; [5]N-nitrosopiperidine; [6]N-nitrosomorpholine. ND denotes the N-nitosamine was not detected.

1.8.1.2 Occupational exposure to AAs

A range of N-nitrosamines could be detected in ambient air or indoor air of the working areas of certain occupations such as rubber industry, metal working, leather tanning and fish or meat curing (169). The highest concentration of N-nitosamines in human environment was reported in ambient air of rubber industry since chemicals used in the process of rubber vulcanization easily decompose to secondary amines that are readily nitrosated due to the presence of nitrosating agents such as nitrous gases in production areas. Moreover NDMA and NMOR were regularly detected not only in ambient air of 19 rubber factories but also in tyres storage and showrooms (170).

Similarly, workers at metal, cutting oils and leather tanning industries are exposed to high level of N-nitosamines which make them at increased risk of multiple organ cancers such as leukemia, bladder, lung and gastric cancer (169). Regulatory control measures have contributed to decreasing occupational exposure to N-nitrosamines in developed countries while in developing countries since working areas are mostly poorly controlled, workers are still exposed to high levels of N-nitosamines.

1.8.1.3 Tobacco-specific N-nitrosamines

Tobacco-specific N-nitrosamines are a group of both human and animal carcinogens formed upon nitrosation of the tobacco alkaloids clarifying the strong association between smoking and cancer (171). They are likely causative factors for multiple cancers including lung, oral cavity, oesophagus, pancreas, liver and urinary bladder in smokers (172). Seven tobacco-specific N-nitrosamines were identified, among those, the most carcinogenic in laboratory animals are 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL), and N′-nitrosonornicotine (NNN) (173,174).

1.8.1.4 Chemotherapeutic agents

Based upon chemical structure chemotherapeutic AAs can be classified to either classical AAs including: nitrogen mustards (alkyl sulfonates, nitrosoureas and azridines) and platinum compounds (cisplatin, oxaliplatin and carboplatin), or non-classical AAs which include: dacarbazine, procarbazine and TMZ (175).

1.8.1.5 Halocarbons

Halocarbons include chloromethane, bromomethane, and iodomethane which are commonly used as methylating agents. The most abundant atmospheric halocarbon is chloromethane (methyl chloride) gas which is generated in many terrestrial environments, e.g. by plants and fungi, but also industrially. Halocarbons were reported as mutagenic and carcinogenic agents that do not require bioactivation (175).

1.8.2 Endogenous formation of NNOC

Endogenous formation of NNOC accounts for 45-75% of the total exposure to NNOC (159). In general, NNOC can be formed *in vivo* in a multistep process (involving formation of both nitrosating agents as well as nitrosable precursors such as amines and amides). Many factors act solely or in combination facilitating *in vivo* N-nitrosation process and the subsequent formation of NNOC; including: bacterial action, acidic pH of the stomach, inflammation as well as simultaneous ingestion of nitrates/nitrites (food preservatives), red/processed meat consumption, nitrogen oxides and ingestion of a nitrosable substrate such as dietary amines (161,176).

1.8.2.1 Endogenous formation of nitrosating agents

Several plausible mechanisms proposed to explain *in vivo* N-nitrosation suggested that endogenous formation of nitrosating agents could be mediated through

(i) chemical acid-catalysed nitrosation (ii) heme-catalysed nitrosation or (iii) Nitric oxide (NO) pathway.

The first mechanism is 'chemical acid-catalysed nitrosation' which relies on nitrates (NO_3) and nitrites (NO_2) ingestion. Usually, nitrates are chemically stable, however, either in food (*in situ*) or inside the human body they are reduced to less stable nitrites. This reduction process occurs under acidic conditions in the stomach or by the action of nitrite reducing bacteria in the oral cavity. Moreover, 80% of gastric nitrites resulting from nitrate reduction while the remaining 20% are due to ingested nitrites (177). Subsequently, nitrites are acidified to form nitrous acid (HNO_2; Equation 1) followed by dimerization and loss of water forming nitrous anhydride (N_2O_3; Equation 2). Nitrous acid can also be protonated to form the nitrous acidium ion ($H_2NO_2^+$; Equation 3). Both nitrous anhydride and nitrous acidium ion act as nitrosating agents or NO donor where they can readily react with amines and/or amides abundant in human gut lumen forming N-nitrosamines ($R_2N–NO$; Equation 4) or N-nitrosamides ($RN(NO)COR'$; Equation 5), respectively (154,178,179).

Equation 1: $NaNO_2 + H^+ \rightarrow HNO_2 + Na^+$
Equation 2: $2\ HNO_2 \leftrightarrows N_2O_3 + H_2O$
Equation 3: $HNO_2 + H^+ \leftrightarrows H_2NO_2^+$
Equation 4: $N_2O_3 + R_2NH \rightarrow R_2N–NO$ (a nitrosamine) + HNO_2
Equation 5: $H_2NO_2^+ + RNHCOR' \rightarrow RN(NO)COR'$ (a nitrosamide) + $H_2O + H^+$

Furthermore, nitrosothiols that are formed from nitrites and thiol groups at low pH of the stomach are possible precursors of NNOC in the colon (180). Notably, processed meat ingestion is the main source of nitrates/nitrites exposure since both sodium nitrate and sodium nitrites are common food preservatives that are added to processed meat to inhibit the growth of *Clostridium botulinum* (181,182); however, pickled and fermented vegetables contain relatively high amounts of nitrates.

Secondly, NNOC could be generated *via* the nitric oxide (NO) pathway at sites of chronic or low grade inflammation by bacteria and activated macrophages which possibly explains the association between ulcerative colitis and increased risk of colorectal cancer (183).

In addition, endogenous N-nitrosation is mediated through heme-catalysed formation of NNOC (180). Generally meat and red meat in particular is rich in haemoglobin which is rich in both heme iron and thiol that could be easily nitrosylated in human intestine forming nitrosyl haem (FeNO) and nitroso thiols (SNO). Both nitrosyl haem and nitroso thiols which act as nitrosating agents (NO donors) that could potentially nitrosate both amines (resulting from bacterial decarboxylation of amino acids) and amides favoring the formation of NNOC in large intestine (180); as illustrated in Figure 1.12.

The ability of red meat to stimulate endogenous intestinal NNOC formation was demonstrated in 1996 (184). Furthermore, a previous study conducted in 2006 demonstrated that red meat diet consumption was associated with significant increase in endogenous formation of NNOC as well as the percentage of the NNOC specific DNA adduct, O^6-CMG detected in exfoliated colonic cells (185).

1.8.2.2 Nitrosable amines in GIT

The body pool of nitrosable amines is maintained by multiple sources either by exogenous supply *via* ingestion of preformed biogenic amines in the diet (in a wide variety of food or beverages) (186) or endogenous sources including gut microbiota (187) or biliary/pancreatic secretions. However, the external supply provides a larger quantity of amines than the endogenous pathway. These dietary nitrosable precursors create an additional influence of dietary factors on determining human colorectal cancer risk.

Biogenic amines could have either aliphatic (spermine, spermidine, cadaverine, agmatine and cadaverine), aromatic (tyramine and phenylethylamine) or heterocyclic (histamine and tryptamine) structures as shown in Figure 1.13 (188,189). Biogenic amines in food and beverages result from bacterial decarboxylation of the corresponding amino acid where histamine, tryptamine, tyramine, cadaverine, agmatine and phenylethylamine produced upon decarboxylation of histidine, tryptophan, tyrosine, lysine, arginine and phenylalanine, respectively. On the other hand, putrescine arises from agmatine and ornithine. Both spermine and spermidine arise from putrescine *via* transfer of aminopropyl moieties catalyzed by spermine synthase and spermidine synthase, respectively (190).

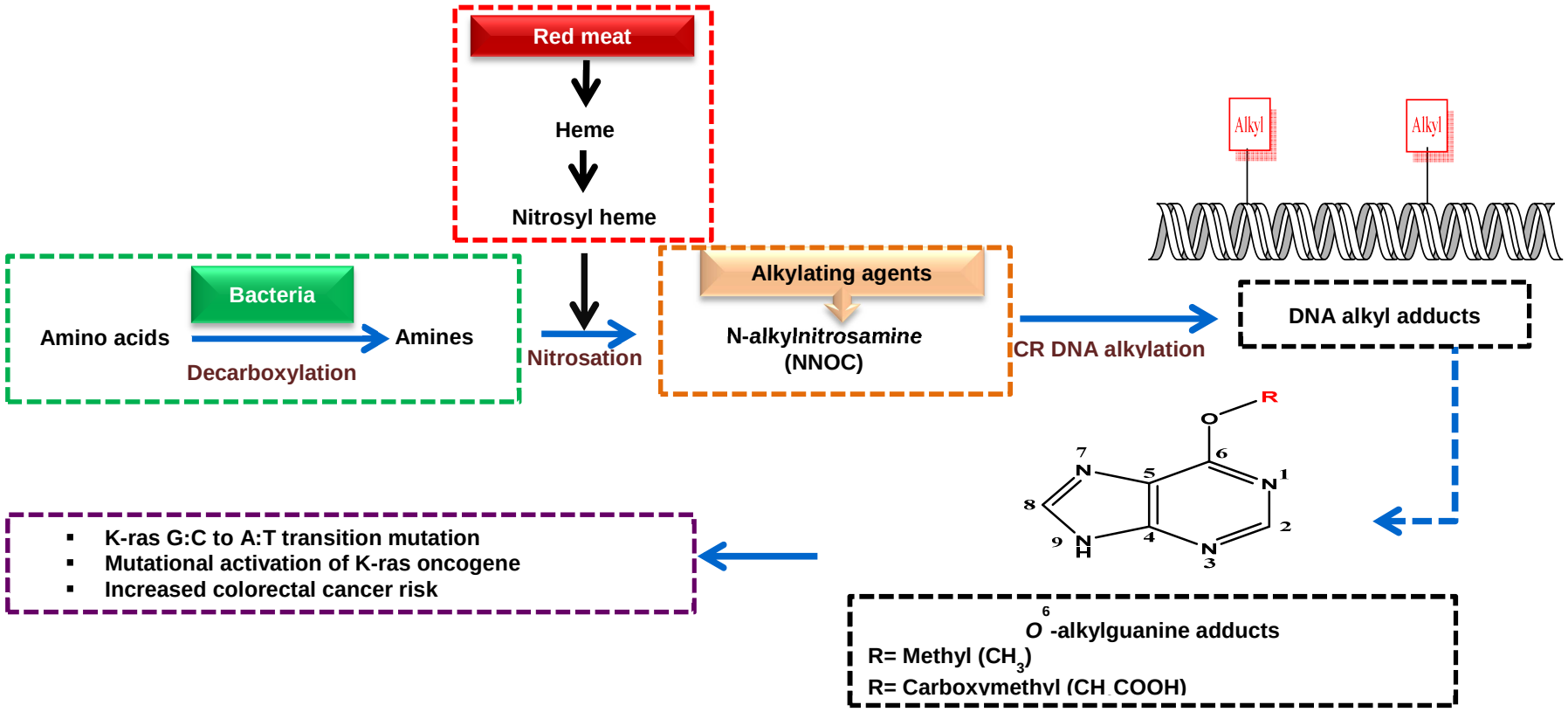


Figure 1.12: Heme-catalysed endogenous formation of NNOC and associated colorectal cancer risk.

Figure adapted with modifications from (157).

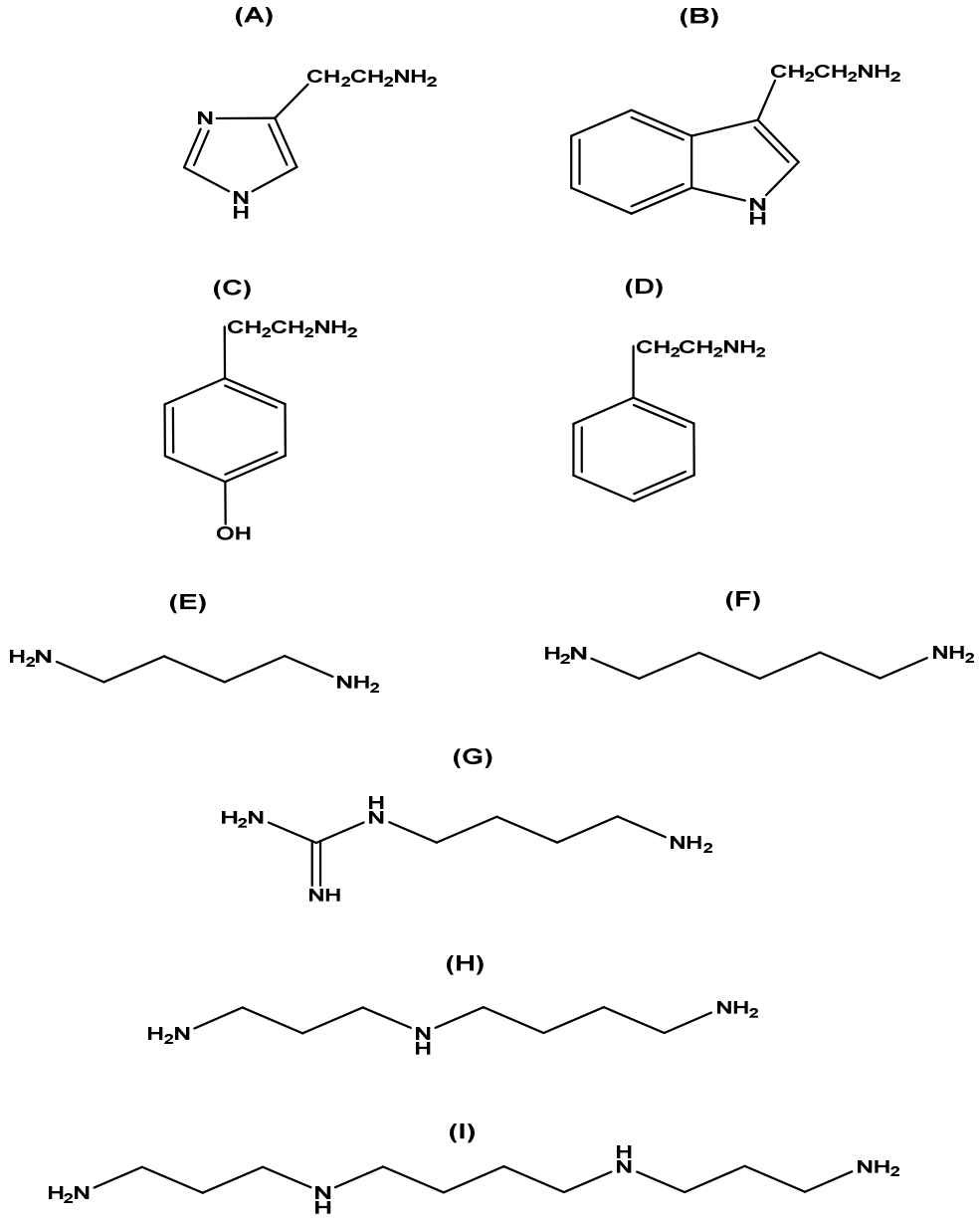


Figure 1.13: Common biogenic amines in food.

(A) Histamine, (B) Tryptamine, (C) Tyramine, (D) Phenylethylamine, (E) Putrescine, (F) Cadeverine, (G) Agmatine, (H) Spermidine, (I) Spermine. Figure adapted with modifications from (188).

Luminal amine concentrations differ in different species and depend on the dietary composition. A study investigated polyamine concentrations in human intestinal lumen of 11 control subjects reported the detection of a range of polyamines including acetylcadaverine, putrescine, cadaverine, spermidine, N8-acetylspermidine, N1-acetylspermidine, spermine and N1-acetylspermine as shown in Table 1.4 (191).

A recent study, published in 2017 (192), has characterised content of amines in food pellets of the Swiss mice then in 2 intestinal segments ileum & caecum as well as in rectal faeces as shown in Table 1.5. They reported the detection of 5 amines in food pellets which were: spermine, spermidine, putrescine, cadeverine and isoamylamine where spermidine was the richest in food followed by putrescine and spermine. The concentrations of all of the detected amines were significantly higher in food pellets with respect to intestinal content, with exception of isoamylamine whose concentration was significantly higher in intestinal content compared to food.

Moreover, in addition to its role in the reduction of nitrates with subsequent production of nitrosating agents, bacterial normal flora plays a pivotal role in the production of nitrosable precursors since NNOC can be formed endogenously in stomach *via* an acid mediated reaction upon bacterial decarboxylation of dietary amino acids into amines followed by nitrosation of the formed amines into N-nitrosoamines (60,193). A previous study primarily focused on exploring the role of normal bacterial flora in endogenous NNOC formation has reported a significant elevation in faecal apparent total NNOC (ATNOC; including nitrosothiols and nitrosyl iron) upon supplementation of drinking water of normal flora rats with nitrite, however, nitrate stimulation of ATNOC formation was not reported in in germ free rats (194).

Furthermore, endogenous N-nitrosation occurs because peptides and amino acids are abundant in the intestine and can be nitrosated to form diazopeptides or N-nitrosopeptides (195). Another important class of endogenous AAs is N-nitroso bile acid conjugates that include two major compounds N-nitrosotaurodeoxycholic acid and N-nitrosotaurocholic acid. N-nitroso bile acid conjugates are formed upon nitrosation of taurocholic acid or deoxytaurocholic acid in gastric juice (196,197).

Table 1.4: Concentrations of polyamines in duodenal aspirates from 11 control human subjects.

Polyamine	Frequency (n=11)	Concentration (μM) Mean±SEM
Acetylcadaverine	8/11	44.2±10.8
Putrescine	8/11	5.0±2.2
Cadaverine	8/11	1.6±0.5
Spermidine	11/11	3.4±0.9
N8-acetylspermidine	1/11	2.0
N1-acetylspermidine	1/11	3.7
Spermine	10/11	3.5±1.0
N1-acetylspermine	1/11	4.6

Table reproduced from (191).

Table 1.5: Concentrations of polyamines and trace amines in pellet food and intestinal content of Swiss mice.

Sample (pmole mg^{-1})	Putrescine		Spermidine		Spermine		Isoamylamine		Cadaverine	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Pellet food	293.9[§]	29.9	779.8[§]	40.0	144.6[§]	12.2	57.4[§]	2.0	26.2[§]	1.3
Chyme-distal ileum	39.5	6.0	193.3	25.8	27.4	4.1	249.8	46.3	9.2	2.3
Chyme-caecum	78.8	8.8	497.3[†]	65.3	35.4	13.2	276.1	47.4	16.3	2.5
Colon faeces	85.7	16.4	351.0	57.7	26.4	5.9	198.5	31.4	18.4	2.8

Values are means ± SE of at least four determinations, expressed as pmole mg^{-1} of the sample. [§]$P < 0.05$ food *vs.* intestinal content; [†]$P < 0.001$ Chyme-caecum *vs.* ileum and colon (ANOVA). Table reproduced from (192).

1.8.2.3 Effect of red and/or processed meat on endogenous NNOC

Endogenous formation of NNOC can be assessed by ATNOC in faeces and there is a large volume of published studies that reported the stimulation of endogenous formation of NNOC with increased red and/or processed meat consumption but not white or fish meat (58,184,198,199). A study conducted by Lewin *et al*. (185), compared the change in faecal ATNOC concentration on changing from a vegetarian to a high red meat diet in 21 volunteers. ATNOC output per day increased significantly from 193±43 µg, on the vegetarian diet, to 667±141 µg on the high-red meat diet (185).

In addition, another study carried out by Kuhnle et al. (200), reported that the concentration of ileal nitrosyl iron compounds significantly increased after a high red meat diet compared to with a non-meat diet (65±6.2 pmole/mg *vs*. 1.6±0.5 pmole/mg). Ileal nitrosothiol concentration also significantly increased after the high red meat diet (11±2.6 pmole/mg *vs.* 0.33±0.2 pmole/mg). Moreover, there was a 400 fold increase in the concentration of faecal nitrosyl iron compounds after a high red meat diet compared with vegetarian diet (91.6±46.2 pmole/mg *vs*. 0.2±0.03 pmole/mg). The concentration of faecal nitrosothiols also increased after a high red meat diet compared with vegetarian diet (54.7±36.9 pmole/mg *vs*. 0.2±0.1 pmole/mg (199).

Furthermore, a study conducted in 2009 highlighted nitrosyl iron to be the main contributor to endogenous nitrosation since both red meat and processed meat diets caused a significant increase in faecal nitrosyl iron (78% and 54%, respectively), and to a lesser extent nitrosothiols (12% and 19% for red and processed meat, respectively) alongside the reported increase in faecal NNOC (201). Moreover, the role of nitrosyl iron in endogenous nitrosation was further confirmed by Jossen *et al*. (202) who also reported increase in both faecal heme and NNOC with increasing meat in diet.

1.8.3 DNA alkyl adduct formation from NNOC

N-Nitrosamines are not carcinogenic unless activated by metabolic enzymes 'CYP450' which are responsible for hydroxylation of N-Nitrosamines at the carbon atom adjacent to the nitroso group forming an intermediate molecule "alpha-hydroxynitrosamine". Such intermediate subsequently undergoes spontaneous loss of an aldehyde molecule *via* cleavage of carbon nitrogen bond which results in the formation of alkyldiazohydroxide. In contrast N-nitrosamides (e.g. alkylnitrosourea) are chemically active under physiologic pH and do not require metabolic activation since they readily decompose yielding the same alkyldiazohydroxide molecule as an intermediate step in the production of diazonium ion.

The diazonium ion is the ultimate carcinogen because its electrophilic nature facilitates its reaction with nucleophilic centres in DNA results in the formation of a range of DNA alkyl adducts including O^6-alkylG adducts as well as N^7-alkylguanine (N^7-alkylG) alongside other adducts (161). DNA alkylation is considered the critical step for NNOC mediated carcinogenicity. The mechanism of NNOC carcinogenicity is illustrated in Figure 1.14.

An example for AAs reaction with DNA is tobacco smoke AAs NNK and NNAL. Both NNK and NNAL require metabolic activation by liver metabolic enzymes, namely, CYP450 generating toxic reactive intermediates which methylate and pyridyloxobutylate DNA, forming DNA alkyl adducts, most notably carcinogenic O^6-MeG and O^6-[4-oxo-4-(3-pyridyl)butyl]guanine (O^6-pobG), as illustrated in Figure 1.15. (203).

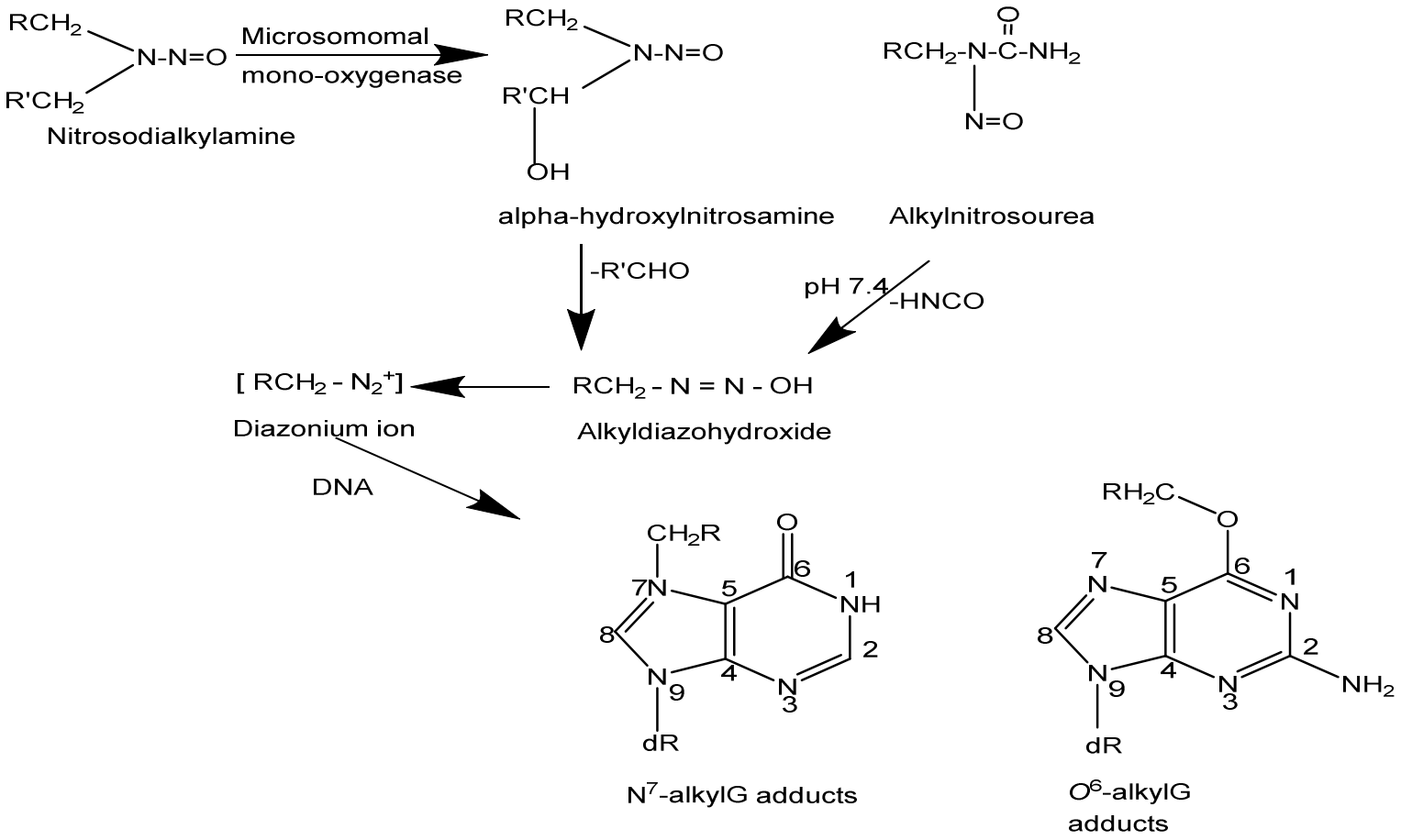


Figure 1.14: DNA alkyl adduct formation from NNOC.

dR denotes 2'-deoxyribose group. Figure adapted with modifications from (161).

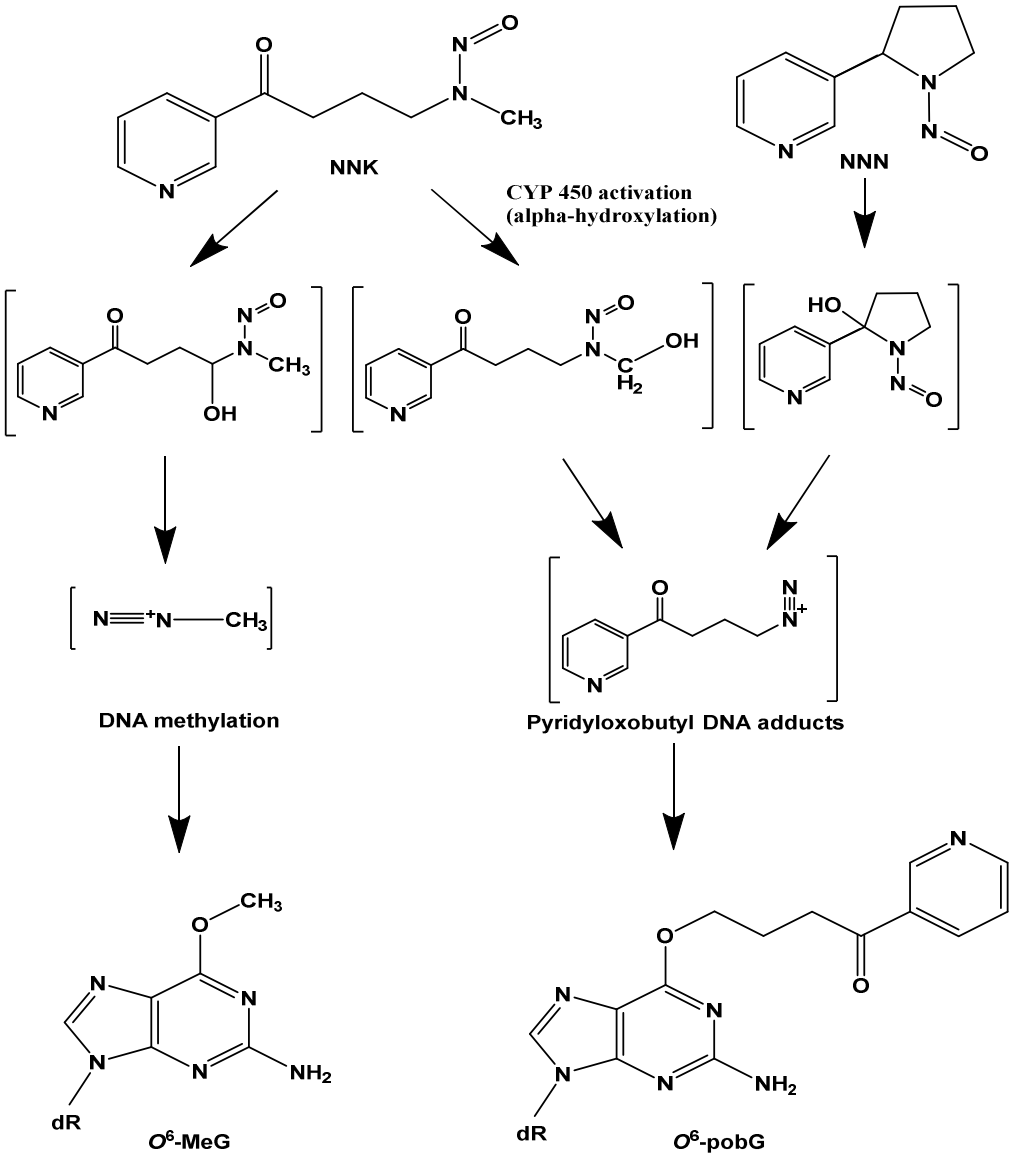


Figure 1.15: NNK and NNN metabolic activation and formation of O^6-MeG and O^6-pobG adducts.

NNK and NNN denote 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and N′-nitrosonornicotine, respectively. Figure adapted with modifications from (171). dR denotes 2'-deoxyribose group.

1.9 DNA alkylation products

AAs react at the 4 exocyclic oxygen sites and 7 ring nitrogen atoms, or the phosphate groups producing 2 possible phosphotriesters (152). The transferred alkyl group could be methyl, ethyl, isopropyl, phenyl or benzyl group as it depends on the AA reacting with the DNA nucleophilic centre. DNA adducts formed due to alkylation of exocyclic oxygen atom include: O^6-alkylG, O^2-alkylthymine (O^2-alkylT), O^4-alkylthymine (O^4-alkylT) and O^2-alkycytosine (O^2-alkylC). In addition, DNA alkyl adducts due to modification of the ring nitrogen are N^1-alkyladenine (N^1-alkylA), N^3-alkyladenine (N^3-alkylA), N^7-alkyladenine (N^7-alkylA), N^3-alkylcytosine (N^3-alkylC), N^3-alkylguanine (N^3-alkylG), N^7-alkylG and N^3-alkylthymine (N^3-alkylT). Possible sites of DNA alkylation are summarised in Figure 1.16.

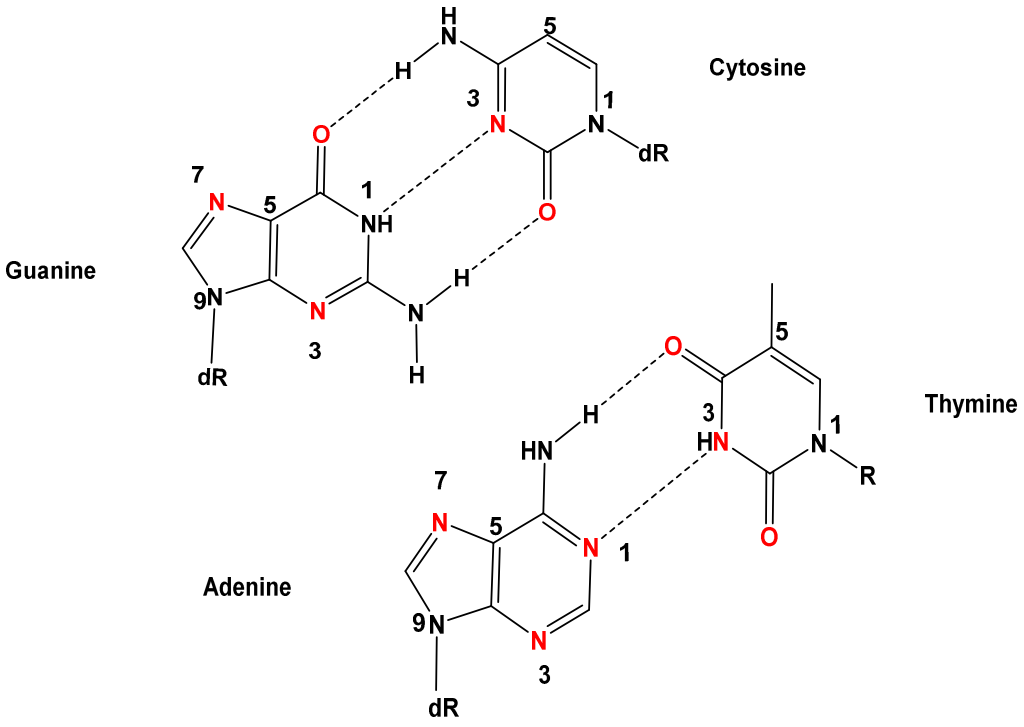


Figure 1.16: DNA Watson Crick DNA base pairs with principal alkylation sites.

Sites of the most common base modifications by alkylating agents are shown in red. dR denotes 2'-deoxyribose group

Different AAs induce different DNA alkylation products, as summarised in Table 1.6; and the primary site of reaction at the DNA depends mainly on the nature of the AA and whether it react through S_N1 or S_N2 mechanism. A previous study identified 90% of the DNA alkyl adducts resulting from metabolic activation of NDMA in rat liver. Alkyl modified guanine adducts were the most commonly occurring adducts representing 73.8% of the total formed DNA alkyl adducts (0.9% N^3-alkylG, 66.8% N^7-alkylG, 6.1% O^6-alkylG) followed by alkyl adenine adducts that represented 3.9% only of the total burden of the detected alkyl DNA adducts (0.9% N^1-alkylA, 2.3% N^3-alkylA, 0.7% N^7-alkylA). Other detected DNA alkyl adducts included: N^3-alkylC (0.6%); O^4-alkylT (trace) and alkyl phosphate triesters (12%) (204). Previous results highlights that NNOC alkylation of DNA leads to the formation of a wide range of DNA alkyl adducts of which O^6-alkylG adducts are of major importance in initiating mutations and carcinogenicity.

Table 1.6: DNA alkylation products following reaction with AAs.

Agent	Source	DNA base alkylation products (% of total products)								
		Guanine			Adenine			Cytosine	Thymine	
		N^7	N^3	O^6	N^1	N^3	N^7	N^3	N^3	O^4
MMS[1]	Salmon sperm (205)	83	0.6	0.3	3.8	10.4	1.8	<1	0.08	ND
MNU[2]	Salmon sperm (206)	75	1.1	7.5	1.5	11	2.6	0.7	0.3	0.1
NDMA[3]	Rat liver (204)	66.8	0.9	6.1	0.9	2.3	0.7	0.6	ND	ND
NNK[4]	Mouse lung (207)	87	ND	13	ND	ND	ND	ND	ND	ND
NOGC[5]	Calf thymus (208)	76	ND	7	ND	17	ND	ND	ND	ND
SAM[6]	Micrococcus luteus (209)	86	ND	1	ND	13	ND	ND	ND	ND

ND indicates that the alkylation product was not detected. [1]MMS: methyl methanesulphonate, [2]MNU: N-methyl-N-nitrosourea, [3]NDMA: N-nitrosodimethylamine, [4]NNK: 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone) [5]NOGC: N-nitrosoglycocholic [6]SAM: S-adenosylmethionine.

Alkylated DNA lesions are able to cause disastrous effect since they contribute to carcinogenicity, mutagenicity, genotoxicity and cytotoxicity (210). Several mechanistic pathways have been proposed to explain DNA alkyl adducts induced mutagenicity and carcinogenicity including formation of DNA strand breaks and mispairing (211). However, the toxic potential of DNA adducts is variable. For example, N^7-methylguanine adduct, which is the principal DNA adduct, is not directly mutagenic. On the other hand, O^6-MeG, which is formed in lesser amounts, is mainly mutagenic and carcinogenic as would be explained in a latter section (158). Eventually, O-alkylations (O^6-alkylG and O^4-alkylthymine) are highly mutagenic and genotoxic, whereas N-alkylations (e.g. N^3-alkylA and N^1-alkylA) are cytotoxic, but relatively less mutagenic as it will be clarified in the following section. (212).

1.9.1 O^6-alkylguanine (O^6-alkylG) adducts

In the spectrum of DNA adducts formed by AAs O^6-alkylG adducts, characterised by the addition of an alkyl group at the O^6 position of guanine base, have attracted particular interest due to their major role in carcinogenesis. O^6-alkylG adducts include a wide range of adducts such as: O^6-MeG, O^6-CMG, O^6-ethylguanine (O^6-ethylG) O^6-butylguanine (O^6-butylG), O^6-propylguanine (O^6-propylG) and O^6-pobG.

O^6-MeG is a characteristic promutagenic and toxic adduct formed due to exposure to environmental or dietary methylating agents such as cigarette smoke alkaloids and/or dietary NNOC (213). However, formation of O^6-MeG adduct could arise from endogenous NNOC (206,209). In addition, various nitrosated glycine derivatives react with DNA *in vitro* to give carboxymethyl adducts (e.g., O^6-CMG) and lesser amounts of methyl adducts (e.g., O^6-MeG); as shown in Figure 1.17 (185,208).

1.9.2 Carcinogenic effect of O^6-alkylG adducts

The positive association between O^6-alkylG adducts and carcinogenesis is believed to be mainly through transition mutations and/or recombination (212).

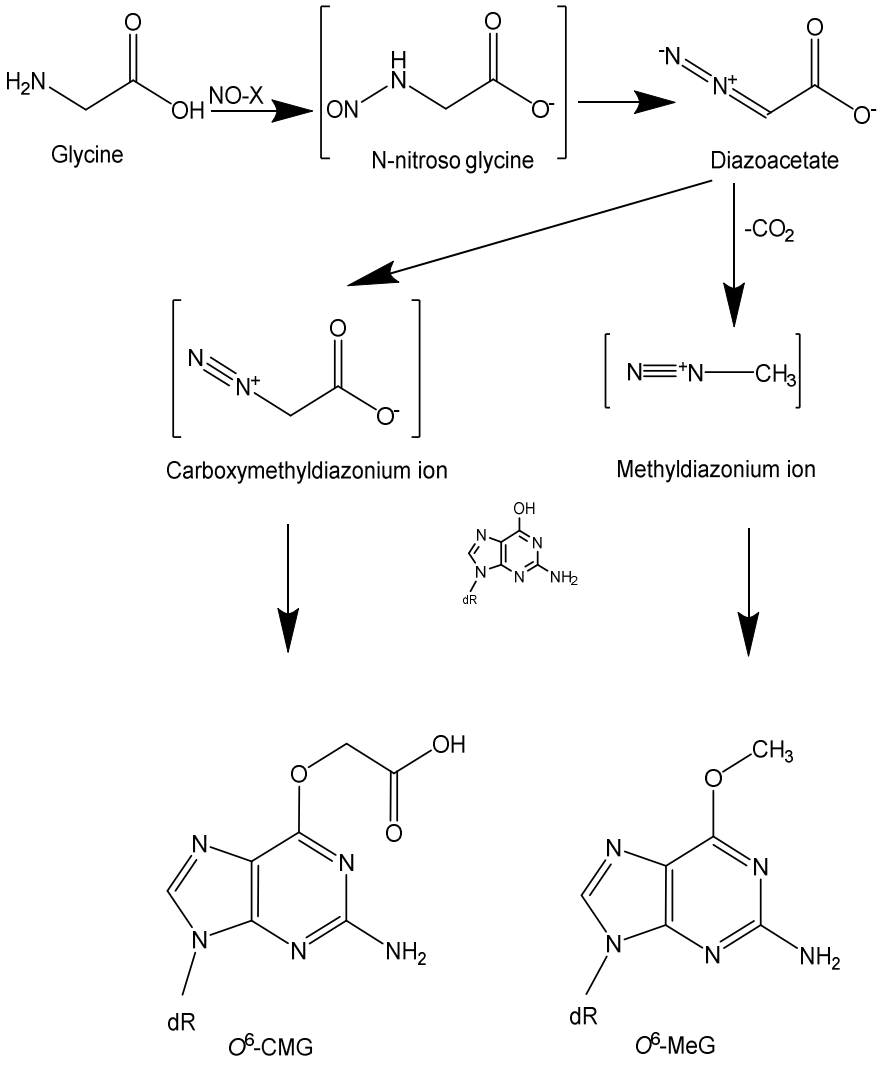


Figure 1.17: Nitrosation of glycine and formation of *O*^{6}-MeG and *O*^{6}-CMG adducts.

Figure adapted with modifications from (214). dR denotes 2'-deoxyribose group.

1.9.2.1 Transition mutation

Earlier publications demonstrated that a range of *O*^{6}-alkylG adducts including *O*^{6}-MeG (215), *O*^{6}-n-propylG, *O*^{6}-n-butylG and *O*^{6}-n-octylguanine (216) induce GC:AT transition mutations after 2 rounds of replication of DNA containing the lesion (210,217,218). Several hypotheses have been proposed to explain the root cause of GC:AT mutation. The first hypothesis proposes that GC:AT transition mutations arise

from the mis-insertion of thymine opposite to O^6-alkylG adduct by DNA polymerases (219,220) and subsequent formation of more stable hydrogen bonded O^6-alkylG:T mispairs. However, this assumption was contradicted by both crystallographic (221) and earlier studies (222) reporting that O^6-alkylG:T mispairs are not more stable than their O^6-alkylG:C counterparts.

Alternatively, Swann *et al*. (223) suggested that O^6-alkylG:T mispairs are attributed to 2 possible mechanisms. DNA polymerases may misread the O^6-alkylG in the template strand for adenine because of the altered hydrogen bonding properties of guanine base caused by the presence of an extra alkyl group at O^6 position and as a result it selects dTTP rather than dCTP. Additionally, the misreading results from the close resemblance of O^6-alkylG to adenine rather than unmodified guanine in terms of bond lengths, angles as well as lipophilic characters (O^6-alkylG is more lipophilic than guanine) (224). This assumption has been further confirmed by a recent study in 2011 (225). Alternatively, O^6-alkylG:T mispairs retain Watson Crick alignment with the N_1 in guanine base juxtaposed to N_3 in thymine base facilitating formation of phosphodiester links whilst the O^6-alkylG:C pairs adopt Wobble conformation (223). However, in a recent crystallographic study (226), a Watson Crick type of structure for the O^6-MeG:C base pair has also been observed; as shown in Figure 1.18.

The association between O^6-alkylG mediated transition mutations (G:A, C:T) and tumorigenesis generally and colorectal cancer specifically is supported by various animal studies demonstrated that rodents exposure to DMH or azoxymethane (AOM) resulted in GC:AT transition mutations in a range of genes (APC, K-ras and β-catenin) as well as induction of colorectal carcinogenesis in the laboratory animals; as summarised in Table 1.7. Whilst, a human study identified mutations in K-ras oncogene in colorectal tumour DNA that contained detectable levels of O^6-MeG adduct. The reported mutation frequencies were 28%, 29% and 42% in cecum, sigmoid colon and rectum, respectively, of which, 72% was in codon 12 and 64% was GC:AT transitions. Whilst, the mutation frequency of K-ras oncogene in macroscopically normal tissue was 12% and 14% for sigmoid and rectal DNA samples, respectively. Notably, previous results highlighted the linkage between O^6-MeG adduct and K-ras oncogene transition mutations and the potential association with colorectal cancer.

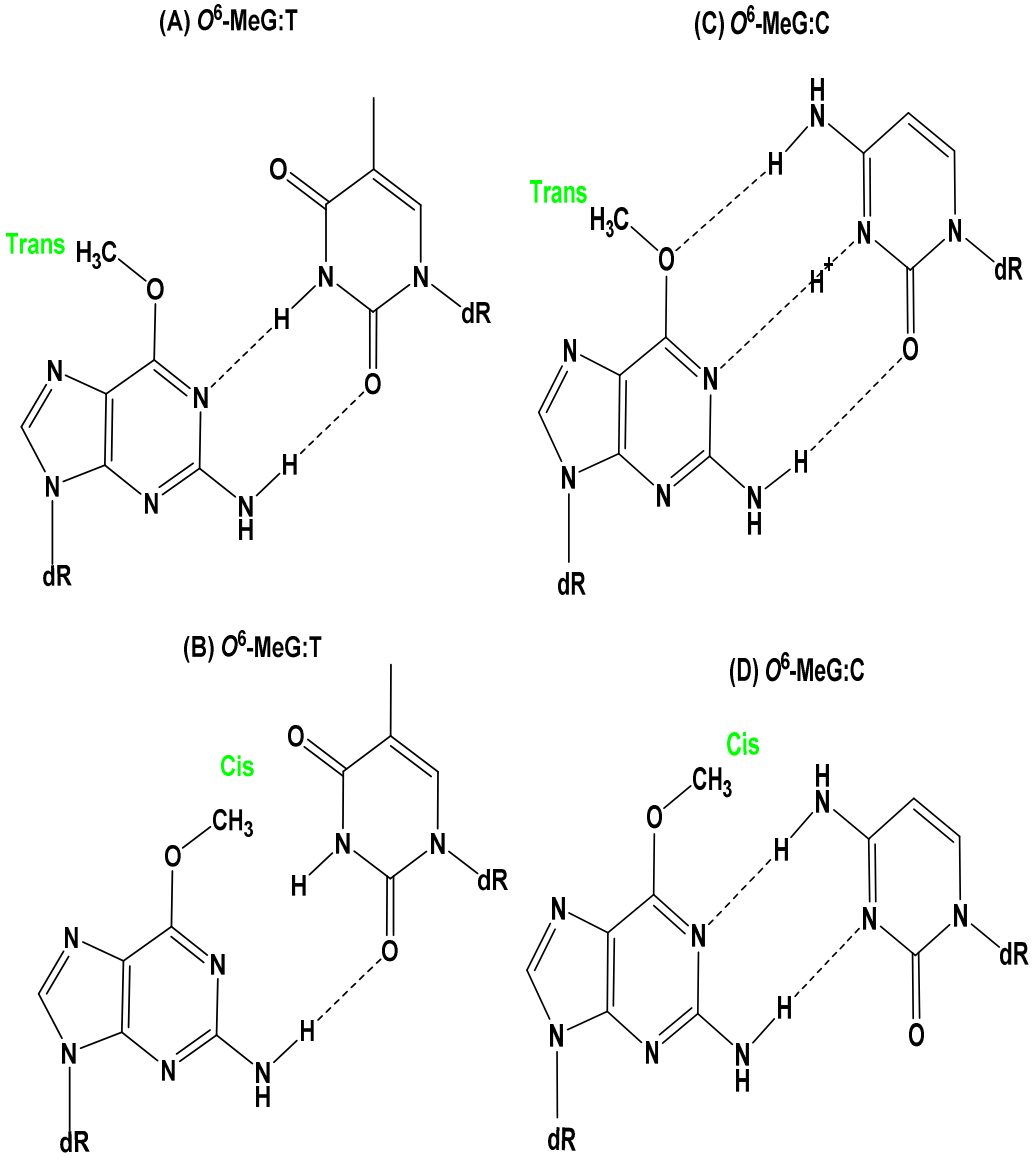


Figure 1.18: Structures of (A) and (B) *O*^{6}-MeG:T, (C) and (D) *O*^{6}-MeG:C base pairs.

Cis and trans denote the two possible stereoisomers of *O*^{6}-MeG. (A), (C) are obtained by the X-ray. (B), (D) are NMR studies in the absence of DNA polymerase. (A-C) Watson Crick type, (D) Wobble type. Figure adapted with modifications from (225,227–231).

Table 1.7: Frequencies of gene mutations found in colorectal tumours induced by DMH/AOM/MNU in rat/mouse models.

Mutated gene	Tumour (n)	Adenoma (n)	Carcinoma (n)	Type of mutation	AA	Ref.
APC	14.3% (27)	18% (17)	10% (10)	80% (C>T) 20% (C>G)	AOM	(232)
	8.8% (57)	ND	ND	80% (C>T) 20% (G>A)	DMH	(233)
	30.4% (23)	33.3% (12)	27.3% (11)	42% G:C>A:T 42% C:G>T:A 15% G:C>T:A	DMH	(234)
K-ras	ND	ND	66% (44)	G>A in codon 12, 13 or 59	DMH	(235)
	22% (50)	4.3% (23)	37% (27)	G>A in codon 12 or 13	AOM	(236)
	ND	ND	60%	G>A in codon 12	AOM/ DMH	(237)
	16.7% (30)	9% (22)	37.5% (8)	C>T in codon 12 or 13	AOM	(238)
	58% (43)	60% (10)	58% (33)	G>A in codon 12	DMH	(239)
	ND	ND	33% (9)	G>A in codon 12 or 13	MNU	(240)
	25% (28)	0% (2) small	8% (12) small 43% (14) large	G:C>A:T in codon 12 or 13	AOM	(241)
β-catenin	ND	ND	100% (10)	G:C>A:T	AOM	(242) *
	68.8% (32)	33% (6)	77% (26)	G:C>A:T	AOM	(241)
	ND	ND	75% (8)	G:C>A:T	AOM	(243)

AA: alkylating agent. DMH: 1,2 dimethylhydrazine, AOM: azoxymethane, MNU: N-methyl-N-nitrosourea. All of the above studies were in rat models. *This study was in mouse not rat model. ND denotes no data.

The second level of relationship between O^6-alkylG adducts and carcinogenicity is through TP53 tumour suppressor gene (125), since G:A transitions were detected in TP53 leading to deactivation of this tumour suppressor gene which facilitate malignant transformation of human cells. TP53 G:A transitions were found in colorectal cancer in a number of locations such as codons 175,204, 213, 248 and 273 (244,245). Consistently, mutational landscape of human colorectal cancer genomes reported that C>T transition mutation is the most predominant nucleotide change, as explained before, highlighting both K-ras and TP53 among the most frequently mutated genes in colorectal cancer.

An *in vitro* study characterised potassium diazoacetate (KDA, stable form of nitrosated glycine) mutational spectrum in a plasmid harbouring human TP53 gene (246), using yeast functional TP53 mutation assay. They reported that 85% of KDA induced mutations were substitutions of which 57% were directed at GC base pairs (52% transitions and 48% transversions) and 28% were directed at AT base pairs. 74% of detected transitions were GC:AT while 26% were AT:GC transition mutations. It's worthwhile to mention that, in addition to the above mentioned mutations, KDA treatment afforded detectable levels of both O^6-MeG and O^6-CMG adducts.

1.9.2.2 Recombination

The recombinational pathway could potentially contribute to O^6-alkylGs induced malignant transformations. Recombination depends on the action of MMR system which causes deletions and/or translocations that could possibly lead to deletions of a tumour suppressor genes or mutational activation of an oncogene (212). O^6-alkylGs could potentially lead to single strand and/or double strand breaks where double strand breaks are processed by recombination repair system ending with either recombination or cell death.

1.9.3 Characterisation of *in vivo* O^6-alkylG adducts in colorectal tissue

Initially, *in vivo* characterisation and detection of O^6-alkylG adducts was based principally on animal studies rather than human samples where experimental animals were exposed to doses of AAs and for durations that are high enough to induce the formation of detectable levels of O^6-alkylG adducts (203,247).

Many of these animal studies have been able to demonstrate the association between colorectal cancer incidence and AAs exposure (247–251); it has been conclusively shown treating SWR mice with DMH, for up to 20 weeks induced colon cancer in mid (43%) and distal colon (87%) whereas no tumours were detected in proximal colon. These results were consistent with the further investigation of O^6-MeG levels in the three colon segments (proximal, mid and distal colon) where the highest O^6-MeG levels were reported in distal colon and the lowest levels were in proximal colon, ranging between < 0.1 to 16.6 fmole of O^6-MeG/µg of DNA (252). Moreover, DMH induced predominantly distal colon tumours in mice (248–250), although DNA alkylation damage was detected in the tissues of colon, liver and kidney, the specific induction of colon tumour was potentially attributed to tissue-specific differences in the persistence of certain DNA adducts (251).

The characterisation and detection of O^6-alkylG that formed *in vivo* not *in vitro*, in human DNA, are even more difficult which is mainly ascribed to the low yield of the adducts, rapid repair rates as well as the limited amounts of human DNA samples. A recent approach has been experimentally able to detect O^6-alkylG (O^6-MeG and O^6-CMG) adducts in human cell lines showing remarkable sensitivity, however, the attained sensitivity was not high enough for the analysis of human DNA (253–255). Only a few studies have been able to detect and/or quantify O^6-alkylG adducts in human DNA samples. Moreover, 4 studies detected and quantified O^6-MeG adduct in human colorectal tissue (256–259) (Table 1.8) and another study detected O^6-CMG in exfoliated human colonocytes (185). In addition, one study detected 4 O^6-alkylG adducts (O^6-MeG, O^6-ethylG, O^6-butylG and O^6-propylG) in human gastric tumour DNA (260); data are summarised in Table 1.9.

Table 1.8: O^6-alkylG adducts levels in human colorectal tissue.

Adduct	Study population (n)	Method	Site (n)	Normal tissue		Tumour tissue		Unit	Ref.
				Frequency (%)	Level	Frequency (%)	Level		
O^6-MedG	UK (62)	RIA	Proximal colon (16) Sigmoid colon (21) Rectum (23) [a]	25 52 48	0.02±0.003 0.044±0.025 0.042±0.020	60 32 63	0.044±0.030 0.050±0.050 0.046±0.048	µmol O^6-MedG/mol dG	(257)
O^6-MedG	UK (11)	RIA	Colorectal region [b]	80	0.083±0.106	15	0.003±0.007	µmol O^6-MeG/mol dG	(258)
O^6-MedG	UK (78)	RIA	Colorectal region	44	0.011±0.94	52	0.011±0.151	µmol O^6-MeG/mol dG	(256)
O^6-MedG	UK (58)	RIA	Cecum Sigmoid colon Rectum	33 52 48	>0.01	58 32 63	>0.01	pmole of O^6-MedG/mol dG	(259)
O^6-CMG	UK (21)	Immuno-histoche-mistry	Colonic exfoliated cells	14±5	-	59±8	-	Mean percentage positive cells	(185)

[a]Normal tissue is taken ~5 cm from the tumour edge, [b]Normal tissue is tumour adjacent mucosa, RIA denotes Radioimmunoassay.

Table 1.9: *O*^{6}-alkylG adducts levels in human gastric tumour tissue.

Adduct	Study population (n)	Method	Site (n)	Normal tissue[b]		Gastric tumour tissue		Unit	Ref.
				Frequency (%)	Level	Frequency (%)	Level		
O^6-MedG *O*^{6}-ethylG *O*^{6}-butylG *O*^{6}-propylG	Italy (24)	HRGC-NICI-SIR[a]	Gastric mucosa	20.8 0 4.2 8.3	0.44 ± 0.26 ND 0.15 ± 0.004 0.09 ± 0.04	12.5[c] 12.5[c] 33.3[c] 4.2[c]	0.37 ± 0.21 0.68 ± 0.33 0.95 ± 0.38 0.08 ± 0.03	fmole/µg DNA	(260)

[a]High-resolution gas chromatography-mass spectrometry with negative-ion chemical ionization and selected ion recording. ND, not detectable. [b]Adjacent area of noninvolved gastric mucosa. [c]Enzyme linked Immunosorbent assay.

1.10 Methods for detection of O^6-alkylG adducts

A variety of techniques have been used for the identification and quantification of O^6-alkylG adducts in DNA isolated from animal or human tissues, cells, and biofluids as well as DNA from *in vitro* studies. These can be grouped as those involving ^{32}P-post-labelling, immunological assays and mass spectrometry based assays.

1.10.1 ^{32}P-post-labelling

^{32}P-post-labelling refers to labelling of the formed DNA adducts with ^{32}P. The labelling process occurs after the adduct formation and is achieved *via* the transfer of ^{32}P containing phosphate group from ATP to adducted deoxyribonucleotides followed by resolving and analysis of labelled adducts. This technique has been used in detecting O^6-MeG adducts (261).

1.10.2 Immunological assays

Immunological assays rely on generating either monoclonal or polyclonal antibodies to specific O^6-alkylG adducts then the produced antibody could be used for quantitative detection of O^6-alkylG adducts *via* a variety of techniques such as immunoslot blot, immunohistochemistry (185), RIA (257) and/or enzyme-linked immunosorbent assay (ELISA) methodology (262).

1.10.3 Mass spectrometry (MS) based assays

Generally MS based assays are being used for detection and quantification, as well as structural identification of O^6-alkylG adducts. MS is an analytical technique the resolves the ions based on their mass to charge ratio (*m/z*) therefore, the first step in mass spectrometric analysis requires the conversion of analyte molecules to gas phase ions. MS based assays could rely on either detecting the O^6-alkylG adducts in the base or in 2'-deoxynucleoside form. Recently, MS based analytical approaches have contributed to the identification and quantification of O^6-MeG adducts as well as O^6-CMG adducts (254,263) and ultra-high pressure liquid chromatography–tandem MS

(UHPLC-MS/MS) has been used to quantify of O^6-MeG and O^6-CMG adducts in intestinal cell line (253).

1.10.3.1 Matrix-assisted laser desorption/ionisation-time of flight (MALDI-TOF)

Generally, every mass spectrometer consists of main integral constituents which are ion source responsible for generating analyte ions, mass analyser that resolves ions according to their *m/z* ratios, detector that detects resolved ions and records their signals and finally data output system that converts the recorded signals to a mass spectrum. MALDI-TOF consists of: MALDI ion source coupled with TOF mass analyser. MALDI ion source relies on laser energy to convert analyte molecules to gas phase ions that could be either positively (such as proteins and peptides that become protonated) or negatively charged (264). Matrix molecules crystallise with the analyte on a target plate and the ionisation process takes place under high vacuum when the matrix transfers the absorbed laser energy to sample peptides or proteins leading to desorption, desolvation and protonation of analyte in the gas phase. Therefore, the function of matrix is to enhance the crystallisation process and to increase the absorption of laser energy. The most commonly used matrix in MALDI is α-cyano-4-hydroxy cinnamic acid. Advantages of MALDI include: it produces singly charged ions and it's a soft ionisation process that conserves the post translation modifications of proteins (265). Table 1.10 summarises the most commonly used matrices and their applications (266).

Table 1.10: Most commonly used MALDI-TOF matrices and their applications.

Matrix	Application
α-cyano-4-hydroxycinnamic acid	Peptides, proteins, organic compounds
3,5-dimethoxy-4-hydroxycinnamic acid	Higher mass polymers including proteins
2,5-dihydroxybenzoic acid	Peptides, proteins, carbohydrates, synthetic polymers
3-hydroxypicolinic acid	Oligonucleotides
Trihydroxyacetophenone	Oligonucleotides, carbohydrates

TOF mass analyser is compatible with the intermittent or discontinuous manner in which analyte ions are produced by MALDI. In this mass analyser ions are driven along a flight path under electric current where ions are being resolved based on the time required to travel the length of the flight path (265). The square of the time is directly proportional to their *m/z* ratios. TOF mass analyser could be operated in two modes either linear or reflective mode (Figure 1.19) using electrostatic reflectors to increase the length of the flight path improving the resolution and to correct disparities in initial velocities of the ions. Advantages of TOF include mass accuracy, good resolution power as well as its tolerance to salts. TOF should be calibrated using a set of peptide standards with known molecular masses for accurate measurement of *m/z* values of sample peptides to be made (267).

The detector is responsible for detecting the resolved ions and generating a signal that could be converted by a computational data output system to a mass spectrum. When ions collide with the detector packets of electrons or photons are generated that could be amplified by electron multipliers. Most commonly used detectors are microchannel plate detectors that are characterised by a plate that contains several parallel channel electron multipliers (264).

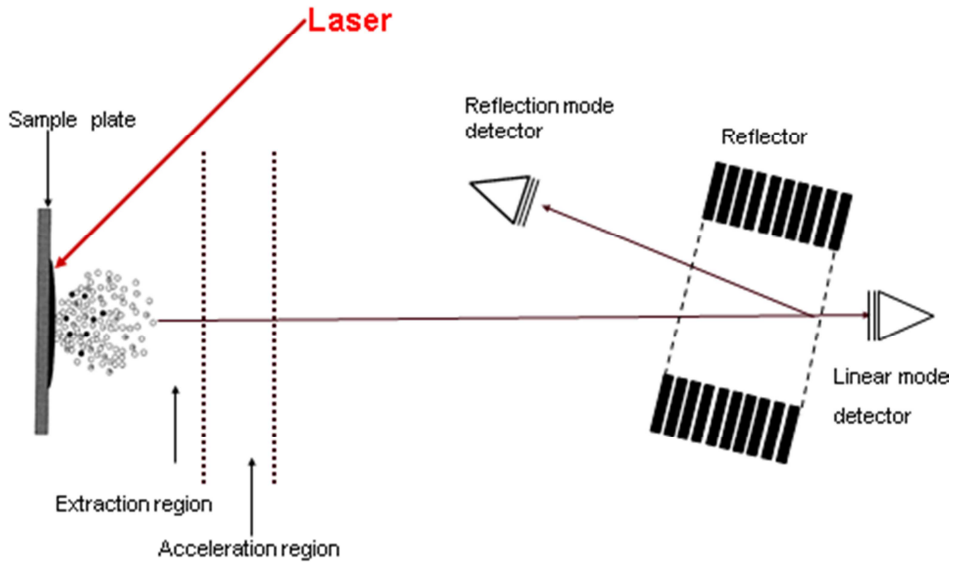


Figure 1.19: A schematic diagram of MALDI-TOF mass spectrometer. Reproduced from (264).

1.11 DNA repair mechanisms

DNA repair mechanisms are considered the first barrier and defence against genotoxic-induced DNA damage. These repair mechanisms aim at maintaining genomic stability and integrity. DNA repair pathways include two major steps. The first step is the recognition of DNA damaged site whilst the second step is either removal of DNA lesion or repairing that lesion (268). Thus, DNA repair mechanisms are essential for the viability and longevity of a healthy organism. The most important DNA repair mechanisms are BER, nucleotide excision repair pathway (NER), mismatch repair (MMR), double strand break repair (DSBR) and direct repair pathway *via* DNA repair protein, namely, O^6-alkylguanine DNA-alkyl transferase (MGMT).

1.11.1 MGMT protein

MGMT, also known as AGT and ATase, is a DNA repair protein or a suicide enzyme that provides protection against the mutagenic, cytotoxic and carcinogenic effects of O^6-alkylG adducts induced by exposure to AAs. MGMT protein functions through transfer of the alkyl group from the O^6 position of the modified guanine base to the cysteine residue within the MGMT active site pocket in an auto-inactivating stoichiometric and irreversible process that results in regeneration of the guanine base; as shown in Figure 1.20 (269). This transfer results in inactivation of the protein followed by its degradation through ubiquitination pathways by the proteasome (270). Therefore, the cellular capacity to withstand or avoid the disasterous effects of O^6-alkylG adducts is directly related to both the total number of active MGMT molecules and the rate of its de-novo synthesis (269).

Increased MGMT activity increases cellular resistance to AAs toxicity whilst depletion of cellular MGMT renders the cell more susceptible to AA-induced toxicity and carcinogenesis. Thus, MGMT is regarded as a key node in the defence against carcinogenic effect of AAs, and a marker of resistance of normal and cancer cells to alkylating chemotherapeutics (271). MGMT protein was discovered in 100 different species including eubacteria and eukaryotes such as *E. coli* bacteria (Ogt) as well as human MGMT. However, MGMT was not detected in plants and *Schizosaccharomyces pombe* (272).

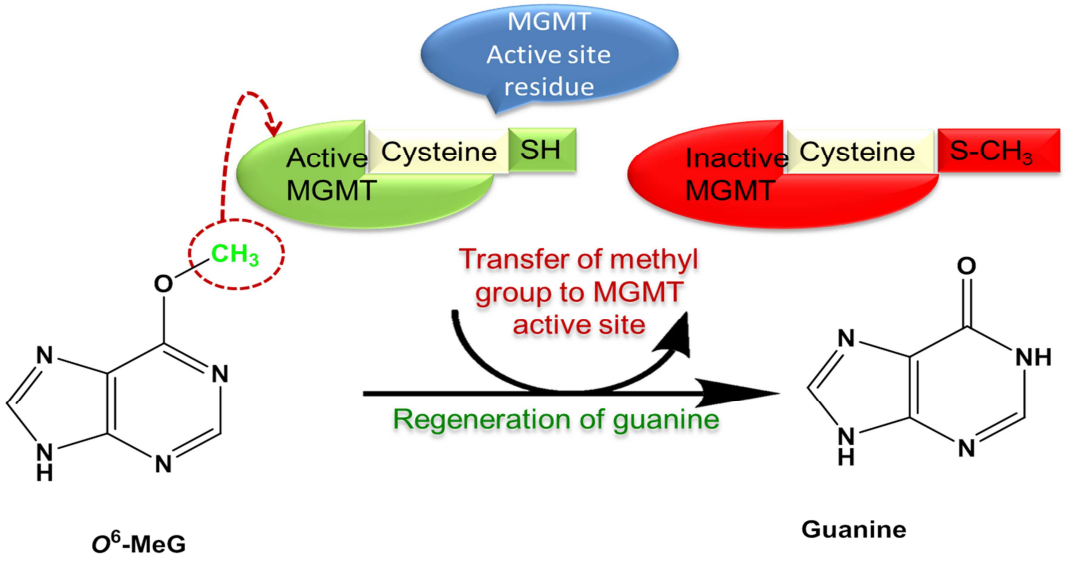


Figure 1.20: MGMT direct reversal repair of O^6-MeG adduct.

1.11.1.1 Human MGMT gene

Human MGMT gene is located on chromosome 10q26 and it is composed of five exons and 4 very large introns (each exceeding 40 kb) and it spans for 300 kb: the last four exons are coding and the second intron is particularly large (170kb) (273). Figure 1.21 illustrates nucleotide sequence and exon structure in MGMT cDNA. Human MGMT gene encodes a protein which is 21.64 kDa and consists of 207 amino acids. MGMT protein amino acid sequence will be shown later. MGMT gene promoter region is located upstream of the first exon and is extremely rich in both G and C nucleotides, a characteristic feature of CpG island (274). Hence, MGMT expression is regulated by methylation of its promoter region “CpG island” (275) where methylation of such region leads to gene silencing and loss of protein expression (276).

```
  1 CCCGCGCCCC GGATATGCTG GGACAGCCCG CGCCCCTAGA ACGCTTTGCG TCCCGACGCC
 61 CGCAGGTCCT CGCGGTGCGC ACCGTTTGCG ACTTGCCCCC CCCGCCCCCC CCGCCGCCCC
121 TTGGTACTTG GAAAAATGGA CAAGGATTGT GAAATGAAAC GCACCACACT GGACAGCCCT
181 TTGGGGAAGC TGGAGCTGTC TGGTTGTGAG CAGGGTCTGC ACGAAATAAA GCTCCTGGGC
241 AAGGGGACGT CTGCAGCTGA TGCCGTGGAG GTCCCAGCCC CCGCTGCGGT TCTCGGAGGT
301 CCGGAGCCCC TGATGCAGTG CACAGCCTGG CTGAATGCCT ATTTCCACCA GCCCGAGGCT
361 ATCGAAGAGT TCCCCGTGCC GGCTCTTCAC CATCCCGTTT TCCAGCAAGA GTCGTTCACC
421 AGACAGGTGT TATGGAAGCT GCTGAAGGTT GTGAAATTCG GAGAAGTGAT TTCTTACCAG
481 CAATTAGCAG CCCTGGCAGG CAACCCCAAA GCCGCGCGAG CAGTGGGAGG AGCAATGAGA
541 GGCAATCCTG TCCCCATCCT CATCCCGTGC CACAGAGTGG TCTGCAGCAG CGGAGCCGTG
601 GGCAACTACT CCGGAGGACT GGCCGTGAAG GAATGGCTTC TGGCCCATGA AGGCCACCGG
661 TTGGGGAAGC CAGGCTTGGG AGGGAGCTCA GGTCTGGCAG GGGCCTGGCT CAAGGGAGCG
721 GGAGCTACCT CGGGCTCCCC GCCTGCTGGC CGAAACTGAG TATGTGCAGT AGGATGGATG
781 TTTGAGCGAC ACACACGTGT AACACTGCAT CGGATGCGGG GCGTGGAGGC ACCGCTGTAT
841 TAAAGGAAGT GGCAGTGTCC TGGGAA
```

Figure 1.21: Nucleotide sequence and exon structure in MGMT cDNA.

Exons 1, 2, 3, 4 and 5 are blue, red, purple, green and black, respectively. The initiation and stop codons are highlighted in yellow. Underlined sequence indicates the coding region. Genbank accession number: NM_002412.3 (274,277).

1.11.1.2 Human MGMT protein structure

X-ray crystallographic studies of unreacted human MGMT protein revealed that it consists of 2 domains and a bound zinc atom. The two domains are an N terminal domain and a C terminal domain, with approximate overall dimensions of 20 Å × 35 Å × 40 Å (278). The N terminal domain spans 1-85 amino acids residues whilst the C terminal domain spans 86-207 residues.

Within the C terminal domain, there is an active site cysteine motif (PCHR, 144-147) that contains the alkyl acceptor residue (Cys-145). In addition, the human MGMT crystal structure revealed a helix-turn-helix (HTH) DNA binding motif which consists of two helices, the first HTH helix is at residues Tyrosine 114- Alanine 127, while the second is called the recognition helix (Alanine127- Glycine136 ; human MGMT numbering) which interacts with DNA (278,279). Although C terminal domains contain all the essential residues for alkyltransferase activity, the C terminal domain retains no activity in absence of N terminal domain (272).

Thus, the role of N terminal domain is potentially related to maintaining the C terminal domain in the active configuration. Occupation of the N terminal domain by zinc increases protein conformation stability and enhances correct folding of the protein conferring a mechanistic enhancement of the repair (280). Crystal structure of MGMT is illustrated in Figure 1.22(A).

1.11.1.3 Repair of O^6-alkylG adducts by MGMT

Human MGMT repair of O^6-alkylG adducts includes a number of distinct steps (Figure 1.22B-C) (281). The first step requires conformational change of MGMT exposing the active site cysteine residue (279). Both DNA binding motifs and MGMT active site are linked *via* an asparagine hinge (Asn-137) which is mainly responsible for stabilizing the two tight turns *via* hydrogen bond formation with the following residues: valine 139, isoleucine 143 as well as the thiol group at the active site cysteine residue 145 (272). During the repair process, MGMT binds to the DNA minor groove *via* the HTH motifs and then the arginine 'finger' (Arg-128) rotates the alkyl modified guanine base outside of the DNA double helix into the active site pocket (279,282).

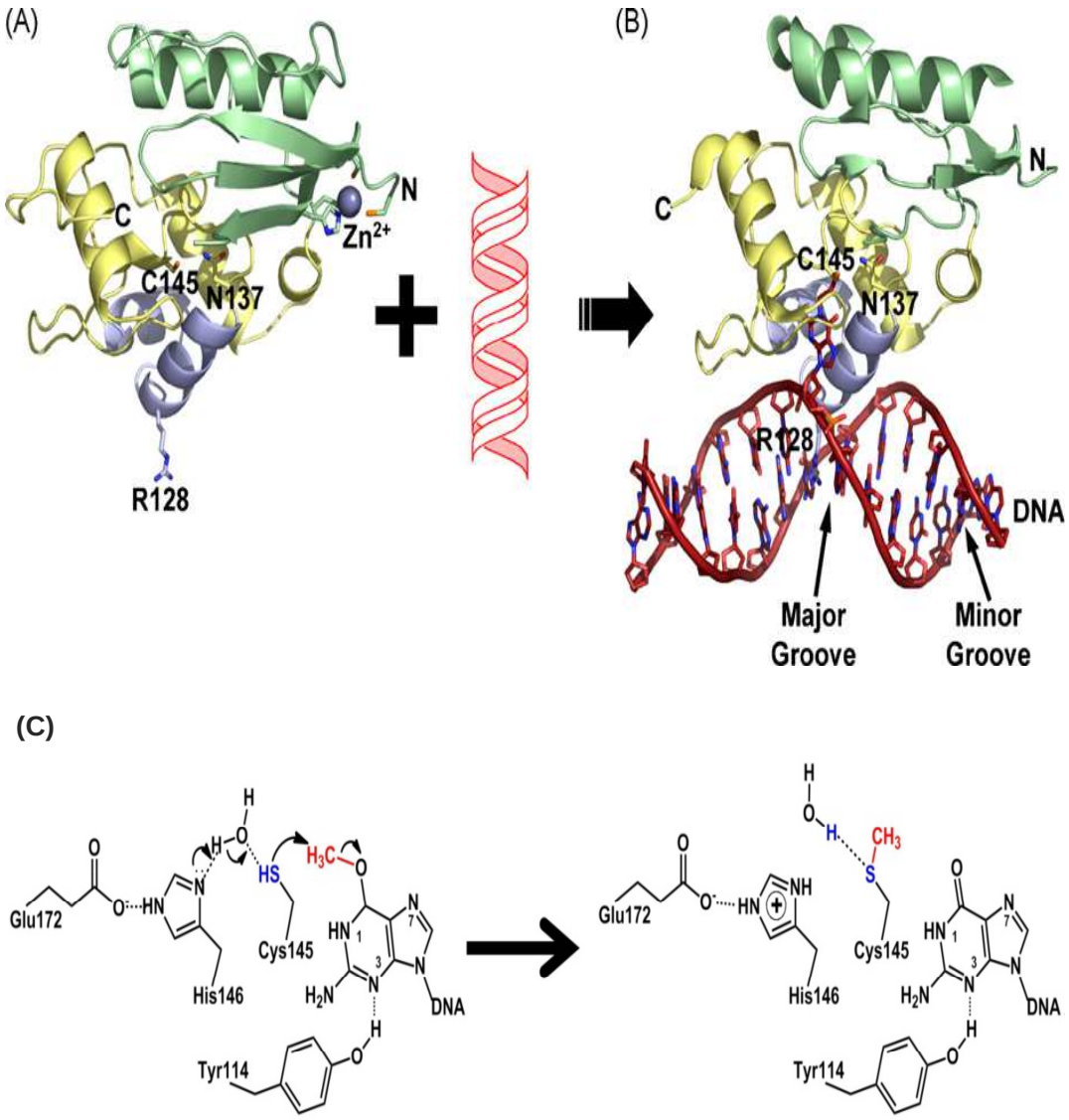


Figure 1.22: Human MGMT X-ray crystallographic structure, zinc site, and DNA binding.

(A) Unreacted MGMT structure. The N-terminal domain is shown in green, the C-terminal domain is shown in yellow, and the HTH motif is shown in blue. The active site cysteine, arginine finger, "Asparginine hinge", and zinc ligands are shown in ball-and-stick representation. (B) DNA-bound MGMT structure. MGMT uses the recognition helix of the HTH motif to bind the minor groove of DNA. (C) Hydrogen bond network and proposed reaction mechanism for MGMT. Figure copied from (279).

This nucleotide flipping is mediated by Arg-128 potentially enters the DNA stack promoting extrusion of target O^6-alkylguanine nucleotides from duplex DNA and stabilizing the gap left by the resultant extrahelical nucleotide (278). Formation of the MGMT-DNA complex does not result in any remarkable structural change to the MGMT protein while it causes the DNA minor grove to widen and bend approximately 15˚ away from the protein to help to flip out the alkylated base from DNA helix (278,282). Then the alkyl group is removed from the O^6 position and transferred to the nucleophilic Cys-145 residue *via* an in-line displacement reaction, rendering MGMT irreversibly inactive. Finally alkylated MGMT dissociates from the repaired DNA duplex and then degraded (283).

MGMT has shown reactivity towards both duplex and SS DNA substrates, however, repair rates seemed to be slower with SS substrates than duplex (284). Moreover, a significant number of publications suggested the role of DNA-mediated conformational change promoting the interaction between MGMT molecules, resulting in cooperative DNA binding which is proposed as the main mechanism for MGMT binding and subsequent repair of SS DNA (284–286). During cooperative binding, MGMT molecules seem to overlap along the DNA contour allowing protein-protein interaction which contributing to cooperativity and DNA-dependant crosslinking of MGMT molecules, as shown in Figure 1.23.

1.11.1.4 MGMT substrate specificity

Human MGMT is almost exclusively specific for repair of O^6-alkylG adducts. Moreover, MGMT demonstrated the ability to repair a wide range of O^6-alkylG adducts containing either aliphatic (methyl, ethyl, hydroxyethyl, carboxymethyl, butyl or propyl) (269), aromatic (benzyl (287) or pyridyloxobutyl (288)) or 4-Bromothenyl (Patrin-2) (289) or 2-chloroethyl alkyl moieties at the modified guanine base. The rates at which MGMT repair O^6-alkylG adducts were found to be influenced by two main factors which are the identity of the alkyl group as well as the surrounding DNA sequence (290,291). Previous reports suggested that MGMT repair efficiency of O^6-alkylG adducts was inversely correlated to the increase in the size of the alkyl group at the O^6 position of the modified guanine which was strongly contradicted by later reports highlighting higher MGMT repair preference for O^6-benzylG adducts relative to O^6-MeG which was ascribed to a potential electronic effect of the benzene ring (275).

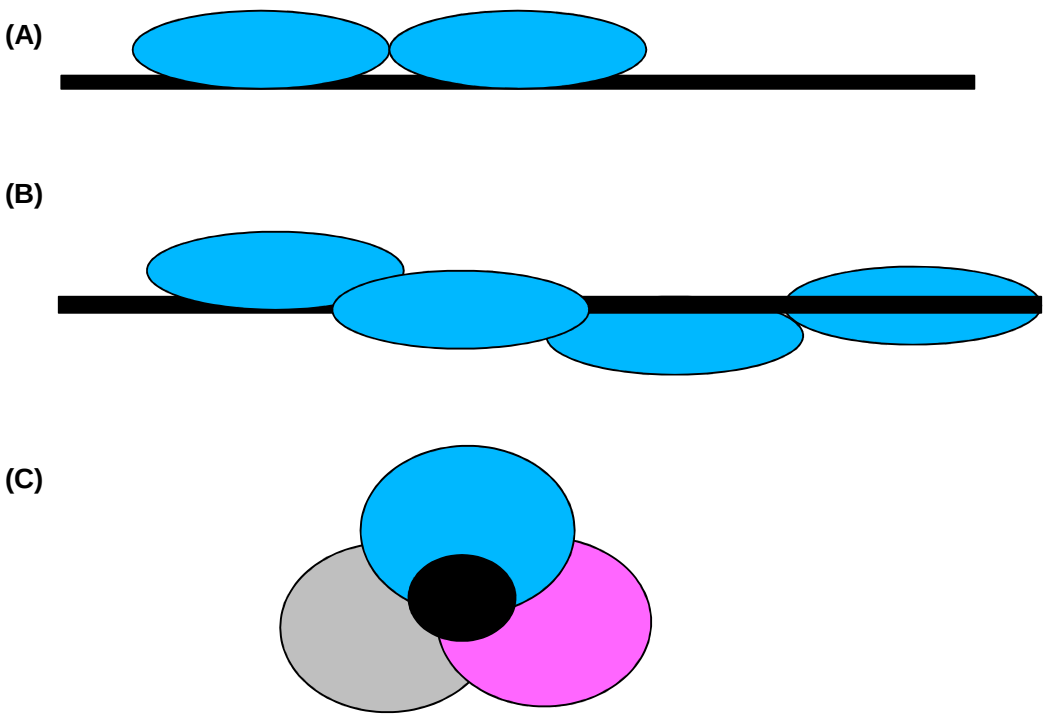


Figure 1.23: MGMT cooperative binding topology.

Black rod and blue filled ovals represent DNA and MGMT, respectively. (A) Non-overlapping adjacent binding. (B) Overlapping binding viewed from perpendicular to the DNA axis, where 2 proteins bind different surfaces of the same DNA segment. (C) Overlapping binding viewed along the DNA axis. The DNA is represented by the black circle in the center. Three proteins, represented by gray, magenta, and blue ovals, are bound with a rotational offset angle of 120°. Figure adapted from (292).

The relative rates of repair are: benzyl > methyl > ethyl > n-propyl, n-butyl (275). Other adducts such as pyridyloxobutyl-, 2-hydroxyethyl, iso-propyl, iso-butyl, tert-butyl are also repaired but at slow rates (293). Table 1.11 summarises the concentrations for both ODNs containing O^6-alkylG adducts and O^6-alkylG free bases required for 50% inactivation of human MGMT (IC_{50} values). These results demonstrate the impact of both the identity of the alkyl group as well as the local DNA sequence context on the ability of O^6-alkylG adducts to inactivate MGMT protein.

Table 1.11: IC_{50} values of oligodeoxyribonucleotides (ODNs) and free base inactivators of human MGMT.

Human MGMT inactivator	IC_{50} (μM)
AAC AGC CAT AT(O^6-MeG) GCC C	0.0157(294)
AAC AGC CAT AT(O^6-benzylG) GCC C	0.0002 (294)
AAC AGC CAT AT[O^6-(4-bromothenylG)] GCC C	0.0001 (294)
AAC AGC CAT AT(O^6-hydroxyethylG(O^6-HOEtG)) GCC C	5.50 (294)
CCG CTA (O^6-MeG)CG GGT ACC GAG CTC GAA T	0.0029 (295)
CCG CTA (O^6-benzylG)CG GGT ACC GAG CTC GAA T	0.0021 (295)
GCC ATG (O^6-MeG)CT AGT A	0.0009 (296)
GCC ATG(O^6-CMG)CT AGT A	0.0017 (296)
O^6-MeG	428 (297), 350 (295)
O^6-benzylG	0.18 (297), 2.7 (295)
O^6-(4-bromothenyl)G	0.0034 (297)

In addition, results shown in table 1.11 highlights that, IC_{50} values for O^6-MeG, O^6-benzylG and O^6-(4-bromothenyl)G bases were higher for free bases than that for the same adducts in ODNs either 13, 16 or 25 mer. The suggested DNA sequence effect on MGMT repair preference is supported by the variation in IC_{50} demonstrated in the results above. The only other possible substrate for MGMT is O^4-methylthymine that is repaired very slowly by human MGMT but at higher rates by other MGMT homologs such as *E.coli* Ogt (298).

1.11.1.5 MGMT expression and activity levels

Both activity and expression levels of MGMT greatly vary among different tissues and tumour types (288). In normal human tissue, the highest MGMT activity levels were found in the liver, colon, esophagus, lung whilst lowest activity levels are reported in brain and myeloid tissue (288–290). Figure 1.24 shows MGMT activities in extracts from human tissue reported by the same study. In addition, previous studies reported a considerable variation in MGMT activity between neoplastic *vs.* non-neoplastic tissues where increased activities were reported in breast, brain and colon tumours relative to adjacent non-malignant tissues (291).

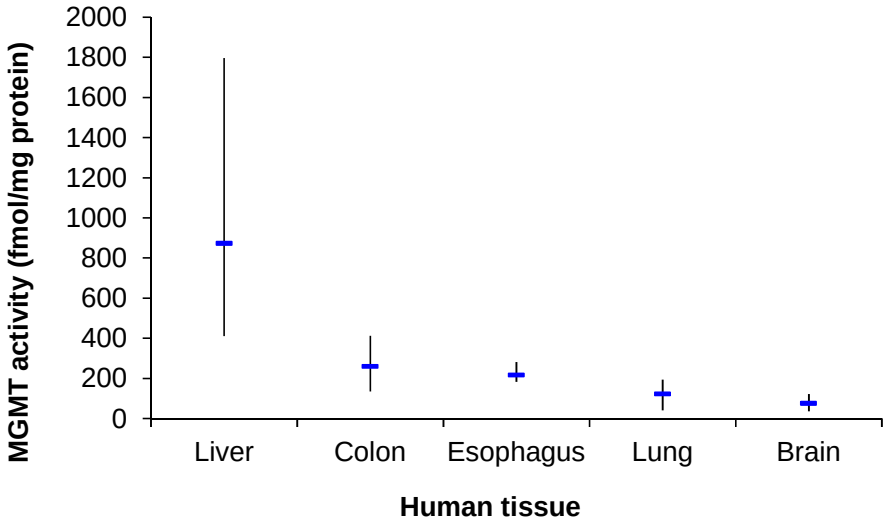


Figure 1.24: MGMT activities in human tissue extracts.

MGMT activity was assessed *via* radioisotopic assay using ^{3}H-methylated DNA. Produced based on data from (299).

Consistently, MGMT activity in both normal and tumour colorectal tissue showed the same pattern of regional disparities where the highest activity was reported in the rectal tissue followed by proximal then sigmoid colon tissues, as shown in Table 1.12 (300). MGMT activity in malignant sigmoid tissue was higher than that of the corresponding non-malignant tissue. Moreover, MGMT activity in normal rectal tissue was significantly higher than that in normal tissues from sigmoid colon. The inter-individual variation in MGMT activity in normal colon and rectal tissues were reported to

be 4-fold and 6-fold, respectively, that rose to 13-fold and 7-fold, respectively, in case of malignancy (300).

Table 1.12: MGMT levels in tumour and macroscopically normal tissue from patients with diseases of the colon and rectum.

Sample set	Normal tissue Mean±SD (n)[1]	Range	Tumour tissue Mean±SD(n)	Range
Benign	98±35 (14)	53-162	-	-
Malignant	107±44 (36)	52-201	146±104 (34)*	35-451
Proximal colon	118±47 (15)	52-201	150±90 (14)	41-320
Sigmoid colon	100±40 (21)	56-194	144±115 (20)*	35-451
Rectum	148±76 (26)**	57-342	170±90 (21)	48-348

[1]Results expressed in fmole/mg protein (n= number of samples). *Activity significantly higher than in normal tissue (P<0.05). **Activity significantly higher than in normal tissue from the sigmoid colon (P<0.05). Reproduced from (300).

Thus tissue heterogeneity in MGMT activity levels mirrors the variation in expression levels as well as the potential presence of expression enhancers (such as diet, cigarette smoke or ionising radiation in human probably *via* induced double strand breaks) or inhibitors (such as low alcohol doses in rats not in humans) (301,302). MGMT expression depends mainly on the methylation status of the CpG islands of the gene promoter region (the presence of 5-methylcytosine in CpG islands). Notably, methylation of these normally unmethylated sites within promoter region of MGMT gene is correlated with MGMT gene silencing and attenuation of MGMT protein expression (276) which is likely to be a result of epigenetic modifications that do not alter genetic information. Furthermore, p53 protein was reported to suppress MGMT transcription in both human tumour cells and animal cells (303,304).

1.11.1.6 MGMT gene silencing and colorectal cancer risk

The epigenetic silencing of MGMT was reported in 30-40% of human colorectal cancers (305). MGMT silencing is particularly important because that results in persistence of O^6-alkylG adducts that induce GC:AT transition mutations which could play a primary role in human neoplasia generally and in colorectal cancer specifically. This finding is supported by a previous study that investigated the MGMT CpG promoter hypermethylation in DNA obtained from 524 primary human tumours, cancer cell lines and normal tissues (306). They reported that all tested normal human tissues were unmethylated at MGMT gene promoter region. However 38% of colon cancer studied samples were methylated (306). Table 1.13 and 1.14 show the frequency of MGMT promoter methylation in cell lines, colorectal cancers, adenoma and normal colorectal tissue reported by previous studies.

Table 1.13: MGMT methylation frequency in microsatellite stable (MSS) and microsatellite instable (MSI) cell lines and human colorectal cancers.

Cell type	MSS	MSI	Total
Colorectal cancer cell lines	5/11 (45%)	5/9 (56%)	10/20 (50%)
Human colorectal cancer	10/25 (40%)	11/28 (39%)	21/53 (40%)

Reproduced from (305).

Table 1.14: Promoter methylation frequency for MGMT gene in colorectal cancer, adenoma, and normal colon.

Tissue	Frequency of MGMT methylation	*P* value
Colorectal cancer	25.5%	0.003[a]
Adenoma	21.1%	NS[b]
Normal	0%	0.012[c]

[a]Colorectal cancer *vs.* normal, [b]Colorectal cancer *vs.* adenoma, [c]Adenoma *vs.* normal. Reproduced from (307).

MGMT silencing is potentially associated with G to A transitions in the K-ras oncogene as well as G:C to A:T transition mutations in TP53 tumour suppressor gene. Previous publication demonstrated that, of 314 colorectal tumours (249 carcinomas and 65 adenomas), 126 demonstrated MGMT gene promoter hyper-methylation of which 43 (34%) showed GC:AT transition mutations in TP53 gene, whereas, MGMT unmethylated tumours only showed GC:AT changes in 37 of 188 (19%) tumours (125). Their data suggested that epigenetic silencing of MGMT by promoter hypermethylation was significantly associated with the presence of G:C to A:T transition mutations in TP53 (*P*= 0.01) (125).

Furthermore, another study showed that, among all studied colorectal cancer cases, 51 cases displayed G to A transition mutation in the K-ras oncogene, 37 displayed other K-ras mutations and 156 cases did not show any mutations in K-ras oncogene. In addition, 71% of the cases displayed G to A transition mutation in the K-ras oncogene had MGMT promoter hypermethylation, whilst only 32% of those with other K-ras mutations demonstrated MGMT methylation (*P*= 0.002) (308).

MGMT gene silencing by promoter hyper-methylation was proposed as a mediator of sporadic colorectal cancer as MGMT promoter methylation was detected in 97%, 37% and 27% of colorectal tumour tissue samples, normal adjacent mucosa (normal mucosa of colon cancer patients) and normal control samples, respectively, by a previous study. In addition, GC:AT transition mutations in K-ras oncogene was detected in 12% and 3% of colorectal tumours with methylation and colorectal tumours without methylation, respectively (309).

Consequently, variation in MGMT activity in human tissues determines tissue susceptibility to GC:AT transition mutations in K-ras oncogene induced by a wide range of AAs either in human tissues or experimental animals. It was reported that MGMT activity was significantly lower in colorectal tissue with K-ras GC:AT transition mutation, not transversion, compared to that in normal colorectal tissue without the specified mutation suggesting MGMT potential protection of colorectal tissue against K-ras GC:AT transition mutation induced by O^6-alkylG adducts (310). Animal studies reported that MGMT overexpression in transgenic rodents provides protection against AA-induced K-ras GC:AT transition mutation (311).

1.11.2 MGMT like proteins

Alkyl transferase like proteins, known as Atl proteins, is another defense mechanism or pathway to protect normal cells against DNA alkylation damage that has been identified in many unicellular species (312). They differ from MGMT in that; their alkyl acceptor putative site is either tryptophan or alanine but not cysteine (302,313,314). Atl proteins can bind to a wide range of O^6-alkylguanine adducts in DNA and can even reverse their binding to MGMT (302,312–314). Atl proteins lack alkyltransferase either from DNA or ODNs, glycosylase, or endonuclease activities (315). However, Atls use nucleotide flipping to access the alkyl modified guanine bases as MGMT. They bind to alkylated DNA in a similar manner to AGT but bend the DNA more extensively by 45° forming a complex recognized by the NER system (Atl-NER pathway).

Most known Atl proteins are from bacteria or fungi but they have not been identified in higher euokaryotes or plants. The three best studied bacterial Atl were obtained from *E.coli* (313) and *Thermus thermophiles* (316) and *Vibrio parahaemoluticus* (317). However, the most characterized fungal Atl was that obtained from fission yeast *Schizosaccaromyces pombe* known as (Atl1) (314). On the other hand, the first Atl in a multicellular organism was that identified in the starlet sea anemone *Nematostella vectensis* (315). It's possible that in organisms such as *E.coli* that has alkyl transferase proteins, Atl proteins are responsible for the recognition of the bulky O^6-alkylG adducts that are not good substrates for alkyl transferases.

1.11.3 Base excision repair (BER)

BER commonly recognizes and repairs DNA base modifications caused by a number of mechanisms including: base deamination, alkylation, oxidation (reactive oxygen species) and hydroxylation (318). BER repairs 7 and 3-alkyl guanine adducts as well as 7 and 3-alkyl adenine adducts (319). BER pathway includes 5 distinct steps. The first is the recognition of DNA damaged base, followed by its removal via glycosylase enzyme. This step leads to the formation of an apurinic or apyrimidinic intermediate. The second step in BER pathway includes cleavage of the phosphodiester backbone 5` to the AP site using AP endonuclease (320). Then, the

third step encompasses removal of 5` sugar fragment. Insertion of a new correct base takes place in the fourth step by DNA polymerase enzyme. The last step is mainly sealing the DNA break by DNA ligase (321). It is noteworthy to mention that, DNA alkyl lesions are removed by lesion-specific DNA glycosylases that excise modified bases to create abasic sites and initiate the BER pathway (322). A previous publication documented the role of BER pathway in protection against AOM-induced colon carcinogenesis in mice as they demonstrated that BER deficient animals were more susceptible to AOM-induced colon carcinogenesis compared to wild type (323).

1.11.4 Nucleotide excision repair (NER) pathway

NER pathway is the most versatile and flexible DNA repair mechanism. Mainly, NER is responsible for repairing base lesions that distort DNA double helix structure, interfere in base pairing and eventually block DNA duplication and transcription. Both UV-induced DNA damage and bulky DNA adducts induced by chemical agents are repaired by NER pathway (324).

The basic NER mechanism includes recognition of the DNA lesion followed by opening of the double helix by helicase enzyme. Then cleavage of the damaged strand a few bases away from the lesion takes place by the action of endonuclease. This is followed by removal of the DNA segment with subsequent gap polymerization using the intact DNA strand as template (325).

NER pathway includes two sub-pathways which are global genome NER (GG-NER) and transcription-coupled NER (TC-NER). The former pathway plays an important role in removing DNA lesions from transcriptionally active genes whilst the later does the same function but for the transcriptionally inert genes (326). The role of NER in protecting cells against the mutagenic and cytotoxic effects of AAs was proposed by a study that reported the ability of NER to act on O^6-butylG adduct in rodent cells (327).

1.11.5 Mismatch repair (MMR) pathway

MMR pathway plays a vital role in repairing DNA replication errors resulting from failure of proofreading by DNA polymerase. In general, DNA polymerase enzymes possess exonuclease activity to excise incorrect bases from the newly formed DNA strand; however, failure of this process leads to mispair persistence. Moreover, MMR is responsible for both recognition and repair of insertions/deletions loops within microsatellites. Therefore, MSI is recognized as MMR failure or impairment. Both recombination and cell death are dependent upon MMR system (328).

MMR recognition of DNA replication errors is mediated by either heterodimer 'MutSα' comprising MSH2–MSH6 or 'MutSβ' heterodimer comprising MSH2–MSH3 based upon the size of the mismatch error. The second step involves recruitment of another heterodimer known as 'MutLα' by the MutS heterodimer. 'MutLα' comprises MLH1 and PMS2 and it forms a complex with MutS at the DNA binding site (329). MutS/ MutL complex start to translocate in both direction of the newly synthesised DNA strand until a strand break is encountered. Then, exonuclease-1 is loaded at the strand break and degrades the strand back towards the mismatch site, while the final stage allows the repair of the gap by DNA polymerase (using the parent strand as a template) and DNA ligase that seals the gap (330).

Notably, MMR system could initially recognize the O^6MeG:T mispair formed following one round of replication. Other components of the MMR system are recruited and a long patch (>1 kb) of the newly replicated strand is removed, as illustrated in Figure 1.25.

1.11.6 Double strand break repair (DSBR) pathway

Double strand breaks are one of the most serious forms of DNA damage. These breaks, if unrepaired, activate disastrous consequences such as genomic instability and cell death. Therefore, DSBR is one of the most critical DNA repair mechanisms. Double strand breaks can result from the effect of reactive oxygen species or ionising radiation on DNA (331).

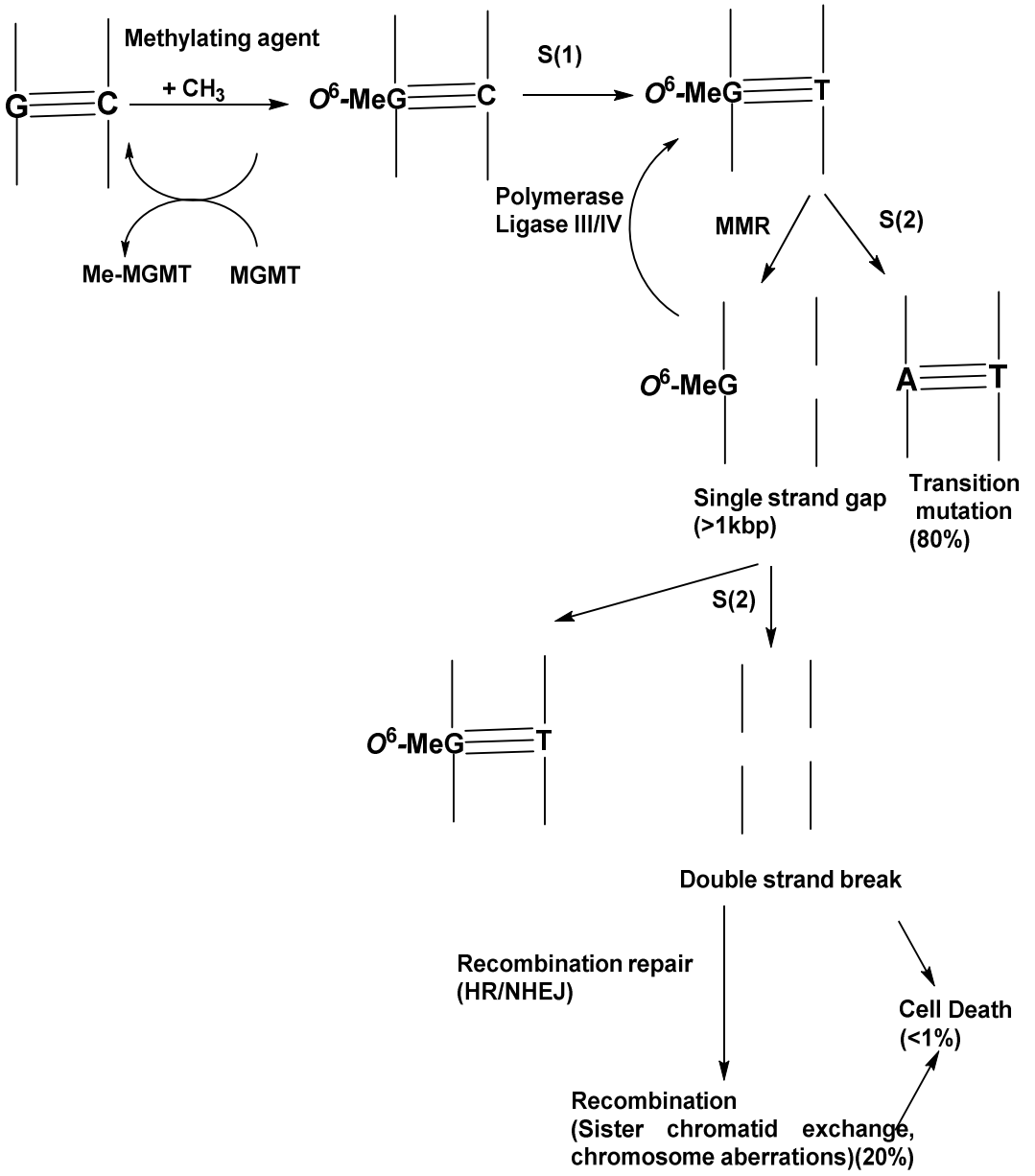


Figure 1.25: Pathways of the biological effects of formation of O^6-MeG adduct in mammalian cells.

S(1) and S(2) indicate the first and second rounds of DNA. MMR corresponds to mismatch repair, MGMT to O^6-methylguanine O^6-methyltransferase, HR and NHEJ to the homologous recombination and non-homologous end joining pathways, respectively. The relative amounts of mutation, recombination and killing events by O^6-MeG are as reported for repair-deficient human cells in culture and are from Rasouli-Nia et al. (332). Produced from (212).

DSBR pathway involves two other sub-pathways which are homologous recombination (HR) and non-homologous end joining (NHEJ). HR operates in dividing cells but NHEJ can function in either dividing or non-dividing cells, regardless the cell cycle (333). HR requires homologous DNA sequence as a template to guide the repair process and gap filling, for this reason, it performs error free repair (334). Whilst, NHEJ mainly provides molecular integrity rather than sequence fidelity this is why it's error prone and could lead to insertion or deletion mutations. NHEJ utilises Ku heterodimer for recognition of the lesion and functions through tethering the two broken ends together (335). HR is the major DSBR mechanism during cell replication as it conserves genome integrity and it's up regulated in S and G2 phases while it's supressed in G1 phase of the cell cycle (336).

In summary, sporadic colorectal cancer is a multifactorial disease in which multiple exposures to either endogenous or dietary carcinogens interact with the individual genetic background in a complex manner to modulate colorectal malignancy risk (337,338). The hypothesis that AAs are risk factors in the aetiology of colorectal cancer was strongly supported by previous reports stating the detection of O^6-MeG and O^6-CMG adducts in human colorectal DNA (185,257–259) which both result from the nitrosation of glycine. In addition to O^6-MeG and O^6-CMG, more O^6-alkylG adducts were believed to be present in human colorectal DNA as human gut lumen contains a diverse assortment of nitrosable precursors (amines and amides) (339) that could be readily nitrosated by endogenously formed nitrosating agents to form NNOC (45). These NNOC can potentially generate a variety of alkyl DNA adducts including pro-mutagenic O^6-alkylG adducts. This alkyl adductome has not yet been characterised. In addition, NNOC formation cannot be accurately estimated from meat consumption alone due to *in situ* processing required for their formation.

1.12 Aims and objectives

For these reasons, the overall aim of this study was to assess NNOC exposure *via* determination of the overall load of O^6-alkylG adducts in human colorectal DNA and then identify and quantify the individual O^6-alkylG adducts as risk factors in the pathogenesis in human colorectal cancer. To investigate it, we have exploited the action of the DNA repair protein, MGMT, which covalently transfers the alkyl group from O^6 position of modified guanine to a cysteine residue at the active site. Following

trypsin treatment, the alkyl modified ASP can be detected by MALDI-TOF MS. The development of this novel approach has been carried out in a multistep manner. The specific objectives of the current work were:

1- Purification, characterization as well as determination of functional activity of both maltose binding protein-MGMT (MBP-MGMT) as well as his tagged-MGMT (his-MGMT) fusion proteins.

2- To investigate whether O^6-alkylG adducts (O^6-CMG and O^6-CEG) are substrates for MGMT and recognized by Atl1.

3- To investigate the mechanism of MGMT inactivation by O^6-MeG, O^6-CMG and O^6-CEG adducts and to confirm that this was a consequence of the transfer of alkyl group from O^6-alkylG containing ODNs to the MGMT-ASP using MALDI-TOF MS.

4- To develop a novel MALDI-TOF MS based methodology that would initially detect and quantify O^6-MeG adduct in TMZ modified CT-DNA using the known action of MGMT.

5- To employ the newly developed MS based assay to assess NNOC exposure by identifying then quantifying O^6-alkylG DNA adducts that result from their exposure in human colorectal DNA.

2 Chapter 2: Materials and Methods

2.1 Materials and chemicals

All chemicals used in this work were purchased from Sigma-Aldrich (Poole, Dorset, UK) unless otherwise stated.

E.coli bacteria harboring MBP-MGMT fusion protein plasmid, purified his-MGMT, purified MBP-Atl1 fusion protein, purified calf thymus DNA (CT-DNA), [^{3}H]methylated CT-DNA (synthesized in house) as well as TMZ-modified CT-DNA (CT-DNA that has been methylated *in vitro* using TMZ) were supplied by Prof. Geoffrey Margison.

CT-DNA was deproteinised by dissolving CT DNA in Tris-EDTA buffer (TE, 10 mM Tris-HCl, 1 mM disodium ethylenediaminetetra acetic acid (EDTA), pH 8) at room temperature (RT), treating with proteinase K and extracting with equal volumes of phenol/chloroform/isoamyl alcohol. DNA was then precipitated by two volumes of ethanol and 0.1 volume of 3M sodium acetate pH 5, then dissolved and diluted in TE buffer.

Our study was conducted using both control G ODNs (with unmodified G bases) as well as O^6-alkylG containing ODNs that contained a single O^6-alkylG adduct in a defined DNA base sequence. Normal control G (13 and 23 mer), O^6-MeG (13 and 23 mer), SIMA labelled complementary ODN (13 mer), biotin labelled complementary ODN (23 mer) and HEX labelled complementary ODN (23 mer) were synthesized by Sigma-Aldrich UK. Whilst, O^6-CMG (13 and 23 mer), O^6-CEG (13 and 23 mer), O^6-[4-oxo-4-(3-pyridyl)butyl]G (O^6-pobG; 13 mer) and O^6-HOEtG; 23 mer containing ODNs were synthesized using phosphoramidite by Dr David Williams, Centre of Chemical Biology, Department of Chemistry, University of Sheffield. Structure of the modified guanine residues in the ODNs are shown in Figure 2.1 and sequence of all used ODNs are listed in Table 2.1. Three different purifications of LH1 were analysed as illustrated in Figure 2.2. They consisted of the original purified ODN (LH1a), and two subsequent re-purified samples (LH1b and LH1c) that were purified from LH1a by Mr Liam Hanson by HPLC.

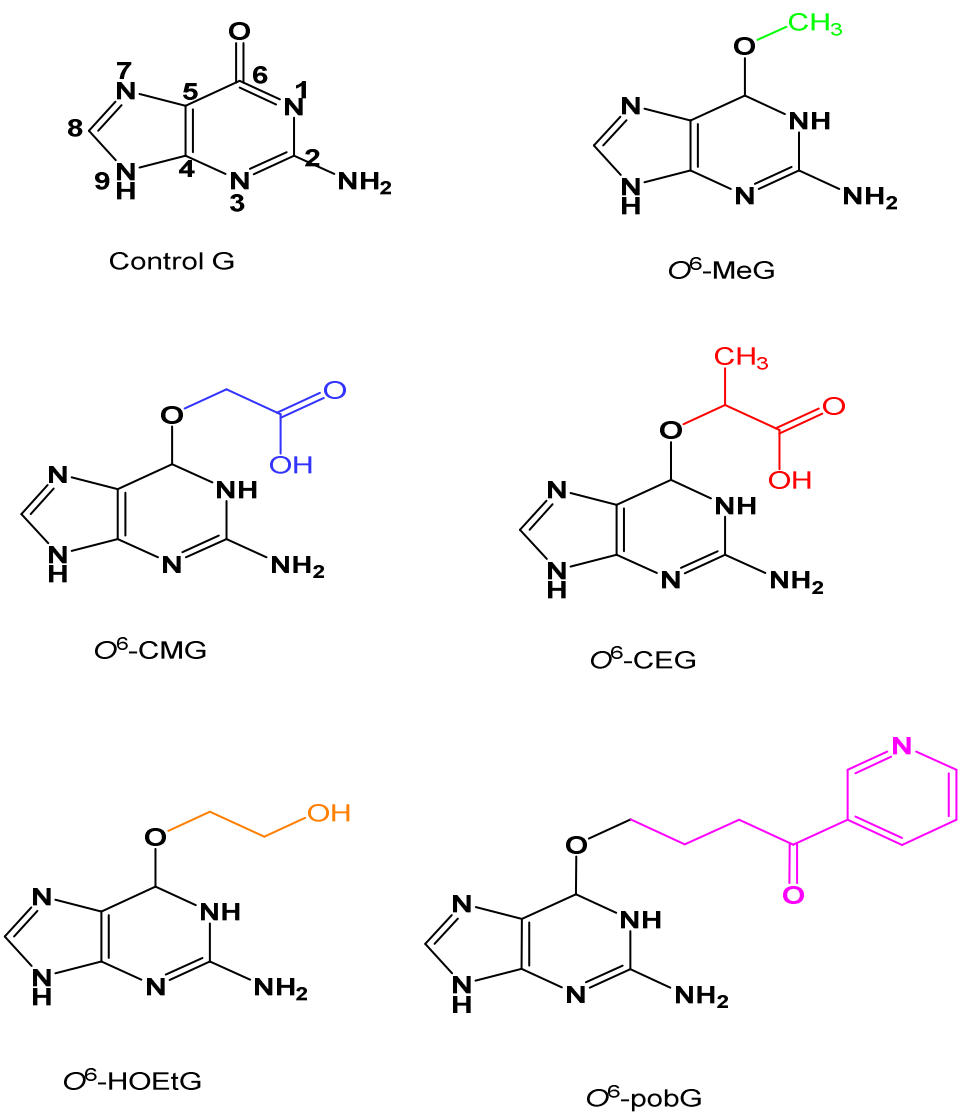


Figure 2.1: Structure of the modified guanine residues in the ODNs used in this study.

Alkyl groups at the O^6 position of the modified guanines are shown in colour.

Table 2.1: Sequences of ODNs used in this study.

Number	Modification	Length	ODN sequence
374	Unmodified Guanine	13 mer	5'- [SIMA]GC CAT GGC TAG TA-3'
16		23 mer	5'-GAA CTG CAG CTC CGT GCT GGC CC-3'
ConG		23 mer	5'-GAA CTG CGG ATC CGT GCT GGC CC-3'
303B	O^6-MeG	13 mer	5'- [SIMA]GC CAT GXC TAG TA-3' (X =O^6-MeG)
15		23 mer	5'-GAA CTX CAG CTC CGT GCT GGC CC-3' (X = O^6-MeG)
GMe		23 mer	5'-GAA CTG CGX ATC CGT GCT GGC CC-3' (X = O^6-MeG)
OW18	O^6-CMG	13 mer	5'- [SIMA]GC CAT GXC TAG TA-3' (X =O^6-CMG)
DW1		23 mer	5'-GAA CTG CGX ATC CGT GCT GGC CC-3' (X = O^6-CMG)
44		23 mer	5'-GAA CTX CAG CTC CGT GCT GGC CC (X = O^6-CMG)
SRO1	O^6-CEG	13 mer	5'- [SIMA]GC CAT GXC TAG TA-3' (X =O^6-CEG)
RA1		23 mer	5'-GAA CTX CAG CTC CGT GCT GGC CC-3' (X =O^6-CEG)

Table 2.1: Sequences of ODNs used in this study (continued).

Number	ODN type	Length	ODN sequence
LH1	O^6-CEG	23 mer	5'-GAA CTX CAG CTC CGT GCT GGC CC-3' (X = O^6-CEG)
OW17	O^6-pobG	13 mer	5'- [SIMA]GC CAT GXC TAG TA-3' (X =O^6-pobG)
57	O^6-HOEtG	23 mer	5'-GAA CTX CAG CTC CGT GCT GGC CC-3' (X =O^6-HOEtG)
113	ODN complements	13 mer	5'-B TAC TAG CCA TGG C-3'
43		23 mer	5'-HEX- GG GCC AGC ACG GAG CTG CAG TTC
C	DW1 complement	23 mer	5'-GG GCC AGC ACG GAT CCG CAG TTC-3'

* SIMA & HEX are fluorophores, and B indicates biotin.

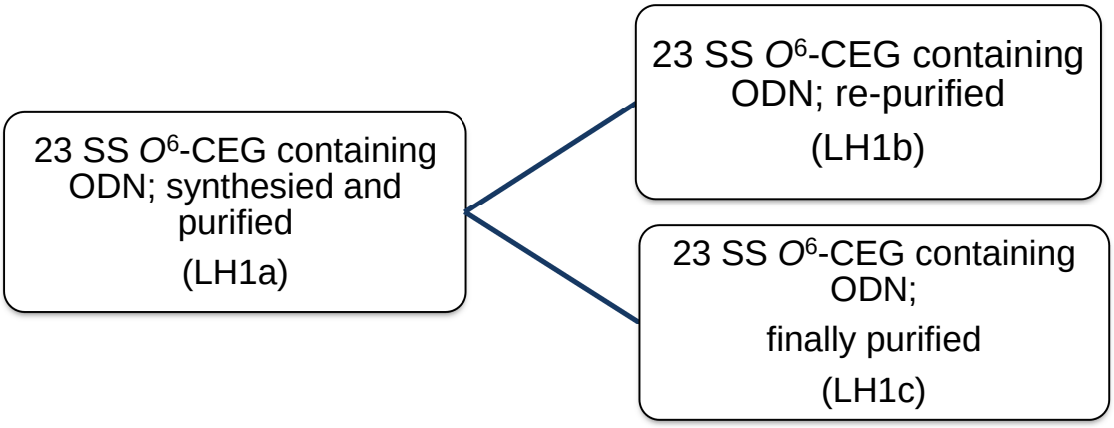


Figure 2.2: LH1 preparation used in the current study.

2.2 Methods

2.2.1 MBP-MGMT plasmid extraction

The human MGMT gene (300 kb) was cloned into a fusion protein expression vector, the MBP-MGMT plasmid (pMAL plasmid – c2X, pMAL – p2X Polylinker, New England BioLabs). MBP-MGMT plasmid constructs were transformed into competent XL-1 blue *E.coli* and streaked on agar plate. Colonies were picked and inoculated in 5 mL Luria broth medium containing ampicillin (100 µg/mL) and glucose (0.2% w/v), and incubated overnight at 37°C in a shaker (Orbital incubator SISO, Stuart). The overnight culture was harvested by centrifugation at 17900g in a cooling centrifuge (Heraeus Biofuge 15R), for 10 minutes at 4°C.

Plasmid DNA was extracted using the QIAprep® Spin Miniprep Kit (Qiagen Ltd, UK) according to the manufacturer's instructions. Pelleted bacterial cells were re-suspended in 250 µL buffer P1 (50 mM Tris-HCl, pH 8; 10 mM EDTA 100 µg/mL RNase A) and lysed by addition of 250 µL buffer P2 (200 mM sodium hydroxide, 1% sodium dodecyl sulfate). After 5 minutes incubation at RT, the mixture was neutralized by the addition of 350 µL buffer N3 (4.2 M guanidinium hydrochloride, 0.9 M potassium acetate, pH 4.8). The sample was centrifuged for 10 minutes at 17900g (Fisher Scientific accuspin™ Microcentrifuges) and the supernatants were transferred to a QIAprep spin column, and allowed to flow through the column. The column was then

washed with 0.75 mL buffer PE (10 mM Tris-HCl, 80% ethanol) and plasmid DNA was eluted with 0.05 ml buffer EB (10 mM Tris-HCl, pH 8.5).

2.2.2 DNA quantification by PicoGreen assay

The MBP-MGMT plasmid DNA concentration was determined by a PicoGreen assay with purified CT-DNA used to prepare a standard curve. CT-DNA was diluted in 1X TE buffer to concentrations between 0-2 µg/mL. The stock PicoGreen dye (Invitrogen, USA) was diluted to a working solution with TE buffer (1/1000). Ten µL of each standard were added to a microtitre plate in duplicate and 100 µL of diluted PicoGreen dye was added to each well. The fluorescence was measured using a TECAN Genios plate reader (excitation wave length 485 nm, emission wave length 535 nm, integration time 40 µseconds). The standard curve was accepted when R^2>0.99. The miniprep DNA concentration values were extrapolated from the CT-DNA standard curve using Magellan v3 software.

2.2.3 MBP-plasmid sequencing

The MBP-MGMT plasmid DNA sequences were obtained using a forward malE primer (5'–GGTCGTCAGACTGTCGATGAAGCC–3') (New England Biolabs Inc., USA). The miniprep DNA (300 ng) was mixed with 0.5 µL of the primer in a total volume of 10 µL. Sequencing was carried out at the Sequencing facility at the University of Manchester. The sequence data obtained was analyzed using Blast, available at http://www.ncbi.nlm.nih.gov/blast/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome.

2.2.4 Expression of MBP-MGMT fusion protein

To express the MBP-MGMT fusion protein, a suspension of the *E-coli* bacteria harboring MBP-MGMT plasmid was inoculated into growth medium composed of 10 mL of Luria broth containing both ampicillin (100 µg/mL) and glucose (0.2% w/v), and incubated overnight at 37°C in a shaker (Orbital incubator SISO, Stuart). Then 5 mL of

the overnight mixture was diluted to 100 mL using the same growth medium and incubated until the optical density was about 0.6 at 595 nm. When this optical density was reached, MBP-MGMT fusion protein expression was induced by the addition of both isopropylthiogalactoside (IPTG, final concentration of 0.4 mM) and Zinc chloride (final concentration 0.1 mM). Three hours later, cells were harvested by centrifugation at 4820 g for 30 minutes at 4°C. After centrifugation, the supernatant was aspirated and cell pellets were harvested and stored at -80°C until they were used for extraction and purification of the MBP-MGMT fusion proteins.

2.2.5 Extraction of MBP-MGMT fusion protein

On ice, bacterial cell pellets were re-suspended in 10 mL of extraction buffer (20 mM Tris-HCl pH 7.6, 25 mM NaCl, 1 mM EDTA, 5 mM dithiothreitol (DTT) containing 5 µg/mL leupeptin and a whole Complete Protease Inhibitor Cocktail tablet (Roche, Germany). The temperature of the bacterial suspension was measured and the sample was sonicated (10 sec pulse, setting 4.5; Sonicator XL; Bandelin Sonoplus Electronics) until the final temperature reached 37°C. After sonication, 100 µL of phenylmethylsulfonyl fluoride (PMSF) (8.7 mg/mL solution in absolute ethanol) was added immediately to the bacterial extracts and then extracts were centrifuged for 10 minutes at 17900g at 4°C cold centrifuge (Heraeus Biofuge 15R). After centrifugation, supernatants of the bacterial extracts were collected and then applied to the amylose column for the purification of MBP-MGMT fusion protein as described in Section 2.2.6.

2.2.6 Purification of putative MBP-MGMT fusion protein

MBP-MGMT fusion protein in the bacterial cell extracts was affinity purified using amylose resin (New England Biolabs Inc., USA) (313). Amylose resin was packed into an empty glass column (5 × 2.0 cm) washed with 500 mL of binding buffer (BB; 20 mM Tris-HCl pH 7.6, 200 mM NaCl, 1 mM EDTA,) at 1.2 mL/minute (P-1 Pump, Pharmacia Biotech) containing 5 mM DTT. UV absorbance of the elute was monitored at 280 nm with a UV detector (Uvicord S II, LKB Bromma) connected to a chart recorder (10 mV, 1 mm per minute, 2210 Chart recorder, LKB Bromma). After washing, bacterial extracts were loaded onto the column at the same flow rate and the column was washed with BB until complete elution of the unbound proteins' peak as

monitored using chart recorder. Putative MBP-MGMT protein was eluted by the addition of 10 mM maltose to either amylose BB (conventional elution buffer; 10mM maltose in BB containing 2 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP)) or 200 mM ammonium bicarbonate (AMBIC) with 2 mM TCEP (MS compatible elution buffer; 10 mM maltose in 200 mM AMBIC containing 2 mM TCEP). The putative MBP-MGMT peak was collected then protein concentration of various fractions was determined as described in Section 2.2.7.

2.2.7 Quantification of protein by Bradford assay

The putative MBP-MGMT protein concentration determined using the Bradford Protein Assay with bovine serum albumin (BSA) used to prepare a standard curve. BSA was diluted in buffer I (50 mM Tris-HCl pH8.3, 1 mM EDTA, 5 mM DTT) to concentrations between 0-0.1 mg/mL. Bradford protein reagent (Bio-Rad) was diluted 1 in 5 in double distilled water (ddH_2O) to a working solution. Forty µL of each standard were added to a microtitre plate in duplicate and 200 µL of Bio-Rad reagent was added to each well. The UV absorbance was measured at 595 nm using a TECAN Genios plate reader. The standard curve was accepted when R^2>0.99. Protein concentration values were extrapolated from the BSA standard curve using Magellan v3 software.

2.2.8 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of purified proteins

Amylose column purified proteins were analyzed using SDS-PAGE. A 12% resolving gel was prepared by mixing 8 mL polyacrylamide protogel acrylamide/bisacrylamide (37.5:1, National Diagnostics UK), 5 mL 1.5 M tris HCl (pH 8.8), 6.6 mL ddH_2O, 200 µL 10% SDS, 200 µL 10% ammonium persulfate (APS) and 20 µL tetramethylethylenediamine (TEMED; Bio-Rad Lab USA). This mixture was then poured into the gel-casting cassette and water-saturated butanol was layered on top. When the resolving gel had polymerized (within 20 minutes), the water saturated butanol layer was removed. An upper (5%) stacking gel was prepared by mixing 3.4 mL polyacrylamide protogel, 2.5 mL 1.0 M Tris HCl (pH 6.8), 13.6 mL ddH_2O, 200 µL 10% SDS, 200 µL 10% APS and 20 µL TEMED and this solution was poured onto the top of the resolving gel in the gel-casting cassette and a comb was carefully inserted so

as to avoid trapping air bubbles. After the stacking gel had set, the plate was attached to the electrode block and wells were washed by moving the comb up and down in running buffer (0.192 M glycine, 25 mM pris-Base, 0.1% w/v SDS). Protein samples (75 µL) were mixed with 25 µL of 4X loading buffer (4XLB, 50mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 1% β-mercaptoethanol, 12.5 mM EDTA, 0.02% bromomphenol blue, NuPAGE™ LDS Sample Buffer; Novex®, Life Technologies USA), heated at 95°C for 5 minutes and centrifuged for 30 seconds at 17900g (Fisher Scientific accuspin™ Microcentrifuges) and layered carefully into the wells. 30 µg of protein load and unbound protein and 10 µg of bound protein peak (amylose purified MBP-MGMT or 5 µg of his-MGMT) were loaded into the gel wells. The protein standard marker (75 µL) (New England Biolabs Inc.) was mixed with 4xLB (25 µL), and 10 µL of this marker mixture was loaded into the gel. Electrophoresis was carried out at 200V, until the bromophenol blue dye was nearly 1 cm from the bottom of the resolving gel. The gel was removed and stained for 30 minutes with Instant blue Coomassie based stain (Expedeon; containing Coomassie dye, ethanol, phosphoric acid and solubilizing agents in water). Destaining of Instant blue stained gels was carried out with ddH_2O for 30 minutes.

2.2.9 Western blot analysis

Western blot analysis was carried out using anti-MGMT (in-house rabbit polyclonal primary) or anti-MBP (mouse monoclonal, Abcam, UK) antibody to confirm that purified protein fractions contained MGMT or MBP. Proteins separated by 12% SDS-PAGE (Section 2.2.8) were transferred by wet blotting onto a nitrocellulose membrane (Whatman, GE Healthcare) by laying the membrane onto the gel for 10 seconds. The membrane was then blocked using 5% PBSTM (5% marvel (Premier Foods, UK) in 0.1% Tween-20 in phosphate buffered saline (PBST; Life Technologies USA)) for 1 hour on a shaker (Orbital incubator SISO, Stuart) at RT, washed once with PBST for 5 minutes. The membrane was incubated for 1 hour on the shaker at RT with primary antibody (anti-MGMT R63 or anti-MBP antibody, diluted (1:500) in 0.5% PBSTM). The primary antibody was decanted, the membrane was washed 3 times with PBST for 5 minutes and then incubated for 30 minutes with the secondary antibody, Polyclonal Goat Anti-Rabbit Immunoglobulin conjugated with HRP (Dako) or Goat Anti-Mouse IgM Secondary Antibody (Thermo Scientific USA)(diluted 1:1000 in 0.5%

PBSTM) on a shaker at RT. The membrane was washed 3 times with PBST for 5 minutes each time.

Enhanced chemiluminescence (ECL) Plus western blot detection reagent (Amersham, GE Healthcare Limited, UK) was prepared by mixing equal volumes of detection reagents 1 (peroxide solution) and 2 (Luminol enhancer solution) ECL reagent (1 mL) was put on a cling film, and the membrane was laid on the cling film with the protein side downwards.

Developer and fixer solutions were prepared by mixing 100 mL of the developer or the fixer with 360 mL ddH_2O. The autoradiography film (Amersham Hyperfilm™ ECL, GE Healthcare Limited, UK) was placed in a cassette, emulsion side up, the membrane was placed on the film (for between 10-30 seconds). The x-ray film was then immersed in the developer solution until the bands were clearly seen, washed briefly with ddH_2O and immersed in fixer for 3 minutes, and washed in ddH_2O. Finally, the developed film was scanned (HP Scanjet 4850).

2.2.10 Determination of MGMT functional activity

MGMT activity was assayed by measuring the transfer of [^{3}H] from N-[^{3}H]-methyl-N-nitrosourea (Hartmann Analysis: 80 Ci/mmole) methylated CT-DNA to MBP-MGMT fusion protein (340).

Varying amounts of MGMT protein were incubated initially at 37°C for 90 mins with 50 µL of [^{3}H]-methylated CT-DNA substrate in a total volume of 200 µL of buffer I containing BSA (1 mg/mL; IBSA). After 1 hour of incubation, 100 µL of 30 mg/mL BSA, 100 µL 4M perchloric acid (PCA), 2 mL of 1M PCA were added and the mixture was incubated for 50 minutes at 75°C to hydrolyze DNA. Samples were then centrifuged for 10 minutes at 2700 g (Sorvall RC-3B Plus) at 20°C and the supernatants were carefully aspirated without disturbing the pellets and 4 mL of 1M PCA was added to the pellets. Centrifugation and aspiration steps were repeated and 300 µL ddH_2O were added to each tube followed by 3 mL Ecoscint scintillation fluid (Mensura Technology) and the radioactivity was determined using scintillation counter (1900 TR Liquid Scintillation Analyzer, Packard Tri-Carb Liquid Scintillation Counter). Then the counts per minute were plotted against amounts of each analyzed fraction and the slope of the line was

determined and used to calculate the MGMT specific activity by calculating the specific activity of the [^{3}H] and the counting efficiency of the machine: 1 fmole of MGMT was equivalent to 34.6 counts per minute (296).

2.2.11 Tandem mass spectrometry analysis

Tandem MS analysis was performed by Bio-MS facility, School of Biological Sciences as service. Protein samples were supplied in SDS-PA gel, as a gel top band, and the service carried out the in-gel digestion as well as LC Orbitrap tandem mass spectrometry analysis.

2.2.12 Annealing of SS ODNs to ODN duplexes

SS ODNs were annealed to complementary strands *via* heating with equimolar amounts of the ODN complement to 95˚C in 50 mM NaCl for 20 minutes then allowed to cool down slowly to RT for >1 hour then stored in -20˚C. Double stranded (DS) ODNs were processed on 20% non-denaturing PAGE analysis.

2.2.13 Non-denaturing (20%) PAGE of DNA fragments

A 20% non-denaturing polyacrylamide gel was prepared by mixing 10 mL polyacrylamide protogel acrylamide/bisacrylamide (19:1, National Diagnostics UK), 9.45 mL ddH_2O, 0.4 mL tris-acetic acid/EDTA (TAE) Buffer (50X; Bio-Rad Laboratories, Inc. USA) (242 g tris free base, 18.61 g disodium EDTA, 57.1 ml glacial acetic acid, ddH_2O to 1 liter), 150 µL 10% APS and 15 µL TEMED; Bio-Rad Lab USA. This mixture was then poured into the gel-casting cassette then a comb was carefully inserted in order to avoid trapping air bubbles.

After the gel was set, the plate was attached to the electrode block and wells were washed by moving the comb up and down in 1X TAE electrophoresis buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA, pH 8.3). ODN samples were mixed with 6X LB (0.4% orange G, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 15% Ficoll® 400, 10mM Tris-HCl (pH 7.5) and 50mM EDTA (pH 8.0), Promega, USA). Either 100

fmole of HEX or SIMA labelled ODNs or 2 pmole of non-labelled ODNs were loaded onto gel. Electrophoresis was carried out using Mini-PROTEAN II apparatus (Bio-Rad) at 150V for 23 mer ODNs and 100V for 13 mer ODNs, until the bromophenol blue dye was nearly 1 cm from the bottom of the gel. The gel was removed and stained for 10 minutes with 1X SYBR Gold gel stain (Invitrogen, Oregon, USA), for non-hex tagged ODNs only. Gels were imaged by phosphoimager (Typhoon 9200 Variable Mode Image, Amersham Biosciences, UK) using HEX setting then the images were processed using ImageJ program (USA).

2.2.14 Inactivation of MGMT by alkylated ODNs

Purified MBP-MGMT (50 fmole; 10 fmole by activity) was incubated with varying concentrations of both SS and DS 23 mer ODNs (control G and O^6-alkylG containing ODNs) (0–250 nM) in a total volume of 150 ml of IBSA buffer (1 mg/mL BSA in buffer I: 50 mM Tris-HCl pH 8.3 containing 1 mM EDTA and 1mM TCEP) for 90 minutes at 37˚C, followed by the addition of excess (50 µL) [^{3}H]-methylated substrate DNA and processing, as described above in Section 2.2.10. This assay examined the following ODNs: O^6-MeG (15), O^6-CMG (DW1 and 44), O^6-CEG (RA1 and LH1) and control G (16). Results were presented as %MGMT activity and it was used to calculate IC_{50} (represents the amount of O^6-alkylG adducts required to decrease MGMT activity to 50%, compared to control) for a range of SS and DS O^6-alkylG containing ODNs. The values presented were the means (±SD) of triplicate experiments and they were used to establish linearity of MGMT inactivation curves with linear regression analysis using Microsoft Excel 2010 (294).

2.2.14.1 Statistical analysis of IC_{50} results

A pairwise comparison using independent t-test was conducted to compare IC_{50} values of SS & DS ODNs as well as to compare IC_{50} values of SS & DS O^6-CMG and O^6-CEG to that of the corresponding O^6-MeG. Independent t-test statistical analysis was performed using IBM SPSS Statistics version 22 (IBM Software NewYork, USA). Statistical significance was considered at $P<0.05$.

Ratios comparing IC_{50} values of SS and DS O^6-alkylG ODNs were calculated using the following equation:

R_{IC50} = IC_{50} for SS O^6-alkylG containing ODN/IC_{50} for corresponding DS O^6-alkylG containing ODN.

2.2.15 Assessment of alkyl group removal by restriction endonuclease site deprotection assay

DS 23 mer control G and O^6-alkylG containing ODNs (0.5 pmole; 5 µL of 0.1 µM) were incubated with either MBP-MGMT (10 active pmole in 5 µL of protein extract) or 5 µL of MBP-MGMT elution buffer in a total volume of 20 µL of IBSA buffer, so that the final protein concentration in each tube was 0.25 mg/mL. The whole mixture was then incubated at 37˚C for 4 hours on a shaker incubator (Orbital Incubator, SISO Stuart), 4 µL of $MgCl_2$ (100 mM) were added and then either PstI (Roche, Applied Sciences, Mannheim, Germany) or SfcI (New England BioLabs, USA) or MboI (New England BioLabs, USA) enzymes were added (2 µL) for subsequent digestion at 37˚C for 1 hour. MboI was used for digesting the DW1, GMe and ConG ODNs whilst PstI and SfcI were used to digest all other O^6-alkylG containing 23 mer ODNs. The samples were processed as described in Section 2.2.13 on a 20% non-denaturing PAGE (296).

However, in order to ensure that new employed MS incubation conditions would not interfere with alkyl group transfer, the assay was carried out in absence of IBSA buffer where MBP-MGMT was incubated with ODNs in a total volume of 20 µL of buffer I, instead of IBSA since BSA tryptic peptides could interfere with MGMT tryptic peptides during MS analysis. Restriction sites for PstI, SfcI and MboI are illustrated in Figure 2.3.

2.2.16 Electrophoretic mobility shift assay (EMSA)

DS 23 mer control G and O^6-alkylG containing ODNs (0.75 pmole; 7.5 µL of 0.1 µM) were incubated with either MBP-Atl1 (prepared in house) (0.5 pmole) in a total volume of 30 µL of buffer I, at 37˚C for 20 minutes. The samples were processed as described in Section 2.2.13 on a 20% non-denaturing PAGE.

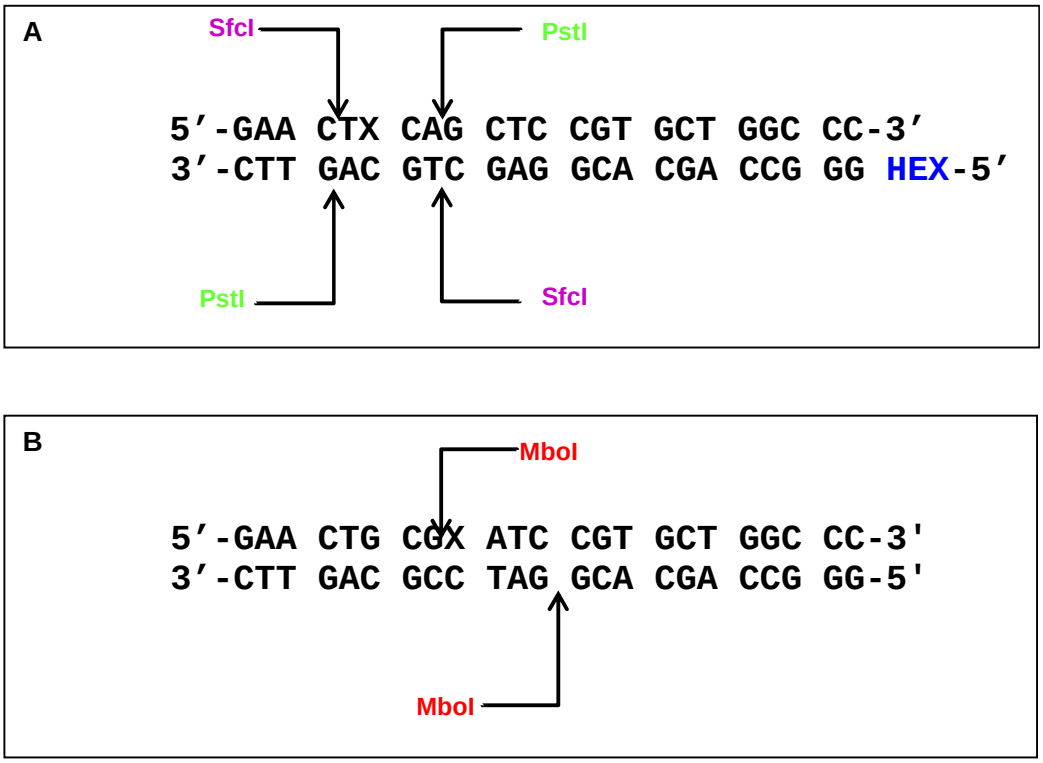


Figure 2.3: Restriction sites for (A) PstI and SfcI (B) MboI.
X denotes O^{6}-alkylG base and HEX is a fluorophore.

2.2.17 MALDI-TOF MS analysis of alkyl group transfer to MBP-MGMT

Both SS and DS Control (G) and O^{6}-alkylG containing ODNs (O^{6}-MeG, O^{6}-CMG or O^{6}-CEG; 13 mer and 23 mer (Table 2.1) (20 pmole) were incubated with MGMT (2 pmole by activity) for 6 hours at 37˚C. Both 23 mer and 13 mer ODNs were used for this assay (Table 2.1). Then trypsin (Roche, Applied Sciences, Mannheim, Germany) (ratio of MBP-MGMT:trypsin, 20:1(wt/wt)) was then added and the reaction was incubated overnight (18 hours) at 37˚C. The digestion was terminated by the addition of 1% formic acid to a final concentration of 0.1% (v/v). Finally, MBP-MGMT tryptic peptides were desalted and concentrated using C18-Ziptips™ (Merck Millipore, Ireland) where 10 µL of 0.1% formic acid in 50% acetonitrile was used to wet the tip then it was washed by 10 µL of 0.1% formic acid (5 times aspirate/dispense cycles). 10 µL of sample was aspirated up and down the Ziptip (10 times). The Ziptip was washed by 10 µL of 0.1% formic acid (10 times aspirate/dispense cycles), to get rid of the salts.

Desalted tryptic peptides were eluted in 5 µL of 0.1% formic acid in 50% acetonitrile/water (10 times aspirate/dispense cycles).

2.2.17.1 Chemical modification of synthetic unmodified MGMT-ASPs

Synthetic MGMT-ASP (20 µL of 1 µM) was reduced *via* incubation with 10 mM DTT at 50°C for 30 min then chemically modified by incubation with 50 mM final concentration of either: Iodoacetamide, Iodoacetic acid, 2-Bromopropionamide or 2-Bromopropionamide then it was processed for MALDI-TOF MS analysis as described below.

2.2.17.2 MALDI-TOF MS analysis and data acquisition

Eventually, 1 µL from each sample was spotted on a MALDI plate (MTP 384 ground steel, Bruker, Germany) together with 1 µL saturated α-cyano-4-hydroxycinnamic acid (Fluka, Buchs, Switzerland) matrix solution (10 mg/mL in 50% ethanol/acetonitrile (v/v)). The MALDI-TOF was calibrated by using either the Peptide Calibration Standard (Bruker, Germany) or J67722 MALDI Certified Mass Spec Calibration Standard (Alfa Aesar, UK). Spectra were acquired over the mass to charge ratio (*m/z)* range 800–2300 using a Bruker (Germany) Ultraflex II operating at 30% laser intensity and 1000 laser shots per spectrum in Reflectron Positive Ion mode. Spectra were generated using FlexAnalysis (Bruker, Germany) software and annotated using the MS-Digest tool of the protein prospector service (http://prospector.ucsf.edu). Data were analysed using Peptide Mass Fingerprint (PMF), a Mascot search tool on the Matrix Science website (http://www.matrixscience.com) with the SwissProt database at peptide mass tolerance of 0.5 Da with 2 allowed missed cleavages. A score greater than the significance level (0.05) in Mascot and Expectation Value less than 0.05 were required. The Mascot Expectation Value is the number of times you could expect to get this score or better by chance. A completely random match has an expectation value of 1 or more. The better the match, the smaller is the expectation value.

2.2.17.3 Label free quantitation of MGMT-ASP (unmodified and alkylated)

Quantitative analyses were conducted measuring R_o ratio (monitoring disappearance of unmodified MGMT-ASP following reaction with alkylated ODNs) and R_1 repair ratio (monitoring extent of alkylation). (R_o) consisted of peak area (PA) of unmodified MGMT-ASP/PA of MGMT TP-8. (R_1) consisted of PA of alkylated MGMT-ASP/PA of MGMT TP-8. R_o and R_1 ratios were determined for all ODNs. As for LH1 and SRO1, an additional R_2 ratio was investigated to compare the relative abundance of *S*-carboxamidoethylcysteine-ASP ion to the contaminant *S*-carboxymethylcysteine-ASP. (R_2) consisted of PA of carbamidoethylated MGMT-ASP/PA of methylated MGMT-ASP. PAs of chosen peptides were measured by FlexAnalysis software (Bruker, Germany).

2.2.17.4 Statistical analysis of R_o, R_1 and R_2

One-way analysis of variance (one way ANOVA) was carried out to assess the significance level between various groups (G, SS and DS O^6-alkylG ODNs); it was performed using GraphPad Prism v7.00. However, pairwise comparisons were performed using unpaired t-test to investigate significance level between O^6-MeG and either O^6-CMG or O^6-CEG, using GraphPad Prism v7.00. Statistical significance was considered at $P<0.05$.

2.2.18 Determination of limit of quantification (LOQ) of synthetic unmodified MGMT-ASP

Synthetic unmodified MGMT peptide containing the active site (Cys-145) of the sequence (GNPVPILIPCHR), corresponds to positions 136-147 of MGMT, was purchased (Cambridge Research Biochemicals, Cleveland, UK). Both non-labelled (light) and heavy isotope labelled ($C^{13}N^{15}$ proline labelled) were used for estimation of the lowest LOQ of unmodified MGMT-ASP in MALDI-TOF MS analysis.

In first protocol, synthetic lyophilized peptide was dissolved in AMBIC buffer (25 mM, pH 8), then the sample was diluted with one volume of 20% acetonitrile in AMBIC

buffer (50 mM, pH 8) and incubated at RT for 30 minutes. Samples were reduced *via* incubation with DTT (10 mM final concentration) at 50°C for 30 minutes followed by capping of unmodified cysteine residues *via* incubation with N-ethylmaleimide (NEM; Thermo Fisher Scientific, USA) (20 mM final concentration) at RT for 45 minutes. Finally dissolved NEM-capped ASP was serially diluted in AMBIC buffer (25 mM, pH 8) to 4 dilutions of concentrations 5 nM, 50 nM, 0.5 µM and 5 µM.

Whilst in the second protocol, synthetic lyophilized peptides were dissolved in 0.1% formic acid in 20 mM Tris-HCl (pH 4), then the samples were processed as described in the first protocol except for that capping of cysteine residues was carried out using NEM (30 mM final concentration). Protocol 3 was similar to protocol 2 except for decreasing pH of AMBIC buffer to 7 instead of 8 where NEM capping step was conducted at pH 7 and diluting the dissolved unmodified ASP with AMBIC buffer (25 mM, pH 7) to 4 dilutions of concentrations 5 nM, 20 nM, 50 nM and 0.5 µM. Protocol 4 was the same as protocol 3 except for skipping the cysteine residues capping with NEM.

After sample preparation, 1 µL of each serial dilution was spotted on MALDI-TOF plate and synthetic peptides were processed for MALDI-TOF analyses on a Bruker Ultraflex^{TM} (Bruker Daltonics, Bremen, Germany) as described in Section 2.2.17. S/N > 10 was required for detected peptides ions.

2.2.19 Determination of LOQ of synthetic methylated MGMT-ASP standard

Both synthetic light and heavy isotope labelled ($C^{13}N^{15}$ Proline labelled) methylated MGMT-ASP (Cambridge Research Biochemicals, Cleveland, UK) were solubilized in 0.1% formic acid in 20 mM Tris-HCl (pH 4), then serially diluted by AMBIC buffer (pH 7). Serial dilutions (5 nM, 20 nM, 50 nM and 0.5 µM) were used for determination of MALDI-TOF MS lowest LOQ as described in Section 2.2.17.

2.2.20 Establishment of linearity of quantitative MALDI-TOF

Five serial dilutions (0.25, 0.5, 1, 1.5 and 2 µM) of the four available synthetic versions of MGMT-ASP were prepared as explained in Section 2.2.18. and 2.2.19.

Then serial dilutions were processed for MALDI-TOF MS analysis as described in Section 2.2.17. Herein, the PAs corresponding to MGMT-ASP standard ions were measured using Flexanalysis software and plotted against amount of MGMT-ASP standard that was analyzed in fmole to establish linearity with linear regression analysis using Microsoft Excel 2010.

2.2.21 On-bead digestion of MBP-MGMT following incubation with CT-DNA

MBP-MGMT (24 pmole; 6 active pmole) was incubated with methylated CT-DNA (TMZ-modified CT DNA; 500 µg containing O^6-MeG adduct (3 pmole; as determine by radioisotopic assay)) or unmodified CT-DNA as negative controls in Buffer I (500 µL; containing 2 mM TCEP) for 6 hours at 37˚C. Following protein reaction with CT-DNA, amylose coated magnetic beads (50 µL; New England BioLabs, USA) were incubated with the protein and CT-DNA mixture at 4˚C overnight to capture MBP-MGMT protein from CT-DNA, according to the manufacturer's instructions. The reaction mixture was then centrifuged at 19000g (Fisher Scientific accuspin™ Microcentrifuges) for 2 minutes and the supernatant was aspirated and decanted. Recaptured proteins from the DNA by the beads were then digested *in situ* (on-beads) by incubation with trypsin (10 µL of 0.1µg/µL) at 37°C for 18 hours. The digestion was terminated by the addition of formic acid to a final concentration of 0.1% (v/v). Protein digests were desalted using C18-ZipTips™ (Section 2.2.17). On-beads tryptic digests were then analysed by MALDI-TOF MS analysis as described in Section 2. 2.17.

2.2.22 In-solution digestion of MBP-MGMT following incubation with CT-DNA

MBP-MGMT (16 pmole; 4 active pmole) was incubated with unmodified CT-DNA (500 µg; sonicated on ice for 20 seconds) in presence of Buffer I with 2 mM TCEP (500 µL) then protein and DNA mixtures were concentrated with Amicon centrifugal tubes (Merck Millipore, Ireland) according to the manufacturer's instructions followed by in-solution trypsin digestion (10 µL of 0.1 µg/µL) via incubation at 37°C for 18 hours. The digestion was terminated by the addition of formic acid to a final concentration of 0.1% (v/v). Eventually protein tryptic peptides were concentrated and desalted using C18-ZipTips™.

Prior to MS analysis, desalted tryptic peptides were spiked with $C^{13}N^{15}$ proline labelled unmodified ASP internal standard (1 pmole) to assess the percentage of protein recovery by means of quantifying unmodified MGMT-ASP. MALDI-TOF MS analysis was carried out as described in Section 2.2.17. PAs of chosen peptides (unmodified MGMT-ASP & internal standard) were measured by FlexAnalysis software (Bruker, Germany).

Moreover, for validating the in-solution digestion protocol MBP-MGMT (16 pmole; 4 active pmole) was incubated with methylated CT-DNA (500 µg) containing 0.5, 1, 2 and 4 pmole of O^6-MeG adduct mg^{-1} CT-DNA; as determined by radioisotopic assay, then protein and methylated CT-DNA mixtures were processed following in-solution protein digestion protocol described above. In addition, desalted tryptic peptides were spiked with synthetic $C^{13}N^{15}$ proline labelled methylated ASP internal standard (1 pmole) for calculating methylated MGMT-ASP recovery. MALDI-TOF MS analysis was carried out as described in Section 2.2.17. PAs of chosen peptides (methylated MGMT-ASP & internal standard) were measured by FlexAnalysis software (Bruker, Germany) and linearity of the assay was established *via* plotting PA of detected methylated MGMT-ASP *vs*. amount of O^6-MeG adduct (pmole) in CT-DNA with linear regression analysis using Microsoft Excel 2010.

2.2.23 MS analysis of his-MGMT following on-bead digestion after incubation with CT-DNA

Purified his-MGMT (50 pmole; 30 active pmole) was incubated with 2 mg of DNA (unmodified CT-DNA), in a 2 mL low protein bind microfuge tube, for 6 hours at 37˚C in presence of buffer I with 2 mM TCEP (600 µL), on a shaker incubator. 100 µL of Ni coated magnetic bead suspension (PureProteome™ Magnetic Beads, MercK Millipore, UK) or Talon® Magnetic Beads (Clontech laboratories, Takara Inc., USA) was transferred into a 1.5 mL microfuge tube that was placed on magnetic stand to collect the beads and storage buffer was carefully removed with a pipette. The beads were re-suspended in 250 µL of equilibration buffer (50 mM sodium phosphate Na_3PO_4, 300mM sodium chloride NaCl, 0.1%Tween 20) and incubated with gentle mixing for one minute at RT and the tube was then placed back into the magnetic stand and the equilibration buffer was aspirated with a pipette. Washed beads were re-suspended in 80 µL of equilibration buffer, mixed by vortexing (for 20 seconds) then

were added to DNA and protein, where the mixture (DNA, his-MGMT and beads) was incubated at 4˚C overnight on a rotor mixer (Blood Tube Rotator, SB1, Stuart Scientific, UK). The following day, the protein, DNA and beads mixture was centrifuged at 19000g (Fisher Scientific accuspin™ Microcentrifuges) for 2 minutes and the supernatant was aspirated to remove the DNA and buffer I. Beads were reconstituted in 40 µL his-MGMT elution buffer (50 mM sodium phosphate Na_3PO_4, 300 mM sodium chloride, 300 mM imidazole, pH 8), 1µg of trypsin was added to the beads and the reaction was incubated overnight (18-20 hours) at 37˚C with shaking. The digestion was terminated by the addition of formic acid to a final concentration of 0.1% (v/v). Finally, his-MGMT tryptic peptides were desalted and concentrated using Millipore® C18-Ziptips as explained in Section 2.2.17. MALDI-TOF analyses were performed as described in Section 2.2.17. S/N > 10 was required for identification of detected alkylated peptides ions.

2.2.23.1 Determination of unmodified MGMT-ASP recovery

With the purpose of examining the recovery of unmodified MGMT-ASP (%Recovery $_{unmodified}$), as an indicator of protein recovery following incubation with unmodified CT-DNA and on-bead digestion; as described in Section 2.2.23, desalted MGMT tryptic peptides were spiked with synthetic $C^{13}N^{15}$ proline labelled unmodified ASP internal standard (10 pmole). MALDI-TOF mass spectral analysis was carried out as described in Section 2.2.17. PAs of chosen peptides (unmodified MGMT-ASP & internal standard) were measured by FlexAnalysis software (Bruker, Germany) and %Recovery $_{unmodified}$ was calculated using the following equation:

$$\%\text{Recovery}_{unmodified} = \frac{(PA_{unmodified} / PA_{STD}) \text{ X pmole of internal standard X } 100}{\text{Amount of his-MGMT (pmole)}}$$

2.2.24 Validation of the novel on-bead tryptic digestion approach

2.2.24.1 LOQ of O^6-MeG adduct in TMZ-modified CT-DNA

In order to investigate the LOQ of O^6-MeG adduct in methylated CT-DNA using his-MGMT, TMZ-modified CT-DNA (2mg) containing various levels of O^6-MeG adduct (0.050, 0.125, 0.250 and 0.50 pmole O^6-MeG adduct mg^{-1} methylated CT-DNA; as determine by radioisotopic assay); was incubated with his-MGMT then the samples were processed as explained in Section 2.2.23. MALDI-TOF analyses were performed as described in Section 2.2.17. S/N > 10 was required for identification of detected methylated ASP peptides ions.

2.2.24.2 Determination of methylated MGMT-ASP recovery

For validation of the developed assay, recovery of the alkylated (methylated) MGMT-ASP (%Recovery $_{methylated}$) after MGMT reaction with methylated CT-DNA was investigated. His-MGMT was incubated with TMZ-modified CT-DNA (2 mg containing 100, 200 or 400 fmole of O^6-MeG adduct; as determine by radioisotopic assay). His-MGMT and methylated CT-DNA mixtures were processed as described in Section 2.2.23 and prior to mass spectral analysis desalted methylated MGMT tryptic peptides were spiked with synthetic $C^{13}N^{15}$ proline labelled methylated ASP internal standard (250 fmole). MALDI-TOF mass spectral analysis was carried out as described in Section 2.2.17. PAs of chosen peptides (methylated MGMT-ASP & internal standard) were measured by FlexAnalysis software (Bruker, Germany) and %Recovery $_{methylated}$ was calculated using the following equation:

$$\%\text{Recovery}_{methylated} = \frac{(PA_{methylated}/PA_{Me\text{-}STD}) \text{ X fmole of internal standard X } 100}{\text{Amount of } O^6\text{-MeG adducts (fmole) in CT-DNA}}$$

In order to establish the linearity of the quantitative assay, amount of quantified methylated MGMT-ASP (fmole) was plotted against the amount of O^6-MeG adduct in methylated CT-DNA (fmole). The values presented were the means (±SD) of triplicate experiments and they were used to establish linearity with linear regression analysis using OriginPro 8.5.1. software (Originlab Corporation, USA).

2.2.25 Characterization of O^6-alkylG adducts in human colorectal DNA

2.2.25.1 Study population

The current study investigated 13 human colorectal DNA samples. Of each 10 DNA samples were extracted from colorectal carcinoma tissue. The other 3 samples, DNA samples were extracted from normal colorectal tissue (macroscopically normal tissue taken ~5 cm from the tumour edge) taken from individuals presenting with colorectal carcinoma at hospitals with in Manchester area. Among the 13 samples studied, 3 paired normal and tumour DNA samples were analyzed. DNA was previously extracted by another published study then stored in -80˚C until they were analyzed by the present study.

2.2.25.2 Ethical approval

REC reference: 15/EM/0505 (East Midlands-Derby Research Ethics Committee, Health research authority, NHS).

2.2.25.3 MS analysis of his-MGMT reacted with human colorectal DNA

His-MGMT fusion protein was incubated with human colorectal DNA (either extracted from normal or tumour tissue) as described in Section 2.2.23 then alkylated peptides were processed for MALDI-TOF MS analysis (Section 2.2.17). Putative alkylated MGMT-ASPs were identified using the criteria illustrated in Table 2.2.

Quantification of O^6-MeG adduct in human colorectal DNA was accomplished by means of quantitation of detected methylated MGMT-ASP following MGMT reaction with human colorectal DNA. MGMT and human colorectal DNA mixtures were processed as described in Section 2.2.23 and prior to mass spectral analysis desalted methylated MGMT tryptic peptides were spiked with synthetic $C^{13}N^{15}$ proline labelled methylated ASP internal standard (250 fmole). MALDI-TOF mass spectral analysis was carried out as described in Section 2.2.17. PAs of chosen peptides (methylated MGMT-ASP & internal standard) were measured by FlexAnalysis software (Bruker, Germany).

Label free quantitation of O^6-CMG adducts was carried out *via* comparing the PA of carboxymethylated ASP to that of methylated ASP and applying the following equation:

Level of O^6-CMG = (PA_{CM}/PA_{Me}) X Nanomole of O^6-MeG X ($R_{1\ of\ 303B}/R_{1\ of\ OW18}$)

Where: $R_{1\ of\ 303B}/R_{1\ of\ OW18}$ corresponds to the R_1 of the 13 mer O^6-MeG (303B) ODN to that of 13 mer O^6-MeG (OW18B) ODN and it equals 1.13. PA_{CM} corresponds to PA of carboxymethylated ASP detected in MGMT tryptic peptides following reaction with human colorectal DNA. PA_{Me} represents PA of methylated MGMT-ASP detected following MGMT reaction with the same human colorectal DNA sample. Nanomole of O^6-MeG corresponds to the amount of O^6-MeG detected in human colorectal DNA (nanomole of O^6-MeG per mole of 2'deoxyguanosine (dG)).

2.2.25.4 Statistical analysis of results of analysis of human colorectal tissue

Linear regression analysis was conducted to investigate the association between O^6-MeG levels and either MGMT activity or O^6-CMG levels. Whilst binary logistic regression analysis was conducted to investigate the association between O^6-MeG levels and K-ras gene mutations. Both tests were performed using IBM SPSS Statistics version 22 (IBM Software NewYork, USA). However, pairwise comparisons were performed using unpaired t-test (GraphPad Prism v7.00) to investigate significance level between O^6-MeG adduct in level in normal and tumour colorectal tissue of the same patient. In addition, unpaired t-test (GraphPad Prism v7.00) was performed to investigate significance level between MGMT activity in normal and tumour colorectal tissue. Statistical significance was considered at $P<0.05$.

Table 2.2: Criteria used for identification of putative alkylated MGMT-ASPs detected in the digest of his-MGMT incubated with human colorectal DNA.

Level of confidence about identity of detected ions	*m/z*			MGMT substrate	O^6-alkyl adduct	Tandem mass spectrometry	
	Known alkyl group	ODN containing the O^6-alkylG adduct	Alkylated active site peptide	RE assay	EMSA	Accurate *m/z*	Amino acid sequencing
Unknown	-	-	-	-	-	-	-
Possible	+	+	-	-	-	-	-
Probable	+	+	+	+	+	-	-
Definite	+	+	+	+	+	+	+

3 Chapter 3: Purification of MGMT fusion proteins

3.1 Introduction

Generally, MGMT is a key node in the defense against commonly found AAs (271). Moreover, several studies have highlighted that MGMT can protect the body from the carcinogenic effect of NNOC that is known to induce cancer by methylating DNA (157). The overall aim of this study was to develop a novel approach that utilized MGMT as an analytical reagent to enable characterization of a wide range of O^6-alkylG adducts in synthetic ODNs and then in human colorectal DNA. Thus, MGMT was expressed as two fusion proteins, namely, MBP-MGMT and his-MGMT. MBP (296,313) and his (282,290,341) are among the most common fusion tags that are used for the expression of human recombinant MGMT as they enable its purification *via* amylose column and Ni column chromatography, respectively .

3.2 Aims and objectives

The overall aim of this chapter focused on purification, characterization as well as determination of functional activity of both MBP-MGMT and his-MGMT fusion proteins.

The specific objectives were:

1. Sequencing of MBP-MGMT plasmid.
2. To purify MBP- and his-MGMT using amylose and Ni-column chromatography, respectively.
3. To characterize both purified MBP- and his-MGMT fractions using both SDS-PAGE and western blot analysis.
4. To determine MBP- and his-MGMT functional activities by radioisotopic assay.

3.3 Results

3.3.1 DNA sequencing of MBP-MGMT fusion protein plasmid

DNA sequencing was performed as described in Section 2.2.3. Results of DNA sequencing show the wild type sequence; no evidence of mixed clones or mutations in the MGMT gene were detected; as shown in Figure 3.1. Appendix 1 shows the nucleotide sequence alignments of MGMT DNA.

← Vector pMAL C2X

TAATTCGAGCTCGAACAACAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAA
GGATGGACAAGGATTGTGAAATGAAACGCACCACACTGGACAGCCCTTTGGGGAAGCTG
GAGCTGTCTGGTTGTGAGCAGGGTCTGCACGAAATAAAGCTCCTGGGCAAGGGGACGTC
TGCAGCTGATGCCGTGGAGGTCCCAGCCCCCGCTGCGGTTCTCGGAGGTCCGGAGCCCC
TGATGCAGTGCACAGCCTGGCTGAATGCCTATTTCCACCAGCCCGAGGCTATCGAAGAG
TTCCCCGTGCCGGCTCTTCACCATCCCGTTTTCCAGCAAGAGTCGTTCACCAGACAGGT
GTTATGGAAGCTGCTGAAGGTTGTGAAATTCGGAGAAGTGATTTCTTACCAGCAATTAG
CAGCCCTGGCAGGCAACCCCAAAGCCGCGCGAGCAGTGGGAGGAGCAATGAGAGGCAAT
CCTGTCCCCATCCTCATCCCGTGCCACAGAGTGGTCTGCAGCAGCGGAGCCGTGGGCAA
CTACTCCGGAGGACTGGCCGTGAAGGAATGGCTTCTGGCCCATGAAGGCCACCGGTTGG
GGAAGCCAGGCTTGGGAGGGAGCTCAGGTCTGGCAGGGGCCTGGCTCAAGGGAGCGGGA
GCTACCTCGGGCTCCCCGCCTGCTGGCCGAAACTGAATTTCAGAATTCGGATCCTCTAG
AGTCGACCTGCAGGCAAGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAA
ACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGT
AATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGA
ATGGCAGCTTGGCTGTTTTGGCGGATGAGATAAGATTTTCAGCCTGATACAGATTAAAT
CAGAACGCAGAAGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGTTC
CCACCTGACCCCATGCCGAACTCAGAA

Figure 3.1: Sequencing results of MGMT cDNA.

The above figure shows the start codon (ATG, in red), the stop codon (TGA, in red) and the vector sequence (underline).

3.3.2 Purification of putative MBP-MGMT

The fusion protein was expressed and extracted (Section 2.2.4 and 2.2.5). The putative MBP-MGMT fusion protein was purified using amylose affinity column chromatography (Section 2.2.6). Two peaks were eluted; the first represents the unbound protein fraction, which was eluted with amylose BB, and a second peak represents the protein fraction that was able to bind to amylose column and it was eluted with the elution buffer (amylose BB supplemented with maltose); as shown in Figure 3.2.

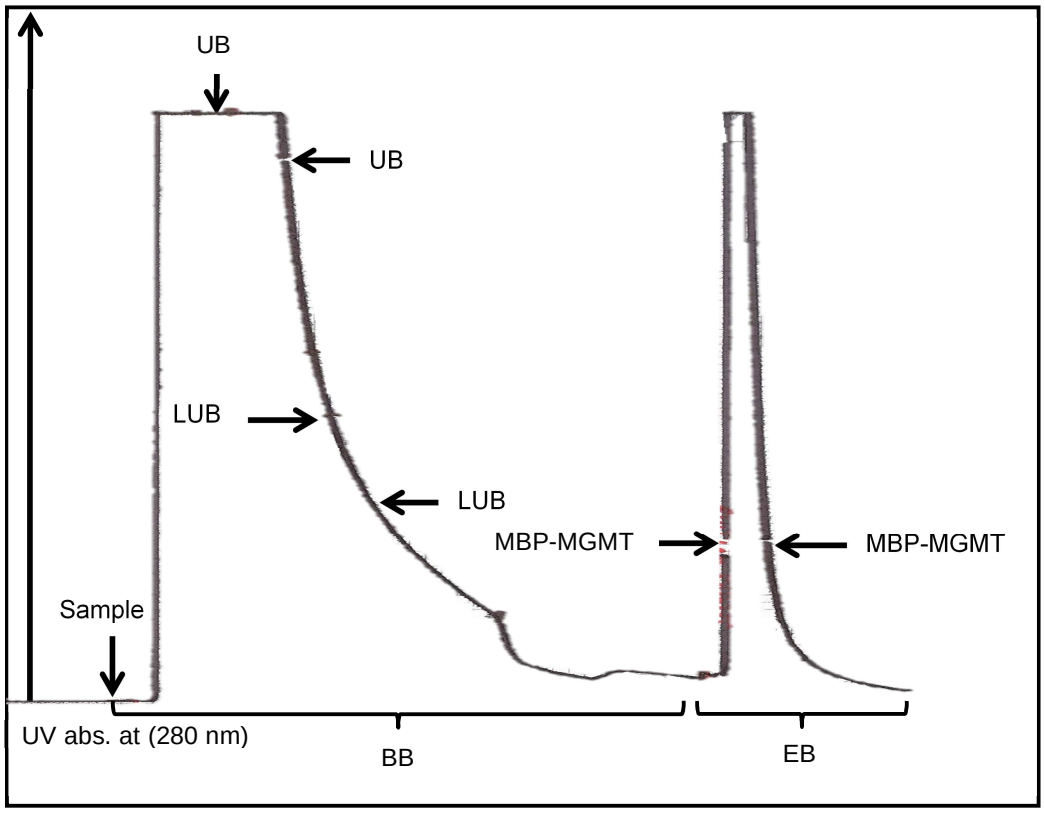


Figure 3.2: Elution profile of MBP-MGMT purification using amylose column.

UB and LUB denote unbound protein fraction and late unbound protein fraction, respectively. MBP-MGMT denotes putative MBP-MGMT fusion protein fraction. BB and EB denote amylose binding and elution buffers, respectively. Amylose column 1.5 x 6 cm. Pump setting x10 (1). Chart speed 1 mm per min, 10 mv.

The protein concentration of the protein extract was 6 mg/mL. In addition, the concentration of proteins in unbound protein and late unbound protein fractions were

2.25 and 0.06 mg/mL, respectively. Moreover, the putative MBP-MGMT peak protein concentration was 0.82 mg/mL. This data demonstrates that the putative MBP-MGMT fusion protein constitutes nearly 14% of the total bacterial protein extract (7 mg of MBP-MGMT out of 50 mg of crude bacterial protein).

3.3.3 SDS-PAGE analysis of putative MBP-MGMT

Bacterial protein extract, unbound protein, late unbound protein and putative MBP-MGMT fractions were analysed by 12% SDS-PAGE followed by Coomassie staining (Section 2.2.8). Figure 3.3 shows that SDS-PAGE analysis of both the protein extract and unbound protein fractions showed several protein bands of a wide range of molecular weight. These proteins were the proteins that did not bind to the amylose column and they were probably the *E.coli* bacterial cell proteins. However, the late unbound protein contained less protein bands, as compared to unbound protein, which might be ascribed to lower amount of proteins resolved on the gel.

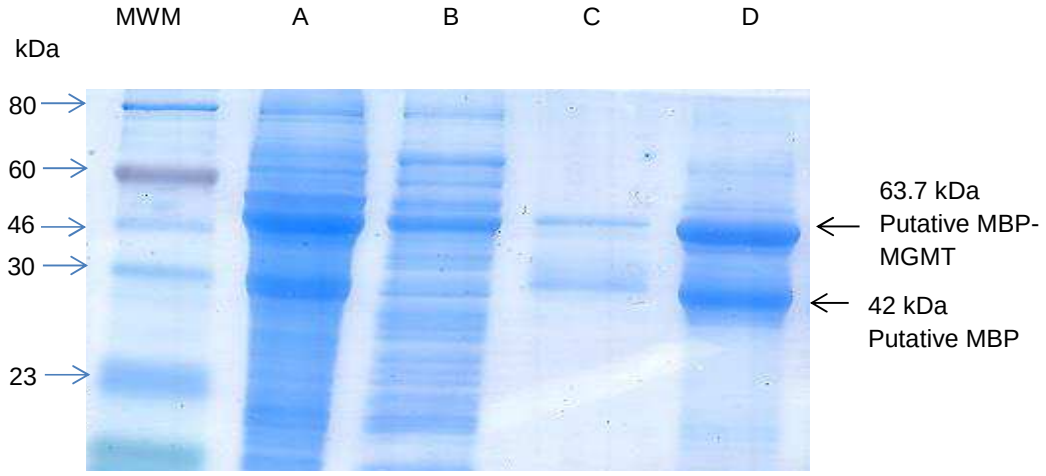


Figure 3.3: SDS-PAGE analysis of fractions obtained by amylose purification of protein extracts obtained from *E.coli* bacteria harboring MBP-MGMT plasmid.

MWM. Molecular weight marker (10 µL) **A.** Bacterial protein extract (30 µg) **B.** Unbound protein fraction (30 µg) **C.** Late unbound protein fraction (1 µg) **D.** Putative purified his-MGMT protein (10 µg). Protein fractions were resolved on 12% SDS-PAGE followed by Instant Blue® Coomassie staining.

Putative MBP-MGMT peak have shown 2 major protein bands and based on marker migration the larger protein band is consistent with MBP-MGMT (MW 63.7 kDa) and the smaller one is consistent with MBP (42 kDa).

3.3.4 Western blot analysis of putative MBP-MGMT

The putative MBP-MGMT fusion protein purified by amylose column (Figure 3.2) was resolved by SDS-PAGE, and the protein was transferred onto a nitrocellulose membrane for western blot analysis using both anti-MGMT R63 and anti-MBP antibodies (Section 2.2.9). Western blot analysis of the resolved proteins using anti-MGMT R63 antibody showed only one band (Figure 3.4B), which indicated that, for amylose purified fusion protein, only one band of the 2 resolved bands contain MGMT. This peak is considered to be the MBP-MGMT fusion protein band (MW 63.7kDa). Whilst western blot analysis of amylose purified putative MBP-MGMT using anti-MBP antibody showed 2 bands, the larger band corresponded to MBP-MGMT (MW 63.7kDa) and the smaller one corresponded to MBP (MW 42 kDa) (Figure 3.4C).

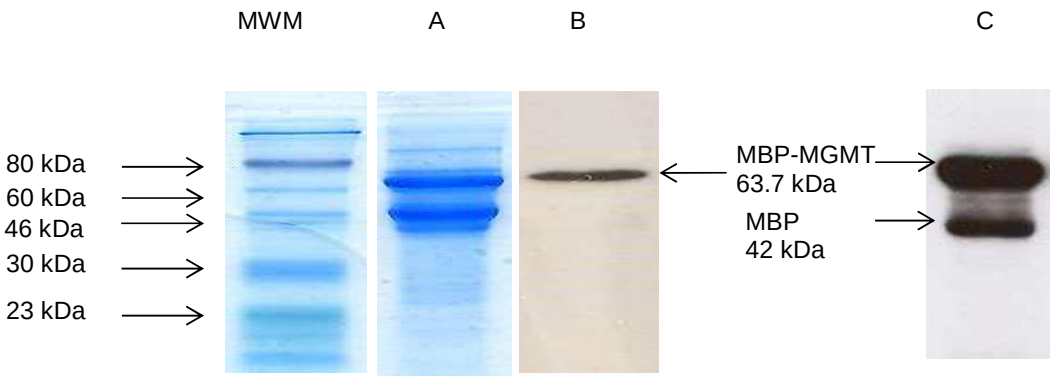


Figure 3.4: SDS-PAGE and western blot analysis of amylose purified putative MBP-MGMT fusion protein.

MWM. Molecular weight marker (10 µL) **A.** 10 µg of putative MBP-MGMT fusion protein obtained from amylose column purification was resolved by 12% SDS-PAGE followed by Instant Blue® Coomassie staining. **B.** Protein was transferred to membrane by wet blot, The membrane was probed with R63 anti-MGMT antibody. **C.** Western blot analysis using anti-MBP antibody.

3.3.5 Determination of functional MGMT activity of putative MBP-MGMT

MGMT activity of the bacterial protein extract, the unbound protein fraction (to monitor the efficiency of the chromatographic purification) and putative MBP-MGMT fusion protein was determined using [^{3}H]-methylated substrate DNA as described in Section 2.2.10. For all tested protein fractions, there was a linear relationship between protein concentration and the MGMT activity (Figure 3.5 and 3.6). Moreover, a plateau was reached at concentrations higher than 80 fmole of MBP-MGMT; as shown in Figure 3.5. Correlation coefficients (R^2) values were 0.994, 0.990 and 0.994, for the bacterial protein extract, the unbound protein fraction and the putative MBP-MGMT fusion protein, respectively. In addition, specific MGMT activities were calculated as pmole/µg of the protein fraction and recorded values were 0.38, 0.02 and 2.98 pmole/µg for the bacterial protein extract, the unbound protein fraction and the putative MBP-MGMT fusion protein, respectively. Approximately 20% of the total purified putative MBP-MGMT fusion protein was functionally active MGMT.

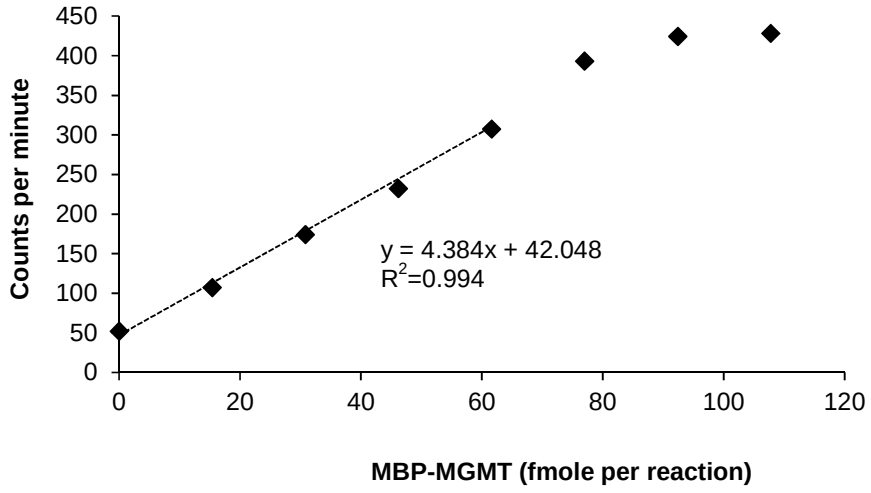


Figure 3.5: Determination of functional MGMT activity for the amylose purified MBP-MGMT fusion protein.

The transfer of [^{3}H]-methyl groups from [^{3}H]-methylated substrate DNA to purified MBP-MGMT with increasing protein concentrations after 90 mins incubation at 37˚C . The linear part of the graph showing R^2= 0.994. Linear point up to 60 fmole.

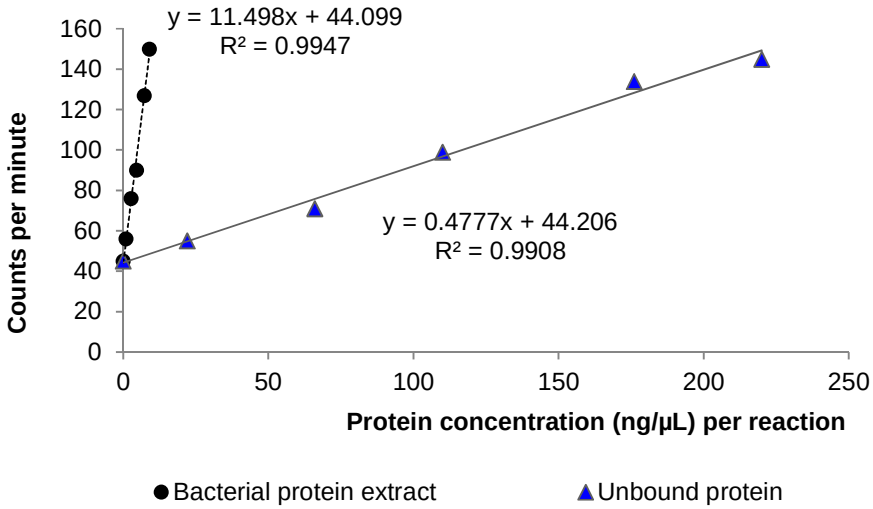


Figure 3.6: Determination of functional MGMT activity for bacterial protein extract and unbound protein fraction.

Bacterial protein extract and unbound protein fraction were diluted in IBSA 1:00 000 and 1:1000 (v/v); respectively and assayed for MGMT functional activity. Transfer of [^{3}H]-methyl groups from [^{3}H]-methylated substrate DNA to diluted proteins increased with increasing protein concentration after 90 mins incubation at 37˚C.

So now with evidence from both western blot analysis and the functional MGMT assay that putative MBP-MGMT fusion protein fraction has MGMT activity and we can legitimately state that this is a MGMT containing fraction. Amylose purified MBP-MGMT was divided into aliquots then stored at -20˚C, for up to 12 months, to avoid exposing the protein to repeated thawing and freezing cycles. Functional MGMT activities of frozen aliquots were re-assayed on the day each aliquot was thawed to calculate post freezing functional activity. Results of post freezing MGMT specific activity was non-significantly different from the specific activity recorded prior to protein freezing. Result of the latest post freezing MGMT functional activity assay of an aliquot that was stored in -20˚C for 12 months showed non-significant 2% decrease in the MGMT specific activity of the stored MGMT aliquot; to 2.93 pmole/µg protein.

MBP-MGMT preparations used for MS analysis were eluted from amylose column in a MS compatible elution buffer which was 200 mM AMBIC buffer supplemented with 10 mM maltose instead of the conventional MBP-MGMT elution buffer (amylose BB supplemented with 10 mM maltose). Both SDS-PAGE and western

blot analysis confirmed the identity of the purified protein (data not shown). To investigate the impact of the new elution buffer on MBP-MGMT activity, functional activity assay was carried out which demonstrated that the new elution buffer retained amylose purified MBP-MGMT functional activity. Recorded MGMT specific activity was 2.81 pmole/µg of the protein and approximately 20% of the amylose purified MBP-MGMT was functionally active, following elution in 200 mM AMBIC buffer supplemented with maltose.

3.3.6 Production and purification of his-MGMT

His-MGMT fusion protein plasmid constructs were prepared by Prof. Geoffrey Margison. Following ligation to generate circular plasmids and transformation into competent XL-1 blue *E.coli,* individual ampicillin resistant colonies were isolated and sequenced. Results of DNA sequencing did not show any evidence of mixed clones or mutations in MGMT gene (Data are not shown). Clones containing his-MGMT plasmid constructs were induced for his-MGMT expression using IPTG and then bacterial cells were harvested. Then his-MGMT fusion proteins were purified using Ni-affinity columns. Purified his-MGMT preparations were supplied by Prof. Geoffrey Margison.

3.3.7 Characterisation of purified putative his-MGMT

3.3.7.1 SDS-PAGE analysis of purified putative his-MGMT

Protein extract, the unbound protein fractions, the late unbound protein and putative his-MGMT fractions were analysed by 12% SDS-PAGE followed by Instant Blue® Coomassie staining as described in Section 2.2.8. SDS-PAGE analysis of both the protein extract and the unbound protein fractions showed several protein bands of a wide range of molecular weight that correspond to *E.coli* bacterial cell proteins that did not bind to the Ni column (Figure 3.7).

However, the late unbound fraction showed less protein bands, as compared to two earlier fractions. Putative his-MGMT peak showed only a single protein band and

based on molecular weight marker migration this protein band is consistent with his-MGMT (MW 25.6 kDa).

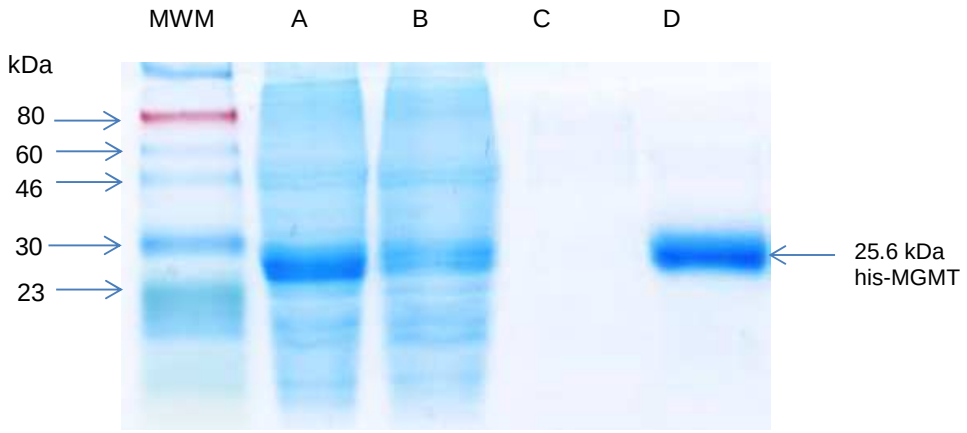


Figure 3.7: SDS-PAGE analysis of fractions obtained by Ni-affinity purification of protein extract obtained from *E.coli* bacteria harboring his-MGMT plasmid.

MWM. Molecular weight marker (10 µL) **A.** Bacterial protein extract (30 µg) **B.** Unbound protein fraction (30 µg) **C.** Late unbound protein fraction (0.6 µg) **D.** Putative purified his-MGMT protein (5 µg). Protein fractions were resolved on 12% SDS-PAGE followed by Instant Blue® Coomassie staining.

3.3.7.2 Western blot analysis of putative his-MGMT

The putative his-MGMT fusion protein purified by Ni column was resolved by SDS-PAGE, and the protein was transferred onto a nitrocellulose membrane for western blot analysis (Section 2.2.9). Western blot analysis of the resolved proteins, using R63 anti-MGMT antibody, showed only one band, which indicated that only one band contain MGMT (Figure 3.8). This peak is considered to be the putative his-MGMT fusion protein band.

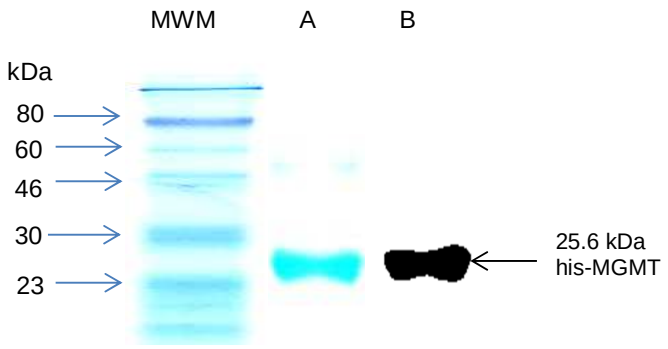


Figure 3.8: SDS-PAGE and western blot analysis of amylose purified putative his-MGMT fusion protein.

MWM. Molecular weight marker (10 µL) **A.** Putative purified his-MGMT protein (5 µg). **B.** Western blot analysis of purified his-MGMT; the protein was transferred to membrane by wet blot and the membrane was probed with R63 anti-MGMT antibody.

3.3.7.3 Determination of MGMT functional activity of putative his-MGMT

MGMT functional activity of purified his-MGMT fraction was determined using [^{3}H]-methylated substrate DNA as described in Section 2.2.10. In addition, MGMT functional activity was determined for both the bacterial protein extract and the unbound protein fraction to monitor the efficiency of the chromatographic purification of his-MGMT. Figure 3.9 (A, B and C) shows, for all assayed protein fractions, there was a linear relationship between protein concentration and the MGMT activity. MGMT titration curves illustrated in Figure 3.9 A and B, were used for calculating the specific MGMT activities for both the bacterial protein extract and the unbound protein fraction, respectively. However, the above results have been used to choose the appropriate his-MGMT protein dilution for further more accurate investigation of the functional activity of the purified his-MGMT protein, as demonstrated in Figure 3.10.

Figure 3.10 shows the results of the titration curve used for determination of the specific activity of purified his-MGMT protein; a plateau was reached at concentrations higher than 30 fmole of his-MGMT. The linear part of this graph has been used to set the linear fitting; showing correlation coefficient (R^2) of 0.999. Specific MGMT activities of assayed protein fractions, calculated as pmole/µg of were 1.27, 0.08 and 17.66 for the bacterial protein extract, the unbound protein fraction and his-MGMT. Approximately 33% of purified his-MGMT protein was functionally active MGMT.

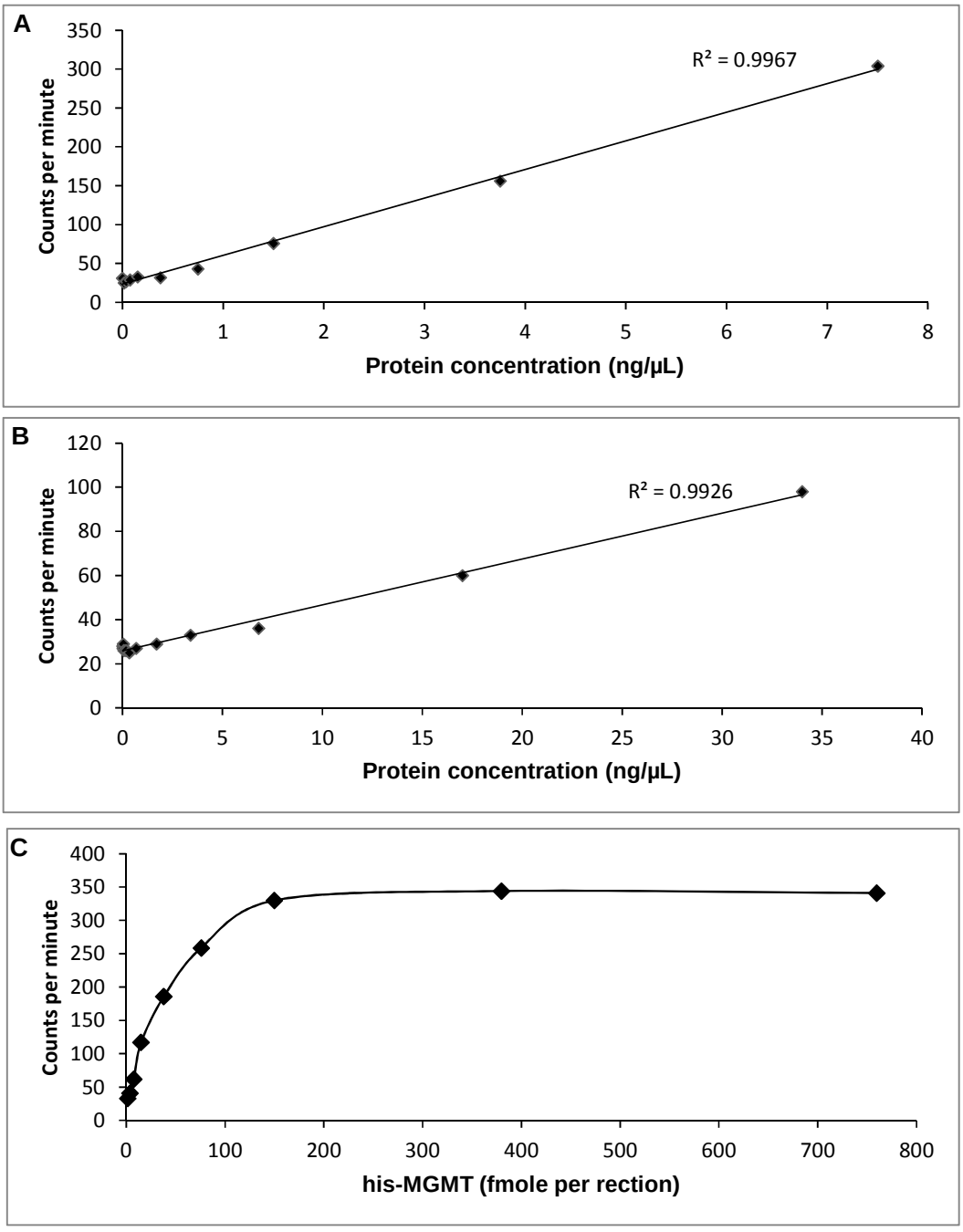


Figure 3.9: Transfer of [^{3}H]-methyl groups from [^{3}H]-methylated substrate DNA to MGMT with increasing protein concentrations.

A. Bacterial protein extract **B**. unbound protein **C**. Putative purified his-MGMT. Examined protein fractions were diluted in IBSA and assayed for MGMT functional activity. Transfer of [^{3}H]-methyl groups from [^{3}H]-methylated substrate DNA to diluted proteins increased with increasing protein concentration after 90 mins incubation at 37˚C.

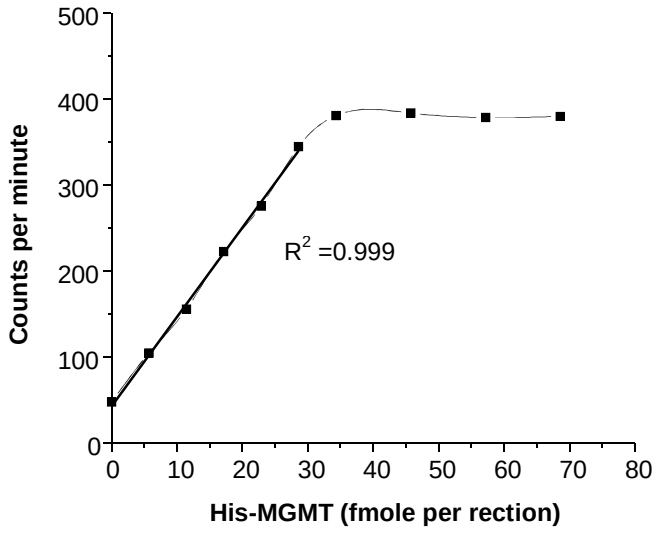


Figure 3.10: Transfer of [^{3}H]-methyl groups from [^{3}H]-methylated substrate DNA to purified his-MGMT.

[^{3}H]-methylated substrate DNA was incubated with increasing his-MGMT concentrations for 90 mins at 37˚C then then radioactivity transferred to MGMT were determined by radioisotopic assay. Radioactivity transferred increased upon increasing protein concentration. Linear point up to 30 fmole.

3.4 Discussion

The work presented in the current chapter indicated the successful production, purification and characterization of two functionally active MGMT fusion proteins, namely, MBP-MGMT and his-MGMT. MBP-MGMT fusion protein would be used to characterize MGMT protein repair of a range of *O*^{6}-alkyG adducts in synthetic ODNs. Whilst, his-MGMT would be used to develop a novel approach that uses MGMT as a reagent to detect and quantify *O*^{6}-alkyG adducts in human DNA. Both MBP-MGMT and his-MGMT fusion proteins demonstrated different functional activities that were 2.98 *vs.* 13.66 pmole/µg, respectively, which may be attributed to the different fusion tags or elution buffers or even both.

MGMT protein was first expressed as an MBP fusion protein to facilitate its purification using amylose affinity chromatography. MBP-MGMT was purified in conventional MBP-MGMT elution buffer (amylose BB supplemented with maltose) according to the previously published procedures used for expression and purification of MBP-MGMT (296,313).

MBP-MGMT conventional elution buffer consisted mainly of amylose BB that contained 200 mM sodium chloride, therefore, this buffer was incompatible with MS since the sensitivity of MS analysis can be significantly lower when some salts, such as sodium chloride, are present in the sample solution (342,343). Recent evidence suggests that even low mM concentrations of some metal ion salts such as sodium can cause severe ion suppression and peak broadening due to cluster and adduct formation (344–346). In addition, sodium ions increase background noise and might interfere with ions detection *via* reducing S/N ratio and increasing spectral complexity (347). Extensive sodium ion adduction to peptide ions could adversely affect the future mass spectroscopic detection of alkylated peptides given the low occurrence of O^6-alkyG adducts in human colorectal DNA and high salt amounts in DNA samples (348). However, MBP-MGMT that would be used for MS based assays was eluted in a novel buffer which was 200 mM AMBIC supplemented with 10 mM maltose. AMBIC buffer was chosen to replace the amylose BB as it was MS compatible buffer (349,350).

Notably, MGMT activity assay, of MBP-MGMT eluted in AMBIC, provided evidence that the replacement of amylose BB buffer with 200 mM AMBIC did not affect MGMT functional activity and the new elution buffer was able to conserve the MGMT activity. To the best of our knowledge this is the first study that eluted MBP-MGMT in 200 mM AMBIC.

3.5 Conclusion

In summary, two functionally active MGMT fusion proteins (MBP-MGMT and his-MGMT) were purified and could be used for accomplishing future aims and objectives of the present study.

4 Chapter 4: O^6-CMG and O^6-CEG adducts are substrates for MGMT and recognised by damage sensing protein Atl1

4.1 Introduction

O^6-alkylG adducts are of major importance in initiating mutations, cytotoxicity and carcinogenicity in human cells (152,210,212,351) and they are formed following exposure to AAs which are common environmental carcinogens that could be even formed endogenously during normal cellular metabolism and possibly inflammation (66,202,352). Repair of these adducts by MGMT plays a crucial role in the defense against AA toxicity (271) and is essential for maintenance of genome integrity. Moreover, slow repair of these adducts in specific DNA sequences may contribute to the mutational spectra observed in human cancer (353,354). Given the toxic and carcinogenic nature of O^6-alkylG adducts, it would be useful to explore if novel O^6-alkylG adducts such as O^6-CMG and O^6-CEG are substrates for MGMT repair to understand the spectrum of the repairable damage.

Previous reports suggested that O^6-CMG is not a substrate for MGMT repair (208), however, a recent study contradicted this reporting inactivation of MGMT by O^6-CMG containing 13 mer ODN (296). Moreover, they reported the alkyl group removal by MGMT as assessed by restriction endonuclease site deprotection assay using 2 REs which are PstI and BsaJI. Whilst, no previous studies investigated the ability of O^6-CEG adduct to inactivate MGMT and if it's a substrate for MGMT repair.

Furthermore, O^6-MeG adduct is recognized by DNA damage sensing protein Atl1 protein (312–314,355). Nevertheless, Atl1 binding to O^6-CMG and O^6-CEG adducts has not been yet characterized nor explored. Generally, unlike MGMT, Atl proteins lack the alkyltransferase activity or any other type of enzymatic activity or catalytic activity (316,355,356). Consequently, Atl proteins are believed to protect cells against the adverse and toxic effects of O^6-alkylG adducts by flagging those lesions for subsequent repair mostly *via* recruiting NER proteins (315,357).

4.2 Aims and objectives

The overall aim of the work described in this chapter was to investigate whether novel O^6-alkylG adducts (O^6-CMG and O^6-CEG) are substrates for MGMT and recognized by Atl1. The current work investigated 2 preparations of O^6-CMG containing 23 mer ODNs (DW1 and 44) as well as 2 preparations of O^6-CEG containing 23 mer ODNs (RA1 and LH1). DW1 and 44 had different DNA base sequence whilst RA1 and LH1 had the same base sequence as indicated in Table 2.1.

The specific aims and objectives are as follows

1- Determination of kinetics of methyl group transfer to MGMT.

2- Determination of IC_{50} (represents the amount of O^6-alkylG adducts required to decrease MGMT activity to 50%, compared to control) for a range of SS and DS O^6-alkylG containing ODNs (O^6-MeG, O^6-CMG and O^6-CEG).

3- Assessment of alkyl group removal by MGMT *via* restriction endonuclease site deprotection assay using 3 REs which were PstI and SfcI for O^6-MeG (15) and O^6-CMG (44) as well as O^6-CEG ODNs (LH1). However, the other O^6-CEG ODN (RA1) was digested with PST1 only. MboI was used to digest O^6-CMG ODN (DW1) alongside the corresponding O^6-MeG and control G ODNs, namely, GMe and ConG.

4- Comparing alkyl group transfer under standard MGMT incubation conditions and under new developed MS conditions by restriction endonuclease site deprotection assay using PstI enzyme.

5- Exploring Atl1 binding to a range of O^6-alkylG (O^6-MeG, O^6-CMG and O^6-CEG) adducts in DS synthetic ODNs using EMSA.

4.3 Results

MGMT activity titration assay was performed as described in Section 2.2.10. and the results were used to calculate protein limiting amounts of MGMT to be used in the MGMT inactivation assay using Excel 2010 software set to linear fitting R^2= 0.99. The protein limiting amounts for MBP-MGMT was 50 fmole (10 active fmole).

4.3.1 Kinetics of methyl group transfer to MGMT

Kinetics of methyl group transfer to MGMT following MBP-MGMT incubation with [^{3}H]-methylated substrate DNA were examined under assay conditions as described in Section 2.2.10. [^{3}H]-methyl group showed rapid transfer even within 30 minutes (74% transfer) with complete transfer after 90 minutes when a plateau is achieved (Figure 4.1). Based on this result, MBP-MGMT pre-incubation with O^6-alkylG adducts was established to be for 90 minutes allowing complete alkyl group transfer to MGMT during MGMT inactivation assays.

4.3.2 MGMT inactivation by O^6-alkylG containing ODNs

The main aim of this experiment is to examine the ability of MGMT to repair O^6-MeG, O^6-CMG and O^6-CEG adducts. MBP-MGMT was pre-incubated at 37˚C with both SS and DS O^6-MeG (15), O^6-CMG (DW1 and 44), O^6-CEG (RA1 and LH1) as well as control G (16) 23 mer ODNs then MGMT inactivation was investigated as described in Section 2.2.14. A pilot experiment was firstly carried out using a very wide range of concentrations to establish the range to use for each O^6-alkylG ODNs (SS and DS).

Results of the MGMT inactivation assay showed that both assayed SS and DS ODNs containing O^6-alkylG adducts (O^6-MeG (15), O^6-CMG (44) and O^6-CEG (RA1 and LH1)) inactivated MGMT and inhibited its action on [^{3}H]-methylated substrate DNA as shown in the inactivation curves in Figure 4.2. MGMT inactivation was in all cases linear with R^2=0.99 or better. However, for DW1 (O^6-CMG ODN), only DS but not SS caused MGMT inactivation. SS DW1 ODN failed to cause 50% MGMT inactivation at increasing concentrations up to 250 nM.

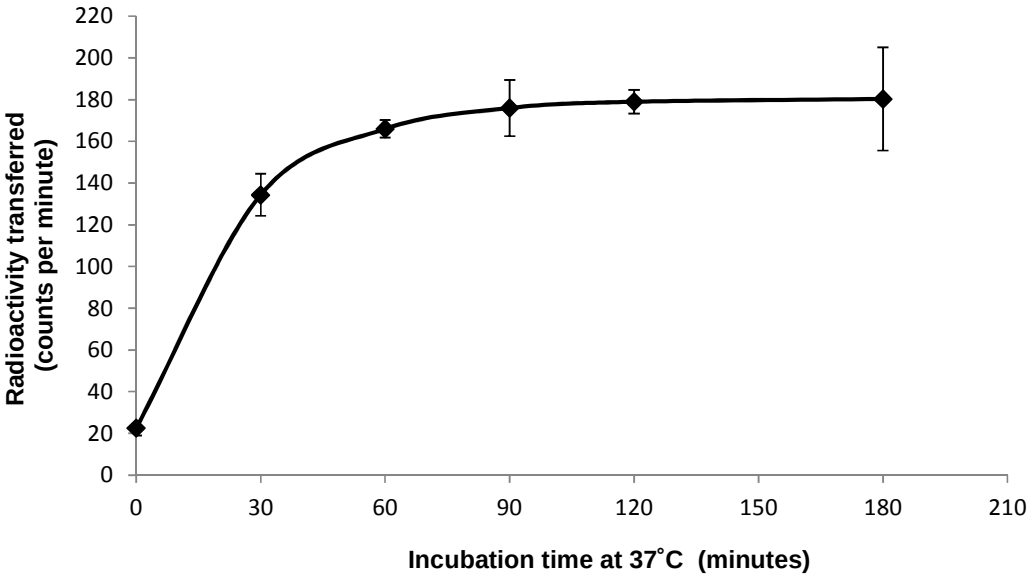


Figure 4.1: Kinetics of methyl group transfer to MGMT

MBP-MGMT (50 fmole; 10 fmole by activity) was incubated with [^{3}H]-methylated substrate DNA (containing tritiated O^6-MeG adduct) for several time intervals (30, 60, 90, 120 and 180 minutes). Then radioactivity transferred to MGMT were determined and plotted against incubation time.

SS control G ODNs did not cause MGMT inactivation at any of the chosen concentrations. Meanwhile, DS G ODN did not cause MGMT inactivation at the lowest chosen concentration (10 nM); however, at higher employed concentrations decrease in MGMT activity was observed, the max reported decrease was 30% of the original MGMT activity. MGMT inactivation curves were used to extrapolate IC_{50} values for each ODN as shown in Table 4.1. Recorded IC_{50} values were 0.17$\pm$0.02, 0.07$\pm$0.01, 0.80$\pm$0.14, 0.42$\pm$0.04, 15.7$\pm$2.1, 2.60$\pm$0.33, 0.98$\pm$0.11, 27.00$\pm$1.70 and 7.50$\pm$0.50 nM for SS 15, DS 15, SS 44, DS 44, DS DW1, SS LH1, DS LH1, SS RA1 and DS RA1, respectively. Notably, recorded IC_{50} values of the two investigated O^6-CEG ODNs, RA1 and LH1, were different. IC_{50} values of SS and DS RA1 ODN were approximately 7-fold the corresponding figures of LH1. Similarly, IC_{50} values of the two investigated O^6-CMG ODNs (DW1 and 44) showed wide variation that reached 37-fold for DS ODNs.

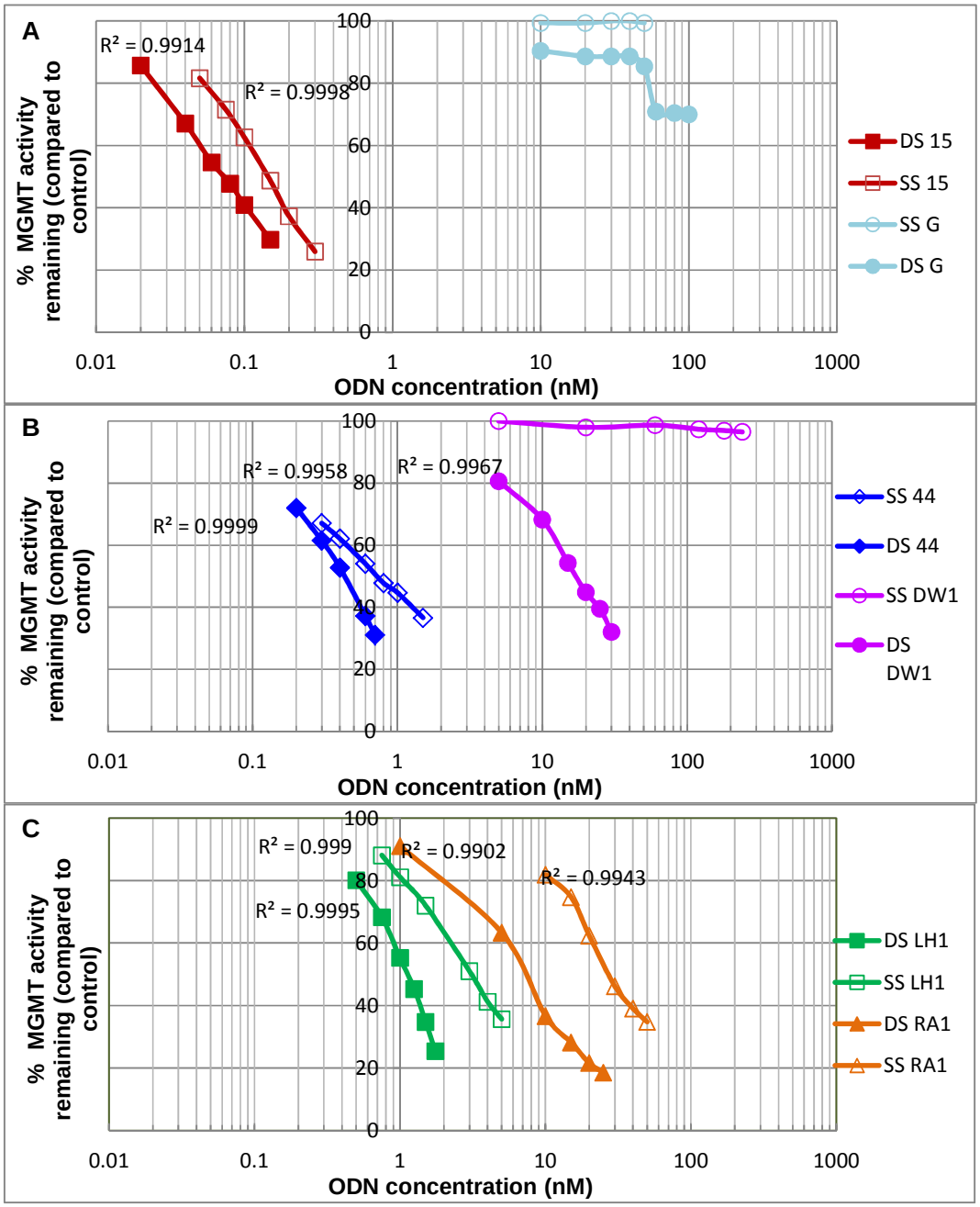


Figure 4.2: Inactivation of MGMT by O^6-alkylG containing ODNS (SS & DS).

MBP-MGMT (50 fmole; 10 fmole by activity) was pre-incubated for 90 minutes at 37˚C with increasing concentrations of SS or DS 23 mer ODNs: (A) unmodified control G ODN (16) or O^6-MeG (15) or (B) O^6-CMG (DW1 or 44) or (C) O^6-CEG (RA1 or LH1) in presence of CT-DNA. Excess [^{3}H]-methylated substrate DNA was then added and incubated with the mixture at 37˚C for further 90 minutes followed by determination of MGMT remaining activity *via* radioisotopic assay. SS ODNs are represented by open symbols while DS ODNs are represented by closed ones.

Table 4.1: IC50 values (nM) for the inactivation of MGMT proteins by O6-alkylG containing ODNs.

ODN	IC_{50} (nM)	*P* value	R_{IC50}
SS O^6-MeG (15)	0.17±0.02	0.007	2.41±0.35
DS O^6-MeG (15/17)	0.07±0.01**		
SS O^6-CMG (44)	0.80±0.14	0.010	1.90±0.20
DS O^6-CMG (44/17)	0.42±0.04*		
SS O^6-CMG (DW1)	ND	ND	ND
DS O^6-CMG (DW1/C)	15.70±2.10		
SS O^6-CEG (LH1)	2.60±0.33	0.008	2.66±0.23
DS O^6-CEG (LH1/17)	0.98±0.11**		
SS O^6-CEG (RA1)	27.00±1.70	0.001	3.62±0.45
DS O^6-CEG (RA1/17)	7.50±0.50**		

Values are mean±SD of three different experiments. *P* value using independent t-test for DS ODNs *vs.* SS ones. R IC_{50}: IC_{50} of SS ODNs/ IC_{50} of corresponding DS ODNs. ND denotes not determined.

For all studied O^6-alkylG containing ODNs except DW1, IC_{50} values for DS ODNs were statistically significantly lower than that of SS ODNs. Ratio of IC_{50} values (R_{IC50}) consisted of IC_{50} of SS ODNs to that of DS ODNs recorded values of 2.41±0.35, 1.90±0.20, 2.66±0.23 and 3.62±0.45 for O^6-MeG (15), O^6-CMG (44) and O^6-CEG (LH1 and RA1) containing ODNs, respectively. R_{IC50} values highlighted that, for all studied O^6-alkylG containing ODNs, DS ODNs were more potent inhibitors of MGMT activity than the corresponding SS ODNs.

Further statistical analysis using independent t-test revealed statistical significant differences between recorded IC_{50} values (nM) for SS 44, LH1 or RA1 *vs.* SS O^6-MeG (*P* values = 0.001, 0.006 and <0.001, respectively), as shown in Table 4.2. By the same token, statistical analysis displayed statistical significant difference

between recorded IC_{50} values (nM) for DS 44, DW1, LH1 and RA1 *vs.* DS O^6-MeG (*P* values = 0.001, <0.001, <0.001 and 0.001, respectively), as shown in Table 4.2.

Table 4.2: Results of statistical analysis for IC_{50} values (nM).

Independent t-test for IC_{50} values (nM)	*P* value
SS O^6-CMG (44) *vs.* SS O^6-MeG	0.001
SS O^6-CEG (LH1) *vs.* SS O^6-MeG	0.006
SS O^6-CEG (RA1) *vs.* SS O^6-MeG	<0.001
DS O^6-CMG (44/17) *vs.* DS O^6-MeG	0.001
DS O^6-CMG (DW1/C) *vs.* DS O^6-MeG	<0.001
DS O^6-CEG (LH1/17) *vs.* DS O^6-MeG	<0.001
DS O^6-CEG (RA1/17) *vs.* DS O^6-MeG	0.001

P values were calculated using independent t-test.

4.3.3 Restriction endonuclease site deprotection assay

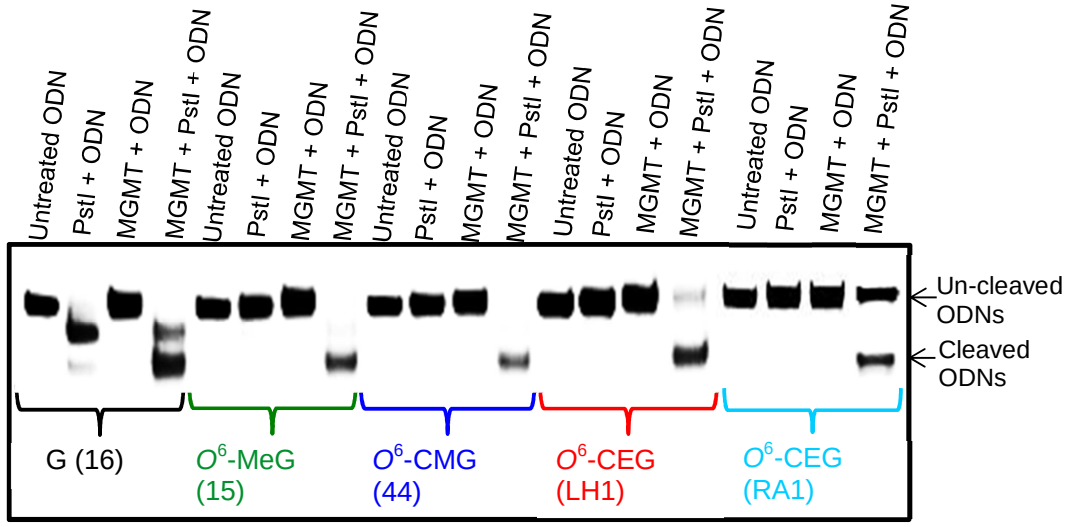
Data shown in previous section demonstrated inhibition of MGMT activity following pre-incubation with O^6-MeG, O^6-CMG and O^6-CEG containing ODNs. Restriction endonuclease site deprotection assay was carried out using three REs; PstI and SfcI for 16, 15, 44 and LH1, whilst, DW1, GMe and ConG were digested by Mob1. RA1 was digested by PstI. The RE sites for 23 mer ODNs used in this study are illustrated in Figure 2.3. Moreover, restriction endonuclease site deprotection assay, using PstI enzyme, was carried out under two different experimental conditions: standard MGMT incubation conditions (O^6-alkylG containing ODNs or control G ODN were pre-incubated with MGMT in presence of Buffer I and BSA prior to subsequent digestion with PstI (Section 2.2.15)) and MS analysis incubation conditions (MGMT was incubated with O^6-alkylG containing ODNs in presence of buffer I only (Section 2.2.15)).

Figure 4.3 demonstrated that, for all investigated ODNs (O^6-alkylG and control G), untreated ODNs have shown a single band corresponding to intact DS ODNs during non-denaturing PAGE. In addition, Figure 4.3 showed that, for each O^6-alkylG containing ODN incubated with either PstI or SfcI only, a single band corresponding to intact DS ODNs was detected by non-denaturing PAGE. This finding showed that O^6-MeG, O^6-CMG (in 44 ODN) and O^6-CEG (in RA1 and LH1 ODNs) adducts totally blocked both PstI and SfcI RE sites as evidenced by inability of both enzymes to digest any of the O^6-alkylG containing ODNs.

However, for all studied REs, blocking of restriction sites was almost completely reversed by MGMT pre-incubation with each of alkylated ODNs. Thus MGMT pre-incubation with either O^6-MeG (15), O^6-CMG (44) and O^6-CEG (LH1 and RA1) ODNs restored the RE sites enabling subsequent PstI and SfcI cleavage of ODNs almost to completion (except for RA1 that was digested with PstI only and not to completion) showing bands of smaller digested ODN fragments during non-denaturing PAGE. Moreover, for each alkylated ODN, incubation with MGMT alone had no remarkable effect as the DS ODNs band migrated unchanged during PAGE. Whilst, PstI or SfcI efficiently digested control G ODN almost to completion and this was not affected by pre-incubation with MGMT. Similar to PstI, Figure 4.4 showed that O^6-CMG in DW1 and O^6-MeG in GMe blocked MboI restriction sites and this effect was reversed by pre-incubation with MGMT. However, the corresponding control G ODN (ConG) was digested by MboI with or without MGMT pre-incubation.

Similar to standard MGMT incubation conditions, the new experimental conditions, MS incubation conditions (in Buffer I only without BSA) enabled PstI RE to differentiate between alkylated and control ODNs. PstI restriction sites were blocked by O^6-alkylG adducts. However, MGMT pre-incubation with alkylated ODNs under MS incubation conditions enabled MGMT repair of O^6-alkylG adducts restoring RE sites enabling PstI RE cleavage of duplex alkylated ODNs as shown in Figure 4.5. Regarding control G ODN, PstI RE efficiently digested it to completion which was not affected by pre-incubation with MGMT.

A



B

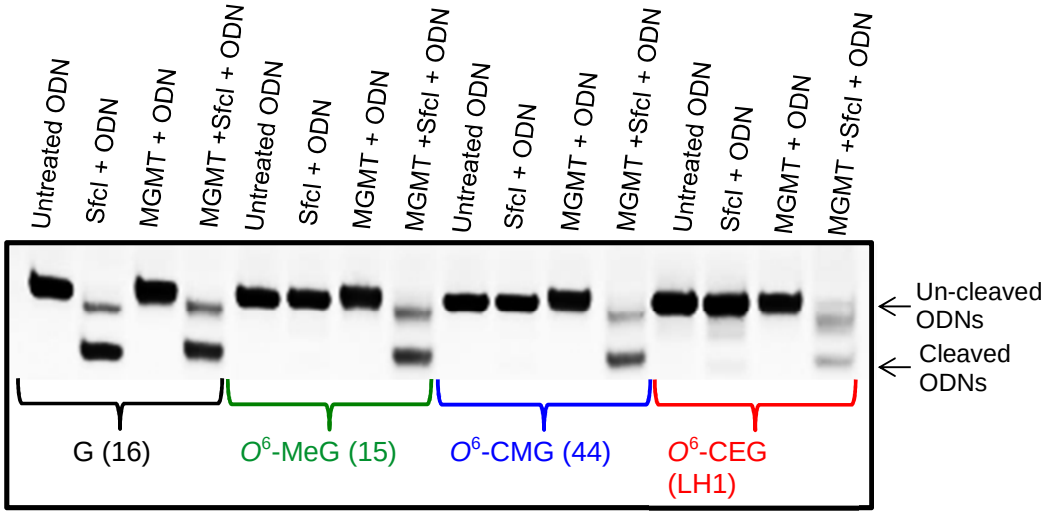


Figure 4.3: Pre-incubation with MGMT of DS ODNs containing *O*^{6}-alkylG adducts enables cleavage by (A) PstI and (B) SfcI.

23 mer ODNs (0.5 pmole), of the sequence 5'-GAA CTX CAG CTC CGT GCT GGC CC-3' where X is either *O*^{6}-MeG (15) or *O*^{6}-CMG (44) or *O*^{6}-CEG (LH1, RA1) or unmodified G base for control G (16) ODN, pre-incubated with MGMT (10 active pmole) at 37˚C for 4 hours (in IBSA buffer) then digested by PstI or SfcI (37˚C for 1 hour) then analyzed by 20% non-denaturing PAGE.

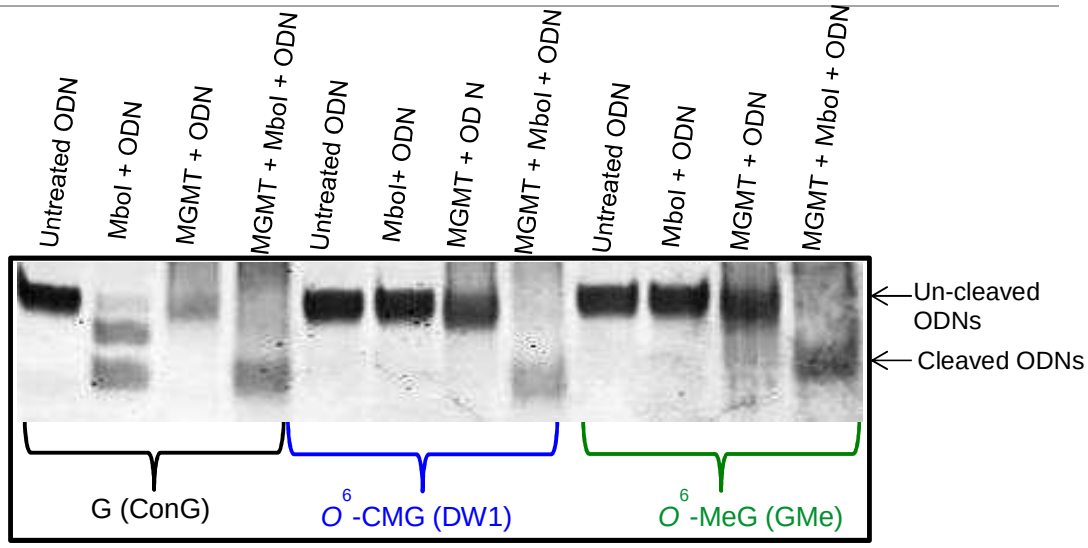


Figure 4.4: Pre-incubation with MGMT of DS ODNs containing *O*^{6}-alkylG adducts enables cleavage by MboI.

23 mer ODNs of the sequence 5'-GAA CTG CGX ATC CGT GCT GGC CC-3' where X is either *O*^{6}-MeG (GMe) or *O*^{6}-CMG (DW1) or unmodified G base for control G ODN. MGMT (10 active pmole) was pre-incubated with ODNs (0.5 pmole) at 37˚C for 4 hours (in IBSA buffer) then digested by MboI (37˚C for 1 hour) then analyzed by 20% non-denaturing PAGE.

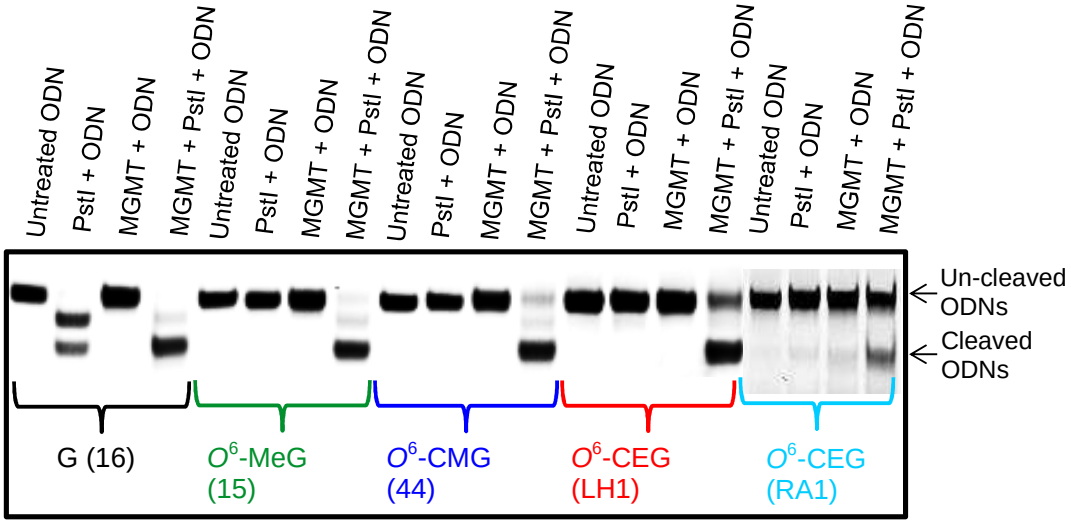


Figure 4.5: Pre-incubation with MGMT of DS ODNs containing *O*^{6}-alkylG adducts enables cleavage by PstI under MS conditions.

O^6-MeG (15), *O*^{6}-CMG (44), *O*^{6}-CEG (LH1, RA1) or unmodified G base for control G ODN 23 mer ODNs (0.5 pmole) pre-incubated with MGMT (10 active pmole) ODNs at 37˚C for 6 hours (in buffer I) then digested by PstI (37˚C for 1 hour) then analyzed by 20% non-denaturing PAGE.

4.3.4 Atl1 binding to O^6-alkylG containing ODNs by EMSA

Following Atl1 incubation with O^6-MeG (15, GMe), O^6-CMG (DW1, 44), O^6-CEG (LH1, RA1) or control G (16, ConG) ODNs, Atl1 protein was able to bind efficiently to O^6-MeG, O^6-CMG and O^6-CEG containing ODNs changing alkylated ODNs mobility through non-denaturing gel matrix causing band shift during PAGE analysis as shown in Figure 4.6 A&B. However, Atl1 couldn't bind to any of the examined control G ODNs since both 16 and ConG DS control G ODN migrated normally following incubation with Atl1 protein.

4.4 Discussion

Results presented in this chapter demonstrated that O^6-MeG, O^6-CMG, O^6-CEG containing ODNs but not control G ODNs were able to inactivate MGMT protein (50% inhibition of MGMT activity). In addition, all investigated O^6-alkylG adducts (O^6-MeG, O^6-CMG, O^6-CEG) blocked examined REs sites that were restored by pre-incubation with MGMT. PstI RE was able to differentiate between O^6-alkylG containing ODNs (O^6-MeG (15), O^6-CMG (44) and O^6-CEG (LH1 and RA1)) and the corresponding normal control G (16) ODN. Similarly, MboI RE differentiated between O^6-alkylG containing ODNs (O^6-CMG (DW1) and O^6-MeG (GMe)) and the corresponding normal control G (ConG) ODN. Moreover, all examined O^6-alkylG adducts (O^6-MeG, O^6-CMG, O^6-CEG) were recognized by Atl1 protein.

4.4.1 MGMT inactivation results

Regarding O^6-CMG adduct, the reported ability of O^6-CMG adduct to inactivate MGMT lends further support to the results by Senthong *et al*., who reported inactivation of MGMT following pre-incubation of the protein with DS 13 mer ODNs containing O^6-CMG adduct (296). Meanwhile, the current study is the first study demonstrating MGMT inactivation curves for 23 mer ODNs containing O^6-CMG adduct. Results of the current study are consistent with many previous studies that reported the ability of O^6-MeG containing ODNs to inactivate MGMT protein (294,296).

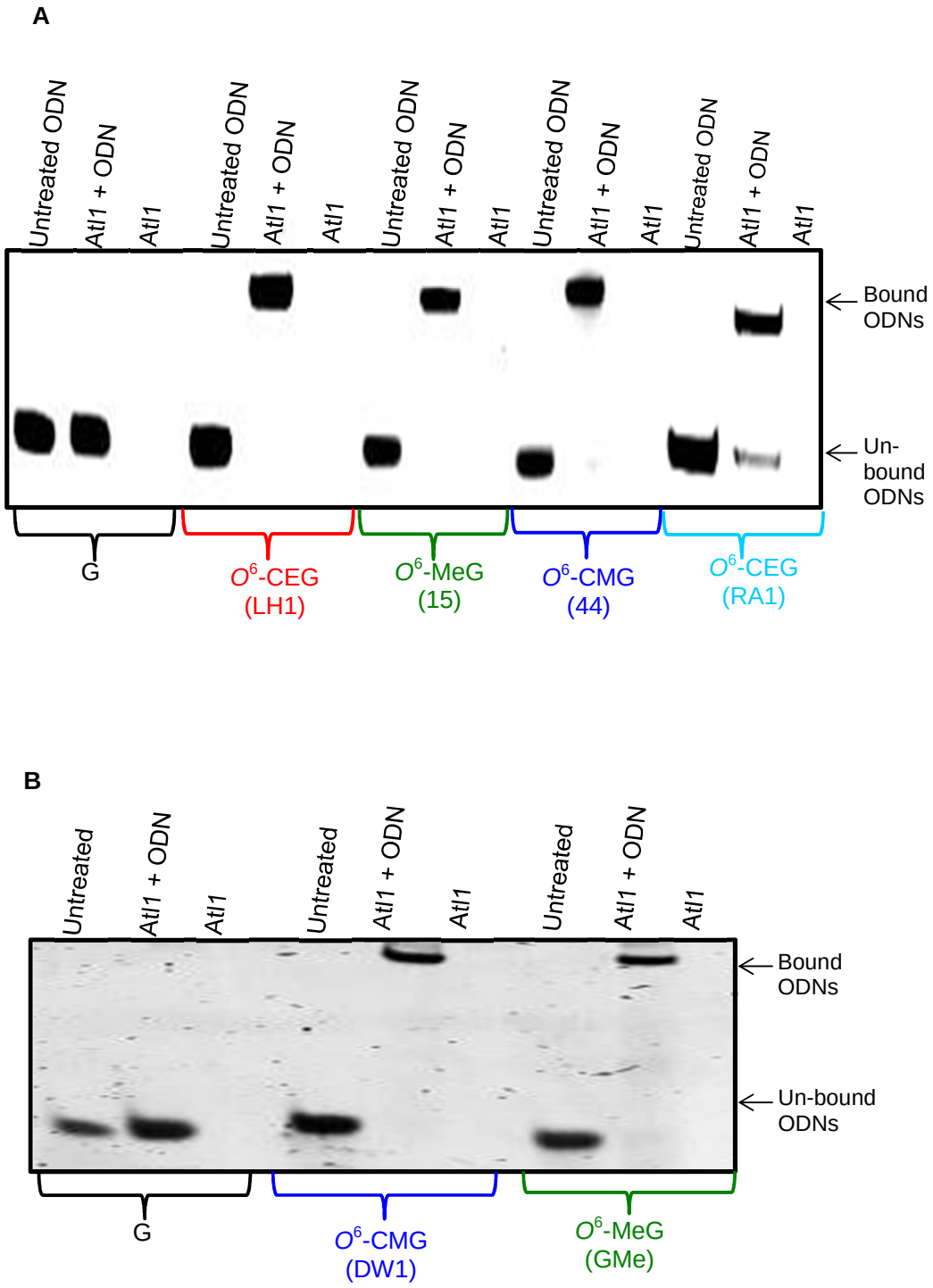


Figure 4.6: EMSA exploring Atl1 interaction with O^6-alkylG containing ODNs

Atl1 (0.5 pmole) was pre-incubated with 23 mer ODNs containing either: (A) O^6-MeG (15) or O^6-CMG (44) or O^6-CEG (LH1) or unmodified control G ODN (16), (B) O^6-MeG (GMe) or O^6-CMG (DW1) or unmodified control G ODN (ConG) at 37˚C for 20 minutes (in buffer I) then analyzed by 12.5% non-denaturing PAGE.

However, our result highlighting the ability of O^6-CMG to inactivate MGMT protein contradicts previous researches stating that O^6-CMG adduct is not a substrate for MGMT (208). The difference in observations between our study and previous work is potentially attributed to different experimental conditions. Our study was conducted using ODNs that contained a single DNA adduct in a defined DNA base sequence whilst previous report investigations relied on CT-DNA reacted with [^{14}C]-carboxy-N-nitrosoglycolic acid that induced the formation of O^6-CMG adducts in CT-DNA alongside other DNA adducts. Although, Margison and Shuker, reported MGMT inactivation following incubation with carboxymethylated DNA, their conclusion that O^6-CMG adduct is not a substrate for MGMT was based on the inability of detection of transfer of ^{14}C-labelled carboxymethyl groups to protein, as assayed by radioisotopic assay. Alternatively, they attributed the inactivation of MGMT to a substantial O^6-MeG adduct, that was formed as a byproduct of DNA carboxymethylation process and potentially interfered with MGMT repair of O^6-CMG adduct. Therefore, further MS analysis of MGMT protein following reaction with O^6-CMG ODNs is critically required to resolve this conflict.

As for O^6-CEG adduct, it is quite a novel adduct that has not been studied before and the current work is the first to provide an evidence on the ability of such adduct to inactivate MGMT protein suggesting that O^6-CEG adduct is a substrate for MGMT repair which was confirmed by further assays.

Numerous mechanisms could be used for explaining the ability of O^6-alkylG containing ODNs to inactivate MGMT protein. Firstly, inhibition of MGMT activity could be attributed to competitive inhibition as a consequence of alkylated ODNs tight binding to MGMT protein which might interfere with ability of MGMT protein to repair any other present O^6-alkylG adducts (269,279). This mechanism could explain the ability of DS control G to achieve a certain level of MGMT inactivation (Highest employed concentration of DS G ODN caused 30% reduction in MGMT activity compared to control MGMT). Secondly, MGMT inactivation of O^6-alkylG containing ODNs is potentially a consequence of alkyl group transfer from the modified guanine base to MGMT active cysteine residue (Cys-145) (275). However, MGMT inactivation assays could not provide enough evidence about mechanisms involved in inhibition of MGMT activity by O^6-alkylG containing ODNs. For O^6-MeG adduct, there is ample literature to support the latter mechanism (358,359). However, as for O^6-CMG and O^6-CEG little or

no evidence is available about the mechanisms of MGMT inactivation to support any of the suggested mechanisms over the other.

Therefore, further MS analysis of MGMT incubated with alkylated ODNs is indispensable for verification of the inactivation mechanism since in the former case alkyl group would remain in the ODN following MGMT incubation, whilst in the latter case the alkyl group would be transferred to MGMT causing mass shift in the MGMT tryptic peptide containing the active site.

Recorded IC_{50} values for both SS and DS O^6-alkylG ODNs (Table 4.1) revealed a statistically significant difference between IC_{50} of DS O^6-alkylG ODNs compared to corresponding SS O^6-alkylG ODNs indicating that DS O^6-alkylG containing ODNs more efficiently inactivated MGMT compared to the corresponding SS ODNs. Moreover, R_{IC50} (IC_{50} of SS ODNs/ IC_{50} of corresponding DS ODNs) was 2.41$\pm$0.35, 1.90+0.20 and 2.66$\pm$0.23 and 3.62$\pm$0.45 for O^6-MeG (15), O^6-CMG (44), O^6-CEG (LH1 and RA1) containing ODNs, respectively. These ratios displayed that IC_{50} values of SS O^6-alkylG containing ODNs were at least 2-fold of that of corresponding DS O^6-alkylG containing ODNs. In other words, SS O^6-alkylG containing ODNs required at least twice the required amount of DS O^6-alkylG containing ODNs to achieve 50% inhibition of MGMT activity.

Moreover, DW1 (O^6-CMG) ODN displayed a more remarkable difference between SS and DS ODNs as SS DW1 failed to achieve 50% MGMT inactivation at concentrations up to 250 nM, whilst, DS DW1 demonstrated IC_{50} at 15.7$\pm$2.1 nM. Notably, the other examined O^6-CMG ODN (44) achieved 50% MGMT inactivation using both SS and DS 44 ODN. IC_{50} value of DS DW1 ODN approximately was 37-fold that of DS 44 ODN. The difference between 44 and DW1 ODNs might be attributed to local DNA sequence context. Similarly, a remarkable difference was reported between the two examined O^6-CEG ODNs, namely, LH1 and RA1, since IC_{50} values of SS and DS RA1 ODN were approximately 7-fold the corresponding figures of LH1 despite both ODNs had the same DNA sequence. These results highlighted a problem, therefore, further MS investigations were required to confirm the identity of the alkyl adducts at the O^6 position of the modified guanine base of ODNs 44, DW1, LH1 and RA1 to investigate the reason behind the recorded differences between 44 *vs*. DW1 and LH1 *vs*. RA1.

The higher ability of DS O^6-alkylG ODNs to inhibit MGMT activity is potentially attributed to number of reasons that may act solely or in combination. First of all, MGMT binds DS DNA with slightly higher affinity than SS DNA (360). Moreover, slower dissociation rates of MGMT and DS DNA substrate complex compared to SS DNA substrates could potentially contribute to better nucleotide flipping and alkyl group displacement. However, in contrast with our results, Mijal *et al*., reported that relative MGMT reaction rates were not significant when performed with DS substrates compared to SS ones, except for the human mutant E166G that repaired O^6-pobG more efficiently when DS substrates were employed (341). The disagreement in the results between the previous study and the present work is attributed to totally different experimental conditions in terms of the type of the assay, nature of alkyl adduct as well as size of DNA substrates employed.

The present work has shown IC_{50} values (nM) of 0.17±0.02 and 0.07±0.01 nM (mean±SD) for SS and DS O^6-MeG ODNs, respectively and 0.80±0.14 and 0.42±0.01 nM (mean±SD) for SS and DS O^6-CMG (44) ODNs, respectively. Whilst, DS DW1 (O^6-CMG ODN) showed IC_{50} of 15.7±2.1 (nM). Previous studies reported that IC_{50} values for O^6-alkylG were 15.7 nM (DS O^6-MeG 16 mer ODNs) (294) and 0.9 nM and 1.7 nM for O^6-MeG and O^6-CMG, respectively (DS 13 mer ODNs) (296). Notably, reported difference in recorded IC_{50} values for O^6-MeG and O^6-CMG ODNs between previous studies and the current work is potentially a consequence of the difference in the size of the DNA substrates as well as the effect of the local DNA sequence context (sequence of the DNA that surrounds O^6-alkylG adduct) (291).

These results highlighted that for both SS and DS alkylated DNA substrates the potent MGMT inactivator was O^6-MeG followed by either O^6-CMG or O^6-CEG. Since the variations in inactivation assay results recorded for the two different preparations of either O^6-CMG (44 & DW1) or O^6-CEG (LH1 & RA1) ODNs interfered with stating which is the more potent MGMT inactivator. Results of the current work are consistent with previous reports stating that MGMT repair preference is markedly dependant on the alkyl group identity. Consistently, they suggested that MGMT repair efficiency diminishes as adduct size increase which could be explained by the generation of steric hindrance in MGMT active site pocket (275,290,361).

4.4.2 Restriction endonuclease site deprotection assay

Restriction endonuclease site deprotection assay provides great potential as a technique to investigate the mechanism of MGMT inactivation by O^6-alkylG containing ODNs. O^6-MeG blocks RE sites, therefore, if MGMT functions through removal of the alkyl group, the blocked RE sites should be restored following pre-incubation of MGMT with O^6-alkylG containing ODNs providing initial insights of MGMT repair mechanisms.

Subsequent investigation of MGMT mediated repair of the three studied O^6-alkylG adducts involved restriction endonuclease site deprotection assay using 3 REs PstI, SfcI and MboI which demonstrated several important findings. Firstly, O^6-MeG (in 15 ODN), O^6-CMG (in 44 ODN) and O^6-CEG (in LH1 and RA1 ODNs) adducts were able to block the RE sites for PstI enzyme. That was evidenced by the inability of PstI to digest any of the O^6-alkylG containing ODNs (15, 44, LH1 and RA1), without MGMT pre-incubation. Secondly, MGMT efficiently repaired the three examined O^6-alkylG adducts restoring RE sites enabling PstI digestion of O^6-alkylG containing ODNs almost to completion (except for RA1 the reaction was not complete), following incubation with MGMT. In addition, similar findings were reported for SfcI RE that was used for digesting 15, 44 and LH1 ODNs. Results of the present work were consistent with previous studies demonstrated similar findings for O^6-MeG and O^6-CMG containing ODNs (296,313).

Conclusively, PstI and SfcI REs were able to differentiate between O^6-alkylG containing ODNs (O^6-MeG (15), O^6-CMG (44) and O^6-CEG (LH1 and RA1, by PstI only)) and the corresponding normal control G (16) ODN. The current study is the first demonstrating the ability of MboI RE to differentiate between O^6-alkylG containing ODNs (O^6-MeG (GMe) and O^6-CMG (DW1) and the corresponding normal control G (ConG) ODN.

The significance of the previous results involves providing initial evidence of the mechanism of MGMT mediated repair of O^6-MeG, O^6-CMG and O^6-CEG adducts suggesting the removal of the alkyl group from the alkylated ODNs. However, further MS analysis of MGMT protein incubated with O^6-alkylG containing ODNs is required to confirm this hypothesis.

Notably, restriction endonuclease site deprotection assay, using the RE PstI, was carried out under two experimental conditions. Initially, it was performed under standard MGMT incubation conditions where MGMT was incubated with alkylated ODNs in presence of both buffer I and BSA (340,362).

However, standard MGMT incubation conditions were not appropriate for MS analysis as BSA tryptic peptides could interfere with MGMT tryptic peptides. To remedy this situation, new MS incubation conditions were developed incubating MGMT with alkylated ODNs in buffer I without BSA and the efficiency of the new incubation conditions were compared to standard ones *via* restriction endonuclease site deprotection assay, using the RE PstI. Results showed that the elimination of BSA from incubation mixtures did not appear to have a major impact on MGMT repair of O^6-alkylG adducts and it was approximately as effective as standard MGMT incubation conditions. This conclusion was evidenced by the fact that the newly developed MS incubation conditions enabled MGMT repair of O^6-MeG, O^6-CMG and O^6-CEG adducts. Our results are in line with previous work that used similar incubation conditions studying MGMT repair of O^6-alkylG adducts *via* MS analysis of alkylated MGMT tryptic peptides (296,341).

4.4.3 EMSA results

Present work investigated the ability of DNA damage sensing protein Atl1, to bind to O^6-MeG, O^6-CMG and O^6-CEG adducts in all investigated ODNs. Results of the current study showed that, Atl1 protein exhibited an efficient ability to bind the three examined O^6-alkylG adducts. Results of the present work are in line with previous studies that have demonstrated the ability of Atl1 to bind specifically to and tightly to a wide range of O^6-alkylG adducts (312,313). Atl1 binds SS DNA containing O^6-MeG, O^6-benzylG and O^6-(4-bromothenyl)G or O^6-HOEtG (314), and DS ODNs containing O^6-MeG as well as O^6-pobG adducts (315). Nevertheless, the current study is the first demonstrating that Atl1 could efficiently bind to O^6-CMG and O^6-CEG causing a visible gel band shift during EMSA analysis. These results confirm the Atl1 overlapping of MGMT substrate specificity which raises the question of the evolutionary relationship between the MGMT and Atl1.

4.5 Conclusion

O^6-MeG, O^6-CMG adducts as well as the novel O^6-CEG adduct are substrates for the DNA repair protein MGMT and recognized by DNA damage sensing protein Atl1. The three studied O^6-alkylG adducts could efficiently inhibit MGMT activity when presented in SS or DS ODNs, whilst, DS O^6-alkylG containing ODNs are more potent MGMT inactivators compared to SS ODNs. Moreover, O^6-MeG, O^6-CMG and O^6-CEG block RE site for the two investigated REs PstI and SfcI. Nevertheless, MGMT pre-incubation with O^6-MeG, O^6-CMG and O^6-CEG containing ODNs restored RE sites enabling PstI and SfcI subsequent digestion of alkylated ODNs. Consistently, mechanisms by which MGMT repair novel adducts O^6-CMG and O^6-CEG are notably intriguing and substantially compelling to MS analysis of MGMT reacted with O^6-alkylG containing ODNs to clarify them.

5 Chapter 5: Repair of O^6-alkylG adducts containing ODNs by MGMT analysed by MS to identify and quantify alkylation

5.1 Introduction

MGMT has been shown to be crucial for cellular protection against the carcinogenicity and mutagenicity of formed O^6-alkylG adducts (269,363). There is a vast literature describing studies on O^6-alkylG adduct repair by human MGMT protein (269,275,364). MGMT is a remarkable protein that, alone in one step, can repair O^6-alkylG adducts *via* the transfer of the alkyl group from the O^6 position of the modified guanine base to the active site cysteine residue (Cys-145) within MGMT (288,296,341). Following this alkyl group transfer, the MGMT peptide containing this active site (MGMT-ASP) now contains the alkyl group and since this results in a mass change it can be detected by MS. A previous study demonstrated both methyl and carboxymethyl group transfer to the MGMT-ASP following incubation with O^6-MeG and O^6-CMG containing ODNs, respectively (296).

Mijal and Kotandeniya demonstrated the transfer of the pob group to the active site (MGMT-ASP) following MGMT reaction with DNA ODNs containing O^6-pobG *via* MS (288,341). Kotandeniya *et al.* (288) developed a MS based approach to investigate the kinetics of MGMT mediated repair of O^6-pob adducts placed within DNA ODNs representing modified rat *H*-ras sequence opposite either C or T, and reported faster rates of alkyl group transfer when O^6-pobG base was paired with T rather than with C.

The efficiency of MGMT mediated repair of O^6-alkylG adducts is affected by a variety of factors including alkyl group structure and local DNA sequence context (291). The relative rates of repair of O^6-alkylG adducts by rat MGMT in alkylated DNA have been reported to decrease with the mass of the alkylation: methyl > ethyl > n-propyl > n-butyl > isopropyl, isobutyl (due to branched alkyl chain) > 2-hydroxyethyl (365). Previous suggestion was later contradicted by a recent study in 2007 stated that the rates of alkyl group transfer to human MGMT active site were as follows: (benzyl > methyl > ethyl > 2-hydroxyethyl > pob) highlighting that the identity of the alkyl group plays a more pivotal role in MGMT repair preference rather than size (290).The effect of annealing of the DNA substrate (SS or DS) on the efficiency of MGMT mediated repair of the O^6-alkylG adducts has not been fully explored.

5.2 Aims and objectives

Data shown previously (Chapter 4) indicated the inactivation of MGMT following incubation with ODNs containing (O^6-MeG, O^6-CMG and O^6-CEG). Hence, the overall aim of this chapter was to investigate the mechanism of MGMT inactivation by O^6-MeG, O^6-CMG and O^6-CEG adducts and to confirm that this was a consequence of the transfer of alkyl group from O^6-alkylG containing ODNs to the MGMT-ASP.

The specific objectives were:

1- To identify/characterise peptides from MBP-MGMT using MALDI-TOF MS.

2- To demonstrate that alkylation of MBP-MGMT can be detected by MALDI-TOF MS following incubation of ODNs containing O^6-MeG, O^6-CMG and O^6-CEG.

3- To compare (qualitatively) the transfer of alkyl groups from two preparations of O^6-CMG (DW1 and 44) and O^6-CEG (RA1 and 2nd preparation (LH1)) 23 mer ODNs so as to investigate the significant difference found in MGMT IC_{50} values (Chapter 4).

4- To compare relative levels of unmodified and alkylated MGMT-ASP using a further MGMT tryptic peptide fragment as the internal control, following MGMT reaction with SS and DS O^6-alkylG ODNs.

5- To compare MGMT repair preference for O^6-MeG, O^6-CMG and O^6-CEG adducts.

6- To perform MS analysis of chemically modified synthetic MGMT-ASP to confirm the identity of the alkyl groups transferred to MGMT following reaction with O^6-CMG and O^6-CEG ODNs multiple preparations.

5.3 Results

MBP-MGMT protein was incubated with excess alkylated ODNs (13 mer and 23 mer) prior to MS investigations (Section 2.2.17).

5.3.1 Identification of MBP-MGMT tryptic peptides

MALDI-TOF MS analysis of MBP-MGMT following digestion with trypsin revealed 26 peptides (10 MGMT; 16 MBP) (Figure 5.1). The peptide which contained the active cysteine residue (Cys145) generated an ion with *m/z* 1315.73 $[M+H]^+$ in the digest of the unmodified protein. Tandem MS (LC-Orbitrap) confirmed that this ion was indeed MGMT-ASP (Data are not shown). The peptide fingerprint spectrum of MBP-MGMT is dominated by two ions (*m/z* 1247.63 $[M+H]^+$ and 1315.73 $[M+H]^+$) corresponding to MGMT-TP 7 (166-175, EWLLAHEGHR) and MGMT-ASP (136-147, GNPVPILIPCHR), respectively.

MGMT tryptic peptides were matched to their constituent protein using MASCOT (UniProtKB/Swiss-Prot; Last modified January 10, 2017). There was a good agreement (MASCOT score 114 and 54% coverage) to MGMT_HUMAN (Accession number: P16455) (113 amino acids out of the total of 207; expect=$8X10^{-8}$), as shown in Figure 5.2. MGMT identified tryptic peptides are shown in Table 5.1.

Only 54 % of the MGMT sequence is detected, which is attributed to the remainder generating tryptic peptides with *m/z* that are either too small – and therefore in the region of the spectrum dominated by matrix species or too large to be successfully detected. MS data was acquired from 800-2300 *m/z* and for the peptides expected in this range there was sequence coverage of 86% (113 out of 131 amino acids).

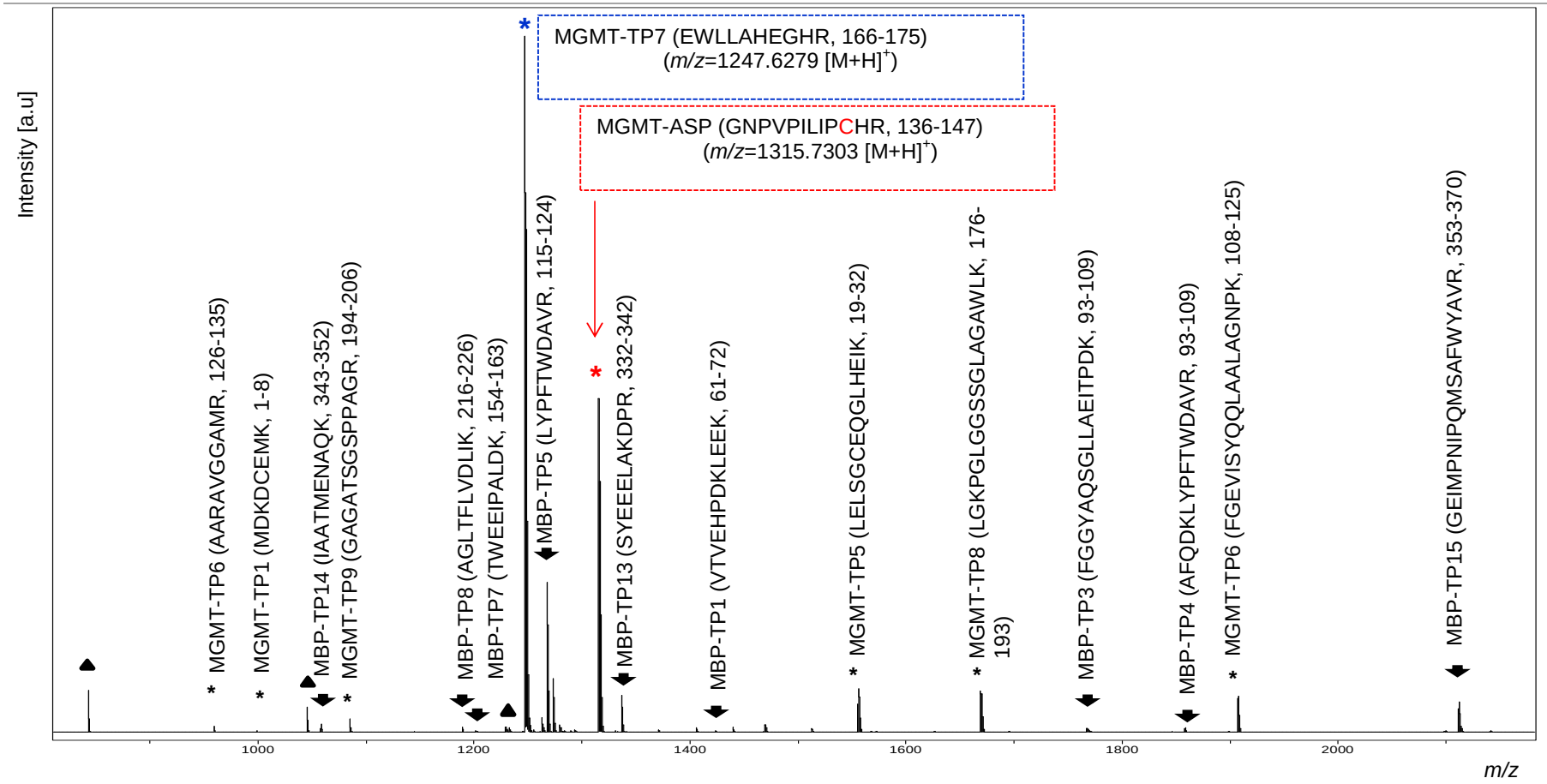


Figure 5.1: MALDI-TOF mass spectra of MBP-MGMT tryptic digest.

Sequences of MGMT and MBP tryptic peptides identified with (*) and (▼), respectively. Peaks with triangles (▲) represent trypsin autolysis products. Peak intensities are shown in arbitrary units on the y-axis.

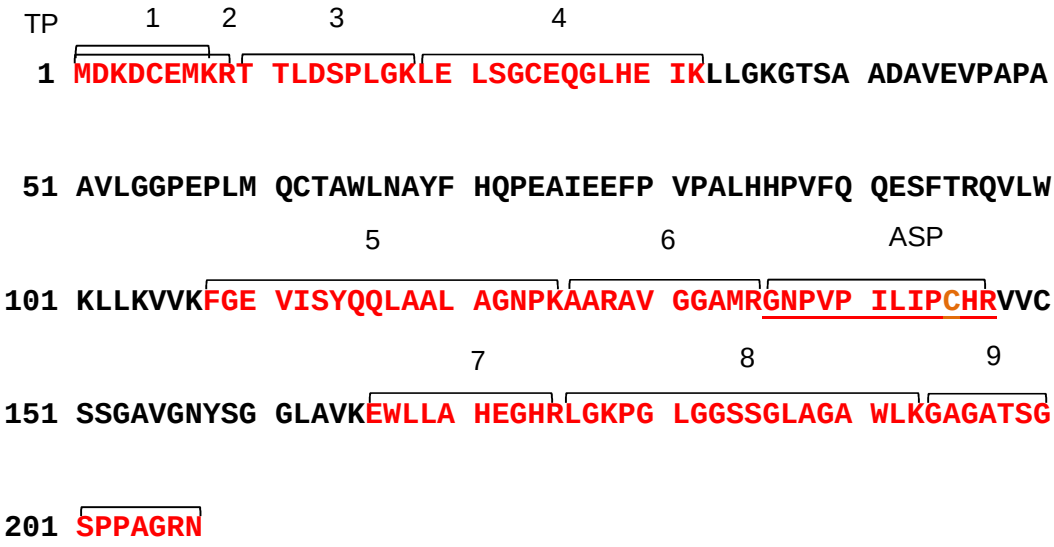


Figure 5.2: MGMT sequences detected using MALDI-TOF MS.

Identified peptides are shown in red; the active site peptide (MGMT-ASP) is underlined. 54% sequence coverage (Peptide tolerance 0.2 and allowed missed cleavages, 2); TP: tryptic peptides.

A single peptide was not detected (VVCSSGAVGNYSGGLAVK, 148-165) (*m/z*=1667.8421) which is perhaps due to the hydrophobic amino acids supressing its release from the C18-Ziptips.

The Mascot search also identified the MBP fusion protein, with a high Mascot score (128) attributed to MALE_ECO57 (Accession number=P0AEX9) and 52% sequence coverage (expect=$9X10^{-8}$) as shown in Figure 5.3. Detected MBP tryptic peptides are shown in Table 5.2.

As for MGMT the lack of full sequence coverage of MBP is attributed to the *m/z* ratios of such peptides falling outside the reasonable detection range. In meeting the first objective, it is evident that MALDI-TOF MS data has confirmed the identity of MBP-MGMT protein that has been identified by western blot analysis (Section 3.3.4).

Table 5.1: MGMT peptides arising from digestion of MBP-MGMT with trypsin and detected using MALDI-TOF MS.

Peptide name	Amino acids	Measured mass	Expected mass	Error (ppm)	Missed cleavages	Peptide
MGMT-TP1	1-8	999.4544	999.3871	67	1	MDKDCEMK
MGMT-TP2	1-9	1155.4780	1155.4882	-9	2	MDKDCEMKR
MGMT-TP3	10-18	931.4867	931.5022	-17	0	TTLDSPLGK
MGMT-TP4	19-32	1555.7580	1555.7712	-8	0	LELSGCEQGLHEIK
MGMT-TP5	108-125	1905.9890	1905.9996	-6	0	FGEVISYQQLAALAGNPK
MGMT-TP6	126-135	959.5071	959.5130	-6	1	AARAVGGAMR
MGMT-ASP	136-147	1315.7160	1315.7231	-6	0	GNPVPILIPCHR
MGMT-TP7	166-175	1247.618	1247.6207	-2	0	EWLLAHEGHR
MGMT-TP8	176-193	1668.9240	1668.9359	-7	0	LGKPGLGGSSGLAGAWLK
MGMT-TP9	194-206	1085.5160	1085.5261	-9	0	GAGATSGSPPAGR

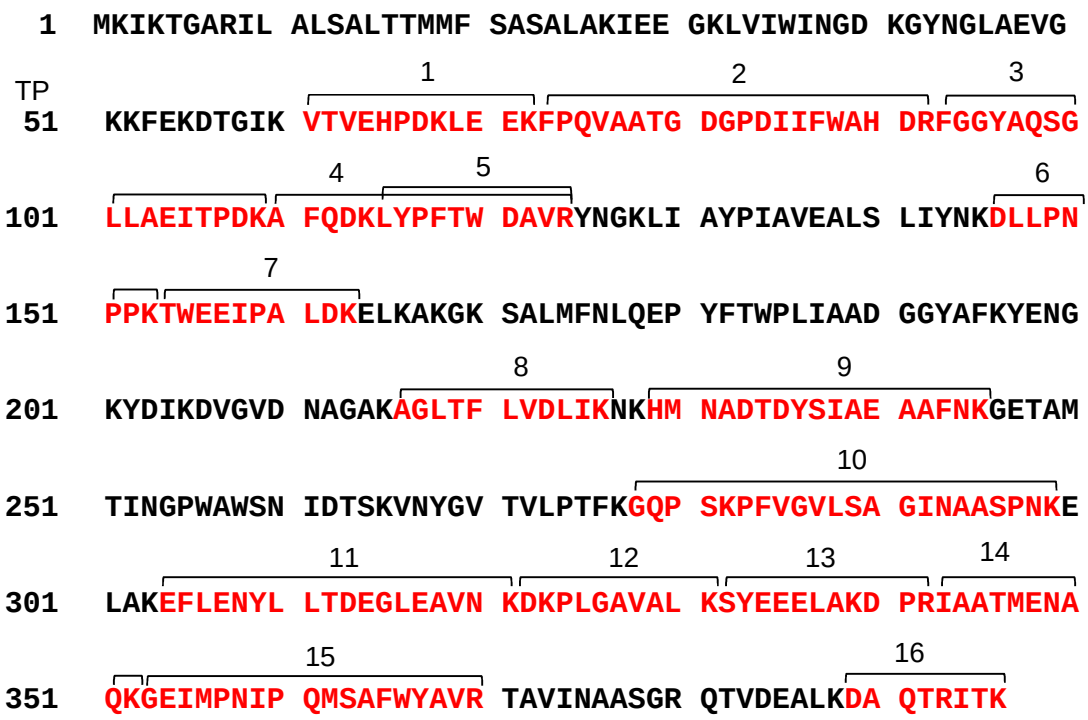


Figure 5.3: MBP sequences detected using MALDI-TOF MS.

Detected peptides are shown in red. 52% sequence coverage (Peptide tolerance 0.2 and allowed missed cleavages, 2). TP: tryptic peptides.

5.3.2 Detection of methyl group transfer to MGMT, *via* the peptide MGMT-ASP following MBP-MGMT repair of O^6-MeG adducts in 13 mer (303B) and 23 mer (15) ODNs

Initially, qualitative MS analysis of tryptic fragments of MBP-MGMT protein revealed the transfer of a methyl group from O^6-MeG to MGMT-ASP, following incubation with both SS and DS O^6-MeG 13 mer (303B) and 23 mer (15) ODNs (Figures 5.4 and 5.5). Both O^6-MeG 13 mer (303B) and 23 mer (15) ODNs were capable of producing methylated MGMT-ASP (GNPVPILIPMe-CHR) that presents at *m/z* 1329.74 $[M+H]^+$ Figure 5.4 I and II (SS & DS O^6-MeG 13 mer, respectively) and Figure 5.5 I and II (SS & DS O^6-MeG 23 mer, respectively). In both cases (13 and 23 mer O^6-MeG ODNs) the recorded mass shift in MGMT-ASP is reported with respect to a methyl group transfer (mass shift = 14.01 Da).

Table 5.2: MBP peptides arising from digestion of MBP-MGMT with trypsin and detected using MALDI-TOF MS.

Peptide name	Start-End	Measured mass	Expected mass	Error (ppm)	Missed cleavages	Peptide
MBP-TP1	61-72	1423.7493	1423.7354	-9	1	VTVEHPDKLEEK
MBP-TP2	73-92	2213.3347	2213.0702	15	0	FPQVAATGDGPDIIFWAHDR
MBP-TP3	93-109	1766.9570	1766.8887	-11	0	FGGYAQSGLLAEITPDK
MBP-TP4	110-124	1857.0355	1856.9257	-1	1	AFQDKLYPFTWDAVR
MBP-TP5	115-124	1267.6460	1267.6397	-5	0	LYPFTWDAVR
MBP-TP6	146-153	893.5003	893.5018	-19	0	DLLPNPPK
MBP-TP7	154-163	1201.5951	1201.6026	-14	0	TWEEIPALDK
MBP-TP8	216-226	1189.7020	1189.7118	-15	0	AGLTFLVDLIK
MBP-TP9	229-245	1897.9399	1897.8312	-7	0	HMNADTDYSIAEAAFNK
MBP-TP10	278-299	2139.3523	2139.1484	1	0	GQPSKPFVGVLSAGINAASPNK
MBP-TP11	304-321	2097.2070	2097.0313	-5	0	EFLENYLLTDEGLEAVNK
MBP-TP12	322-331	1011.6041	1011.6124	-17	0	DKPLGAVALK
MBP-TP13	332-342	1336.6410	1336.6306	-5	1	SYEEELAKDPR
MBP-TP14	343-352	1076.5304	1076.5332	-9	0	IAATMENAQK
MBP-TP15	353-370	2110.2368	2110.0176	13	0	GEIMPNIPQMSAFWYAVR
MBP-TP16	389-396	932.5024	932.5087	-9	1	DAQTRITK

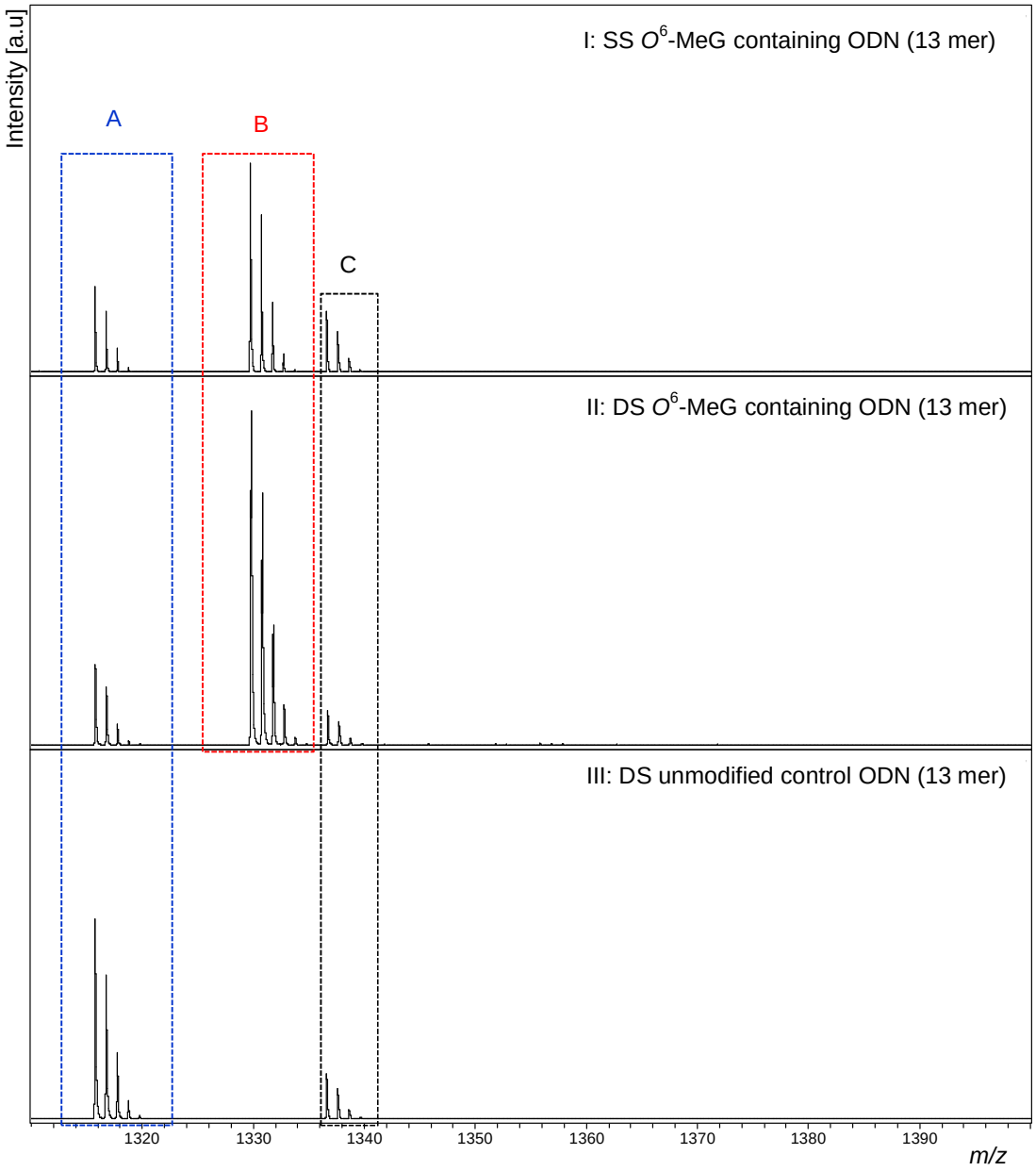


Figure 5.4: Mass Spectra of tryptic digests of MBP-MGMT reacted with ODNs of the sequence GCCATGXCTAGTA where X is (I and II) *O*^{6}-MeG or (III) control G (374/113) ODN.

I & II. SS (303B) and DS (303B/113) *O*^{6}-MeG ODNs, respectively. Multiple peptides were attributable to alkylation as well as the ASP ion (*m/z* 1315.73) observed when MGMT was incubated with unmodified ODN as well as SS and DS 303B ODNs. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147) *m/z* 1315.73. **B**. Methylated MGMT-ASP (GNPVPILIPMe-CHR, 136-147) *m/z* 1329.74. **C**. MBP-TP13 (SYEEELAKDPR, 332-342) *m/z* 1336.54. Peak intensities are shown in arbitrary units on the y-axis and intensity scale is the same for all panels.

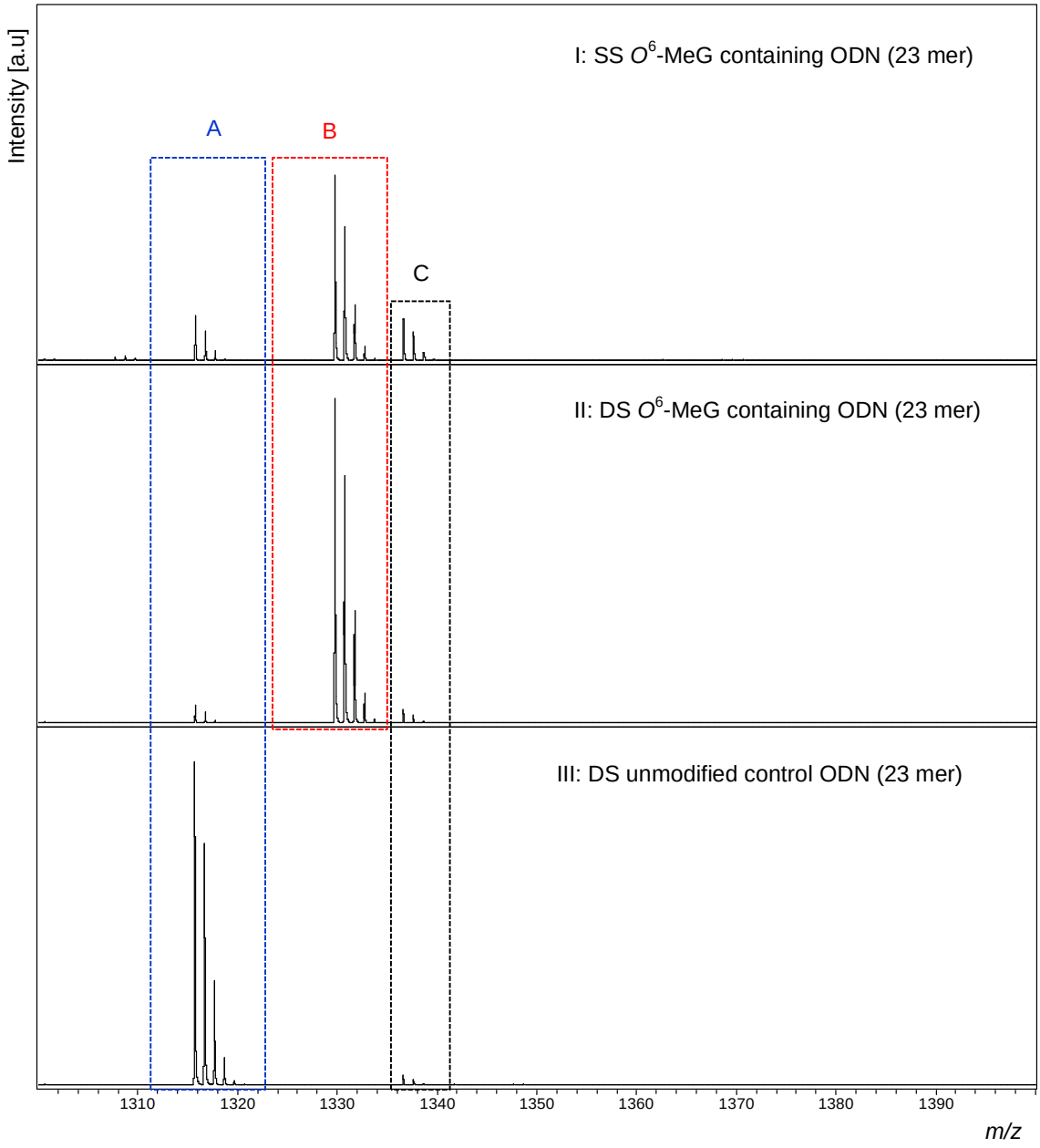


Figure 5.5: Mass Spectra of tryptic digests of MBP-MGMT following reaction with ODNs of the sequence GAA CTX CAG CTC CGT GCT GGC CC where X is (I and II) *O*^{6}-MeG or (III) control G (16/43) ODN.

I & II. SS (15) and DS (15/43) *O*^{6}-MeG ODNs, respectively. Multiple peptides were attributable to alkylation as well as the ASP ion (*m/z* 1315.73) observed when MGMT was incubated with unmodified ODN as well as SS and DS 15 ODNs. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147) *m/z* 1315.73. **B**. Methylated MGMT-ASP (GNPVPILIPMe-CHR, 136-147) *m/z* 1329.74. **C**. MBP-TP13 (SYEEELAKDPR, 332-342) *m/z* 1336.54. Peak intensities are shown in arbitrary units on the y-axis and intensity scale is the same for all panels.

In addition, MS analysis of tryptic digests of MBP-MGMT incubated with control G (13 mer and 23 mer) ODNs generated multiple MGMT and MBP peptides including unmodified MGMT-ASP (GNPVPILIPCHR) but there was no trace of methylated MGMT-ASP as shown in Figure 5.4 III & Figure 5.5 III (13 mer G & 23 mer G, respectively).

5.3.3 MBP-MGMT repair of O^6-CMG adducts in 13 and 23 mer ODNs (SS & DS)

Three O^6-CMG ODN preparations were investigated: O^6-CMG 13 mer (OW18 SS and DS), O^6-CMG 23 mer first preparation (DW1; both SS and DS) and O^6-CMG 23 mer second preparation (44/43; DS only due to unavailability of the SS ODN). Primarily, MGMT mediated repair of O^6-CMG adduct was demonstrated following reaction with both SS & DS O^6-CMG 13 mer ODNs. Figure 5.6 I and II show the ion formed from carboxymethylated MGMT-ASP (*m/z* 1373.74 $[M+H]^+$) following the reaction with SS & DS O^6-CMG 13 mer ODNs, respectively, alongside other ions.

For the 23 mer O^6-CMG ODN, DW1, carboxymethylated MGMT-ASP ion was detected following MGMT incubation with DS O^6-CMG 23 mer ODN (DW1/C) (Figure 5.7 II), but not SS O^6-CMG 23 mer ODN (DW1) (Figure 5.7 I). In both 13 mer (SS & DS-OW18) and DS DW1 (23 mer) O^6-CMG ODNs, the recorded mass shift in MGMT-ASP was reported with respect to a carboxymethyl group transfer (mass shift = 58.01 Da). MS analysis of MGMT incubated with the second preparation of 23 mer O^6-CMG ODN (44/43) did not show any evidence of carboxymethylation, but evidence for alkyl group transfer to MGMT was provided by the significant decrease in the unmodified ASP ion signal, compared to control G ODN as shown in Figure 5.8I. This is associated with the appearance of a new species, completely absent in the control, at *m/z* 1372.77 $[M+H]^+$. This ion with a mass shift of 57.04 from the unmodified peptide is assigned to the transfer of a carboxamidomethyl (acetamide) group putatively. Other MGMT peptides remained unaffected by incubation with DS 44 ODN, also supporting this assignment rather than some non-specific adductation.

Unmodified MGMT-ASP, but not alkylated ASP, was detected in tryptic digests of the unmodified protein reacted with G ODNs (13 and 23 mer Figure 5.6 III, 5.7 III and 5.8 II, respectively). In addition, alkyl group transfer to MGMT-ASP was quantitatively assessed as shown in an upcoming section.

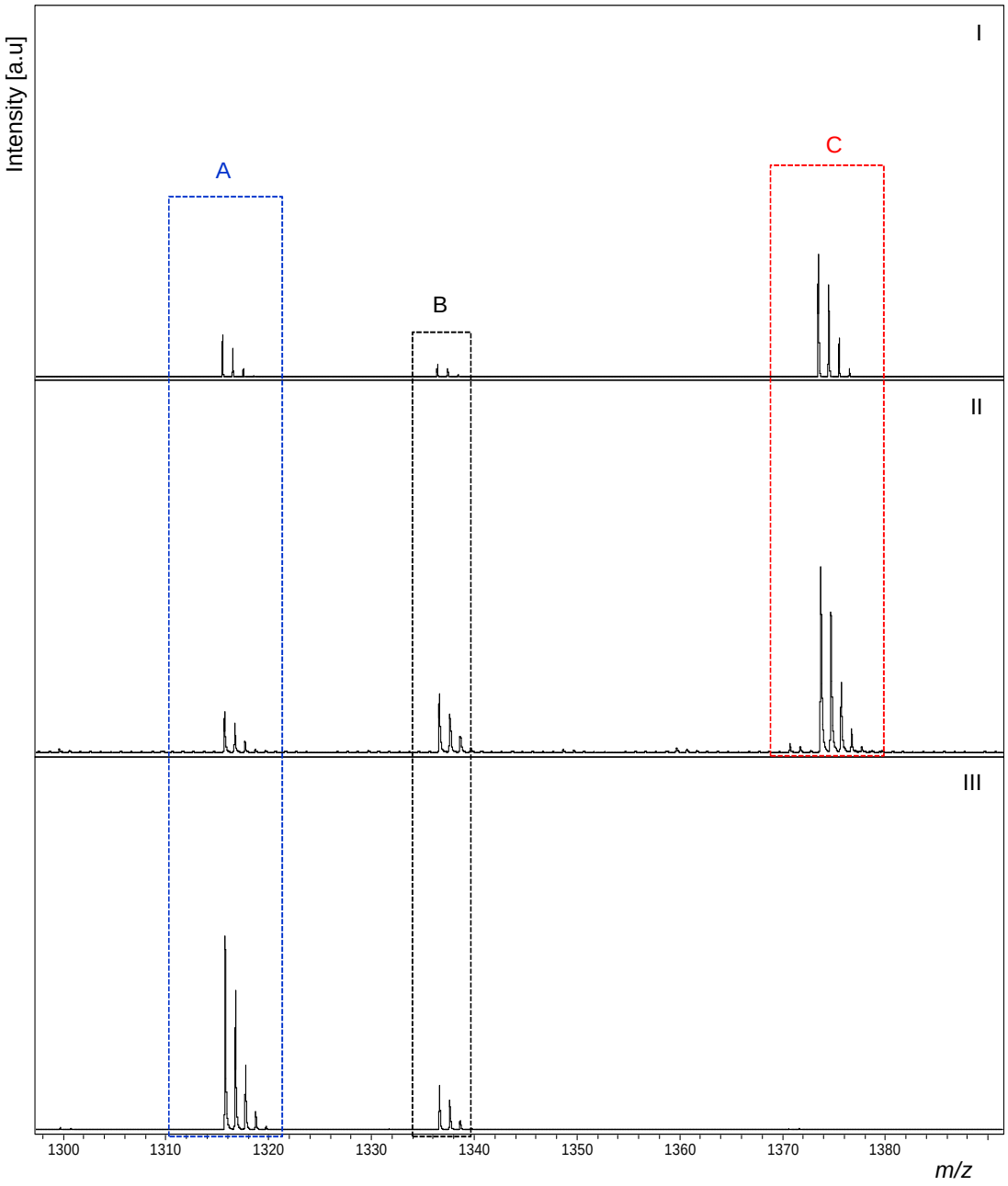


Figure 5.6: Mass Spectra of tryptic digests of MBP-MGMT following reaction with (I) SS *O*^{6}-CMG (OW18), (II) DS *O*^{6}-CMG (OW18/113) or (III) control G (374/113) 13 mer ODNs.

Multiple peptides were attributable to alkylation as well as the ASP ion (*m/z* 1315.73) observed when MGMT was incubated with unmodified ODN as well as SS and DS OW18 ODNs. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147) *m/z* 1315.73. **B**. MBP-TP13 (SYEEELAKDPR, 332-342) *m/z* 1336.54. **C**. Carboxymethylated MGMT-ASP (GNPVPILIPCM-CHR, 136-147) *m/z* 1373.74. Peak intensities are shown in arbitrary units on the y-axis and intensity scale is the same for the three panels (I, II and III).

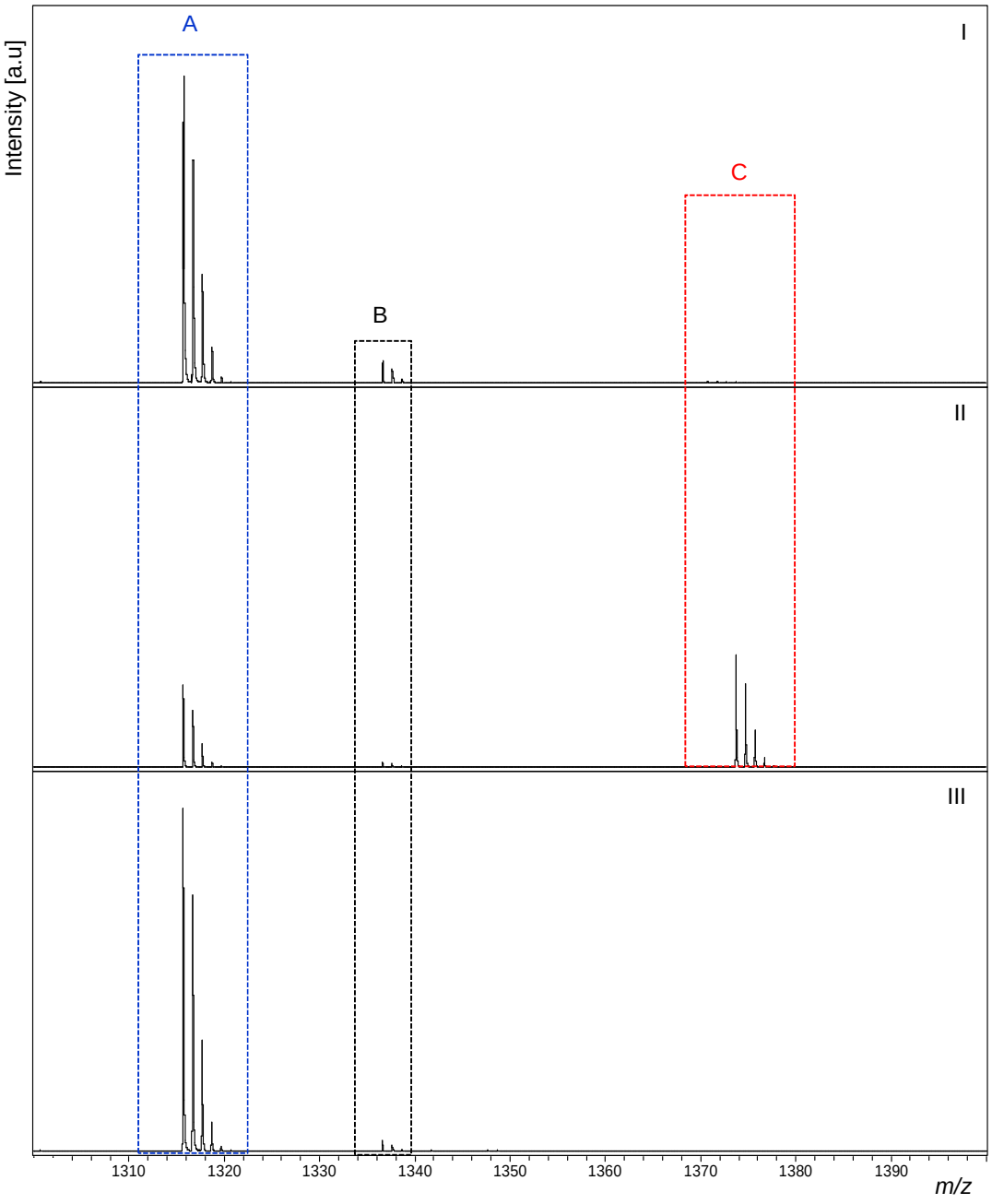


Figure 5.7: Mass Spectra of tryptic digests of MBP-MGMT following reaction with (I) SS O^6-CMG (DW1), (II) DS O^6-CMG (DW1/C) or (III) control G (16/43) 23 mer ODNs.

Multiple peptides were attributable to alkylation as well as the ASP ion (*m/z* 1315.73) observed when MGMT was incubated with unmodified ODN as well as SS and DS DW1 ODNs. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147) *m/z* 1315.73. **B**. MBP-TP13 (SYEEELAKDPR, 332-342) *m/z* 1336.54. **C**. Carboxymethylated MGMT-ASP (GNPVPILIPCM-CHR, 136-147) *m/z* 1373.74. Peak intensities are shown in arbitrary units on the y-axis and intensity scale is the same for the three panels (I, II and III).

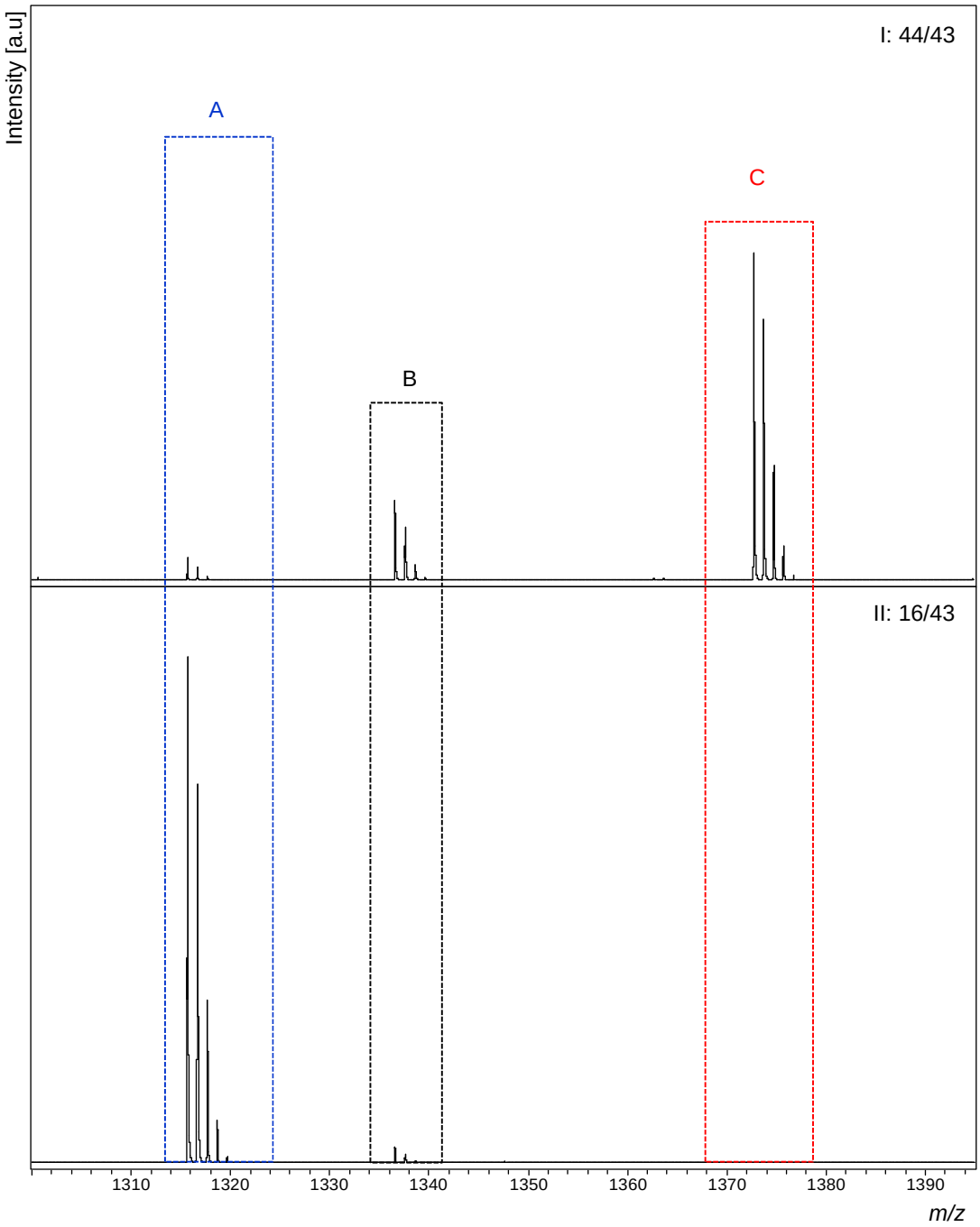


Figure 5.8: Mass spectra of tryptic digests of MBP-MGMT reacted with (I) 44/43, or (II) control G (374/113) ODNs.

Multiple peptides were attributable to alkylation as well as the ASP ion (*m/z* 1315.73), observed when MGMT was incubated with unmodified ODN as well as DS 44 ODN. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147) *m/z* 1315.73. **B**. MBP-TP13 (SYEEELAKDPR, 332-342) m/z 1336.54. **C**. Putative carboxamidomethylated MGMT-ASP (GNPVPILIPcarboxamidomethyl-CHR, 136-147) *m/z* 1372.77. Peak intensities are shown in arbitrary units on the y-axis and intensity scale is the same for the two panels (I and II).

5.3.4 MBP-MGMT repair of O^6-CEG adducts in SS (SRO1) and DS (SRO1/113) 13 mer ODNs

Following incubation of MGMT with SS 13 mer SRO1, MS revealed the modification of the active site by *S*-methylcysteine modified-ASP (*m/z* 1329.74 $[M+H]^+$) only and showed no trace of *S*-carboxyethylcysteine modification (predicted *m/z* 1387.75 $[M+H]^+$) (Figure 5.9 I). Similarly, MS analysis of MGMT reacted with DS SRO1/113 failed to detect carboxyethylated MGMT-ASP, however, methylated MGMT-ASP was revealed at *m/z* 1329.74. Moreover, an additional species was detected at *m/z* 1386.77 following reaction with SRO1/113 (Figure 5.9 II). This ion was not detected in the control and induced the mass shift of 71.04 Da from MGMT-ASP. As with 44 ODN, the incubation of SRO1 with MGMT resulted in the transfer of an alkyl group to MGMT-ASP and formation of an alkylated ASP that was utterly absent in control.

However, the molecular weight of the detected peptide cannot be assigned to a carboxyethylation as it was approximately 0.98 Da less than expected for the transfer of carboxyethyl group to MGMT-ASP. Based on the analogy with carboxamidomethyl, the detected species is putatively assigned to the transfer of carboxamidoethyl to MGMT, following incubation with DS 13 mer SRO1/113 which raises a question about the identity of the alkyl group at the O^6-position of the modified guanine base in SRO1. Furthermore, the unmodified MGMT-ASP (*m/z* 1315.7 $[M+H]^+$) was detected after incubation with G ODNs and after incubation with both SS and DS SRO1 13 mer (Figure 5.9 III).

For SRO1 ODN, a quantitative analysis was conducted to compare the relative abundance of the 2 detected modified MGMT-ASPs with the evaluation of the ratio R_2 corresponding to putative carboxamidoethylated MGMT-ASP PA/methylated MGMT-ASP PA; as explained in section 2.2.17.3. R_2 was zero (mean$\pm$SD, n=3) following MGMT reaction with control and SS SRO1. By contrast, reaction with DS SRO1/113 caused a significant increase in the R_2 to 0.62$\pm$0.13, with respect to the control and SS SRO1 (P<0.001) (Table 5.3).

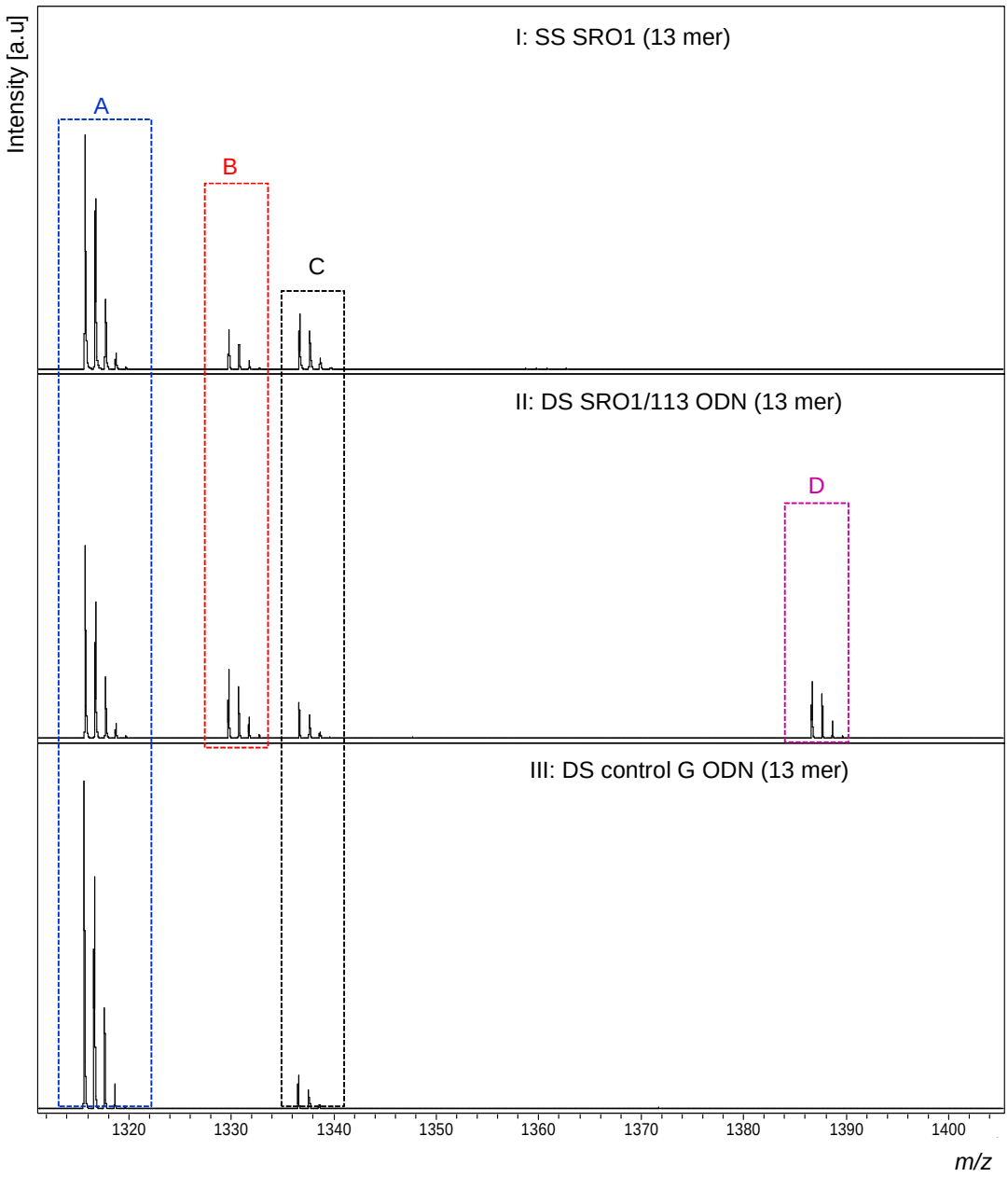


Figure 5.9: Mass Spectra of MBP-MGMT tryptic peptides following incubation with (I) SRO1, (II) SRO1/113 and (III) unmodified G ODN (374/113).

Multiple peptides are attributable to alkylation as well as the ASP ion (*m/z* 1315.73), which was observed when MGMT is incubated with unmodified ODN as well as SS and DS SRO1. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z* 1315.73). **B**. Methylated MGMT-ASP (GNPVPILIPMe-CHR, 136-147; *m/z* 1329.74). **C**. MBP-TP13 (SYEEELAKDPR, 332-342; *m/z* 1336.54). **D**. Putative carboxamidoethylated MGMT-ASP (GNPVPILIPcarboxamidoethyl-CHR, 136-147; *m/z* 1386.77). Peak intensities are shown in arbitrary units on the y-axis and intensity scale is the same for the three panels (I, II and III).

Table 5.3: R_2 for MBP-MGMT reacted with SRO1 and LH1a-c.

ODN	ODN size	R_2	
		SS	DS
G	13-mer	0	0
SRO1	13-mer	0	$0.62\pm0.13^{***\#\#\#}$
G	23-mer	0	0
LH1a	23-mer	0	$1.03+0.03^{***\#\#\#}$
LH1b	23-mer	$0.77\pm0.003^{**}$	$7.2\pm0.34^{***\#\#\#}$
LH1c	23-mer	0.66 ± 0.16	$17.1\pm2.9^{***\#\#\#}$

Values are mean±SD of 3 experiments. $^{**}P<0.01$; $^{***}P<0.001$ (significantly different from control G). $^{\#\#\#}P<0.001$ (significantly different from corresponding SS ODN).

5.3.5 MS analysis of tryptic digests of MBP-MGMT following reaction with 23 mer O^6-CEG containing ODNs

Four 23 mer O^6-CEG containing ODNs were examined (RA1& LH1a-c).

5.3.5.1 MALDI-TOF MS analysis of carboxyethyl group transfer from SS & DS RA1 ODN

Figure 5.10 (I and II) showed that carboxyethyl transfer to MGMT was observed following MGMT reaction with 23 mer O^6-CEG (RA1), either SS or DS, since carboxyethylated MGMT-ASP was detected in tryptic fragments of MGMT (*m/z* 1387.75 $[M+H]^+$). MGMT incubation with the RA1 (SS or DS) induced a shift in the molecular weight of MGMT-ASP=72.02 Da which corresponds to the transfer of a carboxyethyl group. The unmodified MGMT-ASP was present in the tryptic digest of MGMT incubated with DS G as well as SS and DS RA1 (Figure 5.10 III).

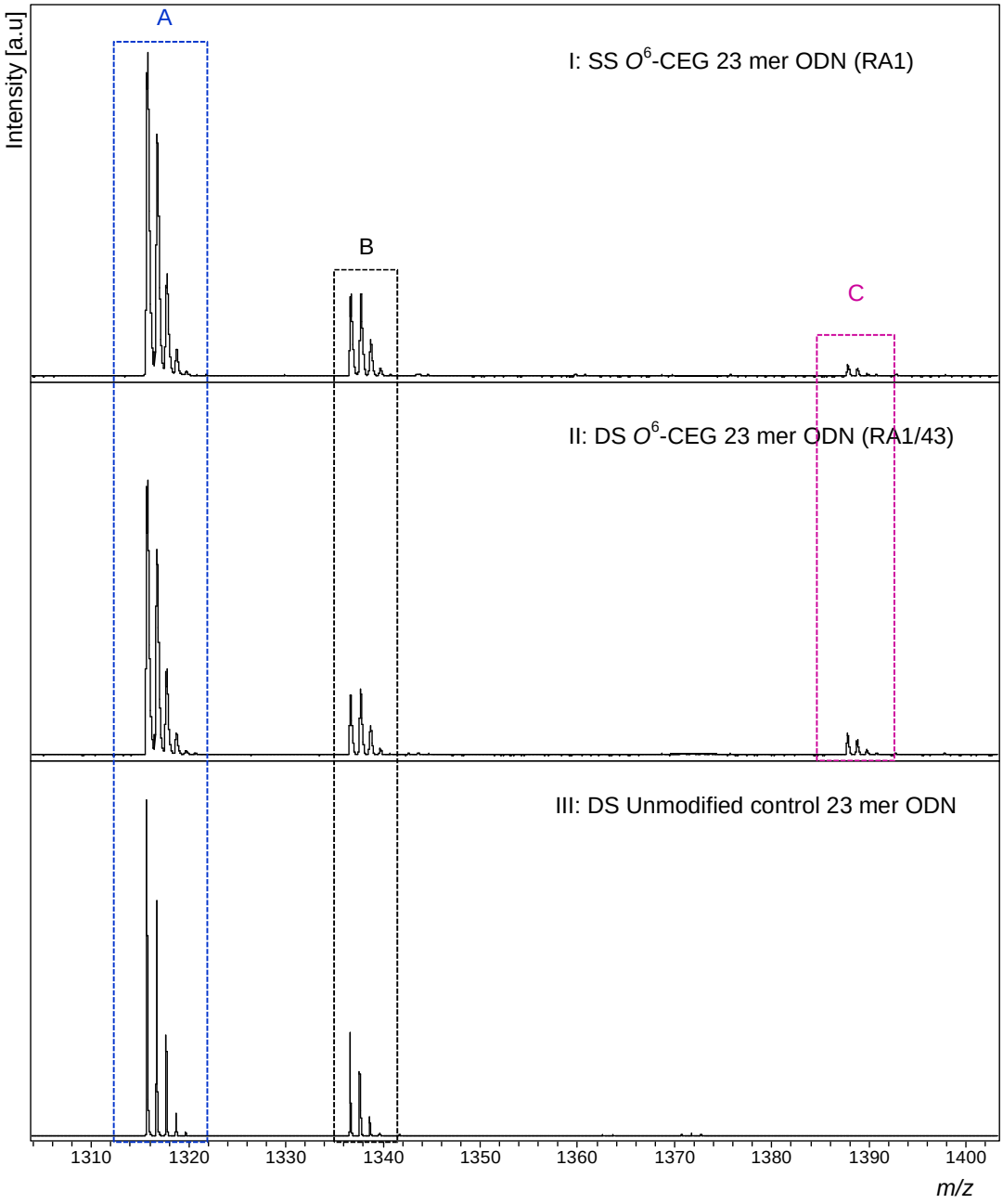


Figure 5.10: Mass Spectra of tryptic digests of MBP-MGMT reacted with ODNs of the sequence GAA CTX CAG CTC CGT GCT GGC CC where X is (I and II) O^6-CEG or (III) control G (16/43) ODN.

I & II. SS and DS RA1, respectively. Multiple peptides were attributable to alkylation as well as the ASP ion (*m/z* 1315.73), which is observed when MGMT was incubated with unmodified ODN as well as SS and DS RA1. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z* 1315.73). **B**. MBP-TP13 (SYEEELAKDPR, 332-342; *m/z* 1336.54). **C**. Carboxyethylated MGMT-ASP (GNPVPILIPcarboxyethyl-CHR, 136-147; *m/z* 1387.75). Peak intensities are shown in arbitrary units on the y-axis and intensity scale is the same for the three panels (I, II and III).

5.3.5.2 MALDI-TOF MS analysis of tryptic peptides of MBP-MGMT incubated with SS LH1a and DS LH1a/43 ODNs

Figure 5.11 (I and II) showed that *S*-methylcysteine modified-ASP was detected in tryptic peptides of MBP-MGMT protein that reacted with either SS or DS LH1a. However, transfer of carboxyethyl group to MGMT-ASP was not demonstrated following incubation of MBP-MGMT with either SS or DS LH1a. In addition, similar to SRO1 ODN, putative *S*-carboxamidoethylcysteine modified-ASP (*m/z* 1386.77 $[M+H]^+$) was detected following reaction with DS LH1a ODN. Therefore, R_2 was calculated for LH1a ODN to compare the relative abundance of the two detected alkylated MGMT-ASPs showing that the PA of *S*-methylcysteine modified MGMT-ASP was approximately the same as that of putative *S*-carboxamidoethylcysteine MGMT modified ASP indicating R_2 to be 1.03±0.03 (mean±SD; n=3), after MGMT reaction with DS LH1a. This result was statistically significant with respect to either control G or corresponding SS ODNs ($P<0.001$) since R_2 was zero for both groups (Table 5.3). Furthermore, the unmodified, but not alkylated, MGMT-ASP (*m/z* 1315.7 $[M+H]^+$) was detected after incubation with control G ODNs (Figure 5.11 III).

5.3.5.3 MS analysis of tryptic digests of MBP-MGMT following reaction with SS LH1b and DS LH1b/43 re-purified ODNs

The presence of *S*-methylcysteine as well as putative *S*-carboxamidoethylcysteine modifications to MGMT following incubation with SS and DS LH1a was unexpected, and to investigate further, LH1a ODN was subjected to a second purification step to investigate if the presence of these ions was due to contamination (Figure 2.2). Purified LH1a was assigned lab number LH1b. Following this purification and subsequent incubation with MBP-MGMT, both ions (*S*-methylcysteine and putative *S*-carboxamidoethylcysteine modified MGMT-ASPs) were still detected after reaction with either SS or DS LH1b (Figure 5.12 I and II). Quantitative analysis showed that, R_2 was 0 after reaction with control G that significantly increased to 0.77±0.003 (mean±SD; n=3) ($P<0.01$; relative to control) and 7.2±0.34 ($P<0.001$, relative to both control & SS LH1b) following reaction with SS and DS LH1b, respectively.

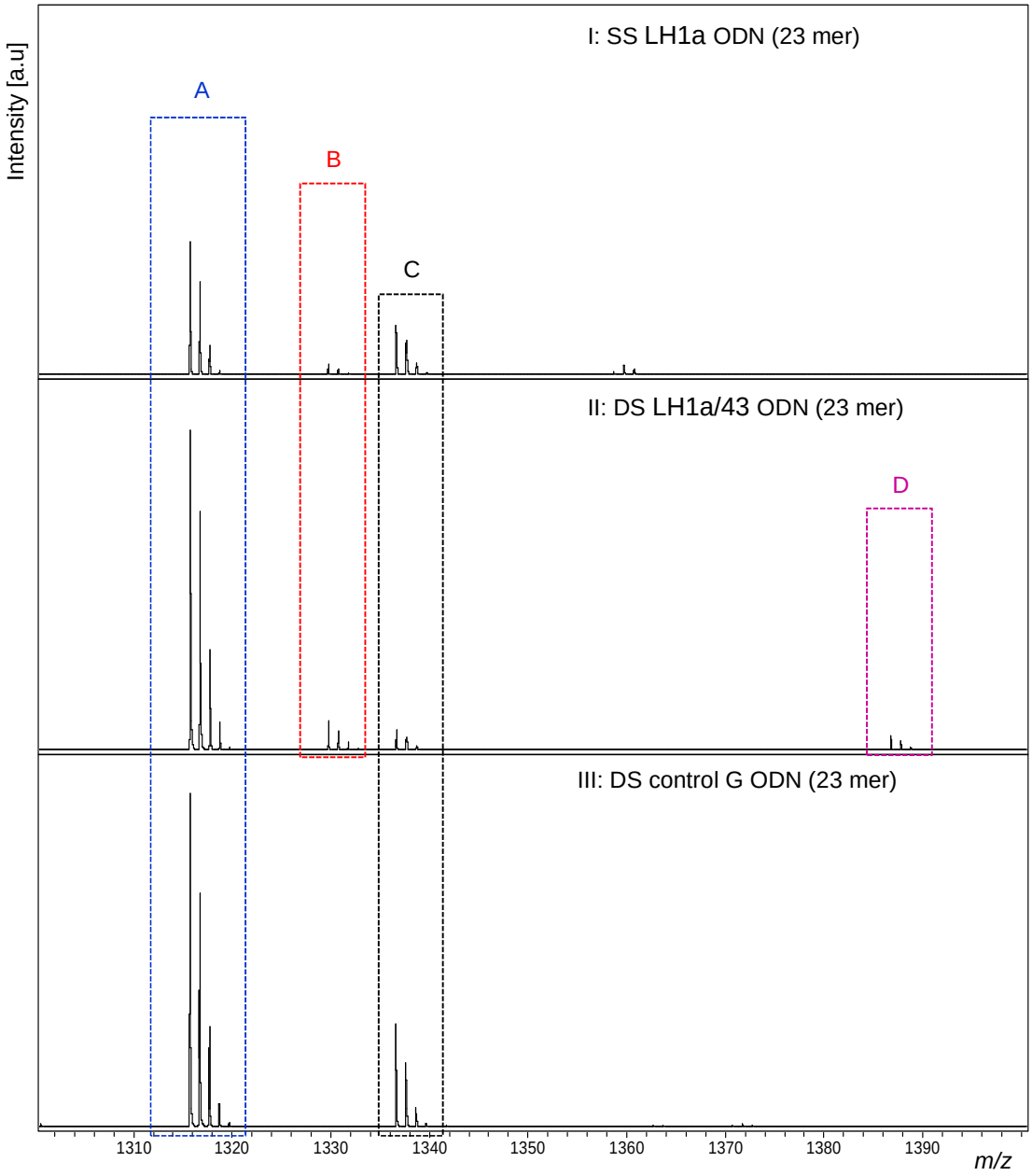


Figure 5.11: Mass Spectra of MBP-MGMT following incubation with (I) SS LH1a, (II) DS LH1a/43 and (III) DS unmodified G ODN (16/43).

Multiple peptides were attributable to alkylation as well as the ASP ion (*m/z* 1315.73), observed when MGMT was incubated with unmodified ODN as well as SS and DS LH1a ODN. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z* 1315.73). **B**. Methylated MGMT-ASP (GNPVPILIPMe-CHR, 136-147; *m/z* 1329.74). **C**. MBP-TP13 (SYEEELAKDPR, 332-342; *m/z* 1336.54). **D**. Putative carboxamidoethylated MGMT-ASP (GNPVPILIPcarboxamidoethyl-CHR, 136-147; *m/z* 1386.77). Peak intensities are shown in arbitrary units on the y-axis and intensity scale is the same for the three panels (I, II and III).

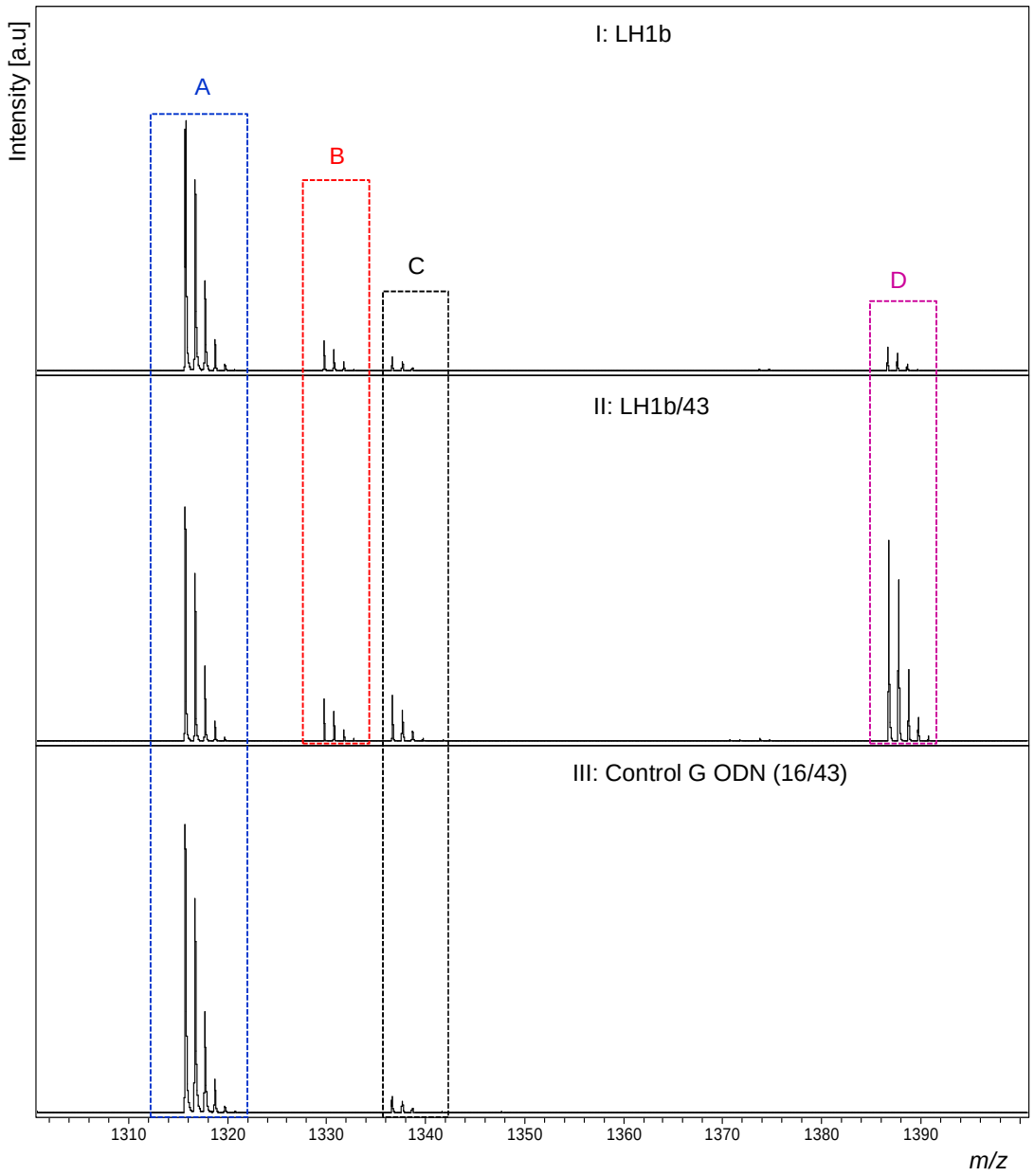


Figure 5.12: Mass spectra of MBP-MGMT reacted with (I) LH1b, (II) LH1b/43 and (III) control 16/43 ODN focusing on locations and masses of modified MGMT active site peptides.

Multiple peptides were attributable to alkylation as well as the ASP ion (*m/z* 1315.73), observed when MGMT was incubated with unmodified ODN as well as SS and DS LH1b ODN. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z* 1315.73). **B**. Methylated MGMT-ASP (GNPVPILIPMe-CHR, 136-147; *m/z* 1329.74). **C**. MBP-TP13 (SYEEELAKDPR, 332-342; *m/z* 1336.54). **D**. Putative carboxamidoethylated MGMT-ASP (GNPVPILIPcarboxamidoethyl-CHR, 136-147; *m/z* 1386.77). Peak intensities are shown in arbitrary units on the y-axis and intensity scale is the same for the three panels (I, II and III).

This result indicates that PA of detected putative carboxamidoethylated MGMT-ASP was seven times that of methylated MGMT-ASP, following MBP-MGMT reaction with DS LH1b/43. From this data, it is clear that the second purification of LH1a did not eliminate *S*-methylcysteine modified MGMT-ASP. Further, the second step of purification of LH1a facilitated the detection of *S*-carboxamidoethylcysteine modified MGMT-ASP and enabled it for the first time after incubation with SS LH1b (Table 5.3).

5.3.5.4 MS analysis of tryptic digests of MBP-MGMT following reaction with SS LH1c and DS LH1c/43 ODNs

A second fraction of 23 mer ODN (LH1a) was subjected to further purification trying to eliminate or at least minimise the unexpected alkyl modified ASPs. Purified LH1a ODNs were assigned the lab code LH1c. Figure 5.13 I and II demonstrated that both *S*-methylcysteine and putative *S*-carboxamidoethyl modified ASPs were detected in tryptic digests of MBP-MGMT incubated with SS (LH1c) or DS (LH1c/43) ODNs.

In addition**,** as with the second purification, final purification of 23 mer LH1a ODNs yielded LH1c ODN still allows the detection of either *S*-methylcysteine or *S*-carboxamidoethylcysteine modified MGMT. Moreover, Figure 5.13 III showed that the unmodified MGMT-ASP was detected after incubating MGMT with unmodified control ODNs, and did not show any traces for alkyl modified MGMT-ASPs.

Additional analysis (R_2) displayed that final purification reduced the intensity of methylated MGMT-ASP PA to seventeen times less than that of putative carboxamidoethylated MGMT-ASP (R_2=17.1$\pm$2.9 (mean$\pm$SD, n=3); P<0.001 with respect to both control G), when MBP-MGMT was incubated with purified LH1c/43. Moreover, R_2 for LH1c/43 was statistically significant with respect to that of SS LH1c that had R_2 of 0.66$\pm$0.16. However, recorded R_2 for LH1c was statistically non-significant with respect to control G ODN (P>0.05); as shown in Table 5.3.

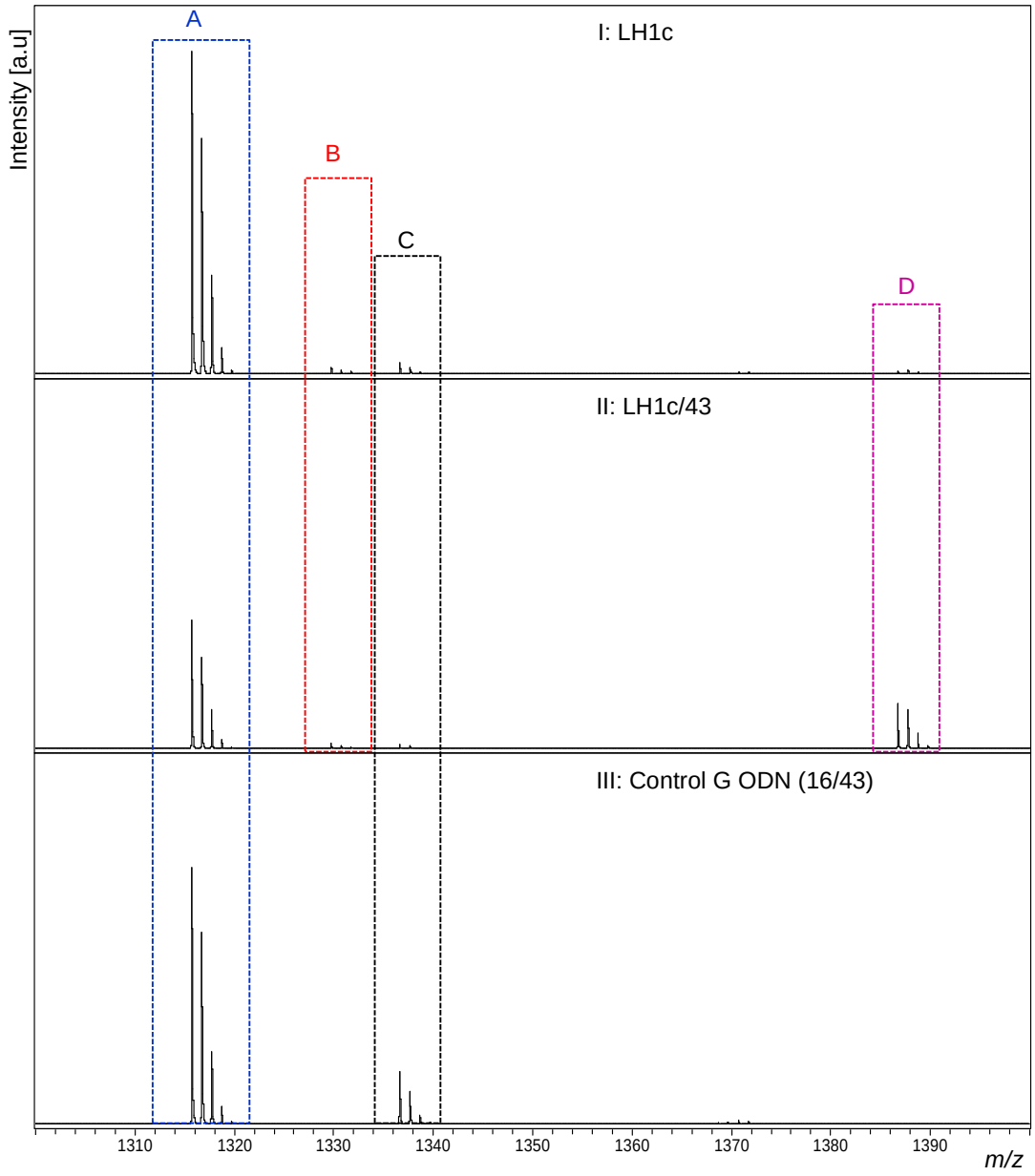


Figure 5.13: MALDI-TOF spectra of MBP-MGMT reacted with (I) SS LH1c, (II) DS LH1c/43 and (III) control G ODN (16/43) focusing on locations of modified MGMT-ASPs.

Multiple peptides were attributable to alkylation as well as the ASP ion (*m/z* 1315.73), observed when MGMT was incubated with unmodified ODN as well as SS and DS LH1c ODN. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z* 1315.73). **B**. Methylated MGMT-ASP (GNPVPILIPMe-CHR, 136-147; *m/z* 1329.74). **C**. MBP-TP13 (SYEEELAKDPR, 332-342; *m/z* 1336.54). **D**. Putative carboxamidoethylated MGMT-ASP (GNPVPILIPcarboxamidoethyl-CHR, 136-147; *m/z* 1386.77). Peak intensities are shown in arbitrary units on the y-axis and intensity scale is the same for the three panels (I, II and III).

5.3.6 MS analysis of chemically modified synthetic MGMT-ASP

In order to help to confirm the identity of transferred alkyl groups to MGMT-ASPs following incubation with various preparations of O^6-CMG and O^6-CEG ODNs, synthetic MGMT-ASP was incubated with iodoacetic acid (carboxymethylating agent), iodoacetamide (carboxamidomethylating agent), 2-bromopropionic acid (carboxyethylating agent) and 2-bromopropionamide (carboxamidoethylating agent). Then chemically modified synthetic MGMT-ASPs were analysed by MS (Section 2.2.17.1).

Figure 5.14 (I, V) shows that iodoacetamide reaction with MGMT caused a 57.03 Da mass shift with the alkylated peptide observed at *m/z* 1372.75 $[M+H]^+$. This result supports the assignment of the alkyl group at the O^6 position of modified G base in this 44 ODN (Figure 5.8) as carboxamidomethyl rather than the carboxymethyl group. Whilst Figure 5.14 (II, VI) shows that iodoacetic acid generated a carboxymethylated ion at *m/z* 1373.73 $[M+H]^+$ (58.01 Da mass shift), consistent with the O^6-CMG ODNs results (DS DW1 and OW18; Figure 5.6 and 5.7).

In addition, the other amide (2-bromopropionamide) caused a mass shift of 71.04 Da *cf.* carboxamidoethylated MGMT-ASP detected at *m/z* 1386.76 $[M+H]^+$; as shown in Figure 5.14 (III, VII). While, 2-Bromopropionic acid incubation with MGMT-ASP generated an ion at *m/z* 1387.74 $[M+H]^+$ (72.02 Da mass shift from the ASP) assigned as carboxyethylation Figure 5.14 (IV, VIII), which is consistent with the O^6-CEG ODN results (RA1; Figure 5.10).

This highlights that both LH1 and SRO1 ODNs do not contain carboxyethyl modification at the O^6 position of the modified G base but that instead they possessed a carboxamidoethyl modification as evidenced by the recorded mass shift (71.04 Da) in MGMT-ASP following LH1 and SRO1 ODNs incubation with MGMT (Figure 5.9, 5.11-13).

In addition, MS analysis of mixture of tryptic peptides of MBP-MGMT that was incubated with 44 and DW1 ODNs, separately, demonstrated the presence of the two peaks at *m/z* 1372.7 and 1373.7 $[M+H]^+$ (Figure 5.15 I-III). The same result was reported for LH1 and RA1; as shown in Figure 5.15 IV-VI.

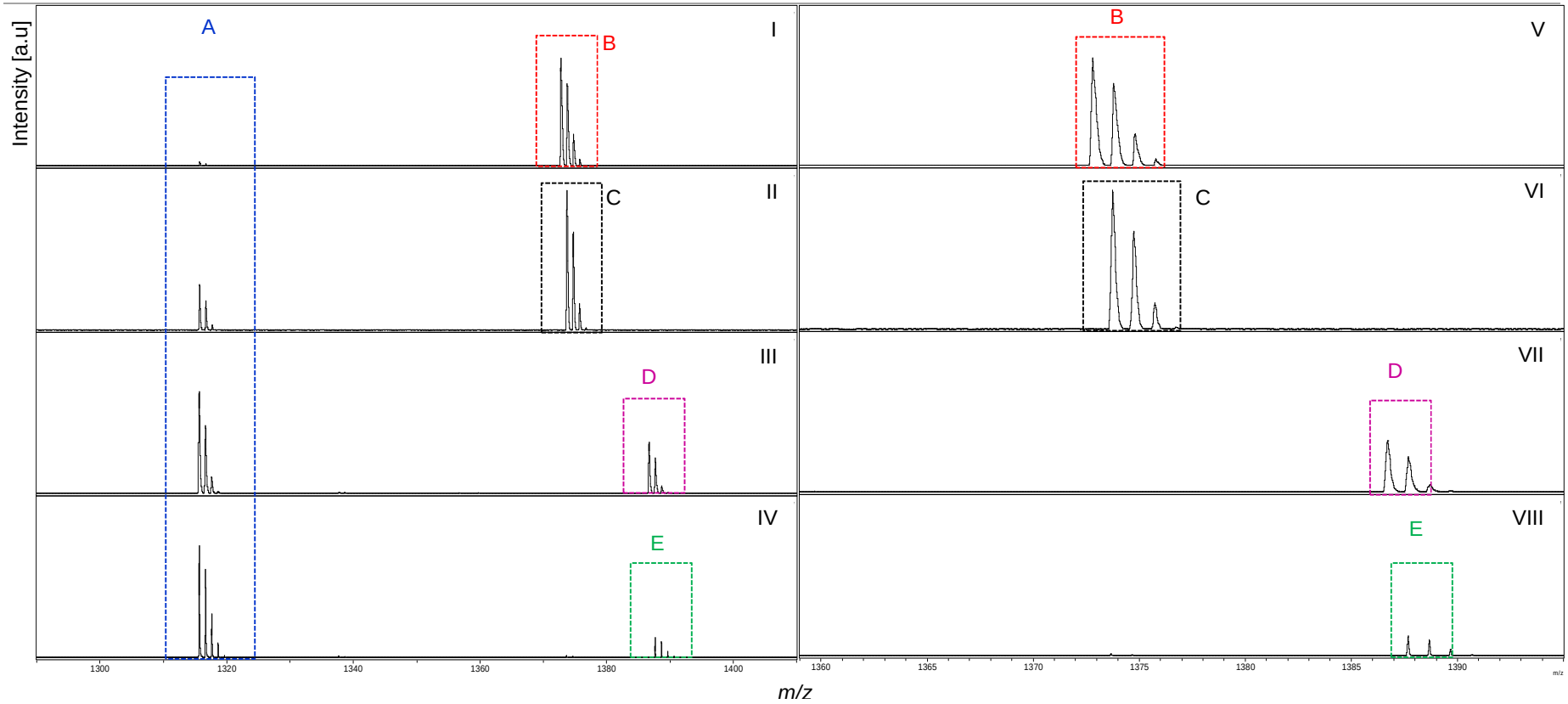


Figure 5.14: Mass spectra of chemically modified synthetic MGMT-ASP.

Synthetic MGMT-ASP (1 µM) was chemically modified *via* incubation with 50 mM: (I, V) Iodoacetamide, (II, VI) Iodoacetic acid, (III, VII) 2-Bromopropionamide or (IV, VIII) 2- Bromopropionic acid. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147) *m/z* 1315.72. **B**. Carboxamidomethylated MGMT-ASP (GNPVPILIPcarboxamidomethyl-CHR, 136-147) *m/z* 1372.75. **C**. Carboxymethylated MGMT-ASP (GNPVPILIPCM-CHR, 136-147) *m/z* 1373.73. **D**. Carboxamidoethylated MGMT-ASP (GNPVPILIPcarboxamidoethyl-CHR, 136-147) *m/z* 1386.76. **E**. Carboxyethylated MGMT-ASP (GNPVPILIPCE-CHR, 136-147) *m/z* 1387.74. Peak intensities are shown in arbitrary units on the y-axis and intensity scale is the same for panels (I, II, III and IV). Panels V-VIII are magnified inserts with the same intensity scale.

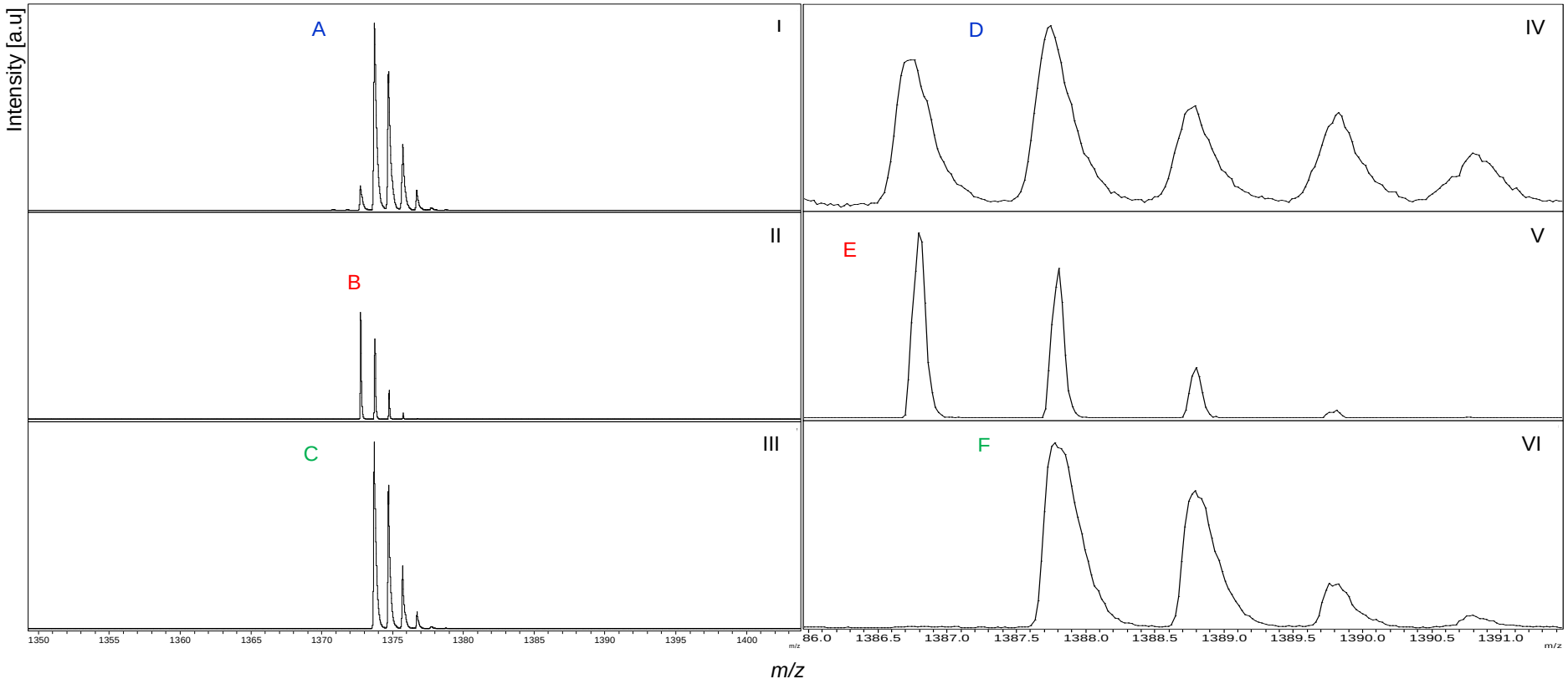


Figure 5.15: Mass spectra of tryptic digests of MBP-MGMT incubated with either *O*^{6}-CMG or *O*^{6}-CEG containing ODNs.

(I) Mixture of tryptic digests of MBP-MGMT incubated with 44 and DW1 (*O*^{6}-CMG) ODNs, (II, III) Tryptic digests of MBP-MGMT incubated with either 44 or DW1, respectively. (IV) Mixture of tryptic digests of MBP-MGMT incubated with RA1 and LH1 (*O*^{6}-CEG) ODNs, (V, VI) MBP-MGMT incubated with LH1c or RA1 ODNs, respectively. **A**. Two peaks at *m/z* 1372.75 &1373.73 **B**. Carboxamidomethylated MGMT-ASP (GNPVPILIPcarboxamidomethyl-CHR, 136-147) *m/z* 1372.75. **C**. Carboxymethylated MGMT-ASP (GNPVPILIPCM-CHR, 136-147) *m/z* 1373.73. **D**. Two peaks at *m/z* 1386.76 &1387.74 **E**. Carboxamidoethylated MGMT-ASP (GNPVPILIPcarboxamidoethyl-CHR, 136-147) *m/z* 1386.76. **F**. Carboxyethylated MGMT-ASP (GNPVPILIPCE-CHR, 136-147) *m/z* 1387.74. Peak intensities are shown in arbitrary units on the y-axis and intensity scale is the same for panels (I, II, III) and (IV, V, VI).

5.3.7 Quantitative comparison of MGMT repair of SS *vs.* DS O^6-alkylG containing ODNs

The ratio (R_o) of the PA of unmodified MGMT-ASP was compared to the PA of MGMT-TP 7 (EWLLAHEGHR) to monitor disappearance of unmodified MGMT-ASP following incubation with alkylated ODNs; as explained in section 2.2.17.3. For LH1 preparations, quantitative analyses were conducted for LH1c merely, because it was the purest available preparation of LH1 containing the least amount of O^6-MeG as evidenced by previous quantitative R_2 results.

Regarding 13 mer ODNs, following incubation of MGMT with the control G 13 mer ODN, R_o was found to be 0.58±0.03 (mean±SD, n=3), whereas R_o significantly decreased to 0.20±0.01 (P<0.001), 0.10±0.01 (P<0.01), 0.26±0.01 (P<0.001) and 0.12±0.03 (P<0.001) following MGMT reaction with 303B, 303B/113, OW18 and OW18/113, respectively, compared to unmodified protein (Table 5.4). Moreover, there was a significant difference between R_o for SS *vs.* DS O^6-MeG (303B) and O^6-CMG (OW18) ODNs (P<0.01 and P<0.05, respectively). Whilst both SS and DS SRO1 reaction with MGMT did not cause a statistically significant change (P>0.05) in R_o pointing out a non-significant change in the abundance of unmodified MGMT-ASP. Recorded R_o values were 0.59±0.02 and 0.48±0.05 for SS and DS SRO1, respectively (Table 5.5).

Quantitative analysis for 23 mer ODNs revealed significant decrease in R_o from 0.63±0.04 for G 23 mer ODN (mean±SD, n=3) to 0.08±0.01, 0.02±0.002, 0.28±0.003, 0.03±0.01, 0.22±0.001, 0.16±0.001, 0.33±0.01 and 0.26±0.01 for 15, 15/43, DW1/C, 44, RA1, RA1/C, LH1c and LH1c/43, respectively. Recorded P values were (P<0.001) for all ODNs except DW1/C (P<0.05) and SS & DS LH1c (P<0.01); with respect to control. However, MGMT incubation with DW1 didn't cause a significant change in R_o (0.54±0.001; P>0.05). R_o results of all DS 23 mer O^6-alkylG ODNs were statistically significantly different from that of the respective SS ODN. P values were P<0.01, P<0.05, P<0.001, P<0.05 for SS *vs.* DS 15, DW1, RA1 and LH1c, respectively (Table 5.4 & 5.5).

Table 5.4: Summary for R_o and R_1 for MGMT reacted with O^6-alkylG ODNs.

Nucleo-base	ODN size (Lab no.)	SS		DS		DS/SS	
		R_o	R_1	R_o	R_1	R_o	R_1
G	13-mer (374)	-	-	0.58±0.03	-	-	-
	23-mer (16)	-	-	0.63±0.04	-	-	-
O^6-MeG	13-mer (303B)	$0.20\pm0.01^{***}$	$0.58\pm0.01^{***}$	$0.10\pm0.01^{**\#\#}$	$0.71\pm0.03^{***\#}$	0.53±0.05	1.2±0.06
	23-mer (15)	$0.08\pm0.01^{***}$	$0.38\pm0.01^{***}$	$0.02\pm0.002^{***\#\#}$	$0.62\pm0.01^{***\#\#\#}$	0.26±0.02	1.6±0.01
O^6-CMG	13-mer (OW18)	$0.26\pm0.01^{***@@@}$	$0.28\pm0.01^{*@@@}$	$0.12\pm0.03^{***\#}$	$0.63\pm0.001^{*\#@}$	0.46±0.10	2.3±0.10
	23-mer (DW1)	$0.54\pm0.001^{@@@}$	-	$0.28\pm0.003^{*\#@@@}$	$0.44\pm0.10^{*\#@@@}$	0.52±0.01	-
O^6-CEG	23-mer (RA1)	$0.22\pm0.001^{***@@@}$	$0.006\pm0.0004^{***@@@}$	$0.16\pm0.001^{***\#\#\#@@@}$	$0.01\pm0.001^{***\#\#\#@@@}$	0.75±0.04	1.7±0.2

R_o: Unmodified MGMT-ASP PA/MGMT-TP 7 PA; R_1: Alkylated ASP PA/MGMT-TP 7 PA. Values are mean±SD of 3 experiments. *P<0.05; **P<0.01; ***P<0.001 (significantly different from control G; one way ANOVA). $^{\#}P$<0.05; $^{\#\#}P$<0.01; $^{\#\#\#}P$<0.001 (significantly different from corresponding SS O^6-alkylG ODN; one way ANOVA). $^{@}P$<0.05; $^{@@@}P$<0.001 (significantly different from corresponding O^6-MeG ODN; non-paired t test).

Table 5.5: Summary for R_o and R_1 for MGMT reacted with SR01, LH1c or 44 ODNs.

ODN no.	ODN size	SS		DS		DS/SS	
		R_o	R_1	R_o	R_1	R_o	R_1
SRO1	13-mer	0.59±0.02	-	0.48±0.05	0.06±0.01$^{*\#}$	0.81±0.1	-
LH1c	23-mer	0.33±0.01**	0.003±0.001	0.26±0.01$^{**\#}$	0.09±0.01$^{**\#\#}$	0.80±0.03	30±10
44	23-mer	-	-	0.03±0.01***	0.63±0.02***	-	-

R_o: Unmodified MGMT-ASP PA/MGMT-TP 7 PA; R_1: Alkylated ASP PA/MGMT-TP 7 PA. Values are mean±SD of 3 experiments. *P<0.05; **P<0.01; ***P<0.001 (significantly different from control G; one way ANOVA). $^{\#}P$<0.05; $^{\#\#}P$<0.01; $^{\#\#\#}P$<0.001 (significantly different from corresponding SS O^6-alkylG ODN; one way ANOVA).

Additional quantitative analyses were conducted to assess the relative repair of O^6-alkylG in SS *vs.* DS ODNs by MGMT by comparing the repair ratio (R_1) of PA of alkylated MGMT-ASP to that of PA of MGMT-TP 7 among all groups (Section 2.2.17.3). Concerning 13 mer ODNs, there were clear differences in R_1 between the differing ODNs being 0 after reaction with control G ODNs, and which increased significantly to 0.58 $\pm$0.01 (mean$\pm$SD, n=3) (P<0.001), 0.71$\pm$0.03 (P<0.001), 0.28$\pm$0.01 (P<0.05), 0.63$\pm$0.001 (P<0.05), 0.06$\pm$0.01 (P<0.05) after incubation of MGMT protein with 303B, 303B/113, OW18, OW18/113 and SRO1/113, respectively. Since carboxamidoethyl modified MGMT-ASP was not detected in tryptic fragments of proteins following incubation with SRO1, therefore, R_1 was zero. R_1s of all DS 13 mer O^6-alkylG ODNs were statistically significantly different from that of the respective SS ODN (P<0.05) (Table 5.4 & 5.5).

Furthermore, there were clear differences in R_1 among 23 mer O^6-alkylG ODNs. R_1 was zero using the control G ODNs, but increased significantly to 0.38$\pm$0.01 (mean$\pm$SD, n=3) (P<0.001) and 0.62$\pm$0.01 (P<0.001), 0.44$\pm$0.10 (P<0.05), 0.63$\pm$0.02 (P<0.001), 0.006$\pm$0.0004 (P<0.001), 0.01$\pm$0.001 (P<0.001) and 0.09$\pm$0.01 (P<0.01) after incubation of MGMT protein with 15, 15/43, DW1/C, 44, RA1, RA1/43 and LH1c/43, respectively, compared to control. However, MGMT reaction with SS LH1c induced a non-significant increase in R_1 (0.003$\pm$0.001; P>0.05), with respect to control. In addition, R_1 was zero for DW1 since no carboxymethylated peptides were detected. Furthermore, R_1 results from incubation with the DS O^6-alkylG ODN were statistically significantly different from those found with the respective SS ODN (P<0.001 for SS *vs.* DS 15 and RA1; P<0.05 for SS *vs.* DS DW1; P<0.01 for SS *vs.* DS LH1c) (Table 5.4 & 5.5).

Further statistical analysis comparing 13 mer O^6-CMG R_o and R_1 (SS & DS) to the corresponding O^6-MeG results displayed significant differences between O^6-MeG and O^6-CMG results (P<0.001 for SS R_o and R_1, P<0.05 for DS R_1). Conversely, there was no statistical difference in R_o between DS O^6-MeG and O^6-CMG (P>0.05). Consistently, statistical analysis for 23 mer ODNs, namely, O^6-MeG, O^6-CMG (DW1) and O^6-CEG (RA1), revealed that R_o and R_1 of the new examined adducts O^6-CMG and O^6-CEG (SS & DS) were statistically significantly different from those of the corresponding O^6-MeG (P<0.001). Both 13 mer and 23 mer ODN results indicated that MGMT efficiently repaired O^6-MeG over corresponding O^6-CMG and O^6-CEG.

5.4 Discussion

The work described above demonstrates that O^6-MeG, O^6-CMG as well as O^6-CEG adducts are repaired by human MGMT in a reaction that results in the transfer of the alkyl groups (methyl, carboxymethyl or carboxyethyl), respectively, from the O^6 position of the modified guanine base to MGMT. The peptide that is the 'proxy' readout of this reaction contains the active cystinyl residue (Cys-145). It has been shown that the newly developed MALDI-TOF based approach was able to differentiate between two preparations of 23 mer ODNs, namely, 44 and DW1, that both were supposed to contain O^6-CMG adduct. Since, the transfer of carboxymethyl group was observed following reaction with DS DW1 but not 44 ODN. As for O^6-CEG containing ODNs, among all assayed 13 mer and 23 mer preparations, only the RA1 of 23 mer O^6-CEG ODN demonstrated the transfer of carboxyethyl group to MGMT-ASP following their reaction.

This study is the first report of the repair of O^6-CEG adduct by MGMT displaying the transfer of carboxyethyl group from a modified guanine base to MGMT. Data shown could be used as evidence for the classification of this O^6-CEG adduct as a MGMT-repairable one. This raises the question about the potential pathways of the formation of this novel adduct and whether this adduct is present *in vivo*.

Following incubation of MGMT with O^6-alkylG containing ODNs, MS revealed ions that corresponded to the both MBP and MGMT tryptic peptide as well as MGMT-ASP accompanied by mass increases consistent with the transfer of methyl (14.01 Da), carboxymethyl (58.01 Da) or carboxyethyl (72.02 Da) to MGMT-ASP. This leads to the conclusion that alkyl group transfer is the primary product as a consequence of O^6-alkylG adducts repair by MGMT. This observation is consistent with a previous report demonstrating alkyl group transfer to MGMT using alkylated DS 13 mer ODNs containing both O^6-MeG and O^6-CMG adducts (296). However, the previous study reported a 57 Da mass shift following reaction with DS O^6-CMG 13 mer ODN. Whilst, the current work finds a mass shift of 58 Da consistent with transfer of carboxymethyl group to MGMT-ASP following incubation with both 13 mer (SS & DS OW18) and 23 mer (DS DW1) O^6-CMG ODNs. According to the Unimod database for protein modifications, the expected mass change due to transfer of carboxymethyl group is 58.005 Da (366), whilst 57 mass shift of a cysteine residue corresponds to the transfer

of a carboxamidomethyl group (367). The discrepancy in results between previous work and the current study may be due to the difference in the identity of the alkyl adduct or in MS conditions or even both.

A major aim of this work was to compare qualitatively the transfer of alkyl groups from the two examined O^6-CMG 23 mer ODN preparations to MGMT; namely DW1 and 44. Results indicate that MGMT incubated with DS (but not SS) DW1 ODN allows the transfer of a carboxymethyl group to MGMT-ASP as evidenced by the detection of a carboxymethylated MGMT-ASP ion. These findings confirm that DW1 ODN was O^6-CMG containing ODN and provided further assurance about the identity of the alkyl group at the O^6 position of the modified G base as a carboxymethyl group. Conversely, qualitative MS results showed that MGMT incubation with DS 44 ODN resulted in generation of an alkylated MGMT-ASP ion peak at 1372.7 m/z $[M+H]^+$ with a recorded mass shift of 57 Da which was not detected in tryptic peptides of the control group. This putative alkylated MGMT-ASP was suggested to be due to the transfer of carboxamidomethyl group instead of a carboxymethyl group to MGMT. These findings were verified by MS analysis of chemically modified (carboxymethylated and carboxamidomethylated) synthetic MGMT-ASP showing that carboxymethylation (CH_2COOH) and carboxamidomethylation (CH_2CONH_2) of synthetic MGMT-ASP generated $[M+H]^+$ ions at *m/z* 1373.7 (58 Da mass shift) and 1372.7 (57 Da mass shift), respectively.

Furthermore, this MS approach was able to distinguish between the three assayed O^6-CEG ODN preparations; namely, 13 mer O^6-CEG ODN (SRO1) as well as 23 mer O^6-CEG ODNs RA1 and LH1 preparations as it revealed major differences in terms of the *m/z* ratio of the detected alkylated ASPs and subsequently the induced change of MGMT-ASP molecular weight. A single preparation of 23 mer O^6-CEG, which is RA1, enabled carboxyethyl group transfer to MGMT confirming the identity of the alkyl modification at the O^6 position of the modified G base as carboxyethyl group (368). However, MS analysis of SRO1 as well as LH1 a, b and c showed the transfer of an alkyl group causing mass shift of 71.04 Da to the active site peptide, and by analogy to carboxamidomethyl, it was tentatively assigned to carboxamidoethylated MGMT-ASP (also known as propionamide) (369,370) which was then supported by control experiments.

MS analysis of MGMT reacted with 44, SRO1 and LH1 ODNs coupled with the synthetic ASP results highlighted that previously mentioned ODNs did not contain the expected alkyl adducts at the O^6 position of the modified G base. However, 44 ODN contained carboxamidomethyl alkyl group instead of carboxymethyl and both SRO1 and LH1 contained carboxamidoethyl instead of carboxyethyl. Mass spectroscopic identification of amidated peptides is relatively straight forward as amidation decreases the mass of the intact ion -0.98 Da which is due to replacement of OH group with NH_2 (371). Possibly that was a result of the use of ammonia for DNA bases de-protection wherein ammonia reacted with the COOH groups forming the $CONH_2$ chemical entities. However, further tandem MS analysis is required to confirm these results.

In addition, further purifications of the initial preparation LH1a minimized the intensity of the methylated ASP, and increased the relative intensity of the carboxamidoethylated species. From this it could be concluded that LH1 ODN contained some containment O^6-MeG and that the repeated cycles of purifications minimised it.

Our qualitative MALDI-TOF MS based approach investigating MGMT repair of the O^6-MeG lesions displayed the same repair preference for both DS and SS ODNs when either presented as 13 or 23 mer contexts since methylated ASP ions was detected in tryptic digests of MGMT incubated with either DS or SS methylated ODNs (13 and 23 mer). Consequently, similar results were observed for O^6-CMG containing 13 mer ODNs (OW18) as well as O^6-CEG containing 23 mer ODNs (RA1) as alkylated ASP ions were detected after reaction with either SS or DS ODNs. In contrast, our qualitative assay of MGMT repair of O^6-CMG lesion (in 23 mer ODN; DW1) highlighted that MGMT preferentially repaired specified adducts when presented in DS ODNs over SS ones. This observation was supported by inability of our MALDI-TOF MS assay to detect the alkylated MGMT-ASP ions (*S*-carboxymethylcysteine) in tryptic peptides of MGMT reacted with SS DW1. However, *S*-carboxymethylcysteine modified MGMT-ASP was successfully detected following incubations with the corresponding DS DW1.

This difference in qualitative results trying to examine MGMT repair preferences urged further quantitative analyses to compare MGMT repair preference for each SS and DS ODN *via* calculating R_o and R_1. Results of quantitative analyses conducted showed that for both O^6-MeG (303B, 15) and O^6-CMG (OW18, DW1) containing ODNs presented as 13 or 23 mer contexts as well as 23 mer O^6-CEG ODN (RA1), MGMT

exhibited significantly higher repair preference for DS ODNs compared to the corresponding SS ones as indicated by lower R_os for DS ODNs compared to SS ones. The significant recorded decline in the R_o indicated a decrease in unmodified-ASP due to the alkylation process. This finding was confirmed by calculating R_1 ratios for SS and DS alkylated ODNs which indicated that R_1 was significantly higher with DS alkylated ODNs compared to the corresponding SS ones. Since the increase in R_1 ratio indicates a higher abundance of alkylated MGMT-ASP, therefore, the observed changes in R_1 confirmed that DS O^6-MeG (303B, 15) and O^6-CMG (OW18, DW1) and 23 mer O^6-CEG (RA1) containing ODNs were more susceptible to MGMT repair than the corresponding SS ODNs.

Our results were consistent with a previous research reported that MGMT reacts more quickly with ODNs that are DS *vs*. SS (360), as MGMT appears to bind DS DNA with slightly higher affinity than SS DNA. Another possible explanation could be due to the stability of MGMT-DNA complex formed that might be more stable with DS DNA rather than SS substrates. However, another study stated insignificant difference in MGMT reaction rates between SS and DS O^6-MeG and O^6-pobG containing ODNs (341). Mechanism of MGMT repair of O^6-alkylG adducts in DS DNA has been extensively examined by previous literature studies (278,279,282) and it is explained in details in chapter 1 of the current study. Moreover, previous researches confirmed the ability of MGMT to bind SS DNA (284), potentially *via* co-operative binding.

Furthermore, poor MGMT repair of O^6-CMG lesion when presented in SS 23 mer ODN was observed, that was in contrast with SS 13 mer ODN containing the same adduct. This finding could be attributed to the contribution of local DNA base sequence context. Several studies reported that MGMT repair efficiency of O^6-alkylG lesions is influenced by the surrounding nucleotide environment (290,372,373).

Notably, quantitative results of 23 mer ODNs (R_o and R_1) highlighted significant differences in MGMT repair preference of the 3 investigated O^6-alkylG adducts namely; O^6-MeG, O^6-CMG and O^6-CEG. Highest MGMT repair preference was observed for O^6-MeG over either O^6-CMG or O^6-CEG. These results are consistent with previous reports stating that MGMT repair preference of O^6-alkylG adducts is mainly dependent on the nature of the alkyl group at the O^6 position of the modified guanine. R_o and R_1 for DS O^6-alkylG ODNs suggest that MGMT preferentially repairs O^6-CMG > O^6-CEG which accord well with a previous study documenting that MGMT repair preference

decreases with the increase of the size of the alkyl adducts (290). This conclusion is contradicted by results of SS O^6-alkylG ODNs indicating higher MGMT repair preference for O^6-CEG *vs.* O^6-CMG which might be ascribed to the DNA sequence. Therefore, more accurate comparison requires 2 ODNs (O^6-CMG and O^6-CEG containing ODNs) of the same DNA sequence.

Similarly for LH1c, R_o and R_1 displayed significant differences between DS and SS ODNs pointing out that MGMT efficiently repaired DS LH1c over corresponding SS one. This result is in accordance with the reported O^6-MeG, O^6-CEG and O^6-CMG results confirming more efficient MGMT repair for DS over SS alkylated ODNs. Whilst, there were no significant differences in R_o calculated for DS *vs.* SS SRO1. Nevertheless, significant elevation in R_1 following MGMT reaction with DS SRO1 indicated a higher repair preference for DS SRO1 *vs.* SS ODN. Both quantitative SRO1 and LH1 results indicated a higher MGMT repair preference for O^6-carboxamidoethylG adduct when presented in DS ODNs relative to SS ones.

The difference in the quantitative results between SRO1 and LH1c ODNs is possibly explained in terms of presence of O^6-MeG adduct contaminant in the ODNs preparations. O^6-MeG is the primary substrate for MGMT therefore it might have interfered with the transfer of the carboxamidoethyl group. A previous study has reported that human wild type MGMT repaired O^6-MeG lesion twice as fast as O^6-pobG whilst other MGMT mutants did not repair any O^6-pobG until all of the O^6-MeG was removed (341).

The main strengths of this MS based approach is that it enabled the detection of the transfer of a wide range of alkyl moieties to MGMT-ASP following MGMT incubation with a wide a range of O^6-alkylG containing ODNs, even some of these alkyl groups were unexpected. However, the major caveats of the current work is lack of tandem MS data to confirm the identity of detected alkylated peptides as well as all of results were one point analysis rather than a kinetic approach.

5.5 Conclusion

A MALDI-TOF MS approach has been developed and shown to successfully detect the presence of O^6-alkylG lesions (O^6-MeG, O^6-CMG and O^6-CEG) in DNA *via*

MS analysis of MGMT tryptic peptides following incubation with alkylated ODNs. In addition, this approach found a preference for MGMT to repair DS O^6-alkylG ODNs over SS ones. This MALDI-TOF MS approach establishes a platform for the identification and quantification of O^6-alkylG adducts in human DNA generally and human colorectal cancer DNA especially.

6 Chapter 6: Detection and quantification of O^6-MeG adducts in methylated CT-DNA

6.1 Introduction

There is compelling evidence from epidemiologic data that endogenously formed NNOC play a pivotal role in colorectal cancer etiology (60,374). NNOC are pro-carcinogens that undergo metabolic activation before they can react with DNA and trigger DNA alkylation forming O^6-alkylG adducts including O^6-MeG (156,184). The formation and persistence of O^6-MeG in cellular DNA has drastic biological consequences that include mutagenicity, clastogenicity and carcinogenicity (212). Moreover, there is a direct mechanistic link between O^6-MeG formation and colorectal carcinogenesis as O^6-MeG adduct is tightly associated with mutational activation of K-ras oncogene (235).

Hence, the detection of O^6-alkylG adducts, more specifically of O^6-MeG adduct is potentially used for monitoring exposure to NNOC. In general, O^6-alkylG adducts are formed in low amounts in DNA and they constitute only 6-8% of total alkylated DNA adducts (156). Therefore, the identification and quantification of such adducts require highly sensitive techniques especially when analyzing human DNA samples where limited amounts of human tissue and/or DNA are available (253). Several methods have been successfully used to detect O^6-MeG adduct including immunoassays (257), Immunoaffinity-HPLC (246,375) and ^{32}P-postlabeling assays (261). Recently, MS based analytical approaches have contributed to the identification and quantification of O^6-MeG adducts as well as O^6-CMG adducts (263,254) and ultra-high pressure liquid chromatography–tandem MS (UHPLC-MS/MS) has been used to quantify of O^6-MeG and O^6-CMG adducts in intestinal cell line (253).

Work in previous chapters (Chapter 5) has shown that alkyl adducts in ODNs can be qualitatively detected as the alkyl modified MGMT tryptic peptide by MALDI-TOF. To the best of our knowledge, no previous studies have reported a methodology that was able to detect O^6-MeG in DNA *via* MGMT.

6.2 Aims and objectives

The aim of this work is to develop a MALDI-TOF MS based methodology to detect and quantify O^6-MeG adducts in TMZ modified CT-DNA using the known function of MGMT. MGMT can irreversibly transfer methyl groups from a modified guanine base to its active cysteine residue. Herein, MS detection of methylated MGMT-ASP in tryptic fragments of MGMT following its reaction with methylated CT-DNA can be used as a tool for O^6-MeG adduct detection and subsequent quantification. As a consequence, it is necessary to determine LOQ and the linearity of detector response for MGMT-ASP in both unmodified as well as methylated forms.

The specific objectives of this work were as follows:

1. To estimate the LOQ of synthetic unmodified and methylated MGMT-ASP standard (both light and heavy isotope labelled).

2. To examine the linearity of MALDI-TOF MS quantification of both unmodified and methylated synthetic MGMT-ASP standards.

3. To develop a novel assay to quantify methylated of MGMT-ASP in MBP-MGMT and his-MGMT tryptic digests following incubation of MGMT with CT-DNA that has been methylated *in vitro* by TMZ.

4. To estimate the LOQ for O^6-MeG adduct in CT-DNA using the newly developed technique.

6.3 Results

6.3.1 LOQ of synthetic MGMT-ASP standards

Here the LOQ is defined as a S/N ratio of 10 (253). Synthetic versions of both unmodified and methylated MGMT-ASP were used for the determination of the lowest LOQ. Stable isotope ($C^{13}N^{15}$) proline labelled MGMT-ASP (unmodified as well as methylated) are used as internal standards for quantification purposes.

6.3.1.1 LOQ of synthetic unmodified MGMT-ASP standard

Preliminary experiments were conducted using unlabelled MGMT-ASP standard to investigate the most appropriate solvent for dissolution of lyophilized peptide standards. Several different protocols were investigated: in the first, MALDI-TOF MS analysis of the 4 serial dilutions of NEM-capped MGMT-ASP (Section 2.2.18) failed to detect below 0.5 pmole/µL (Figure 6.1).

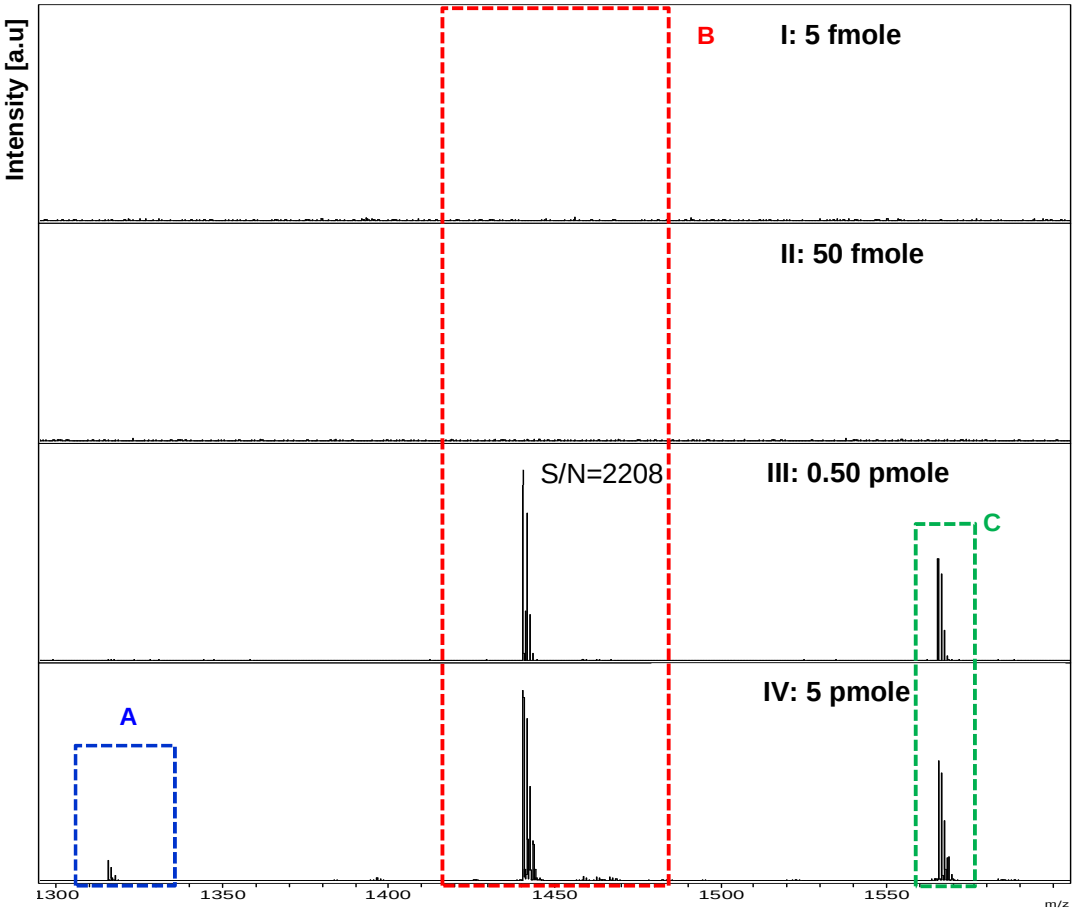


Figure 6.1: MALDI-TOF data to obtain the LOQ for synthetic NEM-capped MGMT-ASP (pH 8).

A. Unmodified synthetic MGMT-ASP (GNPVPILIPCHR, 136-147, *m/z*=1315.73). **B**. Synthetic NEM-capped MGMT-ASP (*m/z* 1440.77). **C**. Putative NEM amino adduct (*m/z* 1565.82). Peak intensities are shown in arbitrary units on the y-axis. The intensity scale of each panel is 20X that of the previous panel except for panels I and II which are at the same scale. Amounts indicate absolute amounts of the synthetic standard spotted on the sample plate and analyzed.

At 5 pmole/µL, two additional species were detected at *m/z* 1315.73 $[M+H]^+$ assigned to unmodified MGMT-ASP and at *m/z* 1565.82 $[M+H]^+$. The latter was also present at 0.5 pmole/µL and assigned to an adduct formed as a result of the non-specific reaction of NEM with amino groups (NEM amino adduct ion). This protocol achieved a LOQ of 500 fmole/µL for unlabelled NEM-capped synthetic MGMT-ASP standard. Due the potential errors from the NEM reaction, this protocol was slightly modified, to remove the cysteine capping step. However, this failed to lower the LOQ (Figure 6.2), although the putative NEM amino adduct ion was not detected at *m/z* 1565.82 supporting the assignment of NEM interference.

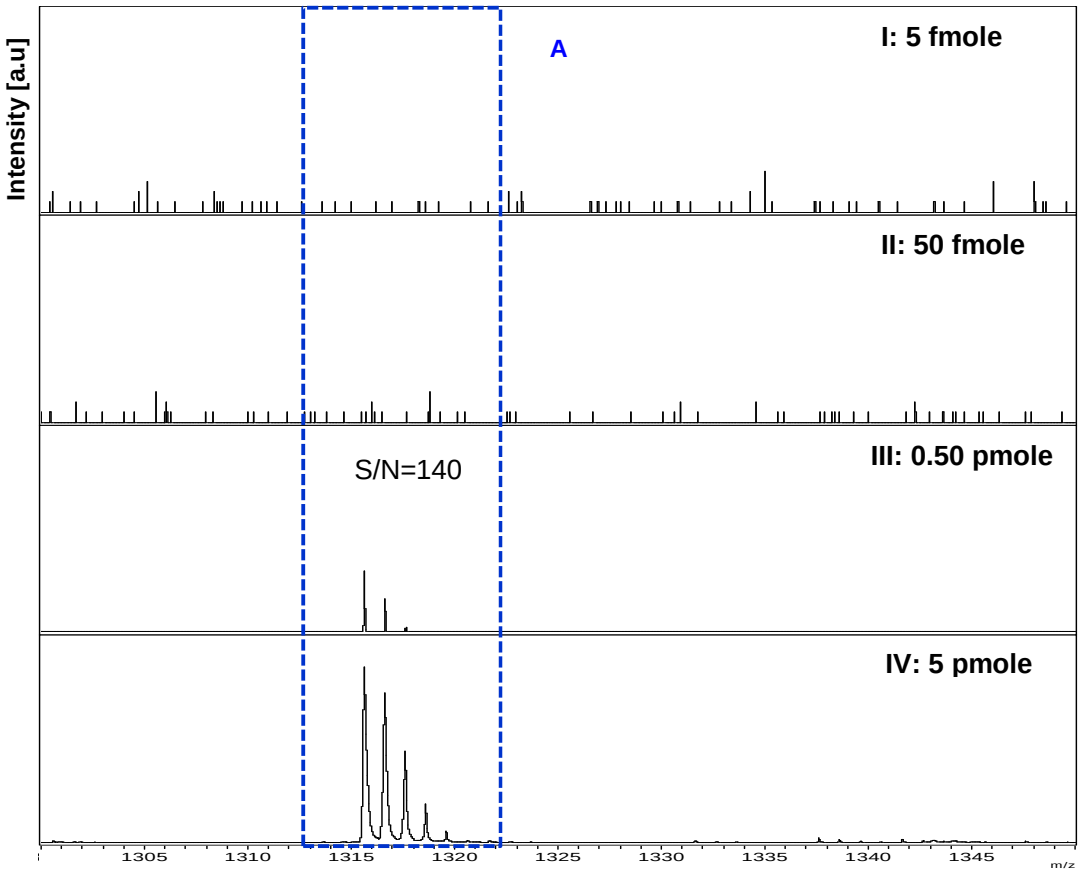


Figure 6.2: MALDI-TOF data to obtain the LOQ for synthetic unmodified MGMT-ASP (pH 8).

A. Unmodified synthetic MGMT-ASP (GNPVPILIPCHR, 136-147, *m/z*=1315.73). Peak intensities are shown in arbitrary units on the y-axis. The intensity scale of each panel is 100X that of the previous panel except for panels I and II which are at the same scale. Amounts indicate absolute amounts of the synthetic standard spotted on the sample plate and analyzed.

In order to increase the sensitivity of the developed assay the protocol was modified by changing the peptide solvent to 0.2% formic acid in 20 mM Tris-HCl pH 4 then NEM capping of reduced cysteine residues at pH 8 (2nd protocol, Section 2.2.18). NEM-capped MGMT-ASP was now detectable at 50 fmole/µL with a LOQ <50 fmole/µL (Figure 6.3). However, this protocol did not remove the putative NEM adduct *m/z* 1565.8. Furthermore, unmodified MGMT-ASP was detected in the serial dilutions of concentrations higher than 50 fmole/µL indicating that the NEM concentration was not sufficient or the reaction was not optimized in the changed solution conditions (Figure 6.3).

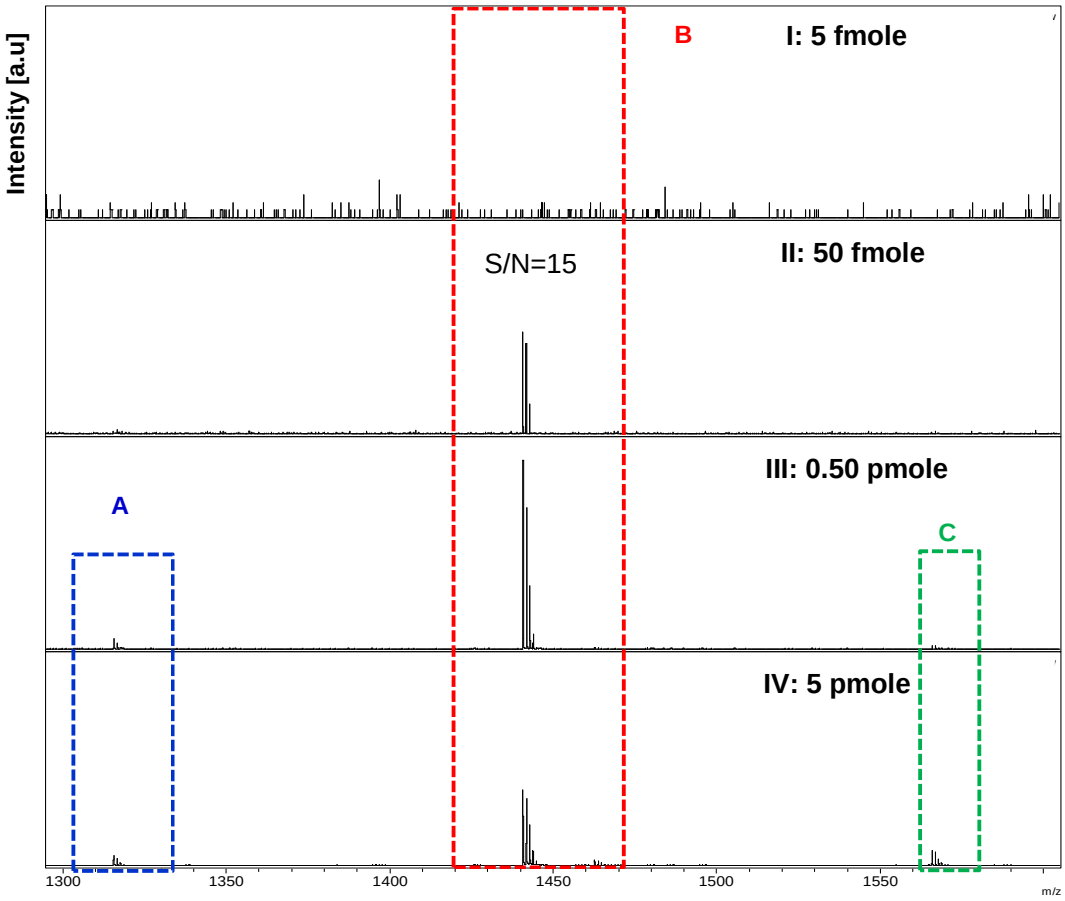


Figure 6.3: MALDI-TOF data to obtain LOQ for synthetic NEM-capped MGMT-ASP (pH 4 then 8).

A. Unmodified synthetic MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z* 1315.73). **B**. Synthetic NEM-capped MGMT-ASP (*m/z* 1440.77). **C**. Putative NEM amino adduct (*m/z* 1565.82). Peak intensities are shown in arbitrary units on the y-axis.The intensity scale of each panel is 10X that of the previous panel. Amounts indicate absolute amounts of the synthetic standard spotted on the sample plate and analyzed.

Further amendments to the protocol were made, the NEM capping step was conducted at pH 7 to avoid the NEM adduct formation and the NEM concentration was increased (30 mM) (Section 2.2.18.). MS analysis showed an improved in the LOQ to < 20 fmole/µL; (Figure 6.4.) However, traces of unmodified MGMT-ASP ion as well as NEM amino adduct ion were still detected with attendant quantitative imprecision.

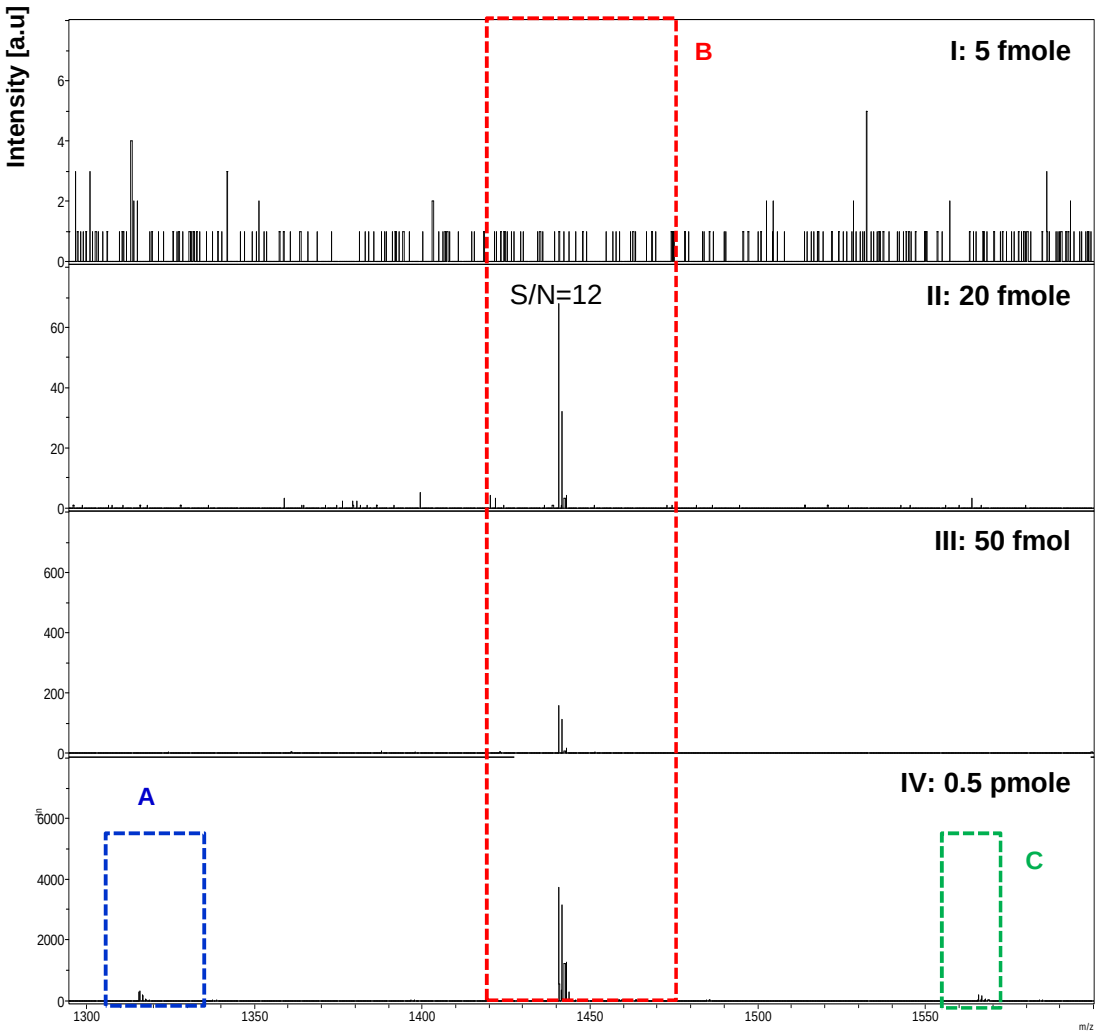


Figure 6.4: MALDI-TOF data to obtain LOQ for synthetic NEM-capped MGMT-ASP (pH 4 then 7).

A. Unmodified synthetic MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z* 1315.73). **B**. Synthetic NEM-capped MGMT-ASP (*m/z* 1440.77). **C**. Putative NEM amino adduct (*m/z* 1565.82). Peak intensities are shown in arbitrary units on the y-axis.The intensity scale of each panel is 10X that of the previous panel. Amounts indicate absolute amounts of the synthetic standard spotted on the sample plate and analyzed.

The next iteration skipped the NEM capping step and achieved LOQ < 20 fmole/µL (Figure 6.5).

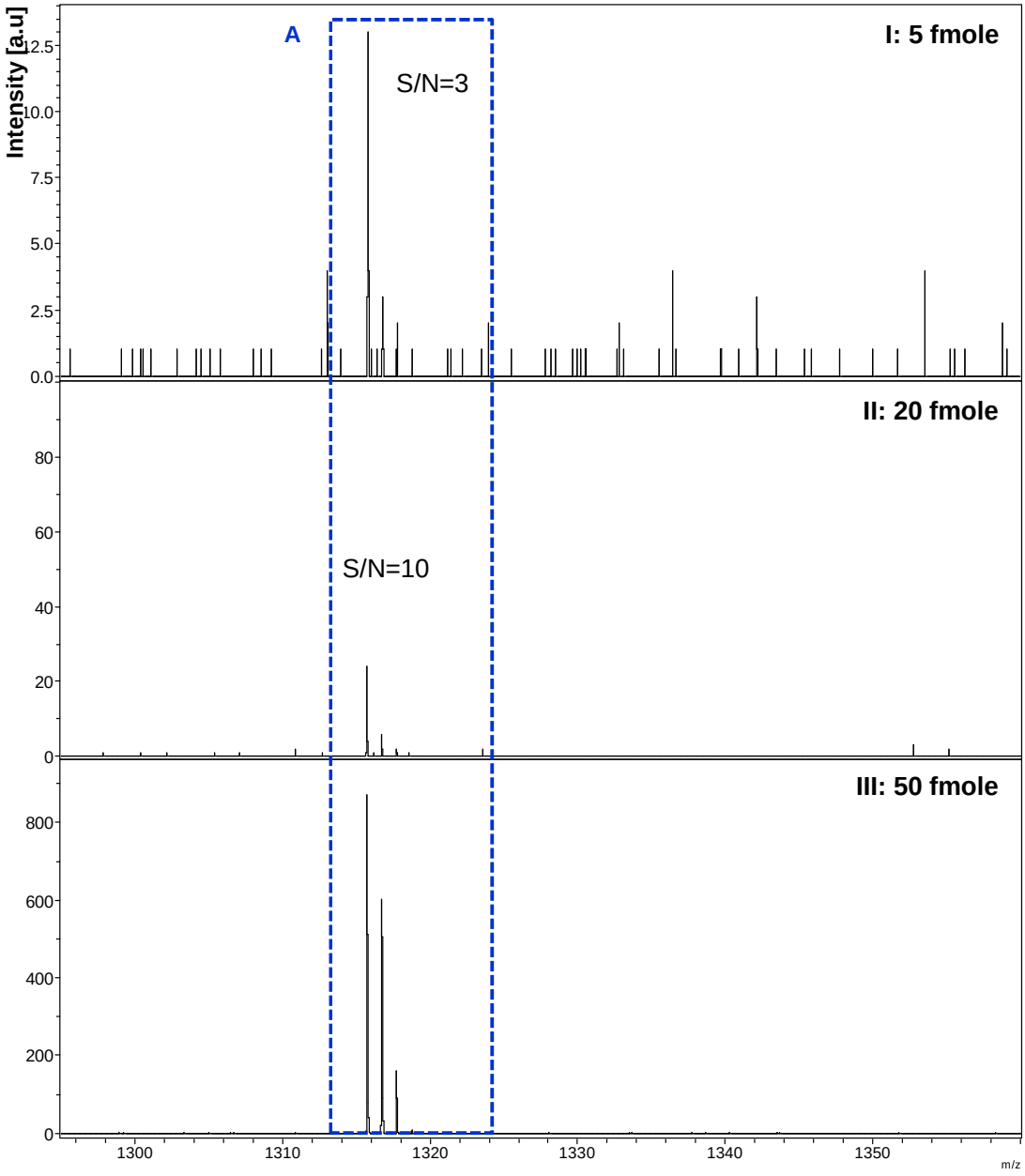


Figure 6.5: MALDI-TOF data to obtain the LOQ for synthetic LOQ of synthetic unmodified MGMT-ASP (pH 4).

A. Unmodified synthetic MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z* 1315.73). Peak intensities are shown in arbitrary units on the y-axis. The intensity scale of each panel is 10X that of the previous panel. Amounts indicate absolute amounts of the synthetic standard spotted on the sample plate and analyzed.

Using $C^{13}N^{15}$ proline labelled unmodified MGMT-ASP, the same protocols were investigated revealing that 2nd protocol achieved a limit of detection < 500 fmole/µL again ions at *m/z* 1571.82 was present (Figure 6.6) and assigned to $C^{13}N^{15}$ labelled NEM amino adduct ion (due to NEM non-specific reaction with amino groups).

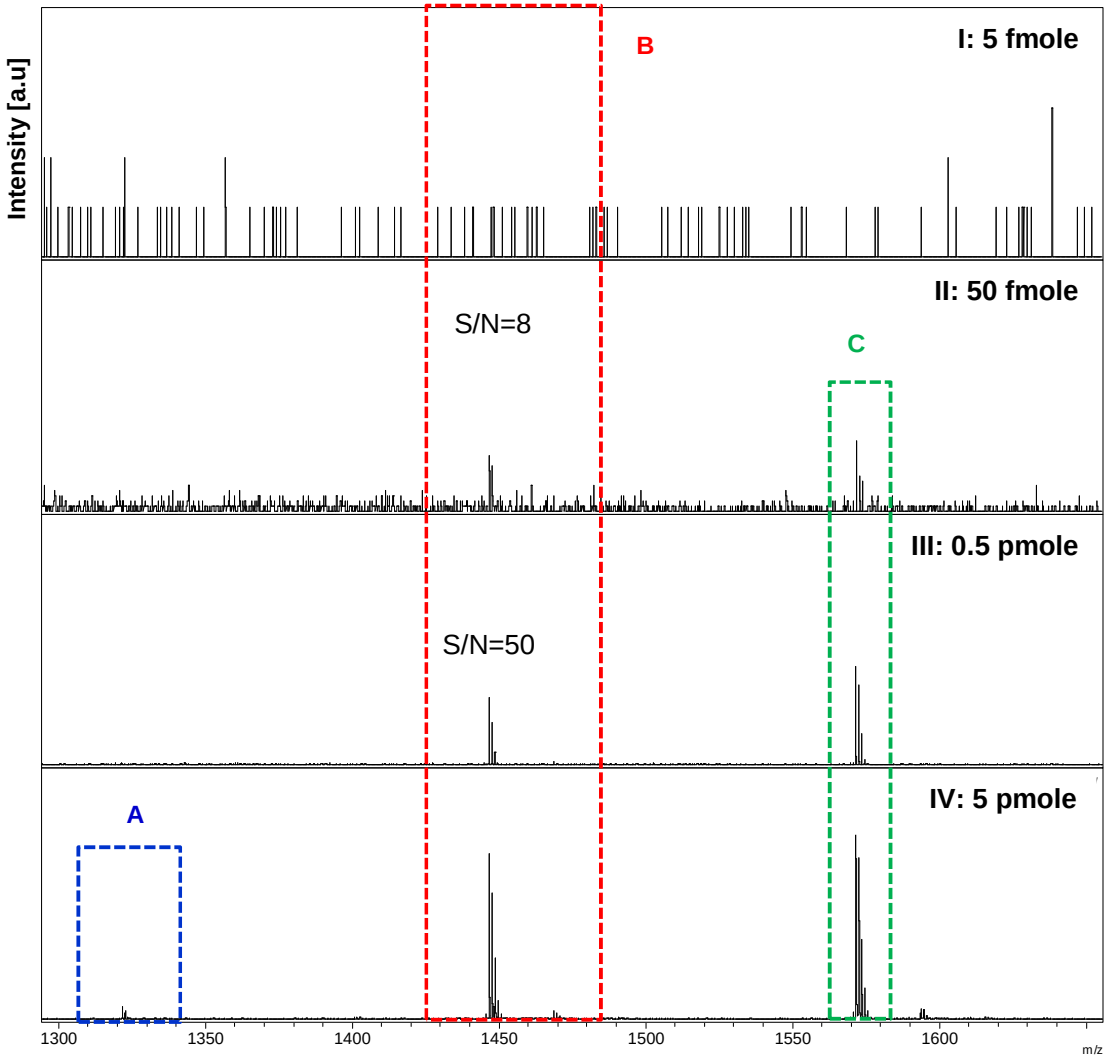


Figure 6.6: MALDI-TOF data to obtain the LOQ for synthetic $C^{13}N^{15}$ proline labelled NEM-capped MGMT-ASP (pH 8).

A. $C^{13}N^{15}$unmodified synthetic MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z* 1321.73). **B**. Synthetic NEM-capped $C^{13}N^{15}$ labelled MGMT-ASP (*m/z* 1446.77). **C**. Putative $C^{13}N^{15}$ labelled NEM amino adduct (*m/z* 1571.82). Peak intensities are shown in arbitrary units on the y-axis. The intensity scale of each panel is 10X that of the previous panel. Amounts indicate absolute amounts of the synthetic standard spotted on the sample plate and analyzed.

Though both protocols 3rd and 4th achieved similar limits of detection <20 fmole/µL (Figure 6.7 and 6.8) the 3rd protocol also had a NEM amino adduct and incomplete cysteine capping (Figure 6.7).

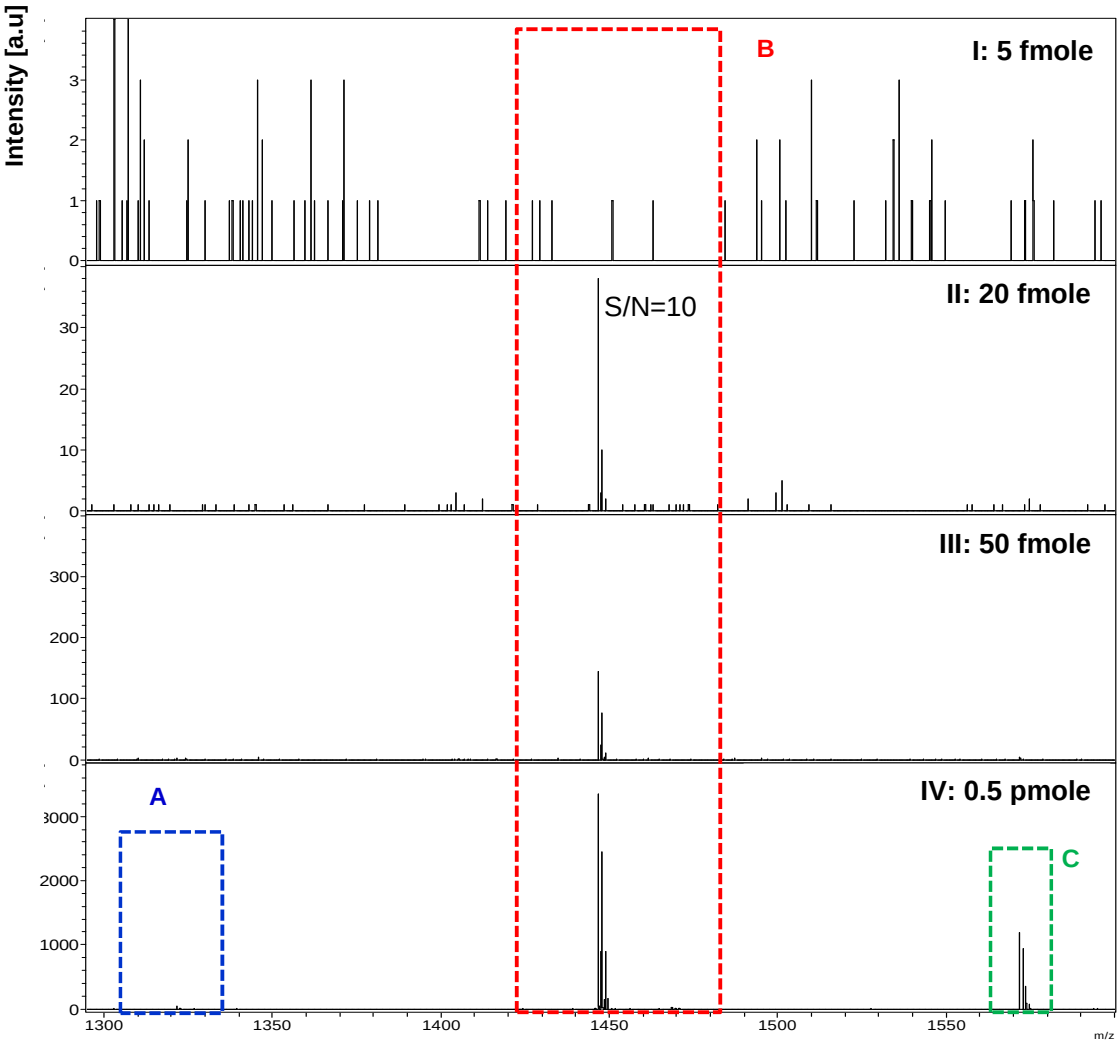


Figure 6.7: MALDI-TOF data to obtain the LOQ for synthetic $C^{13}N^{15}$ proline labelled NEM-capped MGMT-ASP (pH 7).

A. corresponds to unmodified synthetic $C^{13}N^{15}$ labelled MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z*=1321.73). **B**. corresponds to synthetic NEM-capped $C^{13}N^{15}$ labelled MGMT-ASP (*m/z* 1446.77). **C**. Putative $C^{13}N^{15}$ labelled NEM amino adduct (*m/z* 1571.82). Peak intensities are shown in arbitrary units on the y-axis. Max peak intensity of each panel is 10X that of the previous panel. Amounts indicate absolute amounts of the synthetic standard spotted on the sample plate and analyzed.

Based on these results, heavy proline labelled MGMT-ASP internal standard was used for the rest of the study with reduced and unmodified cysteines to ensure the accuracy of quantitation.

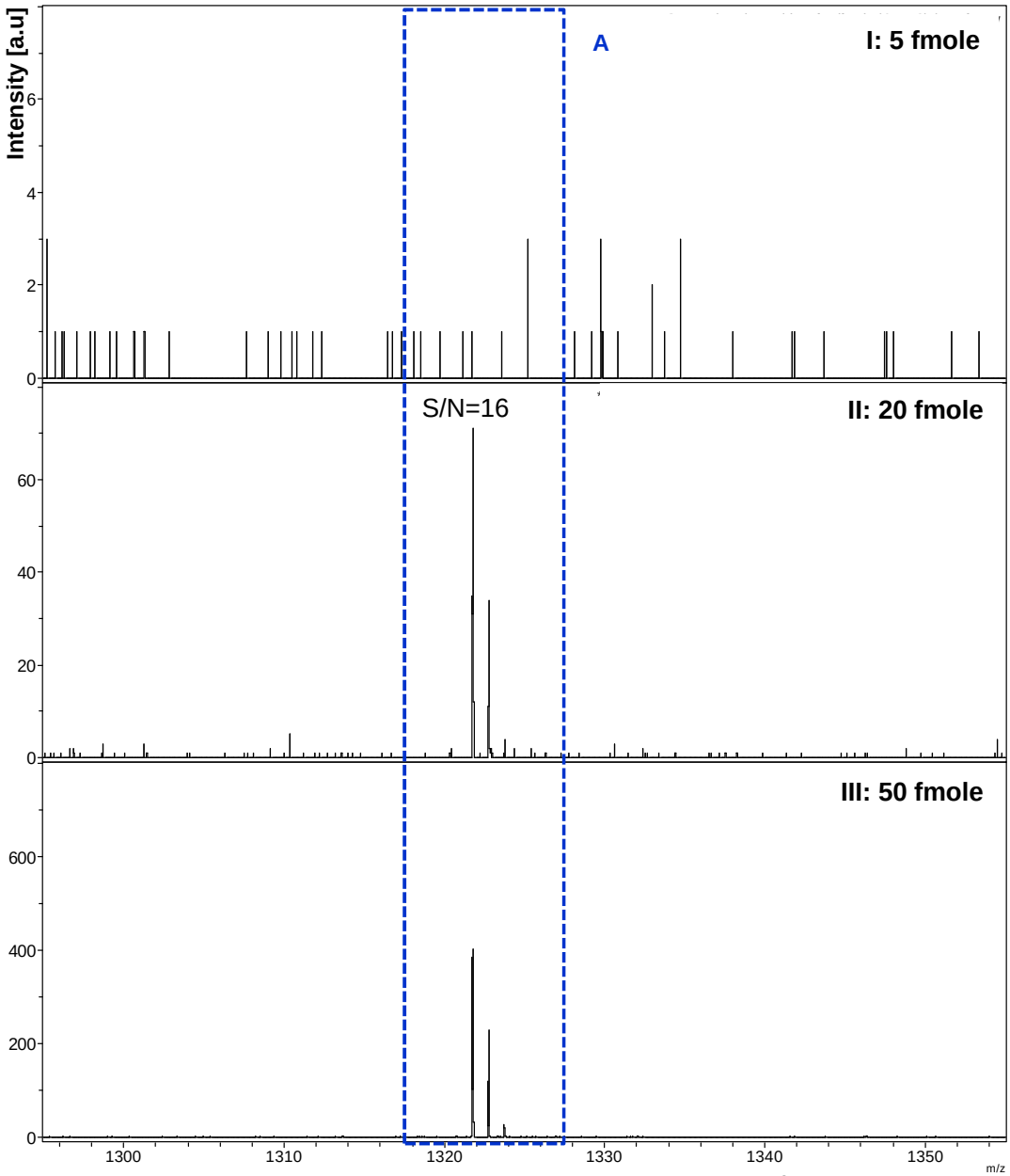


Figure 6.8: MALDI-TOF data to obtain the LOQ for synthetic $C^{13}N^{15}$ proline labelled unmodified MGMT-ASP.

A. Unmodified synthetic $C^{13}N^{15}$ labelled MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z*=1321.73). Peak intensities are shown in arbitrary units on the y-axis. The intensity scale of each panel is 10X that of the previous panel. Amounts indicate absolute amounts of the synthetic standard spotted on the sample plate and analyzed.

6.3.1.2 LOQ of the synthetic methylated MGMT-ASP standard

Figure 6.9 and 6.10 show that the LOQ of synthetic versions of methylated MGMT-ASP (unlabelled *m/z*=1329.74 and $C^{13}N^{15}$ Proline labelled *m/z*=1335.74) was found to be <20 fmole/µL.

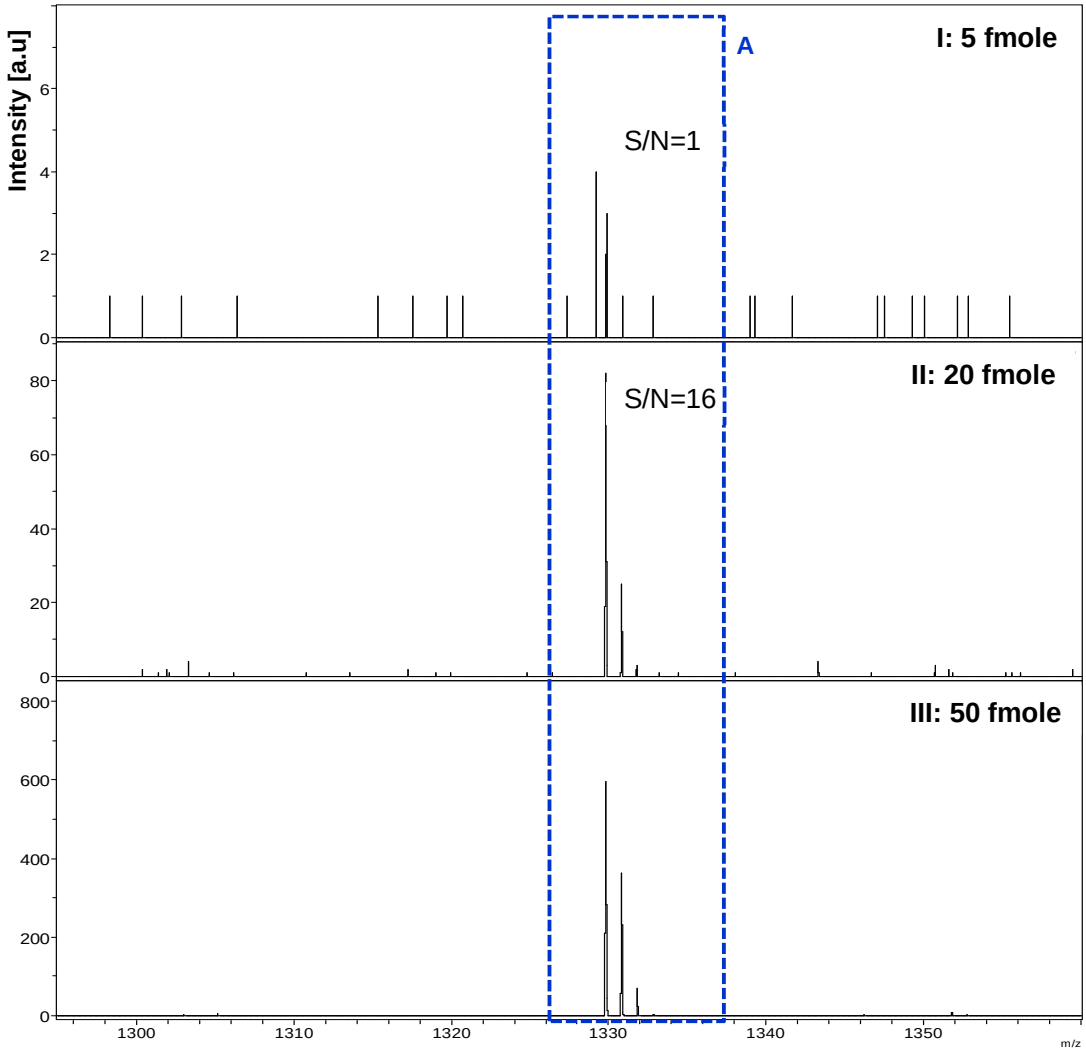


Figure 6.9: MALDI-TOF data to obtain the LOQ for synthetic methylated MGMT-ASP.

A. Methylated synthetic MGMT-ASP (*m/z* 1329.74). Peak intensities are shown in arbitrary units on the y-axis. The intensity scale of each panel is 10X that of the previous panel. Amounts indicate absolute amounts of the synthetic standard spotted on the sample plate and analyzed.

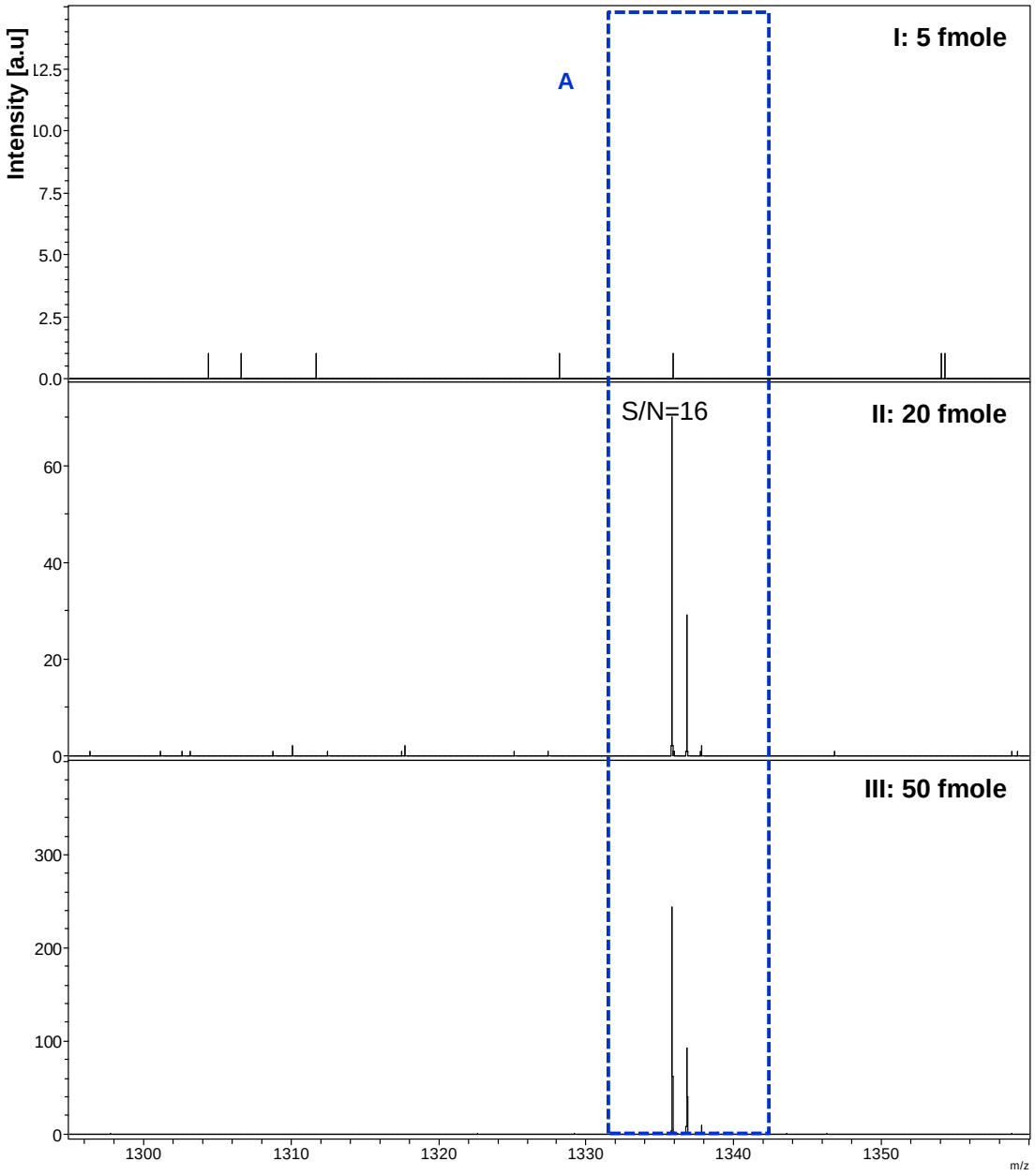


Figure 6.10: MALDI-TOF data to obtain the LOQ for synthetic $C^{13}N^{15}$ proline labelled methylated MGMT-ASP.

A. $C^{13}N^{15}$ proline labelled methylated synthetic MGMT-ASP (*m/z* 1335.74). Peak intensities are shown in arbitrary units on the y-axis. The intensity scale of each panel is 5X that of the previous panel. Amounts indicate absolute amounts of the synthetic standard spotted on the sample plate and analyzed.

6.3.2 Establishment of linearity of quantitative MALDI-TOF MS approach

Serial dilutions of four synthetic versions of MGMT-ASP were prepared. Herein, the PAs corresponding to MGMT-ASP standard ions were measured using MALDI-TOF MS and plotted against absolute amount of MGMT-ASP standard that was analyzed in fmole. Linear correlations with an R^2 value of 0.9947, 0.9982, 0.9985 and 0.9998 were obtained for unmodified unlabelled MGMT-ASP standard, unmodified heavy-isotope labelled MGMT-ASP internal standard, unlabelled methylated MGMT-ASP standard and heavy-isotope labelled methylated MGMT-ASP internal standard, respectively, (Figure 6.11 A-D.)

6.3.3 Detection of methylated MGMT-ASP in tryptic digests of MBP-MGMT following incubation with TMZ-modified CT-DNA and re-captured by amylose coated beads

A new approach using amylose coated magnetic beads was developed to capture MBP-MGMT following its incubation with either TMZ-modified CT DNA or unmodified CT-DNA (negative control) as described in Section 2.2.21. This detected a large number of MBP (20) and MGMT (9) tryptic peptides including MGMT-ASP. A Mascot search of detected tryptic peptides indicated significant matching between theoretical MGMT_HUMAN peptides and recombinant MGMT tryptic peptides masses with 62 hits and 49% of MGMT amino acid sequence coverage (Expect=0.012). Methylated MGMT-ASP was also detected in tryptic digests of MBP-MGMT protein that was incubated with TMZ-modified CT-DNA (Figure 6.12). Methylated MGMT-ASP was not detected in tryptic digests of MBP-MGMT incubated with unmodified CT-DNA. Although initial results in the early protocol stages hinted at some applicability to this approach, it was hindered by lack of reproducibility and inconsistency of the results. Further trials failed to recapture the MBP-MGMT protein from CT-DNA using amylose coated magnetic beads. This result was confirmed by both MS analysis that did not detect MBP-MGMT tryptic peptides (Data not shown).

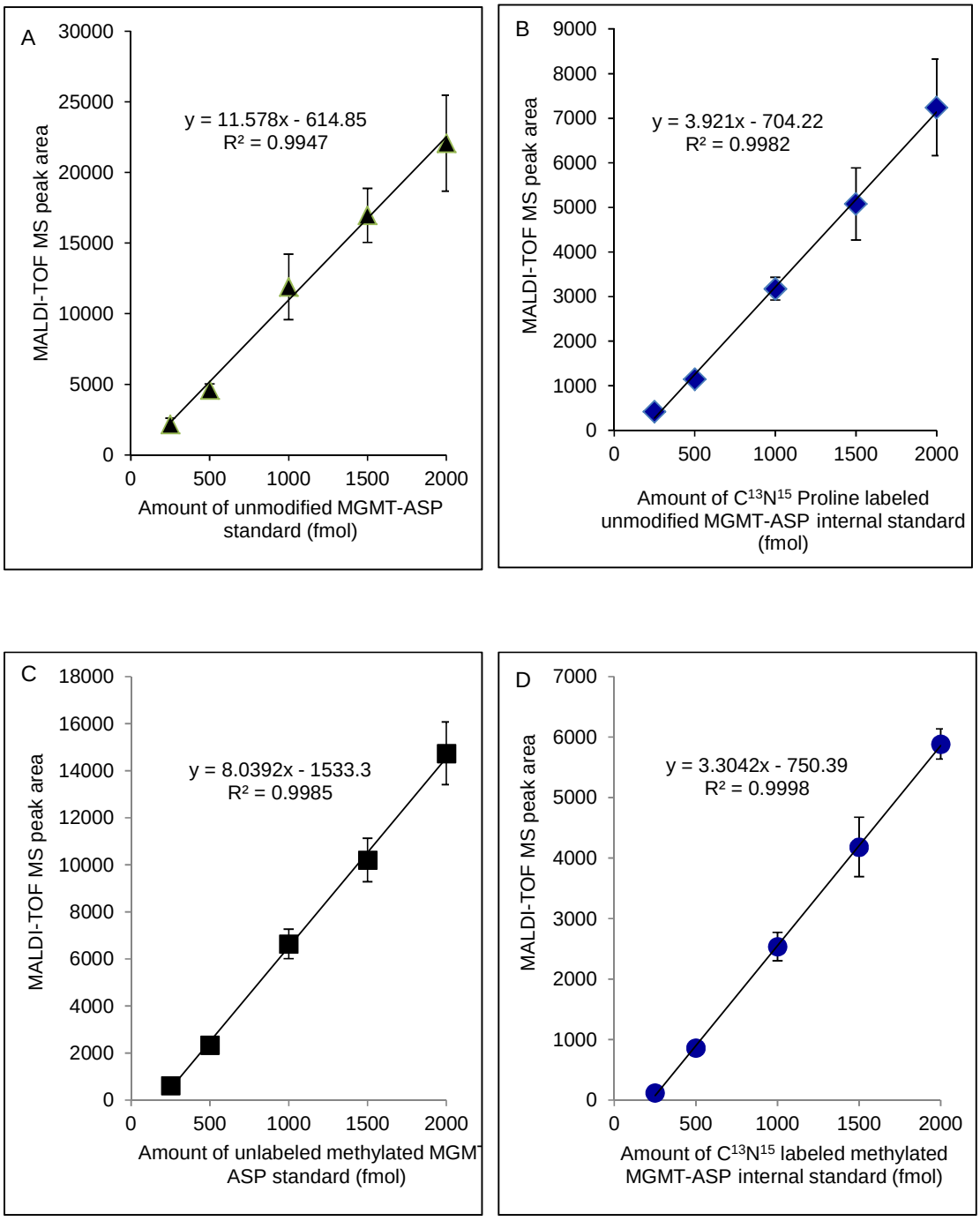


Figure 6.11: Dose response curve of synthetic versions of MGMT-ASP standard against MALDI-TOF MS PA for each ion.

A. Unlabelled unmodified MGMT-ASP standard.
B. $C^{13}N^{15}$ proline labelled unmodified MGMT-ASP internal standard.
C. Unlabelled methylated MGMT-ASP standard.
D. $C^{13}N^{15}$ proline labelled methylated MGMT-ASP internal standard.
(Values are mean±SD, n=3).

In order to confirm MS results, following MBP-MGMT incubation with DNA and recovery with amylose coated magnetic beads, the beads were boiled with 1X SDS-PAGE LB to elute any bound proteins. Eluted protein fractions were resolved on 12% SDS-PAGE and very faint MBP-MGMT bands were detected indicating poor protein recovery by amylose magnetic beads following DNA incubations (Figure 6.13).

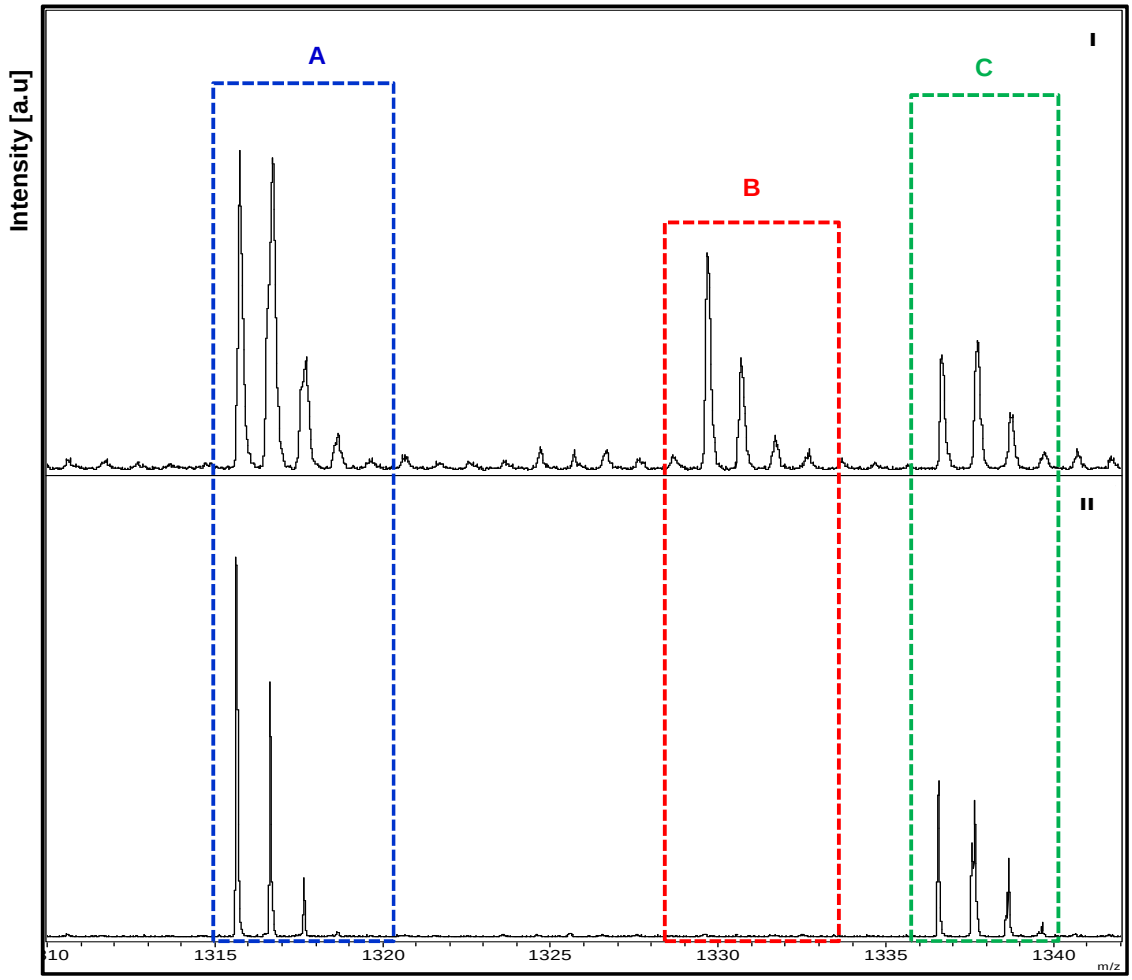


Figure 6.12: MS detection of methylated MGMT-ASP in tryptic peptides of MBP-MGMT incubated with methylated CT-DNA (On-bead digestion).

MBP-MGMT (16 pmole; 4 active pmole) was incubated with (I) TMZ-modified CT-DNA (500 µg containing 3 pmole O^6-MeG adduct) or (II) Unmodified CT-DNA then recovered by amylose beads. Tryptic peptides were desalted followed by MALDI-TOF MS analysis. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z*=1315.73). **B**. Methylated MGMT-ASP (GNPVPILIPMe-CHR, 136-147; *m/z*=1329.74). **C**. MBP-TP13 (SYEEELAKDPR, 332-342; *m/z*=1336.54). Peak intensities are shown in arbitrary units on the y-axis and the intensity scale is the same for both panels.

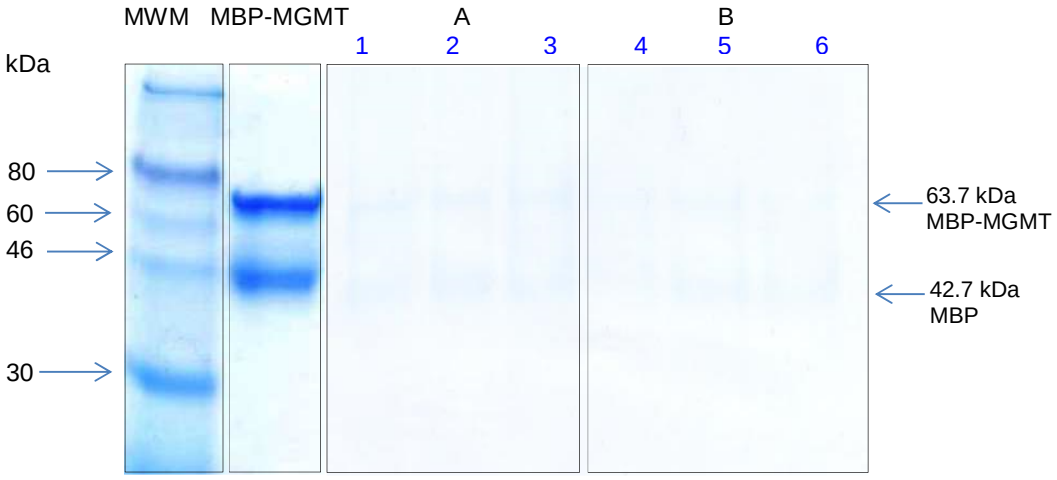


Figure 6.13: SDS-PAGE analysis of MBP-MGMT protein incubated with CT-DNA then re-captured with amylose beads.

From left to right, are lanes for: **MWM.** Molecular weight marker (10 µL) **MBP-MGMT.** protein extract was directly loaded onto the gel (2 µg) **A.** Three replicates of MBP-MGMT protein (2 µg) recaptured by amylose coated beads, following protein incubation with CT-DNA (overnight incubation with beads at 4˚C). **B.** Three replicates of MBP-MGMT protein (2 µg) recaptured by amylose coated beads, following protein incubation with CT-DNA (1 hour incubation with beads at 4˚C). Protein fractions were resolved on 12% SDS-PAGE followed by Instant Blue® Coomassie staining.

6.3.4 Detection of methylated MGMT-ASP in tryptic digests of MBP-MGMT following incubation with methylated CT-DNA followed by in-solution protein digestion

To overcome the inconsistency of amylose coated beads in rescuing the MBP-MGMT a new approach was developed (Section 2.2.22) based on in-solution protein digestion following MBP-MGMT incubation with CT-DNA in presence of buffer I.

6.3.4.1 Validation of the in-solution protein digestion protocol *via* assessment of protein recovery

MBP-MGMT was incubated with unmodified CT-DNA that was previously sonicated, (Section 2.2.22). Tryptic peptides were desalted and spiked with $C^{13}N^{15}$ proline labelled unmodified ASP internal standard to assess the percent of protein recovery prior to analysis. This protocol gave 5.0±0.1% of protein recovery (0.80±0.02 pmole MGMT-ASP was detected in the tryptic digest); Figure 6.14.I. Increasing protein recovery was fundamental for increasing the assay sensitivity; further DNA sonication after incubation with the protein improved protein recovery to 11.0±0.2% (1.7±0.03 pmole of MGMT-ASP were detected); Figure 6.14.II. Although post-incubation sonication successfully increased protein recovery from 5% to 11%, this level of protein recovery was still not adequate.

6.3.4.2 LOQ of O^6-MeG adduct in TMZ-modified CT-DNA

MBP-MGMT was incubated with methylated CT-DNA containing 0.5, 1, 2 and 4 pmole of O^6-MeG adduct per mg of CT-DNA (0.25, 0.5, 1 and 2 pmole absolute amount of O^6-MeG adduct). Methylated CT-DNA was sonicated twice pre and post protein incubation, then the protein was digested in-solution as described in Section 2.2.22. Tryptic digests were further spiked with $C^{13}N^{15}$ proline labelled methylated ASP internal standard. Figure 6.15 shows that methylated MGMT-ASP was detected in tryptic fragments of protein incubated with methylated CT-DNA contained 1, 2 and 4 pmole of O^6-MeG adduct mg^{-1} CT-DNA whilst it was not detected following protein incubation with 0.5 pmole adduct mg^{-1} CT-DNA with a S/N>10.

A non-statistically significant linear correlation was obtained when MS PA of detected methylated MGMT-ASP was plotted against absolute amount of O^6-MeG adducts in methylated CT-DNA (fmole of O^6-MeG adducts incubated with MBP-MGMT determined by radioisotopic assay) (Figure 6.16). In addition to the failure of detection of methylated ASP (Figure 6.14.I), this protocol failed to achieve the next goal which was quantification of the detected methylated MGMT-ASP using the internal standard. Failure of quantification was due to the mass coincidence between $C^{13}N^{15}$ proline

labelled methylated ASP parent ion at *m/z*=1335.74 and the ion due to MBP-TP13 at *m/z*=1336.54. This prevented accurate determination of PAs.

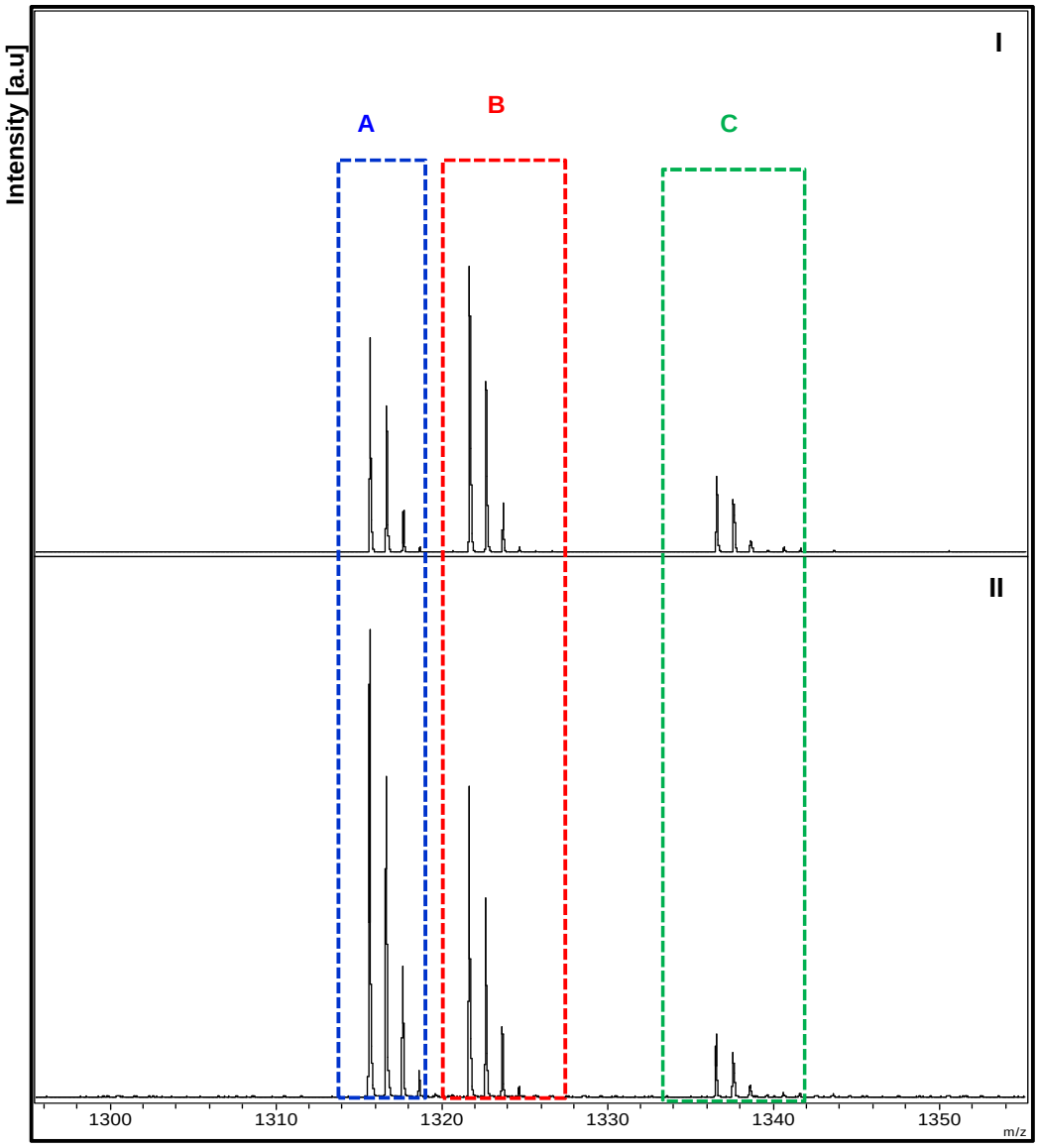


Figure 6.14: MALDI-TOF mass spectra of MBP-MGMT incubated with CT-DNA.

MBP-MGMT (16 pmole; 4 active pmole) was incubated with unmodified CT-DNA (500 µg) then digested *via* in-solution trypsin digestion. Tryptic peptides were desalted then spiked with $C^{13}N^{15}$ proline labelled unmodified MGMT-ASP (internal standard) followed by MALDI-TOF MS analysis. **I**. CT-DNA was sonicated prior to incubation with MGMT only. **II**. CT-DNA was sonicated prior and after incubation with MGMT **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z*=1315.73). **B**. $C^{13}N^{15}$ proline labelled unmodified MGMT-ASP internal standard. *m/z*=1321.73. **C**. MBP-TP13 (SYEEELAKDPR, 332-342) *m/z*=1336.54. Peak intensities are shown in arbitrary units on the y-axis and max peak intensity is the same for both panels.

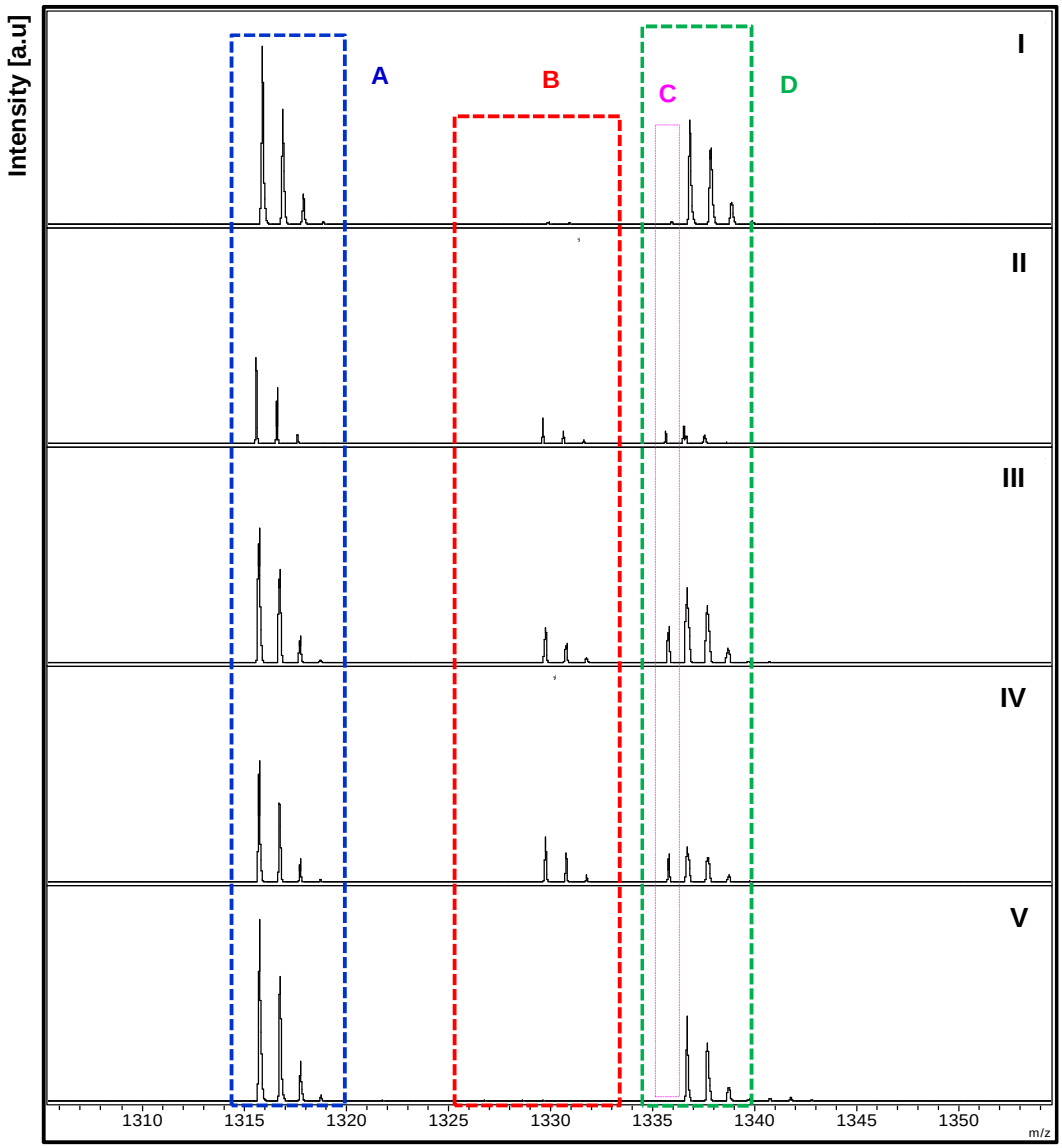


Figure 6.15: MS detection of methylated MGMT-ASP in tryptic peptides of MBP-MGMT incubated with methylated CT-DNA (In-solution digestion).

MBP-MGMT (16 pmole; 4 active pmole) was incubated with TMZ-modified CT-DNA containing 0.5 pmole (I), 1 pmole (II), 2 pmole (III), 4 pmole (IV) O^6-MeG adduct mg^{-1} CT-DNA or unmodified CT-DNA (V) then digested *via* in-solution trypsin digestion. Tryptic peptides were desalted the spiked with $C^{13}N^{15}$ proline labelled methylated MGMT-ASP (internal standard) followed by MALDI-TOF MS analysis. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z*=1315.73). **B**. Methylated MGMT-ASP (GNPVPILIPMe-CHR, 136-147; *m/z*=1329.74). **C**. $C^{13}N^{15}$ proline labelled methylated MGMT-ASP internal standard. *m/z*=1335.74. **D**. MBP-TP13 (SYEEELAKDPR, 332-342) *m/z*=1336.54. Peak intensities are shown in arbitrary units on the y-axis and max peak intensity is the same for all panels.

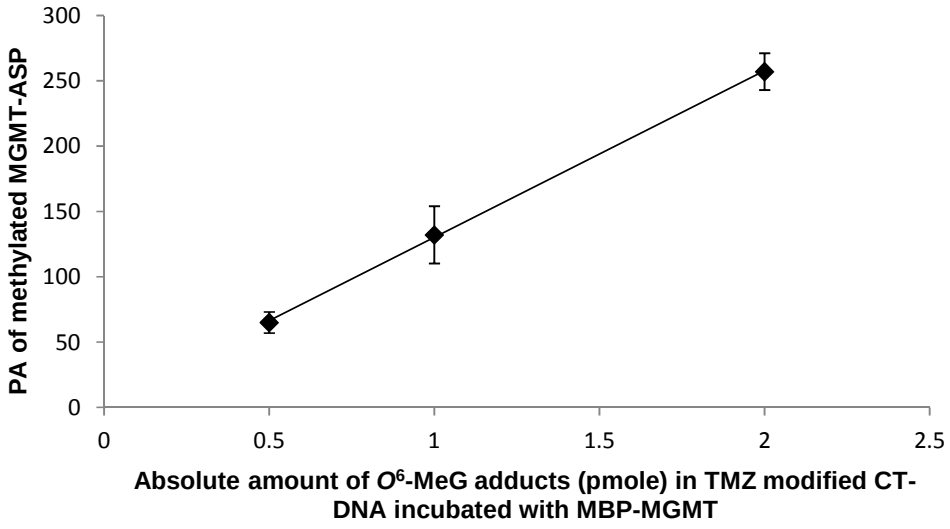


Figure 6.16: Validation results for the MALDI-TOF MS method for detection of methylated MGMT-ASP following MBP-MGMT incubation with methylated CT-DNA.

MBP-MGMT (16 pmole; 4 active pmole) was incubated with TMZ-modified CT-DNA (0.5 mg) containing 0.5, 1 and 2 pmole absolute amounts of O^6-MeG adduct (1, 2 or 4 pmole of O^6-MeG mg^{-1} CT-DNA adducts). Resulting mixtures were subjected to in-solution tryptic digestion. Results are expressed as PA of methylated MGMT-ASP versus amount of adduct incubated with MGMT protein.

6.3.5 MS analysis of his-MGMT following on-bead digestion after incubation with CT-DNA

Work presented previously underlines the necessity of recovering MGMT protein after incubation with DNA to improve protein recovery. Ni-coated magnetic beads were used to recover his-MGMT fusion protein after incubation with unmodified CT-DNA followed by on-beads tryptic digestion (Section 2.2.23). Figure 6.17 shows the MALDI-TOF mass spectra of his-MGMT tryptic peptides, recovered by Ni-coated beads after incubation with CT-DNA, displayed 9 MGMT peptides. The mass spectra contained two prominent ions corresponding to MGMT-TP6 at *m/z* 1247.62 (166-175, EWLLAHEGHR) and MGMT-ASP at *m/z* 1315.73 (136-147, GNPVPILIPCHR). The

amino acid sequence of MGMT-ASP was confirmed by tandem MS (Results are not shown).

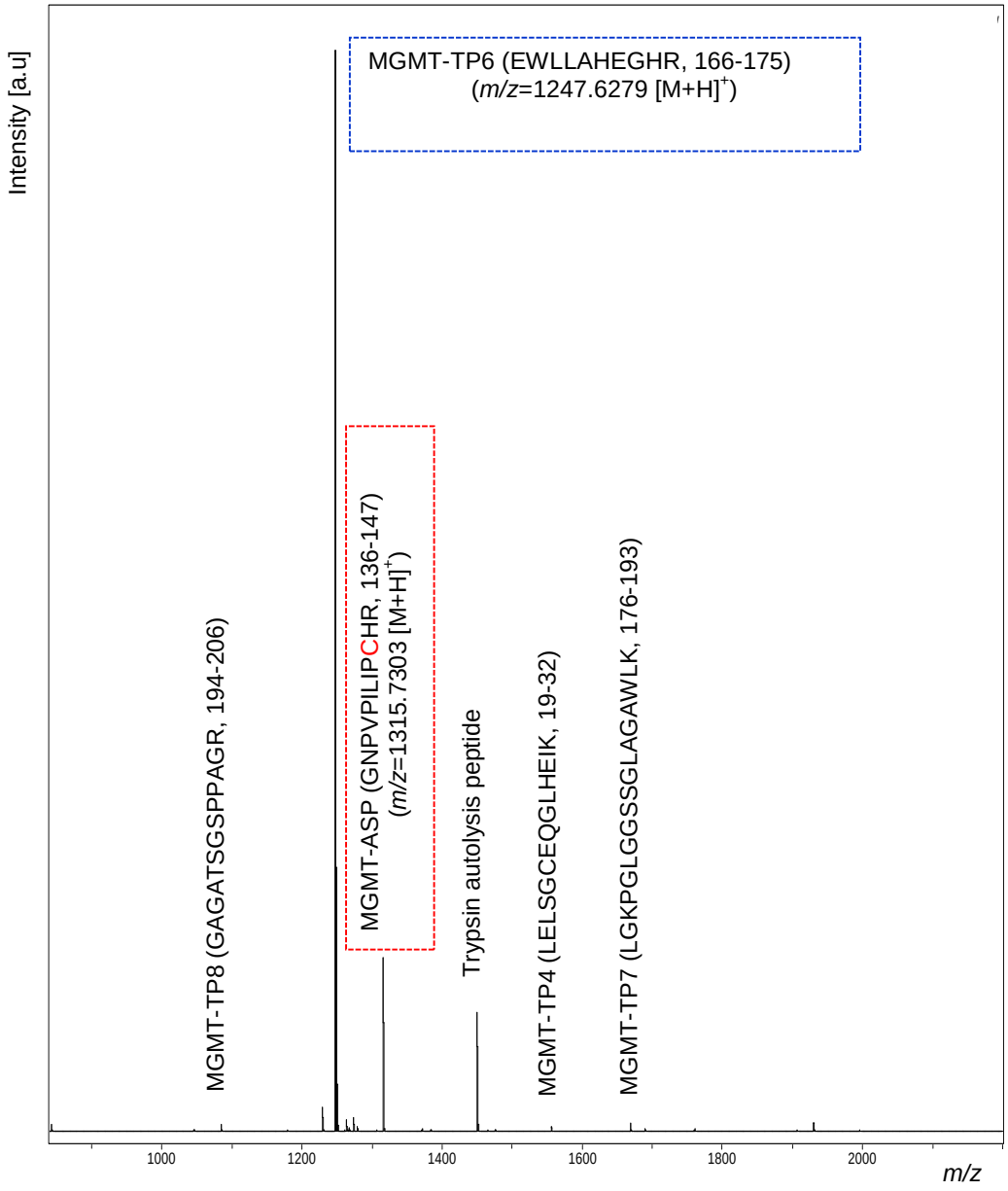


Figure 6.17: MALDI-TOF mass spectra of his-MGMT.

Peak intensities are shown in arbitrary units on the y-axis. Most two prominent ion peaks are MGMT-ASP (in red) and MGMT-TP6 (in blue).

Masses of detected peptides were searched against a human database (UniProtKB/Swiss-Prot; Last modified January 10, 2017). The mascot PMF search of generated tryptic peptides confirmed the identity of MGMT protein due to significant matching between theoretical MGMT_HUMAN peptides and recombinant MGMT tryptic peptide masses with a 92 Mascot score and 49% of MGMT amino acid sequence coverage (102 amino acids sequence coverage out of the total 207; Expect=0.0002). Figure 6.18 shows the MGMT sequence coverage achieved using the on-bead digestion protocol. Of the full sequence (207 amino acids) 76 are undetectable which is attributed to small tryptic peptides in the region of the spectrum dominated by matrix species, or too large to be successfully detected, or inefficient digestion. MALDI-TOF MS analysis achieved 78% amino acid sequence coverage. Identified MGMT peptides are listed in Table 6.1.

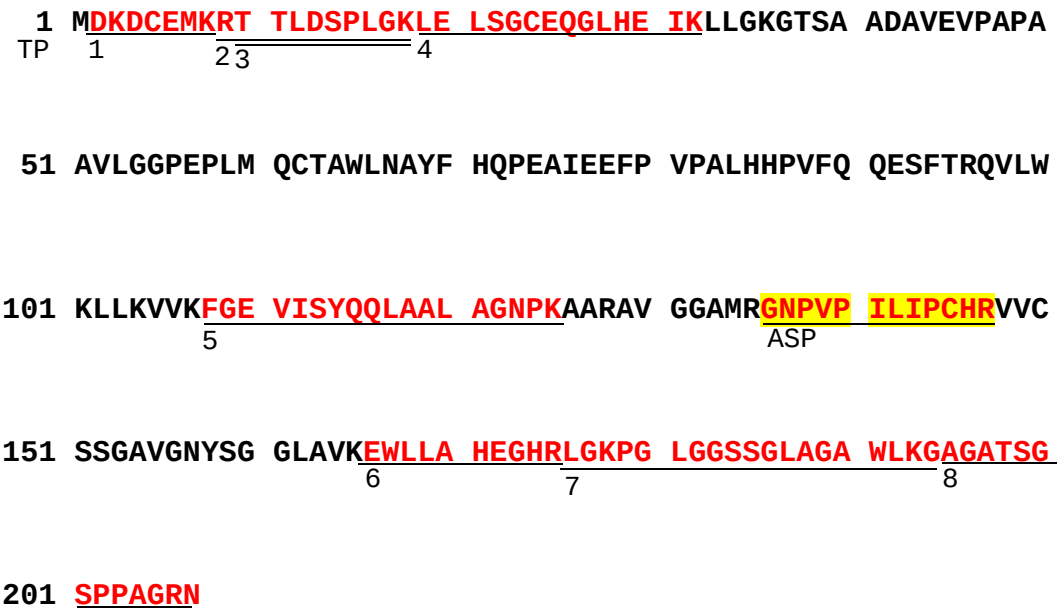


Figure 6.18: MGMT sequences detected using MALDI-TOF MS analysis of his-MGMT tryptic peptides.

Identified peptides are shown in red; highlighted sequence indicates tryptic MGMT-ASP with the active site cysteine residue. 49% sequence coverage (Peptide tolerance 0.5 and allowed missed cleavages, 2).

Table 6.1: MGMT tryptic peptides identified using MALDI-TOF and Mascot database searching following on-bead tryptic digestion of his-MGMT.

Peptide name	Start-End	Measured mass	Expected mass	Error (ppm)	Missed cleavages	Peptide
MGMT-TP1	2-8	868.4295	868.3466	95	1	DKDCEMK
MGMT-TP2	9-18	1087.6010	1087.6033	-2	1	RTTLDSPLGK
MGMT-TP3	10-18	931.4431	931.5022	-63	0	TTLDSPLGK
MGMT-TP4	19-32	1555.8220	1555.7712	32	0	LELSGCEQGLHEIK
MGMT-TP5	108-125	1905.9820	1905.9996	-9	0	FGEVISYQQLAALAGNPK
MGMT-ASP	136-147	1315.7660	1315.7231	33	0	GNPVPILIPCHR
MGMT-TP6	166-175	1247.6650	1247.6207	35	0	EWLLAHEGHR
MGMT-TP7	176-193	1668.9670	1668.9359	19	0	LGKPGLGGSSGLAGAWLK
MGMT-TP8	194-206	1085.5190	1085.5261	-7	0	GAGATSGSPPAGR

6.3.6 Determination of unmodified MGMT-ASP recovery

The next goal was to determine the recovery of unmodified MGMT-ASP following his-MGMT incubation with CT-DNA using internal standard ($C^{13}N^{15}$ proline labelled synthetic unmodified MGMT-ASP (Section 2.2.23.1.) Following incubation of his-MGMT with unmodified CT-DNA, Ni-coated magnetic beads successfully recovered 47.3±0.5 % (n=3) of unmodified MGMT-ASP which was used as an indicator of the protein recovery. Unmodified ASP recovery was calculated using the equation in Section 2.2.23.1. Figure 6.19 shows the region of the MALDI-TOF mass spectra of his-MGMT tryptic digest, following incubation with CT-DNA, indicating both unmodified MGMT-ASP and internal standard ($C^{13}N^{15}$ proline labelled synthetic unmodified MGMT-ASP) ions.

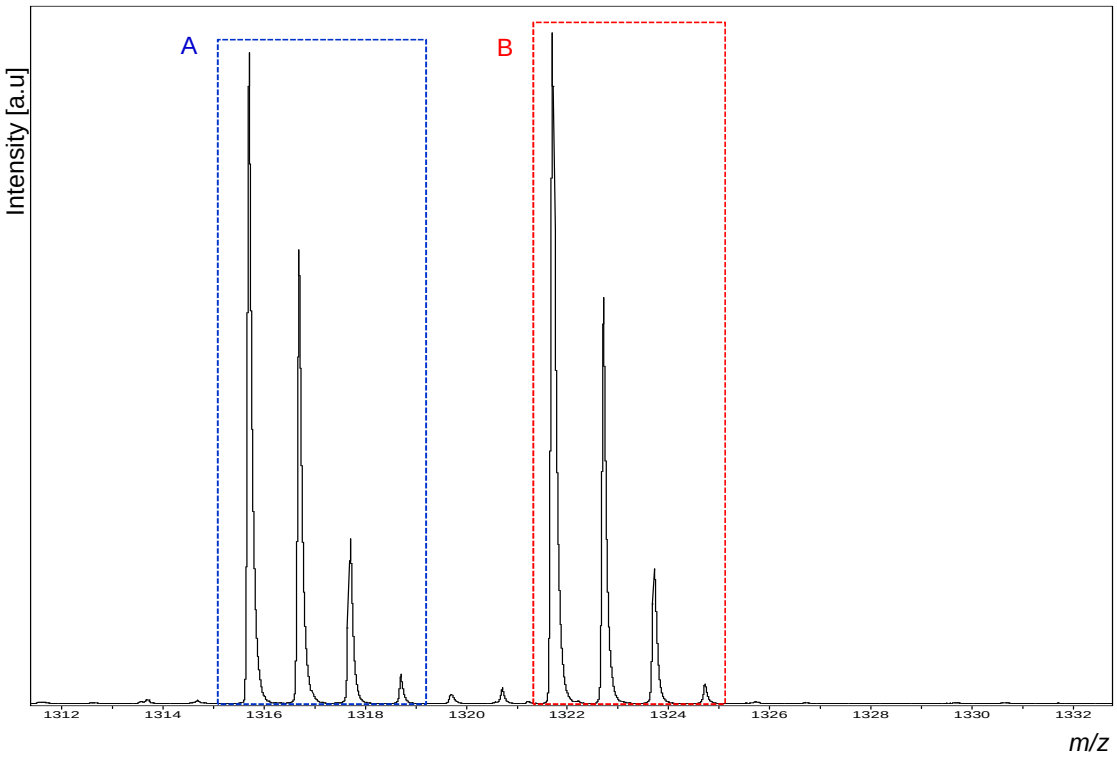


Figure 6.19: Region of mass spectra of his-MGMT incubated with CT-DNA then recovered with Ni-coated magnetic beads indicating the locations of unmodified MGMT-ASP ion and internal standard ion.

A. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147) *m/z*=1315.73. **B**. Synthetic $C^{13}N^{15}$ proline labelled unmodified MGMT-ASP standard (*m/z* 1321.74). Peak intensities are shown in arbitrary units on the y-axis.

In addition, further analysis was conducted comparing protein recovery using PureProteome™ Magnetic Beads and TALON® Magnetic Beads. Figure 6.20 shows detected ions for both unmodified MGMT-ASP and $C^{13}N^{15}$ proline labelled synthetic unmodified MGMT-ASP internal standard using PureProteome™ and TALON® Magnetic Beads.

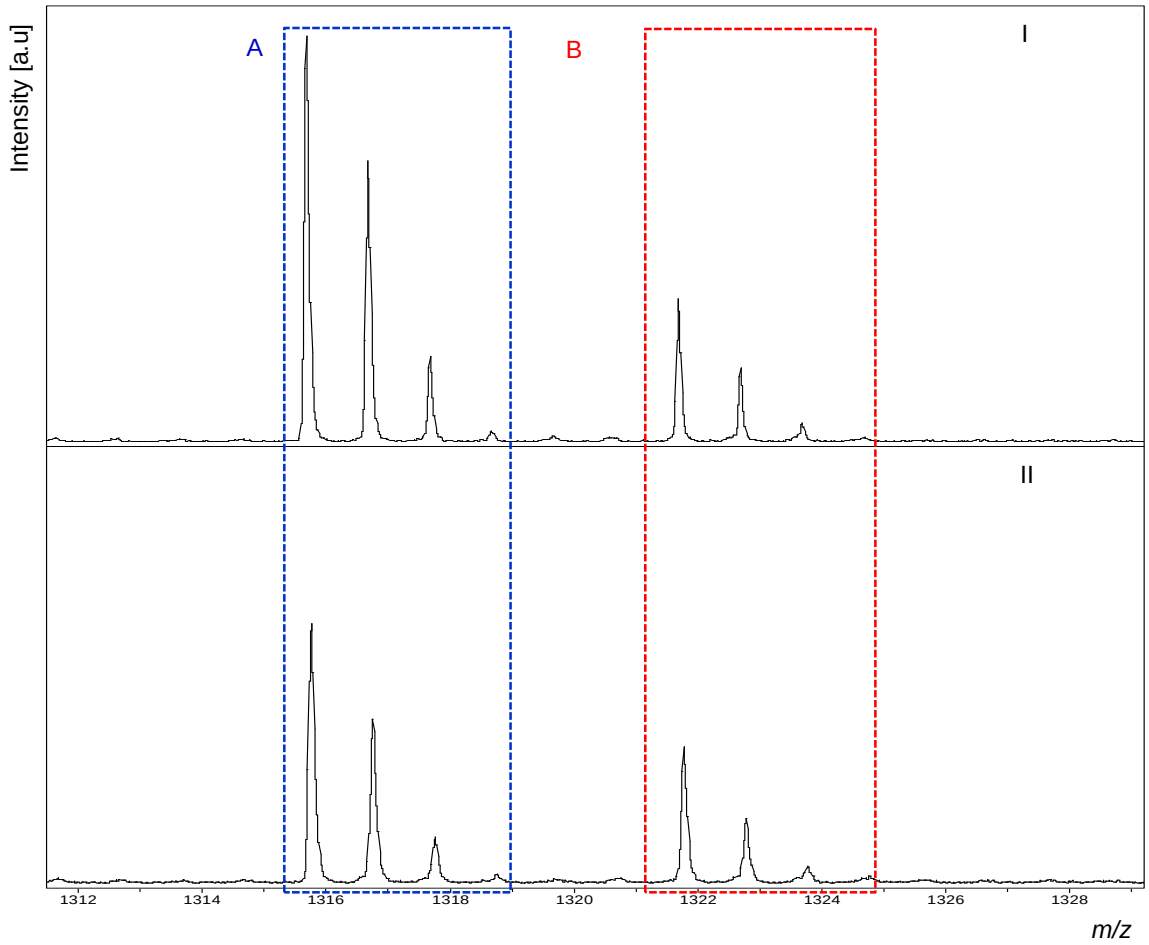


Figure 6.20: Comparing unmodified MGMT-ASP recovery using PureProteome™ and TALON® Magnetic Beads.

His-MGMT (50 pmole) was incubated with CT-DNA (2 mg), recovered by Ni-coated magnetic beads, and treated *in situ* with trypsin. Resulting tryptic peptides were spiked with synthetic $C^{13}N^{15}$ proline labelled unmodified MGMT-ASP (internal standard, 10 pmole). **I**. His-MGMT recovered with PureProteome™ Magnetic Beads. **II**. His-MGMT recovered with TALON® Magnetic Beads. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z* 1315.73) **B**. Synthetic $C^{13}N^{15}$ proline labelled unmodified MGMT-ASP standard (*m/z* 1321.74). Peak intensities are shown in arbitrary units on the y-axis, the intensity scale is the same for both panels I and II.

Table 6.2 shows PAs of MGMT-ASP (PA $_{unmodified}$) and internal STD (PA $_{STD}$) as well as calculated recovery for each tested bead system. %Recovery $_{unmodified}$ was 52.9±6.9 and 47.6±3.8 (mean±SD; n=3) for PureProteome™ Magnetic Beads and TALON® Magnetic Beads, respectively. The difference in recovery percent between the two systems was found to be not statistically significant (P>0.05).

Table 6.2: Comapring %Recovery $_{unmodified}$ of MGMT-ASP by PureProteome™ and TALON® Magnetic Beads following protein incubation with CT-DNA.

Ni coated Beads	PA $_{unmodified}$	PA $_{STD}$	%Recovery $_{unmodified}$
PureProteome™	338.7±19.3	129.0±12.2	52.9±7.0
TALON®	222.3±11.3	93.3±4.0	47.6±3.8

Values are mean+SD (n=3).

6.3.7 Validation of the novel on-bead tryptic digestion approach

6.3.7.1 LOQ of O^6-MeG adduct in TMZ-modified CT-DNA

To test the applicability of the analytical methodology for analysing human DNA samples, we investigated the LOQ of O^6-MeG adduct in TMZ-modified CT-DNA *via* examining the ability of novel MS assay to detect methylated MGMT-ASP in tryptic digest of his-MGMT protein following incubation with TMZ-modified CT-DNA containing various levels of O^6-MeG). His-MGMT was incubated with TMZ-modified CT-DNA containing various levels of O^6-MeG adduct followed by recovery of his-MGMT using Ni-coated beads. On-bead tryptic digestion and desalting of tryptic peptides were conducted as described in Section 2.2.24.1. MS analysis of his-MGMT tryptic peptides demonstrated the detection of methylated MGMT-ASP (GNPVPILIPMe-CHR, 136-147) in MGMT tryptic peptides following protein incubation with methylated CT-DNA and the recorded LOQ for his-MGMT based detection of O^6-MeG adduct in methylated CT-DNA was <0.05 pmole O^6-MeG adduct mg^{-1} methylated CT-DNA. Figure 6.21 shows the region of the mass spectra of his-MGMT incubated with methylated CT-DNA (containing 0.125, 0.25, 0.5 pmole O^6-MeG adducts mg^{-1} methylated CT-DNA) indicating both unmodified as well as methylated MGMT-ASP ions.

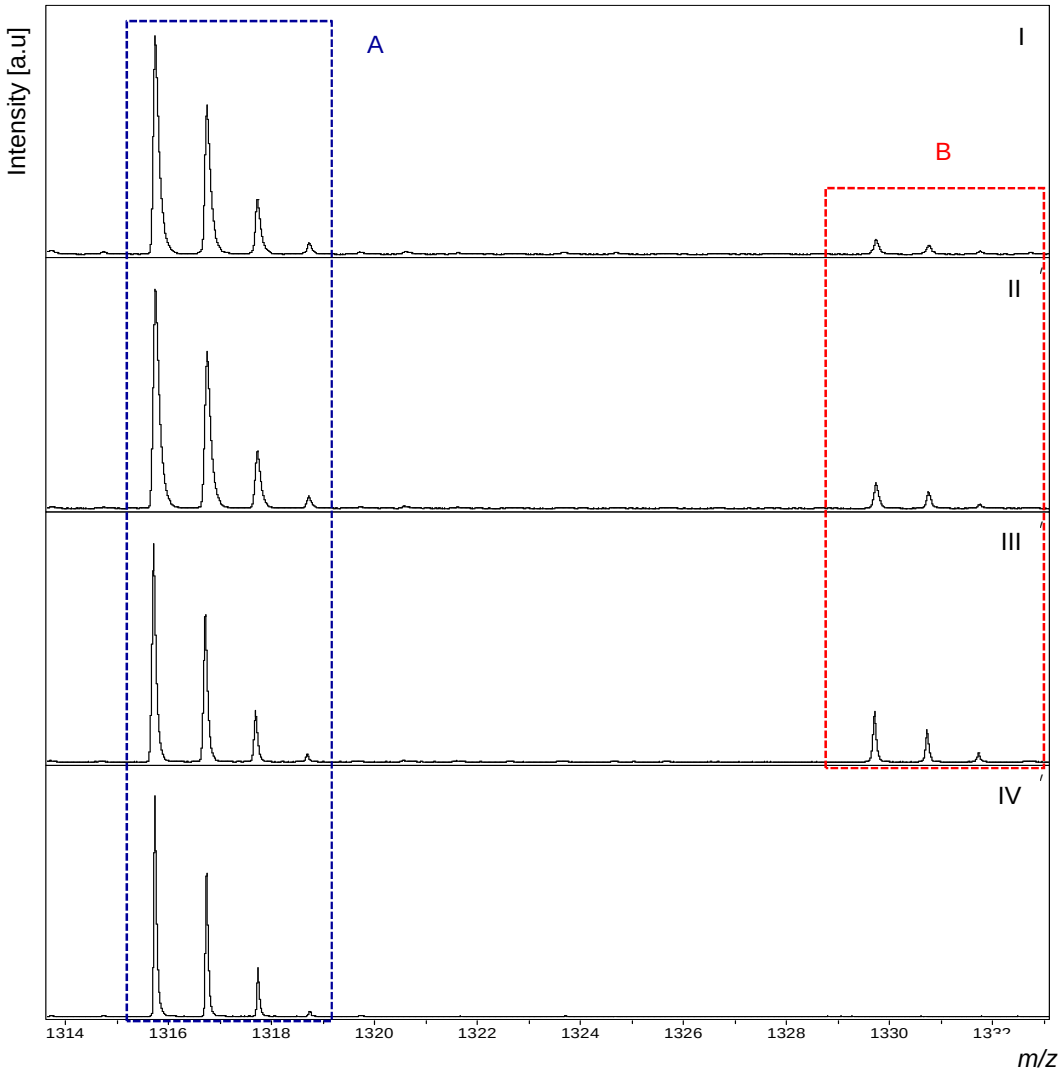


Figure 6.21: Detection of methylated MGMT-ASP in tryptic digests of his-MGMT reacted with TMZ-modified CT-DNA.

His-MGMT (50 pmole) reacted with TMZ-modified CT-DNA (2 mg) that contained (**I**) 0.125 (**II**) 0.25 (**III**) 0.5 pmole O^6-MeG adducts mg^{-1} methylated CT-DNA or with (**IV**) unmodified CT-DNA then protein was recovered by Ni-coated magnetic beads, and treated *in situ* with trypsin. **A**. unmodified synthetic MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z* 1315.7303). **B**. Methylated MGMT-ASP (GNPVPILIPMe-CHR, 136-147) *m/z*=1329.74. Peak intensities are shown in arbitrary units on the y-axis and max peak intensity is the same for all panels.

Moreover, unmodified MGMT-ASP was detected in tryptic peptides of the fusion protein incubated with unmodified CT-DNA and there was no *S*-methylcysteine modified MGMT-ASP (Figure 6.20 IV). Consequently, the observed PAs of methylated MGMT-ASP

were plotted against levels of O^6-MeG adducts indicating a linear correlation but not statistically significant (Figure 6.22).

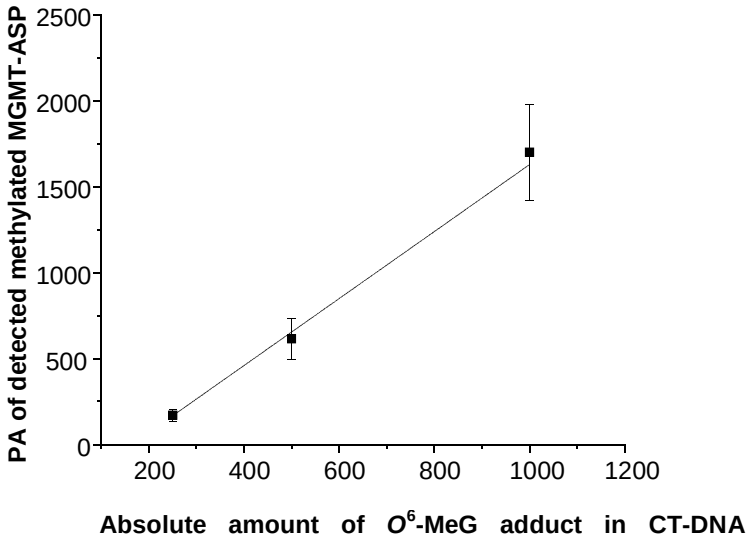


Figure 6.22: Validation results for the MALDI-TOF MS method for detection of methylated MGMT-ASP following his-MGMT incubation with methylated CT-DNA.

His-MGMT (50 pmole) was incubated with TMZ-modified CT-DNA (2 mg) containing 0.25, 0.5 or 1 pmole of O^6-MeG adducts. Resulting mixtures were subjected to on-bead tryptic digestion. Results are expressed as PA of methylated MGMT-ASP *vs.* amount of adduct incubated with MGMT protein. No peaks corresponding to methylated MGMT-ASP (*m/z*=1329.74) were detected in control unmodified CT-DNA with S/N>10.

6.3.7.2 Determination of methylated MGMT-ASP recovery

The optimised approach was verified by considering the recovery of methylated MGMT-ASP (%$Recovery_{methylated}$) following his-MGMT incubation with methylated CT-DNA using the $C^{13}N^{15}$ internal standard. The ratio $PA_{methylated}/PA_{Me\text{-}STD}$ was used to quantify unknown recovered methylated MGMT-ASP (Section 2.2.24.2). Results demonstrated recovery of 39.8±0.9%, 41±3.6% and 48±7.5% (mean±SD; n=3) for methylated MGMT-ASP fragments following his-MGMT incubation with methylated CT-DNA containing 100, 200 and 400 fmole of O^6-MeG adducts, respectively as shown in Figure 6.23 demonstrating a non-statistically significant linear correlation.

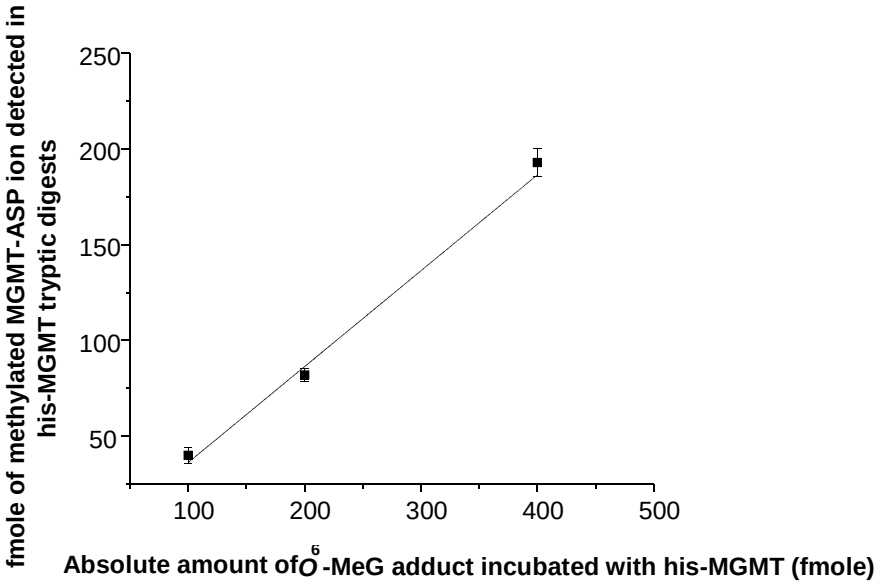


Figure 6.23: Validation results for the quantitative MALDI-TOF MS method for quantification of methylated MGMT-ASP using synthetic $C^{13}N^{15}$ proline labelled methylated ASP internal standard.

His-MGMT (50 pmole) was incubated with TMZ-modified CT-DNA (2 mg) containing 100, 200 or 400 fmole O^6-MeG adducts. Resulting mixtures were subjected to on-bead tryptic digestion; tryptic peptides were spiked with synthetic $C^{13}N^{15}$ proline labelled methylated ASP (internal standard, 250 fmole). Results are expressed as fmole of methylated MGMT-ASP versus amount of adduct incubated with MGMT protein.

6.4 Discussion

Results presented here demonstrated the detection and quantification of O^6-MeG adducts in CT-DNA *via* MS analysis of MGMT tryptic peptides following protein incubation with TMZ-modified CT-DNA. Since MGMT protein has a wide range of O^6-alkylG adducts specificity, this approach could be used to characterise O^6-alkylguanine adducts in human colorectal DNA. The his-MGMT on-bead protocol achieved recovery of 48% and 41% for both unmodified and alkylated MGMT-ASP, *cf.* the MBP-MGMT in-solution digestion protocol (recovery 11%).

Quantitation of proteins and peptides by MS has the potential of lowering cost and time outlays compared to other conventional assays particularly antibody based assays (376). Results here indicate that MBP-MGMT based approaches had major limitations including the inconsistency of protein recovery with amylose coated magnetic beads. In addition, when MBP-MGMT was digested in presence of CT-DNA (in-solution digestion) protein recovery was low. Interferences from MBP-TP13 ion peaks as well as the heavy methylated internal standard ion peak also prevented the quantification of O^6-MeG in DNA. To address this, a his-MGMT on-bead protocol was developed and provided higher recoveries of both unmodified and alkylated MGMT-ASP.

The in-solution digestion protocol (using MBP-MGMT), gave a LOQ for O^6-MeG in CT-DNA of ~1 pmole O^6-MeG mg^{-1} CT-DNA *cf.* 50 fmole O^6-MeG adduct mg^{-1} for on-bead digestion (using his-MGMT). Upadhyaya *et al.,* (203) reported a LOQ for O^6-MeG of 3 fmole mg^{-1} CT-DNA using LC/MS whereas Bussche *et al.*, (253) reported 30 fmole O^6-MeG mg^{-1} CT-DNA using UHPLC-MS/MS. The variations in O^6-MeG LOQs in DNA results between previous studies and the current work is attributable to a variety of reasons. Firstly, different experimental conditions as both previous studies detected O^6-MeG adduct in DNA not *via* MS analysis of MGMT digests. Secondly, previous studies employed different MS approaches either LC/MS or UHPLC-MS/MS.

The his-MGMT approach was based upon the recovery of his-MGMT using Ni-coated magnetic beads from DNA followed by on-bead tryptic digestion. In classical protocols, recovered proteins are eluted from beads by boiling with 1X SDS-PAGE LB then separated on SDS-PAGE followed be staining, destaining, band excision and finally proteins are digested *in situ* in the gels (377). These digestion protocols are laborious and time consuming and furthermore may increase contamination with keratin and other proteins. Additionally, in-gel digestion workflow leads to sample fractionation which could cause loss of analyte that may have a deleterious effect on the identification and quantification of the low abundant O^6-alkylG. All these factors can profoundly affect the quality of the samples and compromise the sensitivity of the protocol. As a consequence, the on-bead digestion strategy was developed to bypass these many steps and provide a valid tool for combined protein elution and digestion which decreased the overall time and handling steps of the experiment. Previously published methods have utilized on-bead digestion protocols coupled with MS analysis for identification of proteins (378–381).

Highest quantitation confidence and measurement precision for peptides in complex samples is usually obtained *via* using a known concentration of synthetic versions of each analyte peptide containing an amino acid labelled with a stable isotope (for example, C^{13} or N^{15}) that are used as internal standards (382,383). The current study used two heavy labelled internal standards which were $C^{13}N^{15}$ proline labelled unmodified MGMT-ASP as well as $C^{13}N^{15}$ proline labelled methylated MGMT-ASP. The latter was used for quantification of methylated MGMT-ASP in tryptic digest of MBP-MGMT protein after reaction with methylated DNA and the former internal standard was used to assess the recovery of unmodified MGMT-ASP in order to investigate the influence of MGMT incubation with DNA, tryptic digestion (either on-bead or in-solution) and sample clean up on recovery. Furthermore, synthetic versions of unlabelled methylated and unmodified MGMT-ASP were obtained to investigate the lower LOQ of these peptides to examine their detectability using MALDI-TOF and to assess the sensitivity of the newly developed approach.

Notably, a linear correlation between O^6-MeG adduct level in methylated CT-DNA and PA of methylated MGMT-ASP ion detected by MALDI-TOF MS analysis was reported using either MBP or his MGMT fusion proteins. This non-statistically significant linear correlation suggests the stoichiometric MGMT repair of O^6-MeG adducts as stated by previous literature studies (275,284).

The presented work explored several approaches or protocols to achieve the lowest possible LOQ for the synthetic peptides that were supplied as lyophilized powder. Results demonstrated poor detection of MGMT-ASP principally because of the improper solubilization of the lyophilized peptide when it was dissolved in basic solvent primarily due to the basic nature of the peptide. Additionally, results displayed NEM non-specific reaction with amino groups within the synthetic peptide forming putative NEM amino adduct at *m/z* 1565.8 which contributed to the poor detection of the NEM capped MGMT-ASP at 1440.7. Maleimides such as NEM are considered fairly specific for SH groups and at pH>7 where reactivity towards primary amines increases (384,385). To remedy that, the primary solvent for all synthetic peptides was changed to 0.1% formic acid in 20 mM Tris-HCl pH 4 which markedly improved peptide solubility achieving lower LOQ ranged between 5-20 fmole for all studied synthetic versions of MGMT-ASP. However, work presented showed that none of the operated protocols succeeded in getting rid of the NEM amino adduct even after either changing the peptide solvent or the pH of the capping reaction or both. Consequently, for both unlabelled and heavy labelled

unmodified MGMT-ASP, it was obligatory to avoid the cysteine capping to ensure quantitation accuracy.

The determination of LOQ of both unmodified and methylated MGMT-ASP either for unlabelled or heavy isotope labelled followed by establishment of the linearity of the measurement of the quantitative MALDI-TOF MS assay established a platform for the identification of O^6-alkylG adducts in DNA as well as the quantification of O^6-MeG adduct using heavy isotope labelled methylated internal standard.

6.5 Conclusion

Proof of concept work has demonstrated the detection and quantification of O^6-MeG adduct *via* MS identification of methylated MGMT-ASP in tryptic peptides of his-MGMT protein reacted with methylated CT-DNA whilst alkylated ASP was not detected following protein incubation with control unmodified CT-DNA. Together with the data shown in previous chapter using O^6-alkylG containing ODNs, this suggests that this newly developed MS approach could be potentially used for combined detection of O^6-MeG adduct as well as a wide range of MGMT repairable O^6-alkylG adducts in human DNA.

7 Chapter 7: Characterization of O^6-alkylG adducts in human colorectal DNA

7.1 Introduction

Colorectal cancer is the third most commonly diagnosed cancer type worldwide (7). The incidence of and mortality from colorectal cancer has been on the rise worldwide among men and women aged younger than 50 years (386). Colorectal carcinogenesis is described as a multistage evolution process that involves both environmental and genetic factors (22,86,387,388). Colorectal cancer genetic alterations have been clearly defined; however, environmental risk factors are still to be fully characterized (389). Genetic factors account for 5-10% of diagnosed colorectal carcinoma cases (388). On the other hand, epidemiologic data highlighted that approximately 30-70% of colorectal cancer cases are attributed to dietary factors (42,390,391).

Colorectal cancer genetic alterations include GC:AT transition mutations in the codons 12 or 13 of K-ras oncogene that are often detected in colorectal tumours (392). O^6-alkylG adducts are known mutagenic and carcinogenic DNA alkyl adducts that, in the absence of repair, can induce GC:AT transition mutations and recombinations in the form of sister chromatid exchanges (212). O^6-alkylG adducts are formed upon human exposure to carcinogenic AAs. Human exposure to AAs is unavoidable and could result from either environmental exposure to pre-formed AAs or *in-vivo* formation of NNOC (60,157,393).

Recent reviews have highlighted a number of dietary factors as colorectal cancer risk factors including the consumption of red and processed meats potentially by heme-catalysed formation of carcinogenic NNOC (58,180,201). NNOC can alkylate colorectal DNA *via* alkylation of O^6-position of guanine base either directly or after their metabolic activation. NNOC formation cannot be accurately estimated from meat consumption alone due to *in situ* processing required for their formation (60).

Furthermore, human intestinal lumen contains a wide range of biogenic amines including polyamines and primary amines (192,394). Body pool of biogenic amines are either due to ingestion of variety of food products (186,395) including fish (396,397), meat (398), cheese (399), wine (400,401) and vegetables (402). The most common dietary biogenic amines are histamine, tyramine, cadaverine, 2-phenylethylamine, spermine,

spermidine, putrescine and tryptamine (189). In addition, biogenic amines are potentially formed endogenously through the microbial decarboxylation of corresponding amino acids (403). Biogenic amines available in the gut lumen are potential precursors for the formation of carcinogenic NNOC. Highly reactive NNOC could be formed by the reaction of amines and/or amides with nitrosating agents subsequently, they could alkylate colorectal DNA (154) contributing significantly to colorectal cancer carcinogenesis. Moreover, both mutagenic O^6-MeG and O^6-CMG adducts formation was reported to arise from nitrosated glycines abundant in human gut lumen (262,404).

Human colorectal DNA contains variable amounts of DNA alkyl adducts generally and O^6-alkylG adducts specifically and these levels are sufficiently high to induce adverse biological effects especially in DNA repair mechanisms deficient cells (185,257,392). Furthermore, three studies detected O^6-MeG adduct in DNA isolated from colorectal and/or gastric tissue samples that were collected from patients in Manchester area (257–259). Isolated DNA samples were digested into nucleosides and fractionated by HPLC then isolated O^6-MeG adduct were quantified by RIA technique using mouse monoclonal α-O^6-MeG antibodies. *In vivo* detection of O^6-CMG adduct was demonstrated in animal studies (197). Similarly, earlier study demonstrated the detection of O^6-CMG adduct in colonic exfoliated cells from healthy individuals who were on high red meat diet *via* immunohistochemistry technique using polyclonal O^6-CMG specific antibodies (185). Both immunohistochemistry and RIA techniques relied on the use of a primary antibody to detect a single specific adduct. Though antibody-based methods are suitable tools for verification, qualified antibodies are generally unavailable especially for new biomarker candidates. No previous techniques provided a valid tool for either detection of the overall load of O^6-alkylG adducts or identification of novel adducts that might be present in human colorectal DNA.

Given the enormous potential for exposure to different AAs, O^6-alkylG adducts (other than O^6-MeG and O^6-CMG) are almost certain to be present. However, and most critically, current methodology does not allow the assessment of overall levels of O^6-alkylG adducts, which could be considered as the potentially pro-mutagenic “load” in colorectal DNA, nor are the identities of adducts other than O^6-MeG and O^6-CMG in colorectal DNA known. This lack of knowledge becomes even more important because both the rate of repair and the mutagenic potential of an O^6-alkylG adduct will depend upon the alkyl group itself, and its local DNA sequence context.

Hence our hypothesis was that colorectal DNA contains a range of hitherto unknown O^6-alkylG adducts and that colorectal cancer risk will depend upon both the type and relative levels of these. Identifying and quantifying these adducts and establishing their mutagenic potential and ease of repair will comprehensively address the role of endogenous nitrosation in human cancer.

7.2 Aims and objectives

The overall aim of work in this chapter is to employ the newly developed MS based assay to assess NNOC exposure by identifying then quantifying O^6-alkylG DNA adducts that result from their exposure in human colorectal DNA, using the known action of the DNA repair protein, MGMT, to irreversibly transfer the alkyl group from the O^6 position of the modified guanine bases to an active site cysteine residue in MGMT.

Specific aims and objectives of are as follows:

1. To determine to what extent there is a range of O^6-alkylG adducts present in human colorectal DNA.

2. To quantify O^6-MeG adducts present in human colorectal DNA using synthetic $C^{13}N^{15}$ labelled methylated MGMT-ASP; internal standard.

3. To determine to what extent known O^6-alkylG adducts (i.e O^6-MeG and O^6-CMG) contribute to total O^6-alkylG adduct levels in human colorectal DNA.

7.3 Results

7.3.1 Analysis of human colorectal DNA samples

7.3.1.1 Demographic characteristics

The current study investigated 13 human colorectal DNA samples (10 tumour samples, 3 normal samples, 3 paired normal and tumour DNA samples) that were

collected from 10 participants of which 6 were men, 2 were women, whilst 2 participants' demographic data were not available. The mean (±SD) age for 8 individuals was 69±14. All study participants were presenting with colorectal carcinoma at hospitals with in Manchester area (caecal carcinoma (2/8; 25%), sigmoid carcinoma (2/8; 25%) and rectal carcinoma (4/8, 50%)); as shown in Table 7.1. In addition, 7 of them were non-smokers (87.5%) and 1 was current smoker (19 cigarettes per day, 12.5%). Regarding, alcohol drinking habits, 4 were non-drinkers, 1 was an occasional drinker and 3 were drinkers (1 consumed <10 units/week; 2 consumed >10 units/week); as shown in Appendix 3. Information regarding participants' medication usage, at the time of sample collection, was obtained from 8 participants as summarized in Appendix 3.

Table 7.1: Demographic characteristics of study participants indicating age, and tumour location.

Human colorectal DNA sample ID	Age	Tumour
25	73	Caecal
31	80	Caecal
39	77	Sigmoid
44	81	Rectal
50	75	Rectal
74	47	Sigmoid
78	73	Rectal
79	43	Rectal

Data are adapted from (256).

7.3.1.2 MGMT activity

MGMT activity was evaluated by a previous study in both normal and tumour colorectal tissues of the same patient and data was available for 7 participants only; as indicated in Table 7.2. Results shows that MGMT activity in tumour tissue (123±29; mean±SD n=7) was statistically significantly higher than that of normal colorectal tissue (90±10; mean±SD n=7) ($P<0.01$).

Table 7.2: Results of MGMT activity as well as K-ras mutations.

Human colorectal DNA sample ID	MGMT activity in normal tissue	MGMT activity in tumour tissue	K-ras normal	K-ras tumour
25	N/A	N/A	Normal	codon 12C/T
31	76	183	Normal	codon 12C/T
39	85	107	Normal	codon 13A
44	95	133	codon 12A	Normal
50	105	103	Normal	Normal
74	81	110	codon 13A	codon 12C/T
78	102	100	Normal	Normal
79	87	128	Normal	codon 12A

N/A denotes data were not available. Data are adapted from (256).

7.3.1.3 K-ras gene mutations

K-ras codon 12 or 13 mutations were investigated by a previous study where their results revealed that, for the same 8 patients, K-ras mutations were detected in 2 out of 8 normal colorectal tissue samples (20%) and 5 out of 8 tumour samples (63%) (Table 7.2).

7.3.1.4 Identification of O^6-alkylG adducts in human colorectal DNA

MALDI-TOF based approach was used for the identification of O^6-alkylG adducts in human colorectal DNA. His-MGMT fusion protein was incubated with human colorectal DNA (either extracted from normal or tumour tissue) then alkylated peptides were processed for MALDI-TOF MS analysis (Section 2.2.25.3).

Mass spectra of tryptic digests of MGMT incubated with human colorectal DNA samples detected multiple MGMT peptides including unmodified MGMT-ASP as well as other ions those ions were believed to correspond to alkylated MGMT active site fragments. This proposal was based upon three main points. Firstly, those ions were not detected in the control sample where MGMT was incubated with unmodified CT-DNA;

second, their *m/z* did not correspond to either any of MGMT tryptic peptides or a known contaminant (such as keratin peptides). Finally, *m/z* of some of those ions, but not all, corresponded to the transfer of known alkyl groups to MGMT-ASP. Putative alkylated MGMT-ASP ions were classified as definite, probable or possible alkylated MGMT-ASPs or even unknown as explained in Table 2.52 and this classification was based on the data attained in previous chapters of the current study regarding the proposed alkyl adduct.

Putative alkylated MGMT-ASP ions *m/z* $[M+H]^+$ were as follows: 1329.7, 1359.7, 1373.7, 1386.7, 1410.7, 1459.7, 1461.7, 1477.7, 1530.7, 1546.7, 1562.7. The *m/z* of putative alkylated MGMT-ASP ions and the frequency of their occurrence across all the investigated (13) human colorectal DNA samples were summarized in Table 7.3. S/N for all detected modified alkylated MGMT-ASP were greater than 10. In addition, our analysis did not detect any alkylated MGMT-ASP ions in control sample (his-MGMT incubated with unmodified CT-DNA).

Among detected putative alkylated MGMT-ASPs ions, probable methylated MGMT-ASP ions (*S*-methylcysteine *m/z* 1329.74 $[M+H]^+$; recorded mass shift 14 Da corresponds to methyl group transfer to MGMT-ASP cysteine residue) were detected in tryptic digests of his-MGMT following incubation with 12 human colorectal DNA samples. These results demonstrated that probable O^6-MeG was detected in 2 out of a total of 3 (66%) normal DNA samples and in all of investigated tumour DNA samples (100%; Table 7.3).

Furthermore, probable carboxymethylated MGMT-ASP fragments (*S*-carboxymethylcysteine *m/z* 1373.73 $[M+H]^+$; recorded mass shift 58 Da corresponds to carboxymethyl group transfer to MGMT-ASP cysteine residue) were detected in tryptic digests of his-MGMT following incubation with 8 human colorectal tumour DNA samples. These results highlighted that probable O^6-CMG adduct was detected in 8 out of a total of 10 (80%; Table 7.3) tumour DNA samples. However, probable O^6-CMG adducts were not detected in any of investigated normal DNA samples since *S*-carboxymethylcysteine modified ASP was not detected in tryptic digests of his-MGMT incubated with any of normal DNA samples. In addition, the novel MALDI-TOF based approach successfully detected an ion at *m/z* 1386.7 $[M+H]^+$ where the recorded mass shift 71 Da corresponds to the transfer of carboxamidoethyl (propionamide) group to the cysteine residue of MGMT-ASP and it was classified as a probable O^6-carboxamidoethylG adduct and it was

detected in 2 out of a total of 10 (20%) tumour DNA samples. This novel adduct was not detected in DNA of any of normal human colorectal DNA samples.

Figure 7.1 showed MS spectra for his-MGMT reacted with 2 human colorectal tumour DNA samples (96T and 25T) as well as control his-MGMT (incubated with unmodified CT-DNA) indicating the detected probable alkylated MGMT-ASPs (methylated, carboxymethylated and carboxamidoethylated). Furthermore, two other ions were detected in digest of MGMT incubated with human colorectal DNA at *m/z* 1359.7 and 1461.7 $[M+H]^+$ which were proposed to be assigned to the transfer of hydroxyethyl and pyridyoxobutyl alkyl group, respectively.

7.3.2 MGMT repair of O^6-HOEtG and O^6-pobG adducts

In order to help to confirm the identity of the two ions at *m/z* 1359.7 and 1461.7 $[M+H]^+$, MGMT was incubated with the corresponding synthetic DS O^6-alkylG ODNs, namely, O^6-HOEtG and O^6-pobG ODNs. Mass spectra of MGMT digests following incubation with O^6-OHEthG and O^6-pobG ODNs revealed the detection of 2 ions at *m/z* 1359.7 and 1462.7 $[M^+ + H^+]$, respectively (Figure 7.2 and 7.3). This result lends further support to the assignment of the ion detected at 1359.7 to the transfer of hydroxyethyl group to MGMT. However, ion detected at *m/z* 1461.7 $[M^+ + H^+]$, following MGMT incubation with human colorectal DNA, is approximately 1 Da less than pyridyloxobutylated MGMT-ASP that was detected at *m/z* 1462.7 $[M^+ + H^+]$; as shown in Figure 7.3. Based upon above result, it could be concluded that possible O^6-HOEtG adduct was detected in one human tumour DNA sample as MALDI-TOF MS analysis of tryptic digests of his-MGMT incubated with tumour DNA sample detected *S*-hydroxyethyl modified MGMT-ASP (*m/z* 1359.7) $[M^+ + H^+]$.

A further seven putative alkylated MGMT-ASPs were detected: of which one ion was found in 5 samples (*m/z* 1410.7 $[M^+ + H^+]$), 1 was found in 4 DNA samples (*m/z* 1461.7 $[M^+ + H^+]$), 1 in 3 (*m/z* 1477.7 $[M^+ + H^+]$), 1 in 2 (*m/z* 1562.7 $[M^+ + H^+]$) and 3 in just one DNA sample (*m/z* 1459.7, 1530.7 and 1546.7 $[M^+ + H^+]$).

Table 7.3: Putative alkylated MGMT-ASP ions detected in tryptic digests of his-MGMT reacted with human DNA that has been extracted from either macroscopically normal or malignant tissue samples from the colon and rectum.

Human colorectal DNA sample ID	Tissue	Putative alkylated MGMT-ASP ions detected in tryptic digests of his-MGMT incubated with human colorectal DNA *m/z*										
		1329.7 Pr-Me	1359.7 Po-HOEt	1373.7 Pro-CM	1386.7 Po- CEt	1410.7 U	1459.7 U	1461.7 U	1477.7 U	1530.7 U	1546.7 U	1562.7 U
19N	Control	+	-	-	-	-	-	-	-	-	-	-
19T	Tumour	+	-	+	-	+	-	+	+	-	-	-
25T	Tumour	+	-	+	-	+	-	+	-	-	-	-
31T	Tumour	+	-	+	-	-	-	-	-	-	-	-
39N	Control	-	-	-	-	-	-	-	-	-	-	-
39T	Tumour	+	-	+	-	-	+	-	+	+	+	+
44N	Control	+	-	-	-	-	-	-	-	-	-	-
44T	Tumour	+	-	-	-	-	-	-	+	-	-	+
50T	Tumour	+	-	+	+	-	-	-	-	-	-	-
74T	Tumour	+	-	+	-	-	-	-	-	-	-	-
78T	Tumour	+	+	-	-	+	-	-	-	-	-	-
79T	Tumour	+	-	+	-	+	-	+	-	-	-	-
96T	Tumour	+	-	+	+	+	-	+	-	-	-	-

(+): putative alkylated MGMT-ASP was detected with S/N ratio >10; (-): putative alkylated MGMT-ASP was not detected (S/N ratio <10). (N) DNA was extracted from macroscopically normal human colorectal tissue taken approximately 5 cm form tumour edge. (T) DNA was extracted from human colorectal tumour tissue. Pr-Me and Pr-CM denote probable methylated and carboxymethylated ASP ions, respectively. Po-HOEt and Po- CEt denote possible hydroxyethylated and carboxamidoethylated ASP ions, respectively. U denotes unknown or unidentified ions.

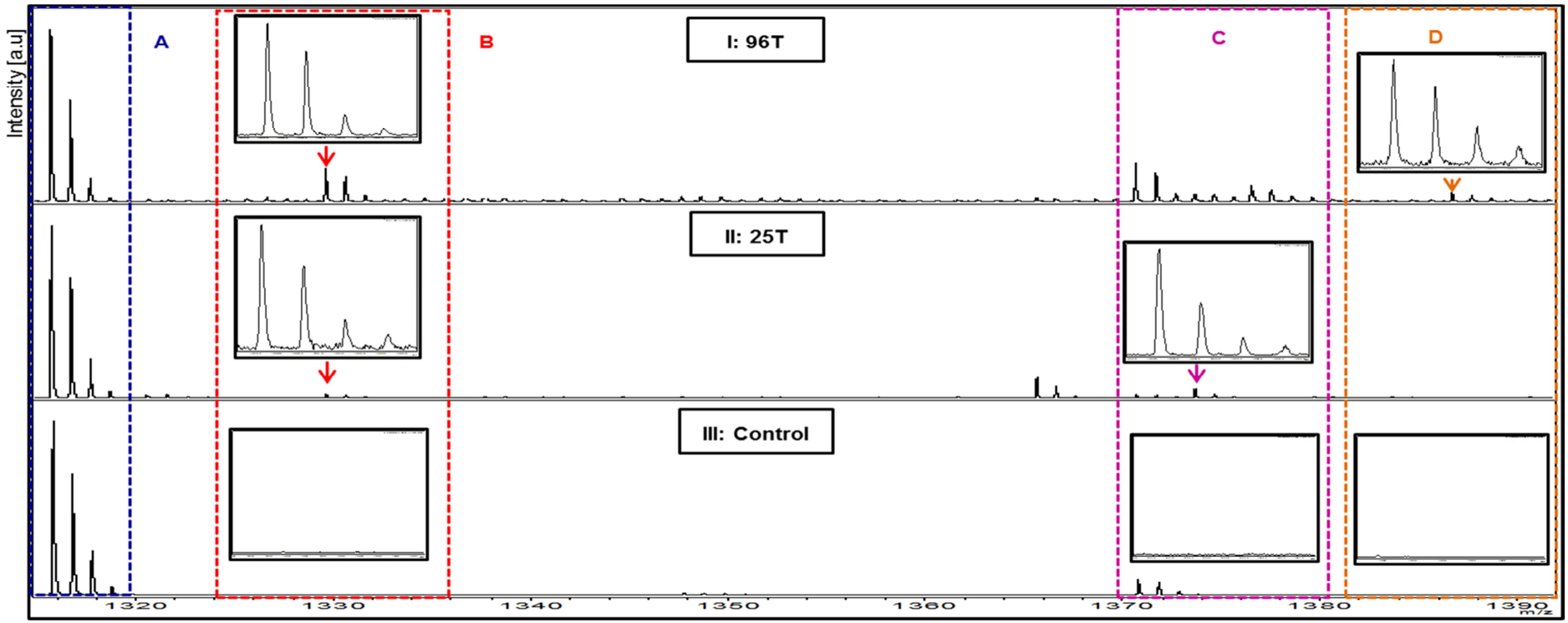


Figure 7.1: Mass spectra of his-MGMT reacted with human colorectal DNA.

His-MGMT (50 pmole) incubated with 2 mg of human colorectal DNA (I) 96T; (II) 25T and (III) unmodified CT-DNA. **A**. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z* 1315.73). **B**. Probable methylated MGMT-ASP (GNPVPILIPMe-CHR, 136-147; *m/z* 1329.74.) **C**. Probable carboxymethylated MGMT-ASP (GNPVPILIPCM-CHR, 136-147) *m/z* 1373.73. **D**. Probable carboxamidoethylated MGMT-ASP (GNPVPILIP carboxamidoethyl -CHR, 136-147; *m/z* 1386.77). Peak intensities are shown in arbitrary units on the y-axis and intensity scale is the same for all panels. S/N >10 for all detected alkylated MGMT.

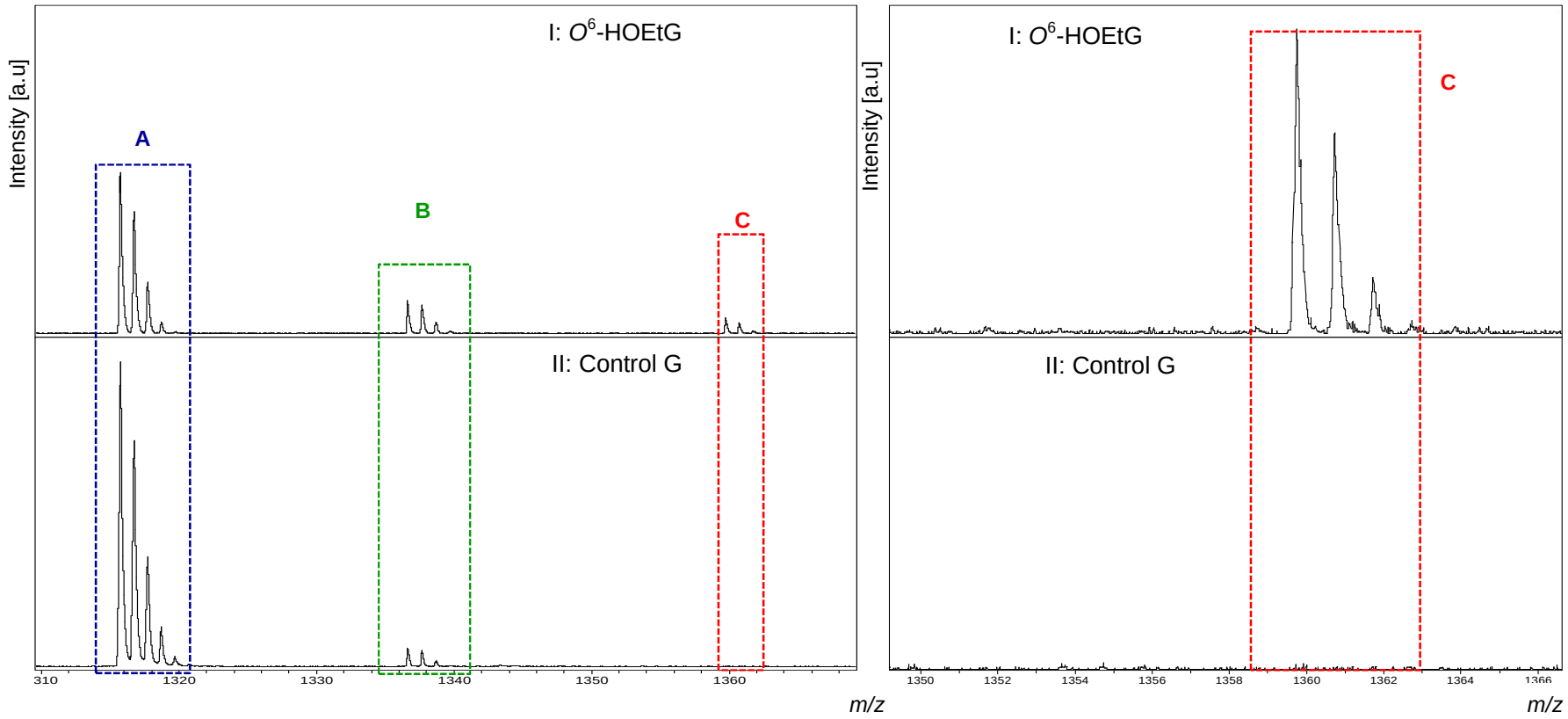


Figure 7.2: Mass spectra of MGMT reacted with (I) DS 23 mer O^6-HOEtG adduct containing ODN and (II) 23 mer control G ODN.

A. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z* 1315.7). **B**. MBP-TP13 (SYEEELAKDPR, 332-342; *m/z* 1336.54. **C**. Hydroxyethylated MGMT-ASP (GNPVPILIPHOEt-CHR, 136-147; *m/z* 1359.7). Peak intensities are shown in arbitrary units on the y-axis and the intensity scale is the same for panels I & II. Right panel is zoomed into left panel.

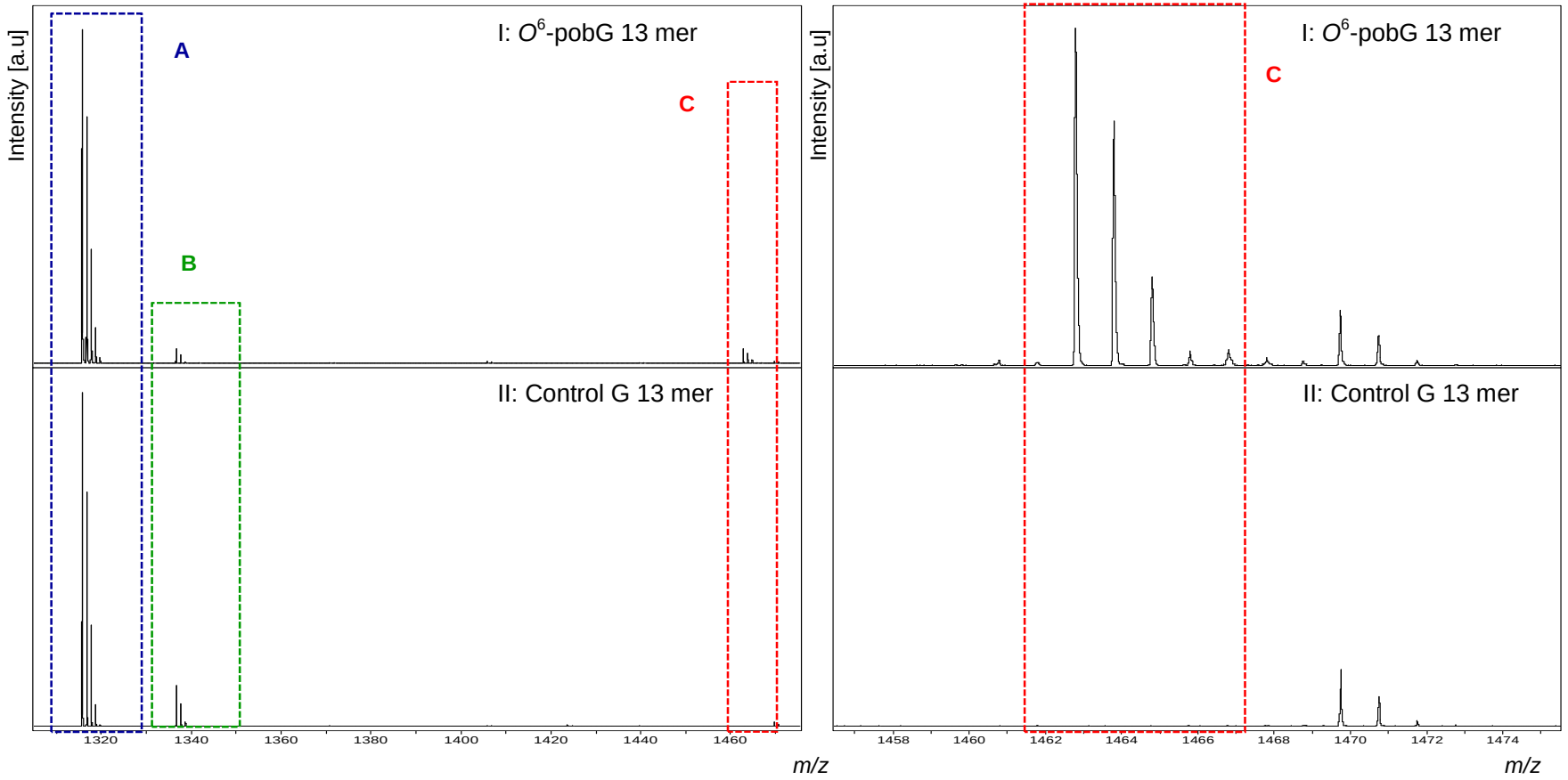


Figure 7.3: Mass spectra of MGMT reacted with (I) DS 13 mer O^6-pob adduct containing ODN and (II) 13 mer control G ODN (II).

A. Unmodified MGMT-ASP (GNPVPILIPCHR, 136-147; *m/z* 1315.7). **B**. MBP-TP13 (SYEEELAKDPR, 332-342; *m/z* 1336.54. **C**. Pyridyloxobutylated MGMT-ASP (GNPVPILIPpob-CHR, 136-147; m/z 1462.7). Peak intensities are shown in arbitrary units on the y-axis and the intensity scale is the same for panels I & II. Right panel is zoomed into left panel.

7.3.3 Quantification of O^6-MeG and O^6-CMG adducts in human colorectal DNA

Following identification of O^6-alkylG adducts present in human colorectal DNA samples, MGMT tryptic digests were spiked with $C^{13}N^{15}$ proline labelled methylated MGMT-ASP (internal standard). O^6-MeG adduct levels in colorectal DNA samples were quantified *via* the quantification of the methylated MGMT-ASP (Section 2.2.25.3). O^6-MeG adduct was present in human colorectal DNA at concentrations of 6.7±0.3-11.1±0.8 and 19.2±0.9-78.2±7.9 nanomole O^6-MeG/mole dG for normal and tumour DNA, respectively. Table 7.4 shows concentrations of O^6-MeG adducts detected in human colorectal DNA that have been determined by the current study (using his-MGMT and MALDI-TOF approach) as well as a previous published study (using RIA technique) (256).

Table 7.4: Comparison of Concentrations of O^6-MeG adducts detected in DNA by the current study and a previous published study.

Human colorectal DNA sample ID	Nanomole of O^6-MeG adduct per mole of dG mean±SD (n=3)			
	Normal Tissue (N)[a]	Tumour tissue (T)[a]	RIA (N)[b]	RIA (T)[b]
19	6.7±0.3	20.7±1.8	N/A	N/A
25	N/A	19.2±0.9	N/A	25
31	N/A	20.2±1.3	N/A	20
44	11.1±0.8	21.8±1.4	5	5
50	N/A	47.5±6.9	N/A	48
74	N/A	21.4±0.8	N/A	26
79	N/A	19.6±0.9	N/A	19
96	N/A	78.2±7.9	N/A	N/A

[a]Data of the current study (adduct was quantified by his-MGMT and MALDI-TOF technique). [b]Data of previous published study (256) (adduct was quantified by RIA). (N) DNA was extracted from macroscopically normal human colorectal tissue taken ~5 cm from the tumour edge. (T) DNA was extracted from human colorectal tumour tissue. N/A denotes data are not available. dG denotes 2'deoxyguanosine.

The levels of O^6-MeG adduct in tumour DNA were significantly higher than those in normal DNA from the same individual for the two analysed pairs (P<0.01) demonstrating significantly higher O^6-MeG adduct in tumour colorectal DNA compared to the normal. The ratio of O^6-MeG adduct in normal DNA to that in tumour DNA was 0.33 and 0.51 for the two investigated pairs 19C&T and 44C&T, respectively.

In addition, label free quantitation of O^6-CMG adducts was carried out as explained in Section 2.2.25.3. Levels of O^6-CMG in human colorectal tumour DNA ranged between 21.53±2.32 and 68.24±5.71 nanomole O^6-CMG/mole dG; as shown in Table 7.5. Further statistical analysis revealed no association between O^6-MeG and O^6-CMG levels in human colorectal DNA (P=0.93). Ratio of levels of O^6-CMG (nanomole of per mole of dG) to that of O^6-MeG was variable among various samples and ranged between 0.44±0.04 and 3.30±0.28.

Table 7.5: Levels of O^6-CMG detected in human colorectal tumour DNA.

Human colorectal DNA sample ID	Nanomole of O^6-CMG adduct per mole of dG (mean±SD (n=3))	Nanomole of O^6-CMG per mole of dG /nanomole of O^6-MeG per mole of dG (mean±SD (n=3))
19T	68.24±5.71	3.30±0.28
25T	26.44±2.75	1.38±0.14
31T	40.96±2.22	2.03±0.11
50T	48.83±3.10	1.03±0.07
74T	21.53±2.32	1.01±0.11
79T	26.81±2.17	1.37±0.11
96T	34.46±2.74	0.44±0.04

7.3.4 Association between O^6-MeG adduct level and MGMT activity or K-ras mutations

There was no significant relationship between O^6-MeG adduct level and either MGMT activity or K-ras mutations (P=0.7 and 0.5 respectively).

7.4 Discussion

Results presented in the current chapter demonstrated successful development of a sensitive and reproducible assay for both identification and quantification of O^6-alkylG adducts in human colorectal DNA samples *via* MALDI-TOF MS analysis of MGMT tryptic peptides following protein incubation with human colorectal DNA samples. These results shows that this novel approach can detect both known and as yet unidentified O^6-alkylGs

and suggests that human colorectal DNA contains a range of O^6-alkylG adducts that probably constitute only a fraction of the DNA alkyl adductome.

Probable O^6-MeG adduct was detected in all 10 tumour colorectal DNA samples and in two out of three of the analyzed normal colorectal DNA samples pointing out 100% and 66% methylation frequencies for tumour and normal colorectal DNA samples, respectively. Results of the current study are consistent with a previous study by Jackson *et al*., (259) who reported higher methylation frequency in tumour DNA samples (32%-62%) compared to normal DNA samples (33%-48%). Furthermore, another study conducted in 2000 in Manchester area reported higher methylation frequency for DNA isolated from tumour tissue (52%) compared to 43% for normal DNA samples (257). Conversely, a published study reported that colorectal tumour DNA was less methylated than DNA from adjacent mucosa with methylation frequencies of 14% and 80%, respectively (258). The reported methylation frequency for DNA isolated from tumour samples in the present study is higher than that for any of other previous studies. A number of different causes could account for this variation in methylation frequency between the current study and previous published studies. Among them the most important is the size of investigated DNA population or number of examined samples since both previous studies as well as the current work were based on relatively small number of patient samples. The current study examined samples that were collected from a smaller number of individuals (13 total samples, 3 normal and 10 tumour; collected from 10 participants only) compared to other studies who examined larger number of DNA samples that ranged from 58-62 normal samples and 55-58 tumour samples (257,259).

Furthermore, probable O^6-CMG adduct was detected in 8 out of a total of 10 (80%) tumour DNA samples. However, O^6-CMG adduct was not detected in any of the examined normal DNA samples. The present work is the first detecting O^6-CMG adduct in human colorectal cancer DNA. However, a previous study reported *in vivo* detection of O^6-CMG adduct in human blood DNA samples (262). Current results accord well with a previous study that documented *in vivo* detection of O^6-CMG adduct in DNA from exfoliated colonic cells collected from stools of human volunteers who consumed high red meet diet for 15 days (185). In addition, a previous research reported the detection of trace levels of O^6-CMG adduct in exfoliated human colonocytes where the obtained low intensity signal was explained by the fact that, volunteers were on a normal not high red meat diet (253). Consistently, *in vivo* O^6-CMG adduct detection was reported in gastric biopsies and in intact small intestine from rats that were treated with N-nitrosopeptide N-acetyl-N^1-prolyl-

N^1-nitrosoglycine and in HT-29 cells treated with diazoacetate demonstrating that O^6-CMG adduct is a potential product of colorectal DNA alkylation damage (185,197). The inability of the current work to detect O^6-CMG adduct in normal human colorectal DNA does not confirm the absence of this adduct from normal colorectal DNA, as it might be present, but in levels below its limit of detection. In addition, the current study has investigated only three normal colorectal DNA samples, thus, there is not enough evidence to make a clear conclusion regarding the detection of this adduct in human normal colorectal DNA. Therefore, more samples of human normal colorectal DNA have to be examined to provide further support to the initial findings of the current study.

In addition to the detection of both probable O^6-MeG and O^6-CMG adducts, present work demonstrated detection of two O^6-alkylG adducts, for the first time in human colorectal tumour DNA. Firstly, possible O^6-HOEtG adduct was detected in a single tumour DNA sample whilst probable O^6-carboxamidoethylG adduct was detected in two tumour DNA samples. No previous researches have detected these two adducts in human colorectal tumour DNA samples. More O^6-alkylG adducts are believed to be present in human colorectal DNA since the present study has detected various peptide ions whose *m/z* ratios did not match any of MGMT tryptic peptides and they were totally absent in control samples. Consistently, unidentified peptide ions are potentially resulting from alkylation of MGMT-ASP.

Current work mainly relied on MALDI-TOF MS analysis for detection of alkylated MGMT-ASPs and the evidence of the detection of the alkylated peptides following MGMT incubation with human colorectal DNA was based on the following three criteria: first, *m/z* of detected ions, second, recorded change in the molecular weight of the ASP (observed mass shift) and the third one is S/N>10. Therefore, further tandem MS analysis is required to confirm identities of detected peptides that were assigned to alkyl group transfer to MGMT, plus, to provide information about the structural identities of detected but unidentified ions.

A number of different mechanisms acting either singly or in combination could account for the detection of O^6-alkylG adducts in human colorectal DNA. Among these, haem-catalysed endogenous formation of highly reactive NNOCs that could alkylate DNA in an electrophilic attack which are formed upon nitrosyl haem mediated endogenous nitrosation of either dietary amines or amides (180,374). Furthermore, DNA alkylation is potentially a result of a nitrosation reaction that is mediated by bacteria within the colon or

by macrophages within chronically inflamed sites (393,405,406). In addition, exogenous exposure to AAs *via* dietary, environmental or occupational sources may contribute to the overall load of O^6-alkylG adducts. By analogy, O^6-CMG adduct potentially arises from the nitrosation of bile component glycocholic acid that could also alkylate DNA (208).

Both O^6-MeG and O^6-CMG are believed to be products of nitrosated glycine. Therefore, the detection of these two adducts in human colorectal DNA supports the hypothesis that nitrosation of glycine occurs in human GI tract contributing particularly to colorectal cancer risk. In addition, data of the current study highlighted that O^6-MeG and O^6-CMG are the most commonly occurring adducts in human colorectal DNA samples, compared to the other two detected adducts O^6-carboxamidoethylG and O^6-HOEtG. This could be related to their glycine precursor. Glycine has a unique chemical structure in terms of having no substituents on α-carbon and that appears to significantly increase DNA damaging potential of the nitrosated amino acid (407). Meanwhile, detection of novel adducts in human colorectal tumour DNA lends further support to the hypothesis that amino acids other than glycine are potentially being nitrosated in human GI tract and their nitrosation products might contribute to colorectal cancer risk.

Nevertheless, further mechanisms might contribute to O^6-HOEtG adduct formation, such as exogenous exposure to the known human carcinogen ethylene oxide (408) which is a common environmental carcinogen, industrial chemical and a component of cigarette smoke (409). *In vivo* detection of O^6-HOEtG adduct has not been reported by previous studies, whilst, N^7-HOEtG adduct was detected in human blood samples as well as human lymphocytes of both smokers and non-smokers (410,411). It's well established that N^7-alkylG adducts are more predominant DNA lesions than O^6-alkylG adducts where the former constitutes 75% of the alkylation products whilst the latter represents only 6-8% of alkylated DNA adducts. This could be used as an explanation for the detection of the N^7-HOEtG prior to its corresponding O^6-alkylG adduct as N^7-HOEtG adduct is far more easier to detect (156).

Regarding the probable O^6-carboxamidoethylG adduct, the alkyl group at the O^6 position of the modified guanine base is carboxamidoethyl group; also known as propionamide a by-product of acrylamide monomer; that is an environmental carcinogen of numerous industrial and laboratory applications (412). Moreover, dietary exposure to acrylamide is well established upon consumption of carbohydrate-rich food cooked at high temperatures (>120 ˚C) such as: baked potatoes, crisps and cookies (413). Acrylamide is

formed in food through a multistep process known as Maillard reaction in which the asparginine amino acid reacts with a reducing sugar such as glucose or fructose (414,415). Based on the evidence of acrylamide carcinogenicity in animals, it has been classified as a probable carcinogen (Group 2A) by IARC and the association between acrylamide exposure and colorectal cancer incidence has been previously suggested as well as other types of human cancers (416–419). Acrylamide adduct detection in human blood was reported by previous studies (420), however, to the best of our knowledge detection of DNA acrylamide adducts has been exclusively reported in animal but not human studies that documented detection of both N^7 and N^3 acrylamide-derived guanine adducts but with different alkyl moieties (421).

The detection of O^6-MeG and O^6-CMG adducts in human colorectal DNA might be indicative risk factor for colorectal cancer (41). O^6-MeG is a known pro-mutagenic adduct that its persistence is associated with carcinogenicity and cytotoxicity. Reported O^6-MeG carcinogenicity is mainly due to its mispairing potential inducing predominantly (>80%) GC:AT transition mutations. However, biological consequences of non-repaired O^6-CMG adducts are less well described with *in vivo* evidence of its pre-mutagenic properties inducing TP53 gene mutations (approximately equal amounts of transitions and transversions) establishing a direct mechanistic link between O^6-CMG adduct and colorectal cancer aetiology (246).

To discuss the benefits of the newly developed approach non-specificity to a single O^6-alkylG adduct comes first as it could be used for the characterisation of the overall load of O^6-alkylG adducts. Additionally, it could be applied for both qualitative and quantitative aspects of human DNA analysis. This technique allows accurate and reproducible quantitation of O^6-alkylG adducts in terms of quantifying detected alkylated peptides using heavy labelled internal standards. Although the current study has only quantified the methylated ASP using the internal standard, accurate quantitation of other alkylated peptides could be achieved using the respective heavy labelled internal standards to shed light on the contribution of their corresponding adducts to the colorectal carcinogenesis. Results highlighted the sensitivity of the assay to the low level of O^6-alkylG adducts as well as low cost. However, the limitations of the current protocol include lack of tandem MS data that are indispensable to confirm the identity of the putative alkylated peptides detected in MGMT digest following incubation with human DNA in order to ascertain that the right O^6-alkylG adducts have been targeted. Notably, the novel assay analysed relatively large amounts of human colorectal DNA which had a positive facet making the

detection of O^6-alkylG occurring in low amounts possible. In order to remedy this, MS analysis could be carried out by further sensitive MS techniques such as Orbitrap or triple quadrupole. In addition, a possible limitation of the newly developed approach is the potential O^6-MeG adduct interference with other O^6-alkylG adducts transfer to MGMT.

7.5 Conclusion

Results of the current work acknowledge the applicability of this novel approach for simultaneous detection of overall DNA alkylation damage in human colorectal DNA; using the action of MGMT protein and coupled to MALDI-TOF MS analysis. Identification of novel O^6-alkylG adducts in human colorectal DNA aids to unravel mechanistic basis of colorectal cancer as well as characterising endogenous nitrosation damage in GI tract. Furthermore, this approach could be useful in investigating the association between O^6-alkylG and other types of human cancer such as breast, ovarian or prostate cancer or even the potential association between DNA alkylation damage and other health problems such as male or female infertility.

8 Overall discussion and future work

8.1 Overall discussion

Human exposure to a plethora of environmental and dietary AAs as well as endogenously formed AAs, namely, NNOC may well generate numerous mutagenic O^6-alkylG DNA adducts potentially causing a complex array of genomic modifications contributing to colorectal cancer carcinogenesis (247,256,257). This study displayed the development of novel MALDI-TOF MS based approach that detected a range of putative O^6-alkylG DNA adducts, namely, O^6-MeG, O^6-CMG and O^6-HOEtG in the DNA of human colorectal cancer tissue. Additionally, O^6-MeG adduct was detected in DNA of macroscopically normal human colorectal tissue from colorectal cancer patients. Moreover, the developed approach enabled the quantification of the detected O^6-MeG. In addition, further seven ions were detected in the digests of MGMT incubated with human colorectal DNA, which could potentially represent alkyl modified MGMT-ASPs. However, the identities of such ions have not yet been defined.

These results accord well with a previous study that demonstrated detection of a diverse assortment of O^6-alkylG DNA adducts in human gut lumen, specifically, gastric DNA from both malignant gastric mucosa and normal mucosa of gastric cancer patients, but they did not examine any human colorectal DNA (260). They also reported a significant positive association between dietary nitrate intake and O^6-alkylG levels in human gastric malignant tissue.

Characterization of O^6-alkylG adducts has been used in biomonitoring studies that aim at examining toxicological and carcinogenic potential of environmental, dietary and occupational chemicals (422). The utility of MS analysis in both detection and quantification of such adducts is previously reported (253,423), where previously established methodologies detected O^6-alkylG adducts in DNA isolated from tissues, cells, blood and saliva as well as DNA from *in vitro* studies. However, the current study utilised the known action of DNA repair protein MGMT to transfer the alkyl group to MGMT-ASP followed by MS detection of the tryptic alkyl modified MGMT-ASP to identify and quantify O^6-alkylGs in human colorectal DNA. This modification enabled using a high amount of human DNA alongside avoiding any possible DNA interference

with the MS analysis since MGMT protein was recaptured from the DNA and analysed by MS individually following its tryptic digestion. The current work demonstrated the detection of different O^6-alkylG adducts in human colorectal DNA; although MGMT repair rates for theses adducts vary. Even the poorly repaired O^6-alkylG adduct, O^6-HOEtG, was detected. This was attributed to using a vast excess of MGMT repair protein as well as long incubation times.

Although the current study was based on a small number of participants, it demonstrated the development of a novel methodology that could potentially contribute to the characterization of the overall load of O^6-alkylG adducts in human colorectal cancer DNA which provide mechanistic information on human colorectal cancer pathogenesis. In addition, characterization of the overall load of O^6-alkylG adducts enables considerable improvements in defining individual colorectal cancer risk, paving the way for more informative monitoring and intervention studies. Moreover, these findings provide further support to the initial hypothesis of the involvement of AAs in human colorectal cancer aetiology as well as their toxicological mode of action. Although the current work did not focus on the identification of the specific AAs that might have contributed to the pathogenesis, it's suggested that NNOC, from endogenous or exogenous sources are involved.

Moreover, among the main contributions of the current work was the demonstration of MGMT ability to repair O^6-CMG and O^6-CEG alongside O^6-MeG in synthetic ODNs *via* MGMT inactivation assay and restriction endonuclease site deprotection assay. Moreover, MALDI-TOF MS analysis of MGMT reacted with O^6-CMG, O^6-CEG and O^6-MeG ODNs indicated the transfer of carboxymethyl, carboxyethyl as well as methyl groups to MGMT-ASP, respectively, confirming the mechanism of MGMT repair of these studied adducts. In addition, it was indicated that O^6-CMG and O^6-CEG alongside O^6-MeG were recognised by damage sensing protein Atl1.

The current study focused on investigating the MGMT repair of a range of O^6-alkylG adducts in ODNs, using an MGMT inactivation assay, to examine whether the developed MALDI-TOF MS based assay would be able to detect them in human DNA or not. These O^6-alkylG adducts included O^6-MeG, O^6-CMG and O^6-CEG. Ability of MGMT to repair O^6-CMG was a matter of controversy with contradicting results reported by previous studies (208,296). The current work examined MGMT ability to

repair O^6-CMG in 2 putative O^6-CMG containing ODNs of different DNA sequence contexts, namely, 44 and DW1. Similarly 2 preparations of putative O^6-CEG 23 mer ODNs (LH1 and RA1) were examined. Results of MGMT inactivation assay highlighted remarkable differences in IC_{50} for 44 ODN and DW1. Firstly, MGMT pre-incubation with DS but not SS DW1 caused 50% MGMT inactivation. Conversely, both SS and DS 44 ODN pre-incubation with MGMT demonstrated 50% MGMT inactivation. The variation in IC_{50} results created some doubt about the identity of the alky groups at the O^6 position of the modified guanine at the two O^6-CMG ODNs 44 and DW1. In addition, it raised a question whether O^6-CMG adduct is a substrate for MGMT repair is SS as well as DS or not. Notably, IC_{50} of DS DW1 ODN was 37-fold that of DS 44 displaying higher MGMT repair preference for 44 ODN *vs*. DW1. Similarly, the two investigated O^6-CEG ODNs LH1 and RA1 displayed unexpected and significant difference in both SS and DS IC_{50} results where IC_{50} of SS and DS RA1 ODN were approximately 7-fold the corresponding figures of LH1 ODN highlighting that the alkyl adduct in LH1 ODN is a better substrate for MGMT repair compared to RA1.

MALDI-TOF MS analysis of MGMT reacted with 44 and DW1 ODN was able to unravel the reason behind the unexpected IC_{50} differences demonstrating that these two ODNs contained 2 different alkyl groups at the O^6 position of the modified guanine bases where DW1 was confirmed to contain a carboxymethyl group whilst 44 ODN turned out to contain the carboxamidomethyl group. In addition, quantitative MS results were consistent with MGMT inactivation assays as it confirmed higher MGMT repair preference for 44 ODN compared to DW1. In addition, MS results assured the inability of MGMT to repair SS DW1 that might have raised an assumption about the inability of MGMT to repair O^6-CMG adducts in SS DNA unless later contradicted by the results of SS 13 mer O^6-CMG ODN (OW18) and attributing this finding in DW1 potentially to the effect of local DNA sequence context. Findings of the current research contribute to resolving the conflict regarding the ability of MGMT to repair O^6-CMG adduct as well as providing knowledge about the mechanism of the repair, however, it significantly highlights the impact of local DNA sequence context on MGMT ability to repair this adduct.

In addition, MALDI-TOF MS analysis of MGMT following incubation with LH1 and RA1 revealed that RA1 was O^6-CEG-containing ODN whilst LH1 contained alternatively a carboxamidoethyl group. Quantitative MS results of RA1 and LH1 accord well with MGMT IC_{50} results confirming higher MGMT repair preference for LH1

compared to RA1. Conclusively, both MGMT inactivation assay results as well as and quantitative MS displayed higher MGMT repair preference for the two investigated amide adducts (carboxamidomethyl (-CH_2CONH_2) in 44 ODN and carboxamidoethyl (-$CH_2CH_2CONH_2$) in LH1 ODN) compared to the corresponding carboxylic acid alkyl adducts (carboxymethyl (-CH_2COOH) in DW1 ODN and carboxyethyl (-CH_2CH_2COOH) in RA1 ODN).

Although, the study objectives did not include investigating MGMT repair of carboxamidoethyl (propionamide) adduct, data obtained from LH1 ODNs paved the way for the identification of this adduct in human colorectal DNA, which was not related to acrylamide gels. Since proteins were subjected to on-bead digestion protocols following reaction with human colorectal DNA not in-gel digestion.

Moreover, MGMT inactivation assay findings highlighted that O^6-MeG is the most potent MGMT inactivator followed by O^6-CEG then O^6-CMG. However, this result could be related to the difference in local DNA sequence between DW1 (O^6-CMG containing ODN) and RA1 (O^6-CEG containing ODN). In addition, both quantitative MS results as well as MGMT inactivation assays demonstrated that DS O^6-alkylG ODNs are better substrates for MGMT repair compared to the corresponding SS ODNs, under the operated experimental conditions.

In conclusion, the results of the present study acknowledge the development of a novel MALDI-TOF MS based approach for characterization of the overall load of O^6-alkylG adduct levels in human colorectal DNA extracted from both normal and tumour tissue *via* analysis of MGMT peptides following protein reaction with alkylated DNA. Furthermore, findings of the study demonstrate a great potential for applying this detection and quantification method in future mechanistic studies evaluating the association between O^6-alkylG adducts and other types of human cancer such as breast cancer.

8.2 Future work

1- Investigating overall load of O^6-alkylG adducts in a larger number of human colorectal DNA samples from malignant colorectal tissue, normal colorectal tissue from a colorectal cancer patient as well as normal colorectal tissue from healthy individuals in order to characterise O^6-alkylG adducts in these three populations.

2- Tandem MS analysis to confirm the identity of putative alkylated MGMT-ASPs as well as to identify the unknown ions detected in the digests of MGMT reacted with human colorectal DNA.

3- Tandem MS analysis to confirm the findings of the MS analysis of MGMT reacted with O^6-alkylG ODNs.

4- Investigating the probable association between: red/processed meat consumption, dietary nitrate intake and/or dietary NNOC intake and O^6-alkylG adduct levels in human colorectal DNA.

5- Investigating the probable association between: MGMT activity in human colorectal tissue and/or methylation of CpG island of MGMT promoter region and O^6-alkylG adduct levels in human colorectal DNA.

6- Investigating the probable association between: smoking, alcohol consumption, gender and/or occupation and O^6-alkylG adduct levels in human colorectal DNA.

References

1. Haggar FA, Boushey RP. Colorectal cancer epidemiology: Incidence, mortality, survival, and risk factors. Clin Colon Rectal Surg. 2009;22(4):191–7.

2. Marley AR, Nan H. Epidemiology of colorectal cancer. Int J Mol Epidemiol Genet. 2016;7(3):105–14.

3. What is colorectal cancer? [Internet]. Cancer.org. 2017 [cited 2017 Aug 8]. Available from: http://www.cancer.org/cancer/colonandrectumcancer/detailedguide/colorectal-cancer-what-is-colorectal-cancer.

4. de Jong UW, Day NE, Muir CS, Barclay TH, Bras G, Foster FH, et al. The distribution of cancer within the large bowel. Int J Cancer. 1972;10(3):463–77.

5. Souglakos J. Genetic alterations in sporadic and hereditary colorectal cancer: Implementations for screening and follow-up. Dig Dis. 2007;25(1):9–19.

6. Cancer Research UK [Internet]. [cited 2017 Aug 8]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence#heading-Four

7. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-tieulent J, Jemal A. Global Cancer Statistics, 2012. CA a cancer J Clin. 2015;65(2):87–108.

8. Brenner H, Kloor M, Pox CP. Colorectal cancer. Lancet. 2014;383(9927):1490–502.

9. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;136(5):E359–86.

10. N. Ireland Cancer Registry [Internet]. [cited 2017 Aug 8]. Available from: http://www.qub.ac.uk/research-centres/nicr/CancerInformation/official-

statistics/BySite/Colorectal/

11. Division IS. Scottish Bowel Screening Programme Statistics [Internet]. 2017 [cited 2017 Aug 8]. Available from: http://www.isdscotland.org/Health-Topics/Cancer/Publications/index.asp?Co=Y

12. Cancer Registration Statistics, England : 2015 [Internet]. [cited 2017 Aug 8]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/previousReleases

13. Cancer Research UK [Internet]. [cited 2017 Aug 8]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence#ref-1

14. Office for National Statistics. Deaths registered in England and Wales Statistical bulletin [Internet]. 2015 [cited 2017 Sep 1]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsregistrationsummarytables/previousReleases

15. Cancer research UK. Bowel cancer mortality statistics [Internet]. 2016 [cited 2017 Sep 20]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/mortality#heading-Zero

16. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, et al. Colorectal Cancer Statistics , 2017. CA Cancer J Clin. 2017;67(3):177–93.

17. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA Cancer J Clin. 2017;67:7–30.

18. Murphy G, Devesa SS, Cross AJ, Inskip PD, McGlynn KA, Cook MB. Sex disparities in colorectal cancer incidence by anatomic subsite, race and age. Int J Cancer. 2011;128(7):1668–75.

19. Meissner HI, Breen N, Klabunde CN, Vernon SW. Patterns of colorectal cancer screening uptake among men and women in the United States. Cancer Epidemiol Biomarkers Prev. 2006;15(2):389–94.

20. Wilmink ABM. Overview of the epidemiology of colorectal cancer. Dis Colon Rectum. 1997;40(4):483–93.

21. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer. 2010;127(12):2893–917.

22. Boyle P, Langman JS. ABC of colorectal cancer: Epidemiology. BMJ Br Med J. 2000;321(7264):805.

23. American Cancer Society. Colorectal Cancer Facts & Figures 2017 - 2019. American Cancer Society. Atlanta Ga; 2017.

24. Howlader N, Noone AM, Krapcho M, Miller D, Bishop K, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ CK. SEER Cancer Statistics Review, 1975-2013. National Cancer Institute. Bethesda, MD; 2016.

25. Cancer research UK. Bowel cancer incidence by age [Internet]. 2016 [cited 2017 Oct 1]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence#heading-One

26. Jasperson KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and Familial Colon Cancer. Gastroenterology. 2010;138(6):2044–58.

27. Burt R. Inheritance of colorectal cancer. Drug Discov Today Dis Mech. 2007;4(4):293–300.

28. Vasen HFA, Tomlinson I, Castells A. Clinical management of hereditary colorectal cancer syndromes. Nat Rev Gastroenterol Hepatol. 2015;12(2):88–97.

29. Barrow E, Hill J, Gareth Evans D. Cancer risk in Lynch Syndrome. Fam Cancer. 2013;12(2):229–40.

30. Lynch HT, de la Chapelle A. Hereditary Colorectal Cancer. N Engl J Med. 2003;348(10):919–32.

31. Rustgi AK. The genetics of hereditary colon cancer. Genes Dev. 2007;21(20):2525–38.

32. Long AG, Lundsmith ET, Hamilton KE. Inflammation and Colorectal Cancer. Curr Colorectal Cancer Rep. 2017;13(4):341–51.

33. Munkholm P. Review article: the incidence and prevalence of colorectal cancer in inflammatory bowel disease. Aliment Pharmacol Ther. 2003;18 Suppl 2:1–5.

34. Bressenot A, Cahn V, Danese S, Peyrin-Biroulet L. Microscopic features of colorectal neoplasia in inflammatory bowel diseases. World J Gastroenterol. 2014;20(12):3164–72.

35. Dyson JK, Rutter MD. Colorectal cancer in inflammatory bowel disease: What is the real magnitude of the risk? World J Gastroenterol. 2012;18(29):3839–48.

36. Eaden JA, Abrams KR, Mayberry JF. The risk of colorectal cancer in ulcerative colitis: a meta-analysis. Gut. 2001;48(4):526–35.

37. Azer SA. Overview of molecular pathways in inflammatory bowel disease associated with colorectal cancer development. Eur J Gastroenterol Hepatol. 2013;25(3):271–81.

38. Herrinton LJ, Liu L, Levin TR, Allison JE, Lewis JD, Velayos F. Incidence and mortality of colorectal adenocarcinoma in persons with inflammatory bowel disease from 1998 to 2010. Gastroenterology. 2012;143(2):382–9.

39. Dulai PS, Sandborn WJ, Gupta S. Colorectal cancer and dysplasia in inflammatory bowel disease: A review of disease Epidemiology, Pathophysiology, and Management. Cancer Prev Res. 2016;9(12):887–94.

40. Castano-Milla C. Chaparro M. Gisbert J.P., Castaño-Milla C, Chaparro M, Gisbert JP, Castano-Milla C. Chaparro M. Gisbert J.P., Castaño-Milla C, et al. Systematic review with meta-analysis: the declining risk of colorectal cancer in ulcerative colitis. Aliment Pharmacol Ther. 2014;39(7):645–59.

41. Norat T, Bingham S, Ferrari P, Slimani N, Jenab M, Mazuir M, et al. Meat, fish, and colorectal cancer risk: The European prospective investigation into cancer and nutrition. J Natl Cancer Inst. 2005;97(12):906–16.

42. Baena R, Salinas P. Diet and colorectal cancer. Maturitas. 2015;80(3):258–64.

43. Xu X, Yu E, Gao X, Song N, Liu L, Wei X, et al. Red and processed meat intake and risk of colorectal adenomas: A meta-analysis of observational studies. Int J Cancer. 2013;132(2):437–48.

44. Bastide NM, Chenni F, Audebert M, Santarelli RL, Taché S, Naud N, et al. A central role for heme iron in colon carcinogenesis associated with red meat intake. Cancer Res. 2015;75(5):870–9.

45. Kim E, Coelho D, Blachier F. Review of the association between meat consumption and risk of colorectal cancer. Nutr Res. 2013;33(12):983–94.

46. Manousos O, Day NE, Trichopoulos D, Gerovassilis F, Tzonou A, Polychronopoulou A. Diet and colorectal cancer: A case control study in Greece. Int J Cancer. 1983;32(1):1–5.

47. Giovannucci E, Rimm EB, Stampfer MJ, Colditz A. Intake of Fat , Meat , and Fiber in Relation to Risk of Colon Cancer in Men. Cancer Res. 1994;54:2390–7.

48. Carr PR, Walter V, Brenner H, Hoffmeister M. Meat subtypes and their association with colorectal cancer: Systematic review and meta-analysis. Int J Cancer. 2016;138(2):293–302.

49. Alexander DD, Weed DL, Miller PE, Mohamed M a. Red Meat and Colorectal Cancer: A Quantitative Update on the State of the Epidemiologic Science. J Am Coll Nutr. 2015;34(6):521–43.

50. WHO | Q&A on the carcinogenicity of the consumption of red meat and processed meat [Internet]. WHO. World Health Organization; 2016 [cited 2017 Aug 12]. Available from: http://www.who.int/features/qa/cancer-red-meat/en/

51. International Agency for Research on Cancer. List of Classifications by cancer sites with sufficient or limited evidence in humans, Volumes 1 to 117 [Internet]. 2017 [cited 2017 Aug 12]. p. 1–12. Available from: http://monographs.iarc.fr/ENG/Classification/Table4.pdf

52. World Cancer Research Fund / American Institute for Cancer Research. Continuous Update Project findings & reports-colorectal cancer [Internet]. 2011 [cited 2017 Aug 12]. Available from: http://www.wcrf.org/int/research-we-fund/continuous-update-project-findings-reports/colorectal-bowel-cancer

53. Parkin D, Boyd L, Walker L. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010 Summary and conclusions. Br J Cancer. 2011;105(10):77–81.

54. Chan DSM, Lau R, Aune D, Vieira R, Greenwood DC, Kampman E, et al. Red and processed meat and colorectal cancer incidence: Meta-analysis of prospective studies. PLoS One. 2011;6(6):1–12.

55. Larsson SC, Wolk A. Meat consumption and risk of colorectal cancer: A meta-analysis of prospective studies. Int J Cancer. 2006;119(11):2657–64.

56. Norat T, Lukanova A, Ferrari P, Riboli E. Meat consumption and colorectal cancer risk: Dose-response meta-analysis of epidemiological studies. Int J Cancer. 2002;98(2):241–56.

57. Ferguson LR. Meat and cancer. Meat Sci. 2010;84(2):308–13.

58. Bingham SA, Hughes R, Cross AJ. Effect of white versus red meat on endogenous N-nitrosation in the human colon and further evidence of a dose response. J Nutr. 2002;132(16):3522S–3525S.

59. Aykan NF. Red meat and colorectal cancer. Oncol Rev. 2015;9(1):38–44.

60. Santarelli RL, Pierre F, Corpet DE, Santarelli L, Pierre F, Corpet DE, et al. Processed Meat and Colorectal Cancer : A Review of Epidemiologic and Experimental Evidence. Nutr Cancer. 2008;60(2):131–44.

61. Bastide NM, Pierre FHF, Corpet DE. Heme iron from meat and risk of colorectal cancer: a meta-analysis and a review of the mechanisms involved. Cancer Prev Res (Phila). 2011;4(2):177–84.

62. Ni W, McNaughton L, LeMaster DM, Sinha R, Turesky RJ. Quantitation of 13 heterocyclic aromatic amines in cooked beef, pork, and chicken by liquid chromatography-electrospray ionization/tandem mass spectrometry. J Agric Food Chem. 2008;56(1):68–78.

63. Sinha R, Rothman N, Salmon CP, Knize MG, Brown DE, Swanson CA, et al. Heterocyclic amine content in beef cooked by different methods to varying degrees of doneness and gravy made from meat drippings. Food Chem Toxicol. 1998;36(4):279–87.

64. Layton DW, Bogen KT, Knize MG, Hatch FT, Johnson VM, Felton JS. Cancer risk of heterocyclic amines in cooked foods: An analysis and implications for research. Carcinogenesis. 1995;16(1):39–52.

65. Zheng W, Lee S-A. Well-done meat intake, heterocyclic amine exposure, and cancer risk. Nutr Cancer. 2009;61(4):437–46.

66. Jägerstad M, Skog K. Genotoxicity of heat-processed foods. Mutat Res - Fundam Mol Mech Mutagen. 2005;574(1–2 SPEC. ISS.):156–72.

67. International Agency for Research on Cancer. IARC. Polynuclear Aromatic Compounds PART 1, Chemical, environmental and experimental data. Vol. 32, IARC, International Agency for Research on Cancer. 1983.

68. Noorwati Sutandyo. Nutritional Carcinogenesis. J Intern Med. 2010;24(1):36–42.

69. Park Y, Hunter DJ, Spiegelman D, Bergkvist L, Berrino F, van den Brandt P a, et al. Dietary fiber intake and risk of colorectal cancer: a pooled analysis of

prospective cohort studies. JAMA. 2005;294(22):2849–57.

70. Cancer research UK. Bowel cancer statistics [Internet]. 2014 [cited 2017 Aug 12]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer#heading-Three

71. Burkitt DP. Epidemiology of cancer of the colon and rectum. Cancer. 1971;28(1):3–13.

72. Slattery ML, Curtin KP, Edwards SL, Schaffer DM. Plant foods, fiber, and rectal cancer 1–4. Am J Nutr. 2004;274–81.

73. Bingham SA, Norat T, Moskal A, Ferrari P, Slimani N, Clavel-Chapelon F, et al. Is the association with fiber from foods in colorectal cancer confounded by folate intake? Cancer Epidemiol Biomarkers Prev. 2005;14(6):1552–6.

74. Schatzkin A. Going against the grain? Current status of the dietary fiber-colorectal cancer hypothesis. Biofactors. 2000;12(1–4):305–11.

75. Fedirko V, Tramacere I, Bagnardi V, Rota M, Scotti L, Islami F, et al. Alcohol drinking and colorectal cancer risk: An overall and dose-Response meta-analysis of published studies. Ann Oncol. 2011;22(9):1958–72.

76. Wang, Y., Duan, H., Yang, H. and Lin J. A pooled analysis of alcohol intake and colorectal cancer. Int J Clin Exp Med. 2015;8(5):6878–89.

77. Moskal A, Norat T, Ferrari P, Riboli E. Alcohol intake and colorectal cancer risk: A dose-response meta-analysis of published cohort studies. Int J Cancer. 2007;120(3):664–71.

78. Na HK, Lee JY. Molecular basis of alcohol-related gastric and colon cancer. Int J Mol Sci. 2017;18(6):1–16.

79. Seitz HK, Stickel F. Acetaldehyde as an underestimated risk factor for cancer development: role of genetics in ethanol metabolism. Genes Nutr.

2010;5(2):121–8.

80. Seitz HK, Stickel F. Molecular mechanisms of alcohol-mediated carcinogenesis. Nat Rev Cancer. 2007;7(8):599–612.

81. Linhart K, Bartsch H, Seitz HK. The role of reactive oxygen species (ROS) and cytochrome P-450 2E1 in the generation of carcinogenic etheno-DNA adducts. Redox Biol. 2014;3:56–62.

82. Sohn OS, Fiala ES, Requeijo SP, Weisburger JH, Gonzalez FJ. Differential effects of CYP2E1 status on the metabolic activation of the colon carcinogens azoxymethane and methylazoxymethanol. Cancer Res. 2001;61(23):8435–40.

83. Tsoi KK, Pau CY, Wu WK, Chan FK, Griffiths S, Sung JJ. Cigarette smoking and the risk of colorectal cancer: a meta-analysis of prospective cohort studies. Clin Gastroenterol Hepatol. 2009;7(6):682–5.

84. Yang C, Wang X, Huang C-H, Yuan W-J, Chen Z-H. Passive Smoking and Risk of Colorectal Cancer: A Meta-analysis of Observational Studies. Asia-Pacific J public Heal. 2016;28(5):394–403.

85. Liang PS, Chen T, Giovannucci E. Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis. Int J Cancer. 2009;124(10):2406–15.

86. Huxley RR, Ansary-Moghaddam A, Clifton P, Czernichow S, Parr CL, Woodward M. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: A quantitative overview of the epidemiological evidence. Int J Cancer. 2009;125(1):171–80.

87. Raimondi S, Botteri E, Iodice S, Lowenfels AB, Maisonneuve P. Gene-smoking interaction on colorectal adenoma and cancer risk: review and meta-analysis. Mutat Res. 2009;670(1–2):6–14.

88. Ding YS, Zhang L, Jain RB, Jain N, Wang RY, Ashley DL, et al. Levels of tobacco-specific nitrosamines and polycyclic aromatic hydrocarbons in

mainstream smoke from different tobacco varieties. Cancer Epidemiol Biomarkers Prev. 2008;17(12):3366–71.

89. Hang B. Formation and repair of tobacco carcinogen-derived bulky DNA adducts. J Nucleic Acids. 2010;2010:1–29.

90. Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K. Body Fatness and Cancer — Viewpoint of the IARC Working Group. N Engl J Med. 2016;375(8):794–8.

91. Jochem C, Leitzmann M. Obesity and colorectal cancer. In: Pischon T, Nimptsch K, editors. Obesity and Cancer Recent Results in Cancer Research. Springer, Cham; 2016. p. 17–41.

92. Samad a K a, Taylor RS, Marshall T, Chapman M a S. A meta-analysis of the association of physical activity with reduced risk of colorectal cancer. Colorectal Dis. 2005;7(3):204–13.

93. Lee KJ, Inoue M, Otani T, Iwasaki M, Sasazuki S, Tsugane S, et al. Physical activity and risk of colorectal cancer in Japanese men and women: the Japan Public Health Center-based prospective study. Cancer Causes Control. 2007;18(2):199–209.

94. Bosman FT, Carneiro F, Hruban RH, Theise ND. WHO Classification of Tumours of the Digestive System, Fourth Edition. In: WHO Classification of Tumours. IARC WHO Classification of Tumours, No 3; 2010. p. 417.

95. Kinzler K, Vogelstein B. Lessons from Hereditary Review Colorectal Cancer. Cell. 1996;87(2):159–70.

96. Dunican DS, McWilliam P, Tighe O, Parle-McDermott A, Croke DT. Gene expression differences between the microsatellite instability (MIN) and chromosomal instability (CIN) phenotypes in colorectal cancer revealed by high-density cDNA array hybridization. Oncogene. 2002;21(20):3253–7.

97. Grady WM, Carethers JM. Genomic and Epigenetic Instability in Colorectal

Cancer Pathogenesis. Gastroenterology. 2008;135(4):1079–99.

98. Lengauer C, Kinzler KW, Vogelstein B. Genetic instability in colorectal cancers. Nature. 1997;386(6625):623–7.

99. Pino MS, Chung DC. The Chromosomal Instability Pathway in Colon Cancer. Gastroenterology. 2010;138(6):2059–72.

100. Rao C V., Yamada HY, Yao Y, Dai W. Enhanced genomic instabilities caused by deregulated microtubule dynamics and chromosome segregation: A perspective from genetic studies in mice. Carcinogenesis. 2009;30(9):1469–74.

101. Pino MS, Chung DC. Microsatellite instability in the management of colorectal cancer. Expert Rev Gastroenterol Hepatol. 2011;5(3):385–99.

102. Lipkin SM, Wang V, Jacoby R, Banerjee-Basu S, Baxevanis AD, Lynch HT, et al. MLH3: a DNA mismatch repair gene associated with mammalian microsatellite instability. Nat Genet. 2000;24(1):27–35.

103. Umar A, Risinger JI, Glaab WE, Tindall KR, Barrett JC, Kunkel TA. Functional overlap in mismatch repair by human MSH3 and MSH6. Genetics. 1998;148(4):1637–46.

104. Arai T, Kasahara I, Sawabe M, Honma N, Aida J, Tabubo K. Role of methylation of the hMLH1 gene promoter in the development of gastric and colorectal carcinoma in the elderly. Geriatr Gerontol Int. 2010;10(SUPPL. 1):S207-12.

105. Fox EJ, Leahy DT, Geraghty R, Mulcahy HE, Fennelly D, Hyland JM, et al. Mutually exclusive promoter hypermethylation patterns of hMLH1 and O6-methylguanine DNA methyltransferase in colorectal cancer. J Mol Diagn. 2006;8(1):68–75.

106. Herman JG, Umar A, Polyak K, Graff JR, Ahuja N, Issa JP, et al. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proc Natl Acad Sci U S A. 1998;95(12):6870–5.

107. Fearon E, Bommer G. Molecular biology of colorectal cancer. In: DeVita VT Jr, Lawrence TS RS, editor. De-Vita, Hellman, and Rosenberg's cancer: principles & practice of oncology Vol 1. Philadelphia: Lippincott Williams & Wilkins; 2008. p. 1218–31.

108. Duval A, Hamelin R. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: Toward a new concept of target genes for instability. Cancer Res. 2002;62(9):2447–54.

109. Lao VV, Grady WM. Epigenetics and colorectal cancer. Nat Rev Gastroenterol Hepatol. 2011;8(12):686–700.

110. Miranda TB, Jones PA. DNA methylation: The nuts and bolts of repression. J Cell Physiol. 2007;213(2):384–90.

111. Li X, Yao X, Wang Y, Hu F, Wang F, Jiang L, et al. MLH1 Promoter Methylation Frequency in Colorectal Cancer Patients and Related Clinicopathological and Molecular Features. PLoS One. 2013;8(3):1–9.

112. Leary RJ, Lin JC, Cummins J, Boca S, Wood LD, Parsons DW, et al. Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. Proc Natl Acad Sci U S A. 2008;105(42):16224–9.

113. Markowitz SD, Bertagnolli MM. Molecular Basis of Colorectal Cancer. N Engl J Med. 2009;361:2449–60.

114. Velho S, Moutinho C, Cirnes L, Albuquerque C, Hamelin R, Schmitt F, et al. BRAF, KRAS and PIK3CA mutations in colorectal serrated polyps and cancer: primary or secondary genetic events in colorectal carcinogenesis? BMC Cancer. 2008;8(255):1–6.

115. Yu J, Wu WKK, Li X, He J, Li X-X, Ng SSM, et al. Novel recurrently mutated genes and a prognostic mutation signature in colorectal cancer. Gut. 2014;64(4):636–45.

116. Wood LD, Parsons DW, Jones S, Lin J, Sjoblom T, Leary RJ, et al. The Genomic Landscapes of Human Breast and Colorectal Cancers. Science (80-). 2007;318(5853):1108–13.

117. Muzny DM, Bainbridge MN, Chang K, Dinh HH, Drummond JA, Fowler G, et al. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012;487(7407):330–7.

118. Gay LJ, Mitrou PN, Keen J, Bowman R, Naguib A, Cooke J, et al. Dietary, lifestyle and clinicopathological factors associated with APC mutations and promoter methylation in colorectal cancers from the EPIC-Norfolk study. J Pathol. 2012;228(3):405–15.

119. Jass JR. Pathogenesis of colorectal cancer. Surg Clin North Am. 2002;82(5):891–904.

120. Nowak MA, Komarova NL, Sengupta A, Jallepalli P V., Shih I-M, Vogelstein B, et al. The role of chromosomal instability in tumor initiation. Proc Natl Acad Sci. 2002;99(25):16226–31.

121. Polakis P. The many ways of Wnt in cancer. Curr Opin Genet Dev. 2007;17(1):45–51.

122. Pietsch EC, Sykes SM, McMahon SB, Murphy ME. The p53 family and programmed cell death. Oncogene. 2008;27(50):6507–21.

123. Green DR, Kroemer G. Cytoplasmic functions of the tumour suppressor p53. Nature. 2009;458(7242):1127–30.

124. Morselli E, Tasdemir E, Maiuri MC, Galluzzi L, Kepp O, Criollo A, et al. Mutant p53 protein localized in the cytoplasm inhibits autophagy. Cell Cycle. 2008;7(19):3056–61.

125. Esteller M, Risques RA, Toyota M, Capella G, Moreno V, Peinado MA, et al. Promoter hypermethylation of the DNA repair gene O(6)-methylguanine-DNA methyltransferase is associated with the presence of G:C to A:T transition

mutations in p53 in human colorectal tumorigenesis. Cancer Res. 2001;61(12):4689–92.

126. Kiaris H, Spandidos DA. Analysis of H-ras, K-ras and N-ras genes for expression, mutation and amplification in laryngeal tumours. Int J Oncol. 1995;7(1):75–80.

127. Leslie A, Carey FA, Pratt NR, Steele RJC. The colorectal adenoma-carcinoma sequence. Br J Surg. 2002;89(7):845–60.

128. Salerno S, Bazan V, Tomasino RM, Migliavacca M, Labianca R, Zanna I, et al. Specific codon 13 K-ras mutations are predictive of clinical outcome in colorectal cancer patients, whereas codon 12 K-ras mutations are associated with mucinous histotype. Ann Oncol. 2002;13(9):1438–46.

129. Poehlmann A, Kuester D, Meyer F, Lippert H, Roessner A, Schneider-Stock R. K-ras mutation detection in colorectal cancer using the Pyrosequencing technique. Pathol Res Pr. 2007;203(7):489–97.

130. Grady WM, Markowitz SD. TGF-β Signaling Pathway and Tumor Suppression. In: Derynck R, Miyazono K, editors. The TGF-β Family. New York: Cold Spring Harbor Monograph Archive; 2008. p. 889–937.

131. Cancer Genom Atlas. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012;487(7407):330–7.

132. Howe JR, Roth S, Ringold JC, Summers RW, Jarvinen HJ, Sistonen P, et al. Mutations in the SMAD4/DPC4 gene in juvenile polyposis. Science. 1998;280(5366):1086–8.

133. Boland CR, Koi M, Chang DK, Carethers JM. The biochemical basis of microsatellite instability and abnormal immunohistochemistry and clinical behavior in Lynch Syndrome: From bench to bedside. In: Familial Cancer. 2008. p. 41–52.

134. Imai K, Yamamoto H. Carcinogenesis and microsatellite instability: The

interrelationship between genetics and epigenetics. Carcinogenesis. 2008;29(4):673–80.

135. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology. 2007;50(1):113–30.

136. Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, et al. Genetic Alterations during Colorectal-Tumor Development. N Engl J Med. 1988;319(9):525–32.

137. Simon K. Colorectal cancer development and advances in screening. Clin Interv Aging. 2016;11:967–76.

138. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz Jr. LA, Kinzler KW. Cancer Genome Landscapes. Science (80-). 2013;339(6127):1546–58.

139. Hamilton S, Bosman F, Boffetta P, Ilyas M, Morreau H. Carcinoma of the colon and rectum. In: World Health Organization Classification of Tumours of the Digestive System. 2010. p. 134–46.

140. Libutti SK, Saltz LB TJ. Colon cancer. In: DeVita VT Jr, Lawrence TS RS, editor. DeVita, Hellman, and Rosenberg's cancer: principles and practice of oncology Vol 1. Philadelphia: Lippincott Williams & Wilkins; 2008. p. 1232–84.

141. Kim T, Lee S, Chung Y. Clinical applications of next-generation sequencing in colorectal cancers. World J Gastroenterol. 2013;19(40):6784–93.

142. Alexandrov LB, Stratton MR. Mutational signatures: The patterns of somatic mutations hidden in cancer genomes. Curr Opin Genet Dev. 2014;24(1):52–60.

143. Petljak M, Alexandrov LB. Understanding mutagenesis through delineation of mutational signatures in human cancer. Carcinogenesis. 2016;37(6):531–40.

144. Alexandrov LB, Nik-zainal S, Wedge DC, Aparicio SAJR. Signatures of mutational processes in human cancer. Nature. 2013;500(7463):415–21.

145. Alexandrov LB, Jones PH, Wedge DC, Sale JE, Peter J. Clock-like mutational processes in human somatic cells. Nat Publ Gr. 2015;47(12):1402–7.

146. Pilati C, Shinde J, Alexandrov LB, Assié G, André T, Hélias-Rodzewicz Z, et al. Mutational signature analysis identifies MUTYH deficiency in colorectal cancers and adrenocortical carcinomas. J Pathol. 2017;242(1):10–5.

147. Marra G, Schär P. Recognition of DNA alterations by the mismatch repair system. Biochem J. 1999;338:1–13.

148. Shrivastav N, Li D, Essigmann JM. Chemical biology of mutagenesis and DNA repair: Cellular responses to DNA alkylation. Carcinogenesis. 2009;31(1):59–70.

149. Singer B. The Chemical Effects of Nucleic Acid Alkylation and Their Relation to Mutagenesis and Carcinogen esis. Prog Nucleic Acid Res Mol Biol. 1975;15(C):219–84.

150. Sturla SJ. DNA adduct profiles: chemical approaches to addressing the biological impact of DNA damage from small molecules. Curr Opin Chem Biol. 2007;11(3):293–9.

151. Warwick GP. The Mechanism of Action of Alkylating Agents. Cancer Res. 1963;23:1315–33.

152. Drabløs F, Feyzi E, Aas PA, Vaagbø CB, Kavli B, Bratlie MS, et al. Alkylation damage in DNA and RNA - Repair mechanisms and medical significance. DNA Repair (Amst). 2004;3(11):1389–407.

153. Soares DG, Escargueil AE, Poindessous V, Sarasin A, de Gramont A, Bonatto D, et al. Replication and homologous recombination repair regulate DNA double-strand break formation by the antitumor alkylator ecteinascidin 743. Proc Natl Acad Sci U S A. 2007;104(32):13062–7.

154. Dietrich M, Block G, Pogoda JM, Buffler P, Hecht S, Preston-Martin S. A review: Dietary and endogenously formed N-nitroso compounds and risk of childhood brain tumors. Cancer Causes Control. 2005;16(6):619–35.

155. Tricker AR, Kubacki SJ. Review of the occurrence and formation of non-volatile N-nitroso compounds in foods. Food Addit Contam. 1992;9(1):39–69.

156. Beranek DT. Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. Mutat Res - Fundam Mol Mech Mutagen. 1990;231(1):11–30.

157. Fahrer J, Kaina B. O6-methylguanine-DNA methyltransferase in the defense against N-nitroso compounds and colorectal cancer. Carcinogenesis. 2013;34(11):2435–42.

158. Povey a C. DNA adducts: endogenous and induced. Toxicol Pathol. 2000;28(3):405–14.

159. Tricker AR. N-nitroso compounds and man: sources of exposure, endogenous formation and occurrence in body fluids. Eur J Cancer Prev. 1997;6:226–68.

160. Mirvish SS. Role of N-nitroso compounds (NOC) and N-nitrosation in etiology of gastric, esophageal, nasopharyngeal and bladder cancer and contribution to cancer of known exposures to NOC. Cancer Lett. 1995;93(1):17–48.

161. Tricker AR, Preussmann R. Carcinogenic N-nitrosamines in the diet: occurrence, formation, mechanisms and carcinogenic potential. Mutat Res Toxicol. 1991;259(3–4):277–89.

162. International Agency for Research on Cancer. IARC. Agents Classified by the IARC Monographs [Internet]. Vols. 1–119. 2017 [cited 2017 Aug 20]. p. 1–36. Available from: https://monographs.iarc.fr/ENG/Classification/ClassificationsAlphaOrder.pdf

163. Park JE, Seo JE, Lee JY, Kwon H. Distribution of Seven N-Nitrosamines in Food. Toxicol Res. 2015;31(3):279–88.

164. Mavelle T, Bouchikhi B, Debry G. The occurrence of volatile N-nitrosamines in French foodstuffs. Food Chem. 1991;42(3):321–38.

165. Pobel D, Riboli E, Cornée J, Hémon B, Guyader M. Nitrosamine, nitrate and nitrite in relation to gastric cancer: A case-control study in Marseille, France. Eur J Epidemiol. 1995;11(1):67–73.

166. Jo C.H., Park H.R., Kim D.S., Lee K.H. KMH. Exposure assessment of N-nitrosamines in foods. Food Sci Biotechnol. 2010;5:541–8.

167. Raoul S, Gremaud E, Biaudet H, Turesky RJ. Rapid Solid-Phase Extraction Method for the Detection of Volatile Nitrosamines in Food. J Agric Food Chem. 1997;45:4706–13.

168. Stuff JE, Goh ET, Barrera SL, Bondy ML, Forman MR. Construction of an N-nitroso database for assessing dietary intake. J Food Compos Anal. 2009;22(SUPPL 1):S42–7.

169. Bartsch, H; Spiegelhalder B. Environmental exposure to N-nitroso compounds (NNOC) and precursors: an overview. Eur J Cancer Prev. 1996;5(Supplement 1):11–8.

170. Spiegelhalder B, Preussmann R. Occupational nitrosamine exposure. 1. rubber and tyre industry. Carcinogenesis. 1983;4(9):1147–52.

171. Xue J, Yang S, Seng S. Mechanisms of cancer induction by tobacco-specific NNK and NNN. Cancers (Basel). 2014;6(2):1138–56.

172. Peterson LA. Formation, repair, and genotoxic properties of bulky DNA adducts formed from tobacco-specific nitrosamines. J Nucleic Acids. 2010;2010:1–11.

173. Hecht SS. DNA adduct formation from tobacco-specific N-nitrosamines. Mutat Res - Fundam Mol Mech Mutagen. 1999;424(1–2):127–42.

174. Hecht SS, Stepanov I, Carmella SG. Exposure and Metabolic Activation Biomarkers of Carcinogenic Tobacco-Specific Nitrosamines. Acc Chem Res. 2016;49(1):106–14.

175. Bordin DL, Lima M, Lenz G, Saffi J, Meira LB, Mésange P, et al. DNA alkylation

damage and autophagy induction. Mutat Res - Rev Mutat Res. 2013;753(2):91–9.

176. Lijinsky W. N-Nitroso compounds in the diet. Mutat Res - Genet Toxicol Environ Mutagen. 1999;443(1–2):129–38.

177. Mirvish SS. The Etiology of Gastric Cancer: Intragastric Nitrosamide Formation and Other Theories. J Natl Cancer Inst. 1983;71(3):629–47.

178. Mirvish SS. N-nitroso compounds: their chemical and in vivo formation and possible importance as environmental carcinogens. J Toxicol Environ Health. 1977;2(6):1267–77.

179. Mirvish SS. Formation of N-nitroso compounds: chemistry, kinetics, and in vivo occurrence. Toxicol Appl Pharmacol. 1975;31(3):325–51.

180. Lunn JC, Kuhnle G, Mai V, Frankenfeld C, Shuker DEG, Glen RC, et al. The effect of haem in red and processed meat on the endogenous formation of N-nitroso compounds in the upper gastrointestinal tract. Carcinogenesis. 2007;28(3):685–90.

181. Keto-Timonen R, Lindström M, Puolanne E, Niemistö M, Korkeala H. Inhibition of toxigenesis of group II (nonproteolytic) Clostridium botulinum type B in meat products by using a reduced level of nitrite. J Food Prot. 2012;75(7):1346–9.

182. Cui H, Gabriel AA, Nakano H. Antimicrobial efficacies of plant extracts and sodium nitrite against Clostridium botulinum. Food Control. 2010;21(7):1030–6.

183. Roediger WE, Lawson MJ, Radcliffe BC. Nitrite from inflammatory cells--a cancer risk factor in ulcerative colitis? Dis Colon Rectum. 1990;33(12):1034–6.

184. Bingham S, Pignatelli B, Pollock JRA, Euul A, Malaveille C, Gross G, et al. Does increased endogenous formation of N - nitroso compounds in the human colon explain the association between red meat and colon cancer ? Carcinogenesis. 1996;17(3):515–23.

185. Lewin MH, Bailey N, Bandaletova T, Bowman R, Cross AJ, Pollock J, et al. Red Meat Enhances the Colonic Formation of the DNA Adduct O6-Carboxymethyl Guanine: Implications for Colorectal Cancer Risk. Cancer Res. 2006;66(3):1859–65.

186. Ali MA, Poortvliet E, Strömberg R, Yngve A. Polyamines in foods: Development of a food database. Food Nutr Res. 2011;55(1):1–15.

187. Jacobs DM, Gaudier E, van Duynhoven J, Vaughan EE. Non-digestible food ingredients, colonic microbiota and the impact on gut health and immunity: a role for metabolomics. Curr Drug Metab. 2009;10(1):41–54.

188. Kalac P, Krizek M. A Review of Biogenic Amines and Polyamines in Beer. J Inst Brew. 2003;109(2):123–8.

189. Ladero V, Calles-Enríquez M, Fernández M, Alvarez MA. Toxicological effects of dietary biogenic amines. Curr Nutr Food Sci. 2010;6(2):145–56.

190. Teti D, Visalli M, McNair H. Analysis of polyamines as markers of (patho)physiological conditions. J Chromatogr B Anal Technol Biomed Life Sci. 2002;781(1–2):107–49.

191. Sawada Y, Pereira SP, Murphy GM, Dowling RH. Polyamines in the intestinal lumen of patients with small bowel bacterial overgrowth. Biochem Soc Trans. 1994;22(4):392S.

192. Sanchez M, Suarez L, Andres MT, Florez BH, Bordallo J, Riestra S, et al. Modulatory effect of intestinal polyamines and trace amines on the spontaneous phasic contractions of the isolated ileum and colon rings of mice. Food Nutr Res. 2017;61(1):1321948.

193. Macfarlane GT, Cummings JH, Macfarlane G, Cummings J. The Colonic Flora, Fermentation, and Large Bowel Digestive Function. In: S. F. Phillips, J. H. Pemberton and RGS, editor. Large Intestine Physiology, Pathophysiology, and Disease. New York, N.Y.: Raven Press; 1991. p. 51–92.

194. Massey RC, Key PE, Mallett AK, Rowland IR. An investigation of the endogenous formation of apparent total N-nitroso compounds in conventional microflora and germ-free rats. Food Chem Toxicol. 1988;26(7):595–600.

195. Sedgwick B. Nitrosated peptides and polyamines as endogenous mutagens in O6-alkylguanine-DNA alkyltransferase deficient cells. Carcinogenesis. 1997;18(8):1561–7.

196. Totsuka Y, Nishigaki R, Enomoto S, Takamura-Enya T, Masumura KI, Nohmi T, et al. Structures and biological properties of DNA adducts derived from N-nitroso bile acid conjugates. Chem Res Toxicol. 2005;18(10):1553–62.

197. Terasaki M, Totsuka Y, Nishimura K, Mukaisho K-I, Chen K-H, Hattori T, et al. Detection of endogenous DNA adducts, O6-Carboxymethyl-2'-deoxyguanosine and 3-ethanesulfonic acid-2'-deoxycytidine, in the rat stomach after duodenal reflux. Cancer Sci. 2008;99(9):1741–6.

198. Hughes R, Cross a J, Pollock JR, Bingham S. Dose-dependent effect of dietary meat on endogenous colonic N-nitrosation. Carcinogenesis. 2001;22(1):199–202.

199. Kuhnle GGC, Bingham S. Dietary meat, endogenous nitrosation and colorectal cancer. Biochem Soc Trans. 2007;35(Pt 5):1355–7.

200. Kuhnle GGC, Story GW, Reda T, Mani AR, Moore KP, Lunn JC, et al. Diet-induced endogenous formation of nitroso compounds in the GI tract. Free Radic Biol Med. 2007;43(7):1040–7.

201. Joosen AMCP, Kuhnle GGC, Aspinall SM, Barrow TM, Lecommandeur E, Azqueta A, et al. Effect of processed and red meat on endogenous nitrosation and DNA damage. Carcinogenesis. 2009;30(8):1402–7.

202. Joosen AMCP, Lecommandeur E, Kuhnle GGC, Aspinall SM, Kap L, Rodwell SA. Effect of dietary meat and fish on endogenous nitrosation, inflammation and genotoxicity of faecal water. Mutagenesis. 2010;25(3):243–7.

203. Upadhyaya P, Lindgren BR, Hecht S. Comparative Levels of O6-Methylguanine, Pyridyloxobutyl-, and Pyridylhydroxybutyl-DNA Adducts in Lung and Liver of Rats Treated Chronically with the Tobacco-Specific Carcinogen 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone. Drug Metab Dispos. 2009;37(6):1147–51.

204. O'Connor,P.J., Saffhill,R. and Margison GP. N-nitrosocompounds: biochemical mechanisms of action. In: Emmelot,P. and Kriek E, editor. Environmental Carcinogenesis. Amsterdam.: Elsevier/North Holland Biomedical Press; 1979. p. 73–96.

205. Singer B, Grunberger D. Molecular Biology of Mutagens and Carcinogens. New York: Plenum Press; 1983. 347 p.

206. Lawley P, Thatcher C. Methylation of DNA by 3H-14C methyl-labelled N-methyl-N-nitrosourea evidencefor transfer of the intact methyl group. Chem Biol Interact. 1973;7:115–20.

207. Peterson LA, Hecht SS. O6-methylguanine is a critical determinant of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone tumorigenesis in A/J mouse lung. Cancer Res. 1991;51(20):5557–64.

208. Shuker DE, Margison GP. Nitrosated glycine derivatives as a potential source of O6-methylguanine in DNA. Cancer Res. 1997;57(3):366–9.

209. Rydberg B, Lindahl T. Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. EMBO J. 1982;1(2):211–6.

210. Saffhill R, Margison GP, O'Connor PJ. Mechanisms of carcinogenesis induced by alkylating agents. BBA - Rev Cancer. 1985;823(2):111–45.

211. Loeb LA, Harris CC. Advances in Chemical Carcinogenesis: A Historical Review and Prospective. Cancer Res. 2008;68(17):6863–72.

212. Margison GP, Santibáñez Koref MF, Povey AC. Mechanisms of

carcinogenicity/chemotherapy by O6-methylguanine. Mutagenesis. 2002;17(6):483–7.

213. Margison G. O6-methylguanine in DNA: bad penny? Cell Cycle. 2010;9(3):440–9.

214. Räz MH, Dexter HR, Millington CL, van Loon B, Williams DM, Sturla SJ. Bypass of Mutagenic O 6 -Carboxymethylguanine DNA Adducts by Human Y- and B-Family Polymerases. Chem Res Toxicol. 2016;29(9):1493–503.

215. Loechler EL, Green CL, Essigmann JM. In vivo mutagenesis by O6-methylguanine built into a unique site in a viral genome. Proc Natl Acad Sci U S A. 1984;81(20):6271–5.

216. Baumgart PM, Kliem H christian, Gottfried-anacker J, Wiessler M, Schmeiser HH. Site-specific mutagenesis induced by single O6-Alkylguanines (O6-n-propyl, O6-n-butyl, O6-n-octyl) in vivo. Nucleic Acids Res. 1993;21(16):3755–60.

217. Ellison KS, Dogliotti E, Connors TD, Basu a K, Essigmann JM. Site-specific mutagenesis by O6-alkylguanines located in the chromosomes of mammalian cells: influence of the mammalian O6-alkylguanine-DNA alkyltransferase. Proc Natl Acad Sci U S A. 1989;86(22):8620–4.

218. Eadie JS, Conrad M, Toorchen D, Topal MD. Mechanism of mutagenesis by O6-methylguanine. Nature. 1984;308(5955):201–3.

219. Woodside AM, Guengerich FP. Effect of the O6 substituent on misincorporation kinetics catalyzed by DNA polymerases at O6-methylguanine and O6-benzylguanine. Biochemistry. 2002;41(3):1027–38.

220. Singh J, Su L, Snow ET. Replication across O6-methylguanine by human DNA polymerase ?? in vitro: Insights into the futile cytotoxic repair and mutagenesis of O6- methylguanine. J Biol Chem. 1996;271(45):28391–8.

221. Eoff RL, Irimia A, Egli M, Guengerich FP. Sulfolobus solfataricus DNA polymerase Dpo4 is partially inhibited by “Wobble” pairing between O6-

methylguanine and cytosine, but accurate bypass is preferred. J Biol Chem. 2007;282(2):1456–67.

222. Gaffney BL, Marky LA, Jones RA. Synthesis and characterization of a set of four dodecadeoxyribonucleoside undecaphosphates containing O6-methylguanine opposite adenine, cytosine, guanine, and thymine. Biochemistry. 1984;23(24):5686–91.

223. Swann PF. Why do O6-alkylguanine and O4-alkylthymine miscode? The relationship between the structure of DNA containing O6-alkylguanine and O4-alkylthymine and the mutagenic properties of these bases. Mutat Res - Fundam Mol Mech Mutagen. 1990;233(1–2):81–94.

224. Yamagata Y, Tomita K. Structural studies of O6-methyldeoxyguanosine and related compounds: a promutagenic DNA lesion by methylating carcinogens. Nucleic Acids Res. 1988;16(19):9307–21.

225. Jena NR, Bansal M. Mutagenicity associated with O6-methylguanine-DNA damage and mechanism of nucleotide flipping by AGT during repair. Phys Biol. 2011;8(4):1–12.

226. Warren JJ, Forsberg LJ, Beese LS. The structural basis for the mutagenicity of O6-methyl-guanine lesions. Proc Natl Acad Sci U S A. 2006;103(52):19701–6.

227. Sriram M, van der Marel GA, Roelen HL, van Boom JH, Wang AH. Structural consequences of a carcinogenic alkylation lesion on DNA: effect of O6-ethylguanine on the molecular structure of the d(CGC[e6G]AATTCGCG)-netropsin complex. Biochemistry. 1992;31(47):11823–34.

228. Patel DJ, Shapiro L, Kozlowski SA, Gaffney BL, Jones RA. Structural studies of the O6meG•T interaction in the d(C-G-T-G-A-A-T-T-C-O6meG-C-G) duplex. Biochemistry. 1986;25:1036–42.

229. Ginell SL, Kuzmich S, Jones RA, Berman HM. Crystal and Molecular Structure of a DNA Duplex Containing the Carcinogenic Lesion O6-Methylguanine. Biochemistry. 1990;29(46):10461–5.

230. Williams LD, Shaw BR. Protonated base pairs explain the ambiguous pairing properties of O6-methylguanine. Proc Natl Acad Sci U S A. 1987;84(7):1779–83.

231. Kalnik MW, Li BFL, Swann PF, Patel DJ. O6-Ethylguanine Carcinogenic Lesions in DNA: An NMR Study of O6etG·C Pairing in Dodecanucleotide Duplexes. Biochemistry. 1989;28(15):6182–92.

232. De Filippo C, Caderni G, Bazzicalupo M, Briani C, Giannini a, Fazi M, et al. Mutations of the Apc gene in experimental colorectal carcinogenesis induced by azoxymethane in F344 rats. Br J Cancer. 1998;77(12):2148–51.

233. Blum CA, Tanaka T, Zhong X, Li Q, Dashwood WM, Pereira C, et al. Mutational analysis of Ctnnb1 and Apc in tumors from rats given 1,2-dimethylhydrazine or 2-amino-3-methylimidazo[4,5-f]quinoline: Mutational “hotspots” and the relative expression of β-catenin and c-jun. Mol Carcinog. 2003;36(4):195–203.

234. Femia A Pietro, Dolara P, Giannini A, Salvadori M, Biggeri A, Caderni G. Frequent mutation of Apc gene in rat colon tumors and mucin-depleted foci, preneoplastic lesions in experimental colon carcinogenesis. Cancer Res. 2007;67(2):445–9.

235. Jacoby RF, Llor X, Teng BB, Davidson NO, Brasitus TA. Mutations in the K-ras oncogene induced by 1,2-dimethylhydrazine in preneoplastic and neoplastic rat colonic mucosa. J Clin Invest. 1991;87(2):624–30.

236. Vivona AA, Shpitz B, Medline A, Bruce WR, Hay K, Ward MA, et al. K-ras mutations in aberrant crypt foci, adenomas and adenocarcinomas during azoxymethane-induced colon carcinogenesis. Carcinogenesis. 1993;14(9):1777–81.

237. Erdman SH, Wu HD, Hixson LJ, Ahnen DJ, Gerner EW. Assessment of mutations in Ki-ras and p53 in colon cancers from azoxymethane- and dimethylhydrazine-treated rats. Mol Carcinog. 1997;19(2):137–44.

238. Luceri C, Filippo C De, Caderni G, Gambacciani L, Salvadori M, Giannini A, et al. Detection of somatic DNA alterations in azoxymethane-induced F344 rat

colon tumors by random amplified polymorphic DNA analysis (AOM) is a useful experimental model as it mimics the human adenoma – carcinoma sequence and allows the study substances . Al. 2000;21(9):1753–6.

239. de Jong T a, Skinner S a, Malcontenti-Wilson C, Vogiagis D, Bailey M, van Driel IR, et al. Inhibition of rat colon tumors by sulindac and sulindac sulfone is independent of K-ras (codon 12) mutation. Am J Physiol Gastrointest Liver Physiol. 2000;278(2):G266–72.

240. Jacoby RF, Alexander RJ, Raicht RF, Brasitus TA. K-ras oncogene mutations in rat colon tumors induced by N-methyl-N-nitrosourea. Carcinogenesis. 1992;13(1):45–9.

241. Takahashi M, Mutoh M, Kawamori T, Sugimura T, Wakabayashi K. Altered expression of beta-catenin, inducible nitric oxide synthase and cyclooxygenase-2 in azoxymethane-induced rat colon carcinogenesis. Carcinogenesis. 2000;21(7):1319–27.

242. Takahashi M, Nakatsugi S, Sugimura T, Wakabayashi K. Frequent mutations of the beta-catenin gene in mouse colon tumors induced by azoxymethane. Carcinogenesis. 2000;21(6):1117–20.

243. Takahashi M, Sugimura T, Wakabayashi K. β-Catenin Is Frequently Mutated and Demonstrates Altered Cellular Location in Azoxymethane-induced Rat Colon Tumors. Cancer Resear. 1998;58:42–6.

244. Hollstein M, Sidransky D, Vogelstein B, Harris CC. P53 Mutations in Human Cancers. Science (80-). 1991;253(5015):49–53.

245. Hiroko Ohgaki, Gordon C. Hard, Norio Hirota, Akihiko Maekawa MT and PK. Selective Mutation of Codons 204 and 213 of the p53 Gene in Rat Tumors Induced by Alkylating N-Nitroso Compounds. Cancer Res. 1992;52(10):2995–8.

246. Gottschalg E, Scott GB, Burns PA, Shuker DEG. Potassium diazoacetate-induced p53 mutations in vitro in relation to formation of O6-carboxymethyl and O6-methyl-2-deoxyguanosine DNA adducts: Relevance for gastrointestinal

cancer. Carcinogenesis. 2007;28(2):356–62.

247. Jackson PE, Cooper DP, Meyer TA, Wood M, Povey AC, Margison GP. Formation and persistence of O6-methylguanine in the mouse colon following treatment with 1,2-dimethylhydrazine as measured by an O6- alkylguanine-DNA alkyltransferase inactivation assay. Toxicol Lett. 2000;115(3):205–12.

248. van Wezel T, Stassen AP, Moen CJ, Hart AA, van der Valk MA, Demant P. Gene interaction and single gene effects in colon tumour susceptibility in mice. Nat Genet. 1996;14(4):468–70.

249. Moen CJ, Groot PC, Hart AA, Snoek M, Demant P. Fine mapping of colon tumor susceptibility (Scc) genes in the mouse, different from the genes known to be somatically mutated in colon cancer. Proc Natl Acad Sci U S A. 1996;93(3):1082–6.

250. Jacoby RF, Hohman C, Marshall DJ, Frick TJ, Schlack S, Broda M, et al. Genetic analysis of colon cancer susceptibility in mice. Genomics. 1994;22(2):381–7.

251. Povey AC, Badawi AF, Cooper DP, Nicholas Hall C, Harrison KL, Jackson PE, et al. DNA alkylation and repair in the large bowel: Animal and human studies. J Nutr. 2002;132(11 SUPPL.):3518–21.

252. Jackson PE, Cooper DP, O'Connor PJ, Povey a C. The relationship between 1,2-dimethylhydrazine dose and the induction of colon tumours: tumour development in female SWR mice does not require a K-ras mutational event. Carcinogenesis. 1999;20(3):509–13.

253. Bussche J Vanden, Moore SA, Pasmans F, Kuhnle GGC, Vanhaecke L. An approach based on ultra-high pressure liquid chromatography–tandem mass spectrometry to quantify O6-methyl and O6-carboxymethylguanine DNA adducts in intestinal cell lines. J Chromatogr A. 2012;1257:25–33.

254. Da Pieve C, Sahgal N, Moore SA, Velasco-Garcia MN. Development of a liquid chromatography/tandem mass spectrometry method to investigate the presence

of biomarkers of DNA damage in urine related to red meat consumption and risk of colorectal cancer. Rapid Commun Mass Spectrom. 2013;27(21):2493–503.

255. Moore SA, Xeniou O, Zeng ZT, Humphreys E, Burr S, Gottschalg E, et al. Optimizing immunoslot blot assays and application to low DNA adduct levels using an amplification approach. Anal Biochem [Internet]. 2010;403(1–2):67–73. Available from: http://dx.doi.org/10.1016/j.ab.2010.04.015

256. Povey AC, Hall CN, Badawi AF, Cooper DP, Guppy MJ, Jackson PE, et al. Host determinants of DNA alkylation and DNA repair activity in human colorectal tissue: O6-methylguanine levels are associated with GSTT1 genotype and O6-alkylguanine-DNA alkyltransferase activity with CYP2D6 genotype. Mutat Res - Genet Toxicol Environ Mutagen. 2001;495(1–2):103–15.

257. Povey AC, Hall CN, Badawi AF, Cooper DP, O'Connor PJ. Elevated levels of the pro-carcinogenic adduct, O(6)-methylguanine, in normal DNA from the cancer prone regions of the large bowel. Gut. 2000;47:362–5.

258. Hall CN, Badawi AF, O 'connor PJ, Saflhill R. The detection of alkylation damage in the DNA of human gastrointestinal tissues. Br J Cancer. 1991;64:59–63.

259. Jackson PE, Hall CN, Badawi AF, O'Connor PJ, Cooper DP, Povey AC. Frequency of Ki-ras mutations and DNA alkylation in colorectal tissue from individuals living in Manchester. Mol Carcinog. 1996;16(1):12–9.

260. Domenico Palli, Calogero Saieva, Claudio Coppi, Giuseppe Del Giudice, Cinzia Magagnotti, Gabriella Nesi, Federica Orsi and LA. O6-Alkylguanines, Dietary N-Nitroso Compounds, and Their Precursors in Gastric Cancer. Nutr Cancer. 2001;39(1):42–9.

261. Phillips DH, Arlt VM. The 32P-postlabeling assay for DNA adducts. Nat Protoc. 2007;2(11):2772–81.

262. Cupid BC, Zeng Z, Singh R, Shuker DEG. Detection of O6-Carboxymethyl-2-deoxyguanosine in DNA Following Reaction of Nitric Oxide with Glycine and in Human Blood DNA Using a Quantitative Immunoslot Blot Assay. Chem Res

Toxicol. 2004;17(3):294–300.

263. Tan SL, Gerber JP, Cosgrove LJ, Lockett TJ, Clarke JM, Williams DB, et al. Is the tissue persistence of O6-methyl-2-deoxyguanosine an indicator of tumour formation in the gastrointestinal tract? Mutat Res Toxicol Environ Mutagen. 2011;721(2):119–26.

264. Westman-Brinkmalm A, Brinkmalm G. A Mass Spectrometer's Building Blocks. In: Ekman R, Silberring J, Westman-Brinkmalm A, Kraj A, editors. Mass Spectrometry: Instrumentation, Interpretation, and Applications. New Jersey: John Wiley & Sons; 2008. p. 15–87.

265. Lewis JK, Wei J, Siuzdak G. Matrix-assisted Laser Desorption / Ionization Mass Spectrometry in Peptide and Protein Analysis. In: Encyclopedia of Analytical Chemistry. Chichester: John Wiley & Sons; 2000. p. 5880–94.

266. Hoffmann E De, Stroobant V. Mass Spectrometry Principles and Applications. second. Mass spectrometry principles and applications. West Sussex, UK: John Wiley & Sons, Ltd.; 2007. 43-55 p.

267. Cotter RJ. Time-of-flight mass spectrometry: an increasing role in the life sciences. Biomed Environ Mass Spectrom. 1989 Aug;18(8):513–32.

268. Christmann M, Tomicic MT, Roos WP, Kaina B. Mechanisms of human DNA repair: An update. Toxicology. 2003;193(1–2):3–34.

269. Pegg AE, Dolan ME, Moschel RC. Structure, Function, and Inhibition of O6-Alkylguanine-DNA Alkyltransferase. Prog Nucleic Acid Res Mol Biol. 1995;51(C):167–223.

270. Xu-welliver M, Pegg AE. Degradation of the alkylated form of the DNA repair protein, O6-alkylguanine-DNA alkyltransferase. Carcinogenesis. 2002;23(5):823–30.

271. Kaina B, Christmann M, Naumann S, Roos WP. MGMT: Key node in the battle against genotoxicity, carcinogenicity and apoptosis induced by alkylating agents.

DNA Repair (Amst). 2007;6(8):1079–99.

272. Fang Q, Kanugula S, Pegg AE. Function of domains of human O6-alkylguanine-DNA alkyltransferase. Biochemistry. 2005;44(46):15396–405.

273. Natarajan AT, Vermeulen S, Darroudi F, Valentine MB, Brent TP, Mitra S, et al. Chromosomal localization of human O6-methylguanine-DNA methyltransferase (MGMT) gene by in situ hybridization. Mutagenesis. 1992;7(1):83–5.

274. Nakatsu Y, Hattori K, Hayakawa H, Shimizu K, Sekiguchi M. Organization and expression of the human gene for O6-methylguanine-DNA methyltransferase. Mutat Res. 1993;293(2):119–32.

275. Pegg AE. Repair of O6-alkylguanine by alkyltransferases. Mutat Res - Rev Mutat Res. 2000;462(2–3):83–100.

276. Baylin SB, Herman JG. DNA hypermethylation in tumorigenesis: Epigenetics joins genetics. Trends Genet. 2000;16(4):168–74.

277. Tano K, Shiota S, Collier J, Foote RS, Mitra S. Isolation and structural characterization of a cDNA clone encoding the human DNA repair protein for O6-alkylguanine. Proc Natl Acad Sci U S A. 1990;87(2):686–90.

278. Daniels DS, Mol CD, Arvai AS, Kanugula S, Pegg AE, Tainer JA. Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical base binding. EMBO J. 2000;19(7):1719–30.

279. Tubbs JL, Pegg AE, Tainer JA. DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by O6-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. DNA Repair (Amst). 2007;6(8):1100–15.

280. Rasimas JJ, Kanugula S, Dalessio PM, Ropson IJ, Fried MG, Pegg AE. Effects of zinc occupancy on human O6-alkylguanine-DNA alkyltransferase. Biochemistry. 2003;42:980–90.

281. Zang H, Fang Q, Pegg AE, Guengerich FP. Kinetic Analysis of Steps in the

Repair of Damaged DNA by Human O6-Alkylguanine-DNA Alkyltransferase. J Biol Chem. 2005;280(35):30873–81.

282. Daniels DS, Woo TT, Luu KX, Noll DM, Clarke ND, Pegg AE, et al. DNA binding and nucleotide flipping by the human DNA repair protein AGT. Nat Struct Mol Biol. 2004;11(8):714–20.

283. Liu L, Xu-welliver M, Kanugula S, Pegg AE. Inactivation and Degradation of O6-Alkylguanine-DNA Alkyltransferase after Reaction with Nitric Oxide 1. Cancer Res. 2002;62:3037–43.

284. Fried MG, Kanugula S, Bromberg JL, Pegg AE. DNA binding mechanism of O6-alkylguanine-DNA alkyltransferase: Stoichiometry and effects of DNA base composition and secondary structure on complex stability. Biochemistry. 1996;35(48):15295–301.

285. Melikishvili M, Rasimas JJ, Pegg AE, Fried MG. Interactions of human O6-alkylguanine-DNA alkyltransferase (AGT) with short double-stranded DNAs. Biochemistry. 2008;47(52):13754–63.

286. Tessmer I, Fried MG. Insight into the cooperative DNA binding of the O6-alkylguanine DNA alkyltransferase. DNA Repair (Amst). 2014;20:14–22.

287. Batts ED, Maisel C, Kane D, Liu L, Fu P, O'Brien T, et al. O6-benzylguanine and BCNU in multiple myeloma: A phase II trial. Cancer Chemother Pharmacol. 2007;60(3):415–21.

288. Kotandeniya D, Murphy D, Seneviratne U, Guza R, Pegg A, Kanugula S, et al. Mass spectrometry based approach to study the kinetics of O6-alkylguanine DNA alkyltransferase-mediated repair of O6-pyridyloxobutyl-2'-deoxyguanosine adducts in DNA. Chem Res Toxicol. 2011;24:1966–75.

289. Ranson M, Middleton MR, Bridgewater J, Lee SM, Dawson M, Jowle D, et al. Lomeguatrib, a potent inhibitor of O6-alkylguanine-DNA-alkyltransferase: phase I safety, pharmacodynamic, and pharmacokinetic trial and evaluation in combination with temozolomide in patients with advanced solid tumors. Clin

Cancer Res. 2006;12:1577–84.

290. Coulter R, Blandino M, Tomlinson JM, Pauly GT, Krajewska M, Moschel RC, et al. Differences in the rate of repair of O6-alkylguanines in different sequence contexts by O6-alkylguanine-DNA alkyltransferase. Chem Res Toxicol. 2007;20(12):1966–71.

291. Guza R, Pegg AE, Tretyakova N. Effects of Sequence Context on O6-Alkylguanine DNA Alkyltransferase Repair of O6-Alkyl-Deoxyguanosine Adducts. In: Structural Biology of DNA Damage and Repair. American Chemical Society; 2010. p. 73–101. (ACS Symposium Series; vol. 1041).

292. Rasimas JJ, Kar SR, Pegg AE, Fried MG. Interactions of Human O6 - Alkylguanine-DNA Alkyltransferase (AGT) with Short Single-stranded DNAs. J Biol Chem. 2007;282(5):3357–66.

293. Pegg AE, Kanugula S, Loktionova NA. O6-alkylguanine-DNA alkyltransferase. Vol. 6, Current Cancer Research. 2011. p. 321–43.

294. Shibata T, Glynn N, McMurry TBH, McElhinney RS, Margison GP, Williams DM. Novel synthesis of O6-alkylguanine containing oligodeoxyribonucleotides as substrates for the human DNA repair protein, O6-methylguanine DNA methyltransferase (MGMT). Nucleic Acids Res. 2006;34(6):1884–91.

295. Terashima I, Kawate H, Sakumi K, Sekiguchi M, Kohda K. Substrate specificity of human O6-methylguanine-DNA methyltransferase for O6-benzylguanine derivatives in oligodeoxynucleotides. Chem Res Toxicol. 1997;10(11):1234–9.

296. Senthong P, Millington CL, Wilkinson OJ, Marriott AS, Watson AJ, Reamtong O, et al. The nitrosated bile acid DNA lesion O6-carboxymethylguanine is a substrate for the human DNA repair protein O6-methylguanine-DNA methyltransferase. Nucleic Acids Res. 2013;41(5):3047–55.

297. McElhinney RS, McMurry TBH, Margison GP. O6-alkylguanine-DNA alkyltransferase inactivation in cancer chemotherapy. Mini Rev Med Chem. 2003;3(5):471–85.

298. Crone TM, Goodtzova K, Pegg AE. Amino acid residues affecting the activity and stability of human O6-alkylguanine-DNA alkyltransferase. Mutat Res - DNA Repair. 1996;363(1):15–25.

299. Grafstrom RC, Pegg AE, Trump BF, Harris CC. O6-alkylguanine-DNA alkyltransferase activity in normal human tissues and cells. Cancer Res. 1984;44(0008–5472):2855–7.

300. Povey AC, Hall CN, Cooper DP, O'Connor PJ, Margison GP. Determinants of O(6)-alkylguanine-DNA alkyltransferase activity in normal and tumour tissue from human colon and rectum. Int J Cancer. 2000;85(1):68–72.

301. Christmann M, Verbeek B, Roos WP, Kaina B. O(6)-Methylguanine-DNA methyltransferase (MGMT) in normal tissues and tumors: enzyme activity, promoter methylation and immunohistochemistry. Biochim Biophys Acta. 2011 Dec;1816(2):179–90.

302. Margison GP, Povey AC, Kaina B, Santibáñez Koref MF. Variability and regulation of O6-alkylguanine-DNA alkyltransferase. Carcinogenesis. 2003;24(4):625–35.

303. Harris LC, Remack JS, Houghton PJ, Brent TP. Wild-type p53 suppresses transcription of the human O6-methylguanine-DNA methyltransferase gene. Cancer Res. 1996;56(9):2029–32.

304. Grombacher T, Eichhorn U, Kaina B. p53 is involved in regulation of the DNA repair gene O6-methylguanine-DNA methyltransferase (MGMT) by DNA damaging agents. Oncogene. 1998;17:845–51.

305. Lee S, Hwang KS, Lee HJ, Kim J-S, Kang GH. Aberrant CpG island hypermethylation of multiple genes in colorectal neoplasia. Lab Investig. 2004;84(7):884–93.

306. Esteller M, Hamilton SR, Burger PC, Baylin SB, Herman JG. Inactivation of the DNA repair gene O(6)-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia. Cancer Res.

1999;59(4):793–7.

307. Lind GE, Thorstensen L, Løvig T, Meling GI, Hamelin R, Rognum TO, et al. A CpG island hypermethylation profile of primary colorectal carcinomas and colon cancer cell lines. Mol Cancer. 2004;3(28):1–11.

308. Esteller M, Toyota M, Sanchez-Cespedes M, Capella G, Peinado MA, Watkins DN, et al. Inactivation of the DNA repair gene O6-Methylguanine-DNA methyltransferase by promoter hypermethylation is associated with G to A mutations in K-ras in colorectal tumorigenesis. Cancer Res. 2000;60(9):2368–71.

309. Shen L, Kondo Y, Rosner GL, Xiao L, Hernandez NS, Vilaythong J, et al. MGMT promoter methylation and field defect in sporadic colorectal cancer. J Natl Cancer Inst. 2005;97(18):1330–8.

310. Jackson PE, Hall CN, O'Connor PJ, Cooper DP, Margison GP, Povey AC. Low O6-alkylguanine DNA-alkyltransferase activity in normal colorectal tissue is associated with colorectal tumours containing a GC-->AT transition in the K-ras oncogene. Carcinogenesis. 1997;18(7):1299–302.

311. Zaidi NH, Pretlow TP, O'riordan MA, Dumenco LL, Allay E, Gerson SL. Transgenic expression of human mgmt protects against azoxymethane-induced aberrant crypt foci and g to a mutations in the k-ras oncogene of mouse colon. Carcinogenesis. 1995;16(3):451–6.

312. Margison GP, Butt A, Pearson SJ, Wharton S, Watson AJ, Marriott A, et al. Alkyltransferase-like proteins. DNA Repair (Amst). 2007;6(8):1222–8.

313. Pearson SJ, Ferguson J, Santibanez-Koref M, Margison GP. Inhibition of O6-methylguanine-DNA methyltransferase by an alkyltransferase-like protein from Escherichia coli. Nucleic Acids Res. 2005;33(12):3837–44.

314. Pearson SJ, Wharton S, Watson AJ, Begum G, Butt A, Glynn N, et al. A novel DNA damage recognition protein in Schizosaccharomyces pombe. Nucleic Acids Res. 2006;34(8):2347–54.

315. Tubbs JL, Latypov V, Kanugula S, Butt A, Melikishvili M, Kraehenbuehl R, et al. Alkylated DNA damage flipping bridges base and nucleotide excision repair. Nature. 2009;459(7248):808–13.

316. Morita R, Nakagawa N, Kuramitsu S, Masui R. An O6-methylguanine-DNA methyltransferase-like protein from thermus Thermophilus interacts with a nucleotide excision repair protein. J Biochem. 2008;144(2):267–77.

317. Aramini JM, Tubbs JL, Kanugula S, Rossi P, Ertekin A, Maglaqui M, et al. Structural basis of O6-alkylguanine recognition by a bacterial alkyltransferase-like DNA repair protein. J Biol Chem. 2010;285(18):13736–41.

318. Krokan HE, Bjørås M. Base excision repair. Cold Spring Harb Perspect Biol. 2013;5(4):1–22.

319. Doetsch PW, Cunningham RP. The enzymology of apurinic/apyrimidinic endonucleases. Mutat Res Repair. 1990;236(2–3):173–201.

320. Abbotts R, Madhusudan S. Human AP endonuclease 1 (APE1): From mechanistic insights to druggable target in cancer. Cancer Treat Rev. 2010;36(5):425–35.

321. Robertson AB, Klungland A, Rognes T, Leiros I. Base excision repair: The long and short of it. Cell Mol Life Sci. 2009;66(6):981–93.

322. Sedgwick B. Repairing DNA-methylation damage. Nat Rev Mol Cell Biol. 2004;5(2):148–57.

323. Wirtz S, Nagel G, Eshkind L, Neurath MF, Samson LD, Kaina B. Both base excision repair and O6-methylguanine-DNA methyltransferase protect against methylation-induced colon carcinogenesis. Carcinogenesis. 2010;31(12):2111–7.

324. Friedberg EC. How nucleotide excision repair protects against cancer. Nat Rev Cancer. 2001;1(1):22–33.

325. Laat WL de, Jaspers NGJ, Hoeijmakers JHJ. Molecular mechanism of nucleotide excision repair. Genes Dev. 1999;13(7):768–85.

326. Sugasawa K. Regulation of damage recognition in mammalian global genomic nucleotide excision repair. Mutat Res - Fundam Mol Mech Mutagen. 2010;685(1–2):29–37.

327. Bol SAM, Van Steeg H, Van Oostrom CTM, Tates AD, Vrieling H, De Groot AJL, et al. Nucleotide excision repair modulates the cytotoxic and mutagenic effects of N-n-butyl-N-nitrosourea in cultured mammalian cells as well as in mouse splenocytes in vivo. Mutagenesis. 1999;14(3):317–22.

328. Abbotts R, Thompson N, Madhusudan S. DNA repair in cancer: Emerging targets for personalized therapy. Cancer Manag Res. 2014;6(1):77–92.

329. Martin SA, Hewish M, Lord CJ, Ashworth A. Genomic instability and the selection of treatments for cancer. J Pathol. 2010;220(2):281–9.

330. Li G-M. Mechanisms and functions of DNA mismatch repair. Cell Res. 2008;18(1):85–98.

331. Mahaney BL, Meek K, Lees-Miller SP. Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining. Biochem J. 2009;417(3):639–50.

332. Rasouli-Nia A, Sigbhat-Ullah, Mirzayans R, Paterson MC, Day RS. On the quantitative relationship between O6-ethylguanine residues, in genomic DNA and production of sister-chromatid exchanges, mutations and lethal events in a Mer- human tumor cell line. Mutat Res Repair. 1994;314(2):99–113.

333. Difilippantonio MJ, Petersen S, Chen HT, Johnson R, Jasin M, Kanaar R, et al. Evidence for Replicative Repair of DNA Double-Strand Breaks Leading to Oncogenic Translocation and Gene Amplification. J Exp Med. 2002;196(4):469–80.

334. Kass EM, Jasin M. Collaboration and competition between DNA double-strand

break repair pathways. FEBS Lett. 2010;584(17):3703–8.

335. Mladenov E, Iliakis G. Induction and repair of DNA double strand breaks: The increasing spectrum of non-homologous end joining pathways. Mutat Res - Fundam Mol Mech Mutagen. 2011;711(1–2):61–72.

336. Sargent RG, Brenneman M a, Wilson JH. Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination. Mol Cell Biol. 1997;17(1):267–77.

337. Landi S, Gemignani F, Moreno V, Gioia-Patricola L, Chabrier A, Guino E, et al. A comprehensive analysis of phase I and phase II metabolism gene polymorphisms and risk of colorectal cancer. Pharmacogenet Genomics. 2005;15(8):535–46.

338. Ishibe N, Sinha R, Hein DW, Kulldorff M, Strickland P, Fretland AJ, et al. Genetic polymorphisms in heterocyclic amine metabolism and risk of colorectal adenomas. Pharmacogenetics. 2002;12(2):145–50.

339. Macfarlane GT, Macfarlane S. Bacteria, colonic fermentation, and gastrointestinal health. J AOAC Int. 2012;95(1):50–60.

340. Watson AJ, Margison GP. O6-Alkylguanine-DNA Alkyltransferase Assay. In: Methods in Molecular Biology. 2000. p. 49–61.

341. Mijal RS, Thomson NM, Fleischer NL, Pauly GT, Moschel RC, Kanugula S, et al. The Repair of the Tobacco Specific Nitrosamine Derived Adduct O6-[4-Oxo-4-(3-pyridyl)butyl]guanine by O6-Alkylguanine-DNA Alkyltransferase Variants. Chem Res Toxicol. 2004;17(3):424–34.

342. Flick TG, Cassou CA, Chang TM, Williams ER. Solution Additives that Desalt Protein Ions in Native Mass Spectrometry. Anal Chem. 2012;84(17):7511–17.

343. Iavarone AT, Udekwu OA, Wiliams ER. Buffer loading for counteracting metal salt-induced signal suppression in electrospray ionization. Anal Chem. 2004;76(14):3944–50.

344. Sterling HJ, Batchelor JD, Wemmer DE, Williams ER. Effects of buffer loading for electrospray ionization mass spectrometry of a noncovalent protein complex that requires high concentrations of essential salts. J Am Soc Mass Spectrom. 2010;21(6):1045–9.

345. Mirza UA, Chait BT. Effects of anions on the positive ion electrospray ionization mass spectra of peptides and proteins. Anal Chem. 1994;66(18):2898–904.

346. Tang L, Kebarle P. Dependence of ion intensity in electrospray mass spectrometry on the concentration of the analytes in the electrosprayed solution. Anal Chem. 1993;65(24):3654–68.

347. Pan P, Mcluckey S a. The Effect of Small Cations on the Positive Electrospray Responses of Proteins at Low pH Solutions consisting of protein and small molecule mix- study the influence of small molecule / cation components. Anal Chem. 2003;75(20):5468–74.

348. Pan P, Gunawardena HP, Xia Y, McLuckey S a. Nanoelectrospray ionisation of protein mixtures: solution pH and protein pI. Anal Chem. 2004;76(4):1165–74.

349. Sterling HJ, Cassou CA, Susa AC, Williams ER. Electrothermal supercharging of proteins in native electrospray ionization. Anal Chem. 2012;84(8):3795–801.

350. Cassou CA, Sterling HJ, Susa AC, Williams ER. Electrothermal supercharging in mass spectrometry and tandem mass spectrometry of native proteins. Anal Chem. 2013;85(1):138–46.

351. Bignami M, O'Driscoll M, Aquilina G, Karran P. Unmasking a killer: DNA O6-methylguanine and the cytotoxicity of methylating agents. Mutat Res - Rev Mutat Res. 2000;462(2–3):71–82.

352. DeMarini DM. Genotoxicity of tobacco smoke and tobacco smoke condensate: A review. Mutat Res - Rev Mutat Res. 2004;567(2–3 SPEC. ISS.):447–74.

353. Gerson SL. MGMT: its role in cancer aetiology and cancer therapeutics. Nat Rev Cancer. 2004;4(4):296–307.

354. Margison GP, Santibáñez-Koref MF. O6-alkylguanine-DNA alkyltransferase: Role in carcinogenesis and chemotherapy. BioEssays. 2002;24(3):255–66.

355. Julie L. and Tainer JAT. Alkyltransferase-like proteins: Molecular switches between DNA repair pathways. Cell Mol Life Sci . 2010;67(22):3749–62.

356. Chen C-S, Korobkova E, Chen H, Zhu J, Jian X, Tao S-C, et al. A proteome chip approach reveals new DNA damage recognition activities in Escherichia coli. Nat Methods. 2007;5(1):69–74.

357. Wilkinson OJ, Latypov V, Tubbs JL, Millington CL, Morita R, Blackburn H, et al. Alkyltransferase-like protein (Atl1) distinguishes alkylated guanines for DNA repair using cation- interactions. Proc Natl Acad Sci. 2012;109(46):18755–60.

358. Wibley JE, Pegg a E, Moody PC. Crystal structure of the human O6-alkylguanine-DNA alkyltransferase. Nucleic Acids Res. 2000;28(2):393–401.

359. Moore MH, Gulbis JM, Dodson EJ, Demple B, Moody PC. Crystal structure of a suicidal DNA repair protein: the Ada O6-methylguanine-DNA methyltransferase from E. coli. EMBO J. 1994;13(7):1495–501.

360. Rasimas JJ, Pegg AE, Fried MG. DNA-binding mechanism of O6-alkylguanine-DNA alkyltransferase: Effects of protein and DNA alkylation on complex stability. J Biol Chem. 2003;278(10):7973–80.

361. Fang Q, Noronha AM, Murphy SP, Wilds CJ, Tubbs JL, Tainer JA, et al. Repair of O6-G-alkyl-O6-G interstrand cross-links by human O6-alkylguanine-DNA alkyltransferase. Biochemistry. 2008;47(41):10892–903.

362. Watson AJ, Middleton MR, McGown G, Thorncroft M, Ranson M, Hersey P, et al. O6-methylguanine-DNA methyltransferase depletion and DNA damage in patients with melanoma treated with temozolomide alone or with lomeguatrib. Br J Cancer. 2009;100(8):1250–6.

363. Sekiguchi M, Nakabeppu Y, Sakumi K, Tuzuki T. DNA-repair methyltransferase as a molecular device for preventing mutation and cancer. J Cancer Res Clin

Oncol. 1996;122(4):199–206.

364. Sekiguchi M, Sakumi K. Roles of DNA repair methyltransferase in mutagenesis and carcinogenesis. Jpn J Hum Genet. 1997;42(3):389–99.

365. Morimoto K, Dolan ME, Scicchitano D, Pegg AE. Repair of O6-propylguanine and O6-butylguanine in DNA by O6-alkylguanine-DNA alkyltransferases from rat liver and E. coli. Carcinogenesis. 1985;6(7):1027–31.

366. Unimod protein modifications for mass spectrometry [Internet]. Unimod, View record [Accession #: 6]. 2017 [cited 2017 Jul 21]. Available from: http://www.unimod.org/modifications_view.php?editid1=6

367. Unimod protein modifications for mass spectrometry [Internet]. Unimod; View record [Accession #: 4]. 2017 [cited 2017 Sep 5]. Available from: http://www.unimod.org/modifications_view.php?editid1=4

368. Unimod protein modifications for mass spectrometry [Internet]. Unimod, View record [Accession #: 378]. 2017 [cited 2017 Jul 21]. Available from: http://www.unimod.org/modifications_view.php?editid1=378

369. Unimod protein modifications for mass spectrometry [Internet]. Unimod, View record [Accession #: 24]. 2017 [cited 2017 Oct 24]. Available from: http://www.unimod.org/modifications_view.php?editid1=24

370. Feng CH, Lu CY. Modification of major plasma proteins by acrylamide and glycidamide: Preliminary screening by nano liquid chromatography with tandem mass spectrometry. Anal Chim Acta. 2011;684(1–2):89–95.

371. Yang H, Zubarev RA. Mass spectrometric analysis of asparagine deamidation and aspartate isomerization in polypeptides. Electrophoresis. 2010;31(11):1764–72.

372. Meyer AS, McCain MD, Fang Q, Pegg AE, Spratt TE. O6-Alkylguanine-DNA Alkyltransferases Repair O6-Methylguanine in DNA with Michaelis-Menten-like Kinetics. Chem Res Toxicol. 2003;16(11):1405–9.

373. Delaney JC, Essigmann JM. Effect of sequence context on O6-methylguanine repair and replication in vivo. Biochemistry. 2001;40(49):14968–75.

374. Cross AJ, Pollock JRA, Bingham SA. Haem, not protein or inorganic iron, is responsible for endogenous intestinal N-nitrosation arising from red meat. Cancer Res. 2003;63(10):2358–60.

375. Harrison KL, Jukes R, Cooper DP, Shuker DEG. Detection of Concomitant Formation of and O 6 -Methyl-2′-deoxyguanosine in DNA Exposed to Nitrosated Glycine Derivatives Using a Combined Immunoaffinity / HPLC Method O6 - Carboxymethyl-. Reactions. 1999;(7):106–11.

376. Gillette MA, Carr S a. Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. Nat Methods. 2012;10(1):28–34.

377. Achour B, Russell MR, Barber J, Rostami-Hodjegan A. Simultaneous quantification of the abundance of several cytochrome P450 and uridine 5???-diphospho-glucuronosyltransferase enzymes in human liver microsomes using multiplexed targeted proteomics. Drug Metab Dispos. 2014;42(4):500–10.

378. Hubner NC, Bird AW, Cox J, Splettstoesser B, Bandilla P, Poser I, et al. Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. J Cell Biol. 2010;189(4):739–54.

379. Collins BC, Gillet LC, Rosenberger G, Röst HL, Vichalkovski A, Gstaiger M, et al. Quantifying protein interaction dynamics by SWATH mass spectrometry: application to the 14-3-3 system. Nat Methods. 2013;10(12):1246–53.

380. Lambert J-P, Ivosev G, Couzens AL, Larsen B, Taipale M, Lin Z-Y, et al. Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. Nat Methods. 2013;10(12):1239–45.

381. Turriziani B, Garcia-Munoz A, Pilkington R, Raso C, Kolch W, von Kriegsheim A. On-Beads Digestion in Conjunction with Data-Dependent Mass Spectrometry: A Shortcut to Quantitative and Dynamic Interaction Proteomics. Biology (Basel).

2014;3(2):320–32.

382. Nilsson T, Mann M, Aebersold R, Yates JR, Bairoch A, Bergeron JJM. Mass spectrometry in high-throughput proteomics: ready for the big time. Nat Methods. 2010;7(9):681–5.

383. Domon B, Aebersold R. Options and considerations when selecting a quantitative proteomics strategy. Nat Biotechnol. 2010;28(7):710–21.

384. Wong SS, Jameson DM. Chemistry of Protein and Nucleic Acid Cross-Linking and Conjugation, Second Edition. London: CRC Press; 2011. 37-38 p.

385. Brewer F, Riehm P. Evidence for possible nonspecific reactions between N-ethylmaleimide and proteins. Anal Biochem. 1967;18:248–55.

386. Siegel RL, Miller KD, Jemal A. Cancer statistics. CA Cancer J Clin. 2016;66(1):7–30.

387. Johnson IT, Lund EK. Review article: Nutrition, obesity and colorectal cancer. Aliment Pharmacol Ther. 2007;26(2):161–81.

388. Lynch HT, Lynch JF, Lynch PM, Attard T. Hereditary colorectal cancer syndromes: Molecular genetics, genetic counseling, diagnosis and management. Fam Cancer. 2008;7(1):27–39.

389. Song M, Garrett WS, Chan AT. Nutrients, foods, and colorectal cancer prevention. Gastroenterology. 2015;148(6):1244–60.

390. Willett WC. Diet, nutrition, and avoidable cancer. Environ Health Perspect. 1995;103(SUPPL. 8):165–70.

391. Potter JD. Nutrition and colorectal cancer. Cancer Causes Control. 1996;7(1):127–46.

392. Lees NP, Harrison KL, Hall CN, Margison GP, Povey AC. Reduced MGMT activity in human colorectal adenomas is associated with K-ras GC???AT

transition mutations in a population exposed to methylating agents. Carcinogenesis. 2004;25(7):1243–7.

393. Lundberg JO, Weitzberg E. Biology of nitrogen oxides in the gastrointestinal tract. Gut. 2013;62(4):616–29.

394. Kalač P. Health effects and occurrence of dietary polyamines: A review for the period 2005-mid 2013. Food Chem. 2014;161:27–39.

395. Larqué E, Sabater-Molina M, Zamora S. Biological significance of dietary polyamines. Nutrition. 2007;23(1):87–95.

396. Prester L. Biogenic amines in fish, fish products and shellfish: a review. Food Addit Contam Part A. 2011;28(11):1547–60.

397. Al Bulushi I, Poole S, Deeth HC, Dykes G a. Biogenic amines in fish: roles in intoxication, spoilage, and nitrosamine formation--a review. Crit Rev Food Sci Nutr. 2009;49(4):369–77.

398. Ruiz-Capillas C, Jiménez-Colmenero F. Biogenic amines in meat and meat products. Crit Rev Food Sci Nutr. 2004;44(7–8):489–599.

399. Mayer HK, Fiechter G, Fischer E. A new ultra-pressure liquid chromatography method for the determination of biogenic amines in cheese. J Chromatogr A. 2010;1217(19):3251–7.

400. Anli, R.E, Bayram M. Biogenic amines in wines. Food Rev Int. 2009;25(1):86–102.

401. Smit AY, du Toit WJ, du Toit M. Biogenic amines in wine: Understanding the headache. South African J Enol Vitic. 2008;29(2):109–27.

402. Zaborskieke G, Garmiene G, Jasutiene I, Bundiniene O, Jankauskiene J. the Influence of Environmental Conditions on Biogenic Amines, Nitrogen, Nitrite and Nitrate Content in Leafy Vegetables. Vet Ir Zootech. 2009;45(67):97–103.

403. De Las Rivas B, González R, Landete JM, Muñoz R. Characterization of a second ornithine decarboxylase isolated from Morganella morganii. J Food Prot. 2008;71(3):657–61.

404. Harrison KL, Jukes R, Cooper DP, Shuker DE. Detection of concomitant formation of O6-carboxymethyl- and O6-methyl-2'-deoxyguanosine in DNA exposed to nitrosated glycine derivatives using a combined immunoaffinity/HPLC method. Chem Res Toxicol. 1999;12(1):106–11.

405. Kimura H, Hokari R, Miura S, Kimura H, Hokari R, Miura S, et al. Increased expression of an inducible isoform of nitric oxide synthase and the formation of peroxynitrite in colonic mucosa of patients with active ulcerative colitis Increased expression of an inducible isoform of nitric oxide synthase and the formation o. Gut. 1998;42:180–7.

406. Hussain SP, Amstad P, Raja K, Ambs S, Nagashima M, Bennett WP, et al. Increased p53 mutation load in noncancerous colon tissue from ulcerative colitis: A cancer-prone chronic inflammatory disease. Cancer Res. 2000;60(13):3333–7.

407. Rana RR. Development and application of a fluorescent postlabelling assay for the detection of N7-alkylguanines. University of Leicester; 1997.

408. IARC. Acrylamide. Int Agency Res Cancer (IARC) Some Ind Chem IARC Monogr Eval Carcinog Risks to Humans. 1994;60(73):389–433.

409. De Bont R, van Larebeke N. Endogenous DNA damage in humans: A review of quantitative data. Mutagenesis. 2004;19(3):169–85.

410. Kumar R, Hemminki K. Separation of 7-methyl- and 7- (2-hydroxyethyl) -guanine adducts in human DNA samples using a combination of TLC and HPLC. Carcinogenesis. 1996;17(3):485–92.

411. Bolt HM, Leutbecher M, Golka K. A note on the physiological background of the ethylene oxide adduct 7-(2-hydroxyethyl)guanine in DNA from human blood. Arch Toxicol. 1997;71(11):719–21.

412. Kocadağli T, Gökmen V. Metabolism of Acrylamide in Humans and Biomarkers of Exposure to Acrylamide. In: Acrylamide in Food: Analysis, Content and Potential Health Effects. 2015. p. 109–28.

413. Tareke E, Lyn-Cook B, Robinson B, Ali SF. Acrylamide: A dietary carcinogen formed in vivo? J Agric Food Chem. 2008;56(15):6020–3.

414. Mottram DS, Wedzicha BL, Dodson AT. Food chemistry: Acrylamide is formed in the Maillard reaction. Nature. 2002;419(6906):448–9.

415. Tareke E, Heinze TM, Da Costa GG, Ali S. Acrylamide formed at physiological temperature as a result of asparagine oxidation. J Agric Food Chem. 2009;57(20):9730–3.

416. Mucci L a, Adami H-O, Wolk A. Prospective study of dietary acrylamide and risk of colorectal cancer among women. Int J Cancer. 2006;118(1):169–73.

417. Virk-Baker MK, Nagy TR, Barnes S, Groopman J. Dietary acrylamide and human cancer: a systematic review of literature. Nutr Cancer. 2014;66(5):774–90.

418. Larsson SC, Akesson A, Bergkvist L, Wolk A. Dietary acrylamide intake and risk of colorectal cancer in a prospective cohort of men. Eur J Cancer. 2009;45(4):513–6.

419. Hogervorst JGF, De Bruijn-Geraets D, Schouten LJ, Van Engeland M, De Kok TMCM, Goldbohm RA, et al. Dietary acrylamide intake and the risk of colorectal cancer with specific mutations in KRAS and APC. Carcinogenesis. 2014;35(5):1032–8.

420. Fennell TR, Sumner SCJ, Snyder RW, Burgess J, Spicer R, Bridson WE, et al. Metabolism and hemoglobin adduct formation of acrylamide in humans. Toxicol Sci. 2005;85(1):447–59.

421. Doerge DR, Gamboa Da Costa G, McDaniel LP, Churchwell MI, Twaddle NC, Beland FA. DNA adducts derived from administration of acrylamide and

glycidamide to mice and rats. Mutat Res - Genet Toxicol Environ Mutagen. 2005;580(1–2):131–41.

422. Brown K. Methods for the Detection of DNA Adducts. In: Parry J, Parry E, editors. Genetic Toxicology Methods in Molecular Biology (Methods and Protocols). New York, NY: Springer; 2012. p. 207–30.

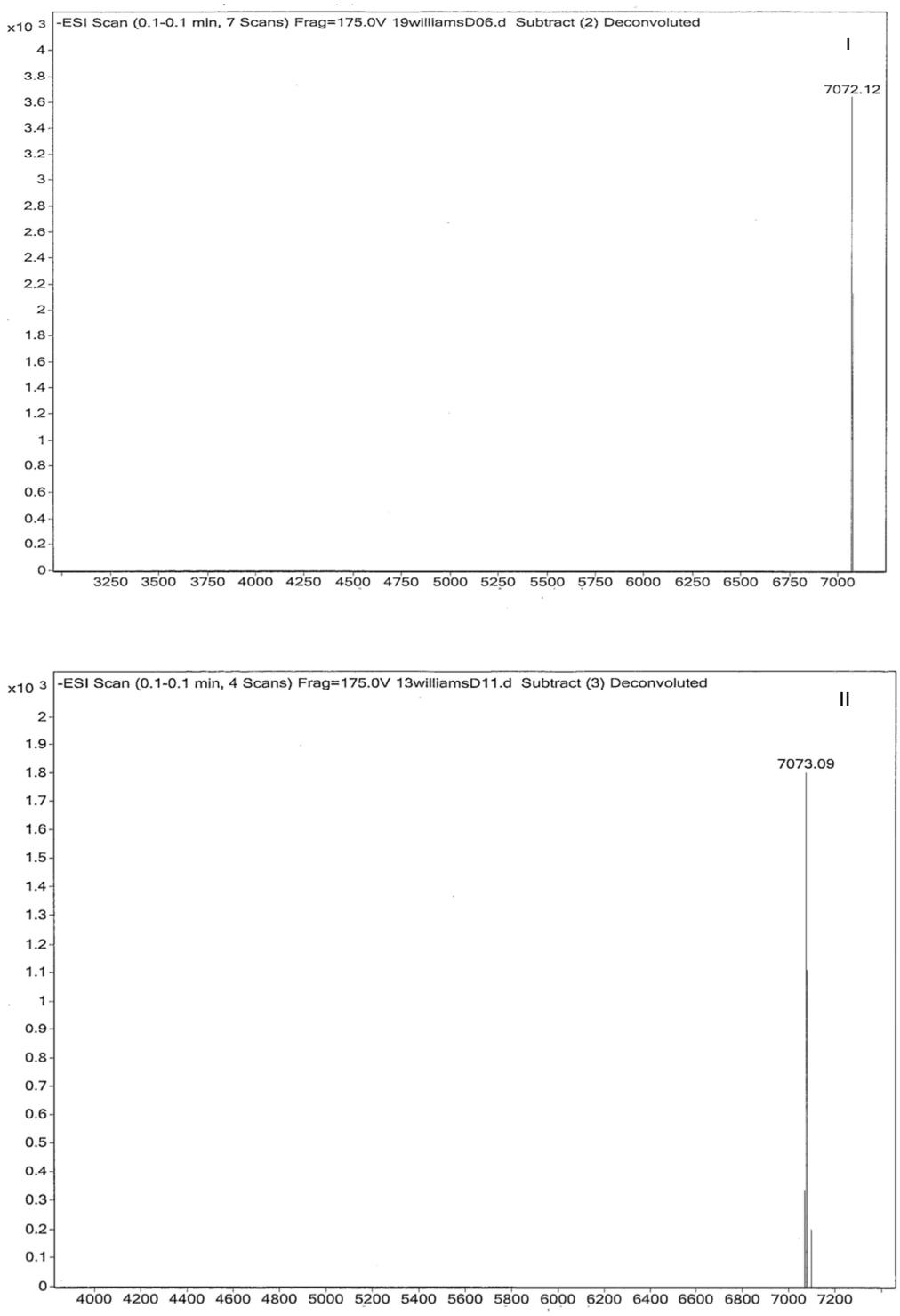
423. Bussche J Vanden, Hemeryck LY, Van Hecke T, Kuhnle GGC, Pasmans F, Moore SA, et al. O6-carboxymethylguanine DNA adduct formation and lipid peroxidation upon in vitro gastrointestinal digestion of haem-rich meat. Mol Nutr Food Res. 2014;58(9):1883–96.

Appendices

```
Query  81   ATGGACAAGGATTGTGAAATGAAACGCACCACACTGGACAGCCCTTTGGGGAAGCTGGAG  140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  10   ATGGACAAGGATTGTGAAATGAAACGCACCACACTGGACAGCCCTTTGGGGAAGCTGGAG  69

Query  141  CTGTCTGGTTGTGAGCAGGGTCTGCACGAAATAAAGCTCCTGGGCAAGGGGACGTCTGCA  200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  70   CTGTCTGGTTGTGAGCAGGGTCTGCACGAAATAAAGCTCCTGGGCAAGGGGACGTCTGCA  129

Query  201  GCTGATGCCGTGGAGGTCCCAGCCCCCGCTGCGGTTCTCGGAGGTCCGGAGCCCCTGATG  260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  130  GCTGATGCCGTGGAGGTCCCAGCCCCCGCTGCGGTTCTCGGAGGTCCGGAGCCCCTGATG  189

Query  261  CAGTGCACAGCCTGGCTGAATGCCTATTTCCACCAGCCCGAGGCTATCGAAGAGTTCCCC  320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  190  CAGTGCACAGCCTGGCTGAATGCCTATTTCCACCAGCCCGAGGCTATCGAAGAGTTCCCC  249

Query  321  GTGCCGGCTCTTCACCATCCCGTTTTCCAGCAAGAGTCGTTCACCAGACAGGTGTTATGG  380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  250  GTGCCGGCTCTTCACCATCCCGTTTTCCAGCAAGAGTCGTTCACCAGACAGGTGTTATGG  309

Query  381  AAGCTGCTGAAGGTTGTGAAATTCGGAGAAGTGATTTCTTACCAGCAATTAGCAGCCCTG  440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  310  AAGCTGCTGAAGGTTGTGAAATTCGGAGAAGTGATTTCTTACCAGCAATTAGCAGCCCTG  369

Query  441  GCAGGCAACCCCAAAGCCGCGCGAGCAGTGGGAGGAGCAATGAGAGGCAATCCTGTCCCC  500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  370  GCAGGCAACCCCAAAGCCGCGCGAGCAGTGGGAGGAGCAATGAGAGGCAATCCTGTCCCC  429

Query  501  ATCCTCATCCCGTGCCACAGAGTGGTCTGCAGCAGCGGAGCCGTGGGCAACTACTCCGGA  560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  430  ATCCTCATCCCGTGCCACAGAGTGGTCTGCAGCAGCGGAGCCGTGGGCAACTACTCCGGA  489

Query  561  GGACTGGCCGTGAAGGAATGGCTTCTGGCCCATGAAGGCCACCGGTTGGGGAAGCCAGGC  620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  490  GGACTGGCCGTGAAGGAATGGCTTCTGGCCCATGAAGGCCACCGGTTGGGGAAGCCAGGC  549

Query  621  TTGGGAGGGAGCTCAGGTCTGGCAGGGGCCTGGCTCAAGGGAGCGGGAGCTACCTCGGGC  680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  550  TTGGGAGGGAGCTCAGGTCTGGCAGGGGCCTGGCTCAAGGGAGCGGGAGCTACCTCGGGC  609

Query  681  TCCCCGCCTGCTGGCCGAAAC  701
            |||||||||||||||||||||
Sbjct  610  TCCCCGCCTGCTGGCCGAAAC  630
```

Appendix 1: Nucleotide sequence alignments of MGMT cDNA.

Nucleotide sequence alignments were generated by Blast database; as explained in Section 2.2.3.



Appendix 2: Electrospray ionization mass spectra of (I) LH1c and (II) RA1 ODNs.

This work was conducted by Centre of Chemical Biology, Department of Chemistry, University of Sheffield.

Appendix 3: Demographic characteristics of study participants indicating smoking, alcohol and medication usage.

Human sample ID	Smoking	Alcohol	Anti-histamine	Anti-muscarinic	Benzo-diazepine	Cortico-steroid	H_2-receptor antagonist	Iron supplement	Laxative	Non-opioid pain killers	Opioid pain killers
25	Non-smoker	Non-drinker	Non-user	User	Non-user	Non-user	Non-user	Non-user	Non-user	Non-user	User
31	Non-smoker	Non-drinker	Non-user	User	Non-user	Non-user	Non-user	Non-user	Non-user	Non-user	User
39	Non-smoker	Non-drinker	Non-user	User	Non-user	Non-user	Non-user	Non-user	Non-user	User	User
44	Non-smoker	Non-drinker	Non-user	Non-user	Non-user	Non-user	Non-user	User	Non-user	Non-user	Non-user
50	Non-smoker	<10units/ week	Non-user	User	User	Non-user	Non-user	Non-user	Non-user	Non-user	User
74	Non-smoker	>10 units/week	Non-user	User	Non-user	Non-user	Non-user	Non-user	User	Non-user	User
78	Non-smoker	Occasional drinker	Non-user	User	Non-user	Non-user	Non-user	Non-user	Non-user	Non-user	User
79	Smoker (1-19 cigarettes/ day)	>10 units/week	Non-user	User	Non user	User	Non user	Non-user	Non user	Non user	User