Oxidative demethylation of DNA damage by *Escherichia coli* AlkB and its human homologs ABH2 and ABH3

A thesis submitted for the degree of Ph.D.

by

Sarah Catherine Trewick

Clare Hall Laboratories

Cancer Research UK London Research Institute

South Mimms, Potters Bar

Hertfordshire, EN6 3LD

and

Department of Biochemistry

University College London

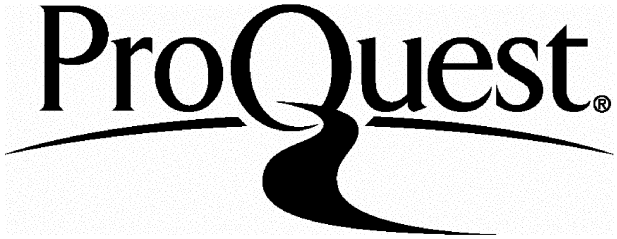
Gower Street, London, WC1E 6BT

ProQuest Number: U642489

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.



ProQuest U642489

Published by ProQuest LLC(2015). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

The *E. coli* AlkB protein was implicated in the repair or tolerance of DNA methylation damage. However, despite the early isolation of an *E. coli alkB* mutant, the function of the AlkB protein had not been resolved (Kataoka *et al.*, 1983). The *E. coli alkB* mutant is defective in processing methylated single stranded DNA, therefore, it was suggested that the AlkB protein either repairs or tolerates lesions generated in single stranded DNA, such as 1-methyladenine (1-meA) or 3-methylcytosine (3-meC), or that AlkB only acts on single stranded DNA (Dinglay *et al.*, 2000). However, despite extensive testing, no enzymatic activity could be assigned to the AlkB protein. Recently, theoretical protein fold recognition suggested that the AlkB protein resembles members of the α-ketoglutarate-Fe(II) dependent dioxygenase superfamily (Aravind and Koonin, 2001). Here, the biochemical function of the enigmatic *E. coli* AlkB protein and its two human homologs ABH2 and ABH3 are elucidated. An *in vitro* assay was developed for the AlkB, ABH2 and ABH3 proteins and it was demonstrated that the activities of these proteins are dependent on α-ketoglutarate and Fe(II) and are stimulated by ascorbic acid. The requirement of AlkB, ABH2 and ABH3 for these distinctive co-factors strongly supports the proposal that these proteins are members of the α-ketoglutarate-Fe(II) dependent dioxygenase superfamily, which employ iron-oxo intermediates to oxidise chemically inert compounds. The AlkB, ABH2 and ABH3 proteins are shown to act specifically on 1-meA and 3-meC in both double and single stranded DNA. It was demonstrated that these proteins convert 1-meA and 3-meC to their unsubstituted parent residues in DNA and therefore act by a direct reversal mechanism. It is proposed that the *E. coli* AlkB protein and its human homologs directly revert DNA damage by oxidative demethylation, an unprecedented mechanism of DNA repair. This discovery may have implications for the treatment of cancer because antagonists of the human ABH2 and ABH3 proteins could be useful adjuncts to cancer chemotherapy that uses simple alkylating agents.

TABLE OF CONTENTS

LIST OF FIGURES AND APPENDIXES

LIST OF TABLES

ABBREVIATIONS

1-meA	1-methyladenine
3-meA	3-methyladenine
N^6-meA	6-methyladenine
7-meA	7-methyladenine
3-meC	3-methylcytosine
N^4-meC	4-methylcytosine
5-meC	5-methylcytosine
O^2-meC	O^2-methylcytosine
3-meG	3-methylguanine
7-meG	7-methylguanine
8-meG	8-methylguanine
O^6-meG	O^6-methylguanine
3-meT	3-methylthymine
O^2-meT	O^2-methylthymine
O^4-meT	O^4-methylthymine
3-meU	3-methyluracil
A	Adenine
AP	Apurinic / apyrimidinic
ATP	Adenosine 5'-triphosphate
BCNU	Bis-chloroethylnitrosourea
BER	Base excision repair
BICINE	N,N-Bis(2-hydroxyethyl)glycine
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
C	Cytosine
CIP	Calf intestinal alkaline phosphatase
dCTP	Deoxycytidine 5'-triphosphate
DMS	Dimethyl sulphate
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
dRP	5'-deoxyribosephosphate
DSBH	Double-stranded β-helix
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FaPy	Formamidopyrimidine

G	Guanine
HEPES	N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
HhH	Helix-hairpin-helix
HIV	Human immunodeficiency virus
HNPCC	Hereditary non-polyposis colon cancer
HPLC	High performance liquid chromatography
HtH	Helix-turn-helix
ICLs	Interstrand crosslinks
IPTG	Isopropyl β-D-thiogalactoside
IVD	Isovaleryl-CoA dehydrogenase
MeI	Methyl iodide
MMC	Mitomycin C
MNNG	N-methyl-N'-nitro-N-nitrosoguanidine
MMR	Mismatch repair
MMS	Methylmethanesulfonate
MNU	N-methyl-N-nitrosourea
MTases	Methyltransferases
NCBI	National Centre for Biotechnology Information
NER	Nucleotide excision repair
NTA	Nitrilotriacetic acid
P(me)s	Sp-diastereoisomer of DNA methyl phosphotriesters
PCR	Polymerase chain reaction
PEG	Poly(ethylene glycol)
Pfu	Plaque forming units
PNK	Polynucleotide kinase
RNA	Ribonucleic acid
SAM	*S*-Adenosylmethionine
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
T	Thymine
TLS	Translesion synthesis
Tris	Tris(hydroxymethyl)aminomethane
UV	Ultraviolet
XP	Xeroderma pigmentosum

ACKNOWLEDGEMENTS

I would particularly like to acknowledge my supervisors Dr. Tomas Lindahl and Dr. Barbara Sedgwick for their help and guidance. I would like to thank my UCL supervisor Prof. Peter Swann. I am grateful to the members of the Mutagenesis laboratory, both past and present for their assistance and advice. I would also like to thank my friends at Clare Hall for many useful discussions.

A massive thank you goes to my family, especially Graham for all his love and understanding. Thank you Mum, Dad and Stephen for all your love and support. I couldn't have done it without you lot!

DEDICATION

This thesis is dedicated to the memory of my Father, Michael John Trewick (1944-2003).

PUBLICATIONS

Dinglay, S., Trewick, S. C., Lindahl, T. and Sedgwick, B. (2000) Defective processing of methylated single-stranded DNA by *E. coli alkB* mutants. *Genes Dev*, **14**, 2097-2105.

Trewick, S.C., Henshaw, T.F., Hausinger, R.P., Lindahl, T. and Sedgwick, B. (2002) Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. *Nature*, **419**, 174-178

*Duncan, T., *Trewick, S.C., Koivisto, P., Bates, P.A., Lindahl, T. and Sedgwick, B. (2002) Reversal of DNA alkylation damage by two human dioxygenases. *Proc Natl Acad Sci U S A*, **99**, 16660-16665. * = joint first author.

CHAPTER 1

INTRODUCTION

CHAPTER 1. INTRODUCTION

1.1 Preface

Fifty years ago, this year, the structure of the deoxyribonucleic acid (DNA) molecule was published. Initially, it was assumed that in order to accomplish the high fidelity required to accurately maintain the genetic information, DNA would be an extremely stable molecule. However, we now know that our DNA is in a dynamic state. Several of DNA's chemical bonds are unstable under physiological conditions, allowing spontaneous alterations to occur in the DNA structure. Our DNA is constantly under siege from endogenous and environmental agents such as ionising radiation, ultraviolet (UV) light and alkylating agents that chemically modify the DNA. Also, errors are introduced into the DNA sequence during replication, recombination and repair. Collectively these modifications to the DNA structure are known as DNA damage. DNA damage can be toxic to the cell as it can block the biological activity of DNA, or it can corrupt the information encoded by DNA leading to mutations, which are heritable changes in the DNA sequence.

The perpetual onslaught of DNA damage has exerted a strong selective pressure for cells to evolve systems that safeguard their genetic information. All known free living organisms and many viruses have devised a myriad of mechanisms for tolerating and repairing DNA damage. As the principle threats to the genome have been present since the origin of life, many DNA repair enzymes appeared early in evolution and show strong conservation from microbes to man. Therefore, investigation of the DNA repair mechanisms of the bacteria

Escherichia coli has yielded substantial insights into the DNA repair processes of human cells.

DNA repair is of considerable clinical interest because the inactivation of certain DNA repair genes can be an important event in carcinogenesis. Loss of DNA repair efficiency can lead to genomic instability, which accelerates the rate of genetic change and allows the accumulation of mutations in tumour suppressor genes and proto-oncogenes, leading to uncontrolled cell growth. This is well illustrated by the identification of several cancer prone syndromes that are caused by heritable mutations in DNA repair genes. For example, individuals with hereditary non-polyposis colon cancer (HNPCC) have an 80 % lifetime risk of developing colorectal cancer, compared to 6 % in the general population. Patients with xeroderma pigmentosum (XP) have a 10,000-fold increased risk of skin cancer associated with sunlight exposure. However, DNA damaging agents are also important in the treatment of cancer because radiotherapy and some forms of cancer chemotherapy employ DNA damage to kill rapidly dividing cells. The ability of cancer cells to repair DNA damage is often an important factor in determining the susceptibility or resistance of tumours to these cancer treatments.

Alkylating agents cause DNA damage by the addition of alkyl groups such as methyl or ethyl groups to the DNA structure. These agents are present in the environment, are generated endogenously by cellular metabolism and are used as cancer chemotheraputic agents. As DNA alkylation damage can be cytotoxic and / or mutagenic, cells have developed several efficient strategies to repair or tolerate unrepaired DNA alkylation damage. The *E. coli* AlkB protein has been implicated in the repair or tolerance of alkylation damage, but the function of the AlkB protein has remained an enigma for over twenty years. The aim of this work is to develop an *in vitro* assay for the *E. coli* AlkB protein and determine its

biochemical function. As alkylating agents are used in cancer chemotherapy, it was also of interest to investigate human homologs of the AlkB protein.

1.2 DNA alkylating agents

Environmental, endogenous and chemotheraputic agents

The ubiquitous presence of DNA alkylation repair activities in pro- and eukaryotes indicates that cells are universally exposed to alkylating agents. The precise origins of these alkylating agents are not clear, however, the increased spontaneous mutation frequency of cells unable to repair DNA alkylation damage suggests that DNA alkylation damage can be produced endogenously, as a consequence of normal cellular metabolism (Rebeck and Samson, 1991; Xiao and Samson, 1993). *S*-Adenosylmethionine (SAM), a ubiquitous intracellular methyl donor, can act as a weak alkylating agent (Barrows and Magee, 1982; Rydberg and Lindahl, 1982). However, variation of SAM levels over 100-fold range had no significant effect on spontaneous mutagenesis in *E. coli* (Posnick and Samson, 1999a). The most likely candidates for endogenous alkylators are nitrosated amino compounds, formed by nitrosation of the amino groups of peptides and amino acids by endogenously generated nitric oxide (Taverna and Sedgwick, 1996; Sedgwick, 1997; Shuker and Margison, 1997; Garcia-Santos Mdel *et al.*, 2001). Also, several alkylating agents have been identified in the environment, for example, nitrosamines are found in cigarette smoke (Hecht *et al.*, 1988) and at low levels in some foods (Lai and Arcos, 1980). Methylchloride is synthesised by certain fungi and algae and is released into the environment by biomass burning (Wuosmaa and Hager, 1990; Vaughan *et al.*, 1991). Since the 1950s a number of alkylating agents, such as the nitrogen mustards, have been used in cancer

chemotherapy treatment. The ability of these alkylating agents to kill rapidly dividing cells was first identified after they were used as chemical weapons.

The chemistry of alkylating agents

Alkylating agents are a diverse group of chemicals. Directly or via metabolically activated intermediates, alkylating agents transfer alkyl groups, such as methyl or ethyl groups, to electron-rich nucleophilic sites (for example S: O: or N:) in biological molecules including DNA, RNA, lipids, carbohydrates and proteins.

Alkylating agents can attack and alkylate virtually all of the oxygen atoms and nitrogen atoms in DNA (Singer, 1975; Singer and Grunberger, 1983; Lawley, 1984). The major sites of alkylation of the DNA bases are shown in Figure 1. In general, the sites alkylated by an agent depend upon the electrophilicity of the alkylating agent, the nucleophilicity of the DNA site and the secondary structure of the DNA.

Traditionally, alkylating agents are designated as either S_N1 or S_N2 agents, the term referring to the reaction kinetics of the alkylation reaction. The general reaction mechanisms of S_N1 and S_N2 alkylating agents are shown in Figure 2. S_N1 alkylating agents, such as, N-methylnitrosourea (MNU) and N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) are strong electrophiles and react with DNA in a first order reaction. The rate-limiting step of the reaction is the formation of a reactive electrophilic carbocation intermediate, for example, MNU and MNNG hydrolyse to form the methyl diazonium ion, which is attacked by nucleophilic centres in DNA. S_N2 alkylating agents, such as methyl methanesulphonate (MMS), dimethylsulphate (DMS) and methyl iodide (MeI), are weaker electrophiles and react with nucleophilic centres in biological molecules in a one-step mechanism

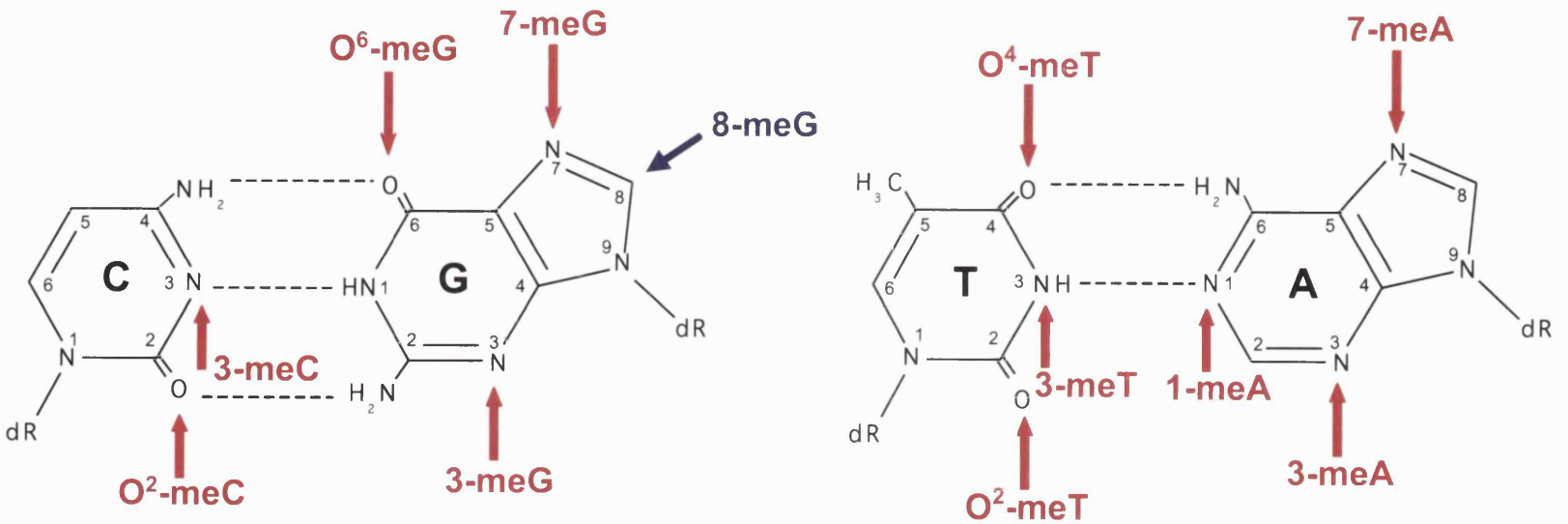
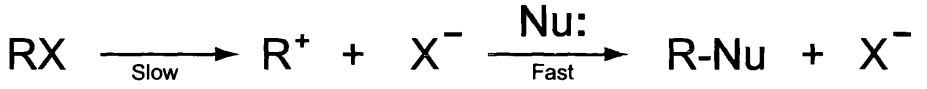


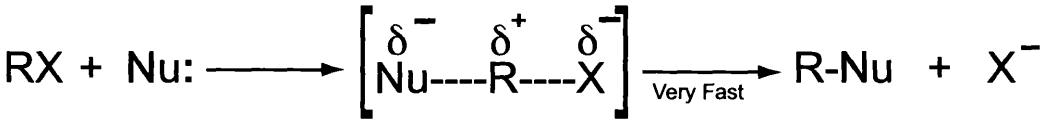
Figure 1. Sites of methylation in the DNA bases. Red arrows indicate sites of DNA methylation by S_N1 and S_N2 agents. The blue arrow indicates an additional site alkylated by methyl radicals

S_N1 alkylation reaction



Alkylating Agent **Carbocation Intermediate** DNA

S_N2 alkylation reaction



Alkylating Agent DNA Transition State

Figure 2. Reaction mechanisms of S_N1 and S_N1 alkylating agents.

that obeys second order kinetics (Figure 2). Although the S_N1 and S_N2 classification of alkylating agents is useful, the mechanisms of most alkylating agents fall somewhere in-between these two extremes (Swain and Scott, 1953).

There is a large variation in the nucleophilicity and therefore reactivity of the potential alkylation sites in DNA, in general the nitrogen atoms in the purine and pyrimidine rings are more nucleophilic than the exocyclic base oxygens. The S_N2 alkylating agents are weak electrophiles and therefore react almost exclusively with the strongly nucleophilic ring nitrogen atoms of the purine and pyrimidines. For example, in double stranded DNA the S_N2 agents primarily modify the 7 position of guanine and the 3 position of adenine (Table 1). However, the strongly electrophilic S_N1 alkylating agents act at a broad range of DNA sites (Table 1) (Lawley, 1984), including ring nitrogens, such as the 7 position of guanine, the exocyclic oxygens such as the O^6 position of guanine, as well as oxygen atoms in the sugar-phosphate backbone producing diastereoisomers in the R and S configuration.

There are slight differences in the relative abundance of the methylated lesions produced in single and double stranded DNA (Table 1). The 1 position of adenine and the 3 position of cytosine are much less reactive in double stranded DNA. This is because in base-paired DNA, the nitrogen atoms at position 1 of adenine and position 3 of cytosine use their lone pair of electrons as hydrogen bond acceptors. Some alkylation is observed at these positions in double stranded DNA as, even at 37 °C, there is some thermal denaturation of the DNA (Bodell and Singer, 1979; Singer and Grunberger, 1983; Lawley, 1984). In contrast to base-paired nitrogen atoms, the extent of reaction of base-paired oxygen atoms is not a function of strandedness, since the oxygen at position O^6 of guanine, O^4 of thymine and O^2 of cytosine all possess an unreacted lone pair of electrons even

		% of Total Alkyation			
		Double Stranded DNA		Single Stranded DNA	
Modification	Repair	MNU (S_N1)	MMS (S_N2)	MNU (S_N1)	MMS (S_N2)
Adenine					
N1		1.3	3.8	2.8	18
N3	Tag AlkA	9	10.4	2.6	1.4
N7	AlkA	1.7	(1.8)	1.8	3.8
Guanine					
N3	(Tag) AlkA	0.8	(0.6)	0.4	~1
O6	Ogt Ada	6.3	(0.3)	3	nd
N7	AlkA	67	83	69	68
Uracil / Thymine					
O2	AlkA	0.11	-	-	-
N3		0.3	-	-	-
O4	Ogt (Ada)	0.4	-	-	-
Cytosine					
O2	AlkA	0.1	nd	-	-
N3		0.6	(<1)	2.3	10
Diester	Ada	17	0.8	~10	2

Table 1. DNA base lesions induced by an S_N1 (MNU) and an S_N2 (MMS) alkylating agent.
nd = lesion not detected; ( ) = estimated value [Adapted from: Singer and Grunberger, 1983]

while base-paired. The N3 position of adenine, which is not involved in base-pairing, is alkylated more efficiently in double rather than single stranded DNA.

Bifunctional alkylating agents such as mitomycin C (MMC), bis-chloroethylnitrosourea (BCNU) and nitrogen mustards generally contain two reactive groups and therefore can form DNA intrastrand and DNA interstrand crosslinks and also DNA-protein crosslinks (Singer and Grunberger, 1983). DNA can also be methylated by methyl radicals generated by agents such as 1,2-dimethylhydrazine, diazoquinones and tert-butylhydroperoxide, which reacts with several sites in DNA including the 8 position of guanine (Hix *et al.*, 1995).

1.3 DNA alkylation damage

DNA alkylation damage can be innocuous, mutagenic and / or cytotoxic to the cell. Promutagenic lesions can mispair during DNA replication for example, O^6-methylguanine (O^6-meG) can mispair with thymine and O^4-methylthymine (O^4-meT) can mispair with guanine, resulting in G:C to A:T and A:T to G:C transition mutations respectively (Loechler *et al.*, 1984; Saffhill and Hall, 1985). O^2-methylthymine (O^2-meT) and O^2-methylcytosine (O^2-meC) exhibit weaker miscoding properties than O^6-meG or O^4-meT (Singer *et al.*, 1979; Singer and Grunberger, 1983). 3-methylcytosine (3-meC) is a cytotoxic lesion because its methyl group blocks the formation of DNA base-pairs, which stalls the replicative DNA polymerase (Saffhill, 1984). The methyl groups of 3-methyladenine (3-meA) and 3-methylguanine (3-meG) physically block the interaction of the replicative polymerase with DNA and are therefore cytotoxic lesions (Beard *et al.*, 1996; Eom *et al.*, 1996). Interstrand crosslinks (ICLs) are cytotoxic lesions as they prevent DNA strand separation during replication and transcription (Dronkert and Kanaar, 2001). However, some DNA alkylation damage is relatively innocuous,

for example, 7-methylguanine (7-meG) as well as methylation of the DNA phosphodiester backbone.

Alkylation of DNA can lead to the formation of secondary lesions, for example, methylation at the 7 position of guanine destabilises the imidazole ring of the purine, making it prone to alkali catalysed ring opening, to produce the cytotoxic lesion formamidopyrimidine (FaPy) (O'Connor *et al.*, 1988). Also, alkylation weakens the N-glycosidic linkage of several DNA bases, particularly 7-meG which is six orders of magnitude more likely to depurinate than unmodified guanine, producing an abasic site (Lindahl, 1993). Abasic sites can react to form DNA-DNA and DNA-protein crosslinks as well as single strand breaks (Goffin and Verly, 1983).

1.4 Repair of DNA alkylation damage

To counteract the potential adverse biological effects of this diverse array of DNA alkylation damage, cells have evolved several genetically and functionally distinct activities for repairing DNA alkylation damage. Several alkylated DNA lesions are repaired by O^6-methylguanine-DNA methyltransferases (MTases), which act by a direct reversal mechanism. Other alkylated lesions are repaired by DNA excision repair mechanisms, in which the alkylated base or an oligonucleotide containing the alkylated base is cut from the DNA and the gap is refilled using the undamaged DNA strand as a template. The base excision repair (BER) pathway acts on small alkylated lesions, whereas, the nucleotide excision repair (NER) pathway works on large or bulky alkylation damage.

O^6-methylguanine-DNA methyltransferases

The MTases repair DNA alkylation damage by direct reversal, that is, they restore the DNA to its undamaged state in a single-step reaction. MTases transfer alkyl groups from DNA lesions onto an active site cysteine residue. As alkyl cysteine is very stable, the MTase is inactivated. Because each molecule of MTase is consumed during the reaction, MTases are not true catalysts and thus are not true enzymes (Lindahl *et al.*, 1982). MTases are found almost universally in both prokaryotes and eukaryotes. *E. coli* has a constitutively expressed MTase, Ogt, that repairs low levels of damage (Margison *et al.*, 1990). The expression of a second MTase, Ada (Sedgwick, 1982) is induced in response to DNA alkylation damage. Human cells have only one MTase, MGMT.

The *E. coli* MTases, Ada and Ogt, as well as the human MGMT protein repair the mutagenic lesion O^6-meG. Ogt also repairs O^4-meT, whereas Ada and MGMT are inefficient at repair of O^4-meT. Attempts by MGMT to repair O^4-meT may shield the lesion from repair by NER (Samson *et al.*, 1997). MTases are able to repair longer chain alkyl groups and benzyl adducts, however, there is considerable species variation in the efficiency of these reactions (Goodtzova *et al.*, 1997). In addition to averting the production of G-C to A-T transition mutations caused by endogenous and environmental alkylating agents, MTases also repair the O^6-meG generated by anticancer chemotherapy drugs that methylate DNA, such as dacarbazine, procarbazine and temozolomide, or that chloroethylate DNA, for example, bis-chloroethylnitrosourea, clomesone and fotemustine. As O^6-meG forms the cytotoxic basis of such chemotherapy, increased MGMT expression is a common mechanism of resistance in human tumours to these anti-neoplastic agents (Margison and Santibanez-Koref, 2002).

Crystallographic studies of MTases from the three kingdoms of life reveal that despite low sequence homology, the two domain α/β fold has been remarkably conserved. The human MGMT protein contains an atom of Zn(II) that acts structurally to stabilise the interface between the two domains. The current model for MGMT action (Daniels and Tainer, 2000) indicates that MGMT uses a carboxyl-terminal helix-turn helix (HTH) motif to bind DNA (Vora *et al.*, 1998), an arginine finger is used to flip the damaged base into the active site. The active site cysteine that participates in an extensive hydrogen-bond network, drives the dealkylation reaction (Daniels and Tainer, 2000). The alkyl group is transferred onto the active site cysteine in an irreversible and stoichiometric reaction. This causes a conformational change that dissociates the MTase from the DNA and targets the MGMT for degradation (Srivenugopal *et al.*, 1996). Due to the stoichiometric nature of its activity the expression of human MGMT is tightly controlled, protein kinase C and glucocorticoid hormone both regulate transcription of the human MGMT gene (Boldogh *et al.*, 1998; Biswas *et al.*, 1999).

The amino-terminal domain of the *E. coli* Ada protein contains a second methyltransferase domain that is not present in either the Ogt or MGMT proteins. This MTase domain uses an atom of Zn(II) for correct protein folding and also to activate the nucleophilicity of the accepting Cys-38 residue (Myers *et al.*, 1994) (Verdine, G.L. personal communication). The amino-terminal MTase domain of Ada transfers alkyl groups from the Sp-diastereoisomer of DNA methyl phosphotriesters (P(me)s) (McCarthy and Lindahl, 1985). Transfer of methyl groups from P(me)s turns Ada into a transcriptional activator of its own gene and also that of the *alkA*, *alkB* and *aidB* genes (Sedgwick and Lindahl, 2002) (see section 1.7).

Base excision repair

In general, the base excision repair (BER) pathway acts on damaged DNA bases which cause only minor disturbances to the DNA helix, such as deaminated or oxidised bases (Lindahl, 2001). Several methylated bases, including 3-meA, 3-meG, 7-meG, O^2-methylthymine (O^2-meT) and O^2-methylcytosine (O^2-meC) as well as secondary lesions such as formamidopyrimidine and abasic sites are repaired by BER.

The general mechanism of BER is shown in Figure 3. BER is initiated by a battery of DNA glycosylases, which each recognise a different damaged base or sets of damaged bases. Aberrant bases are flipped-out from the DNA helix and inserted into the active site of the DNA glycosylase, which then cleaves the N-glycosidic bond between the base and the deoxyribose, leaving an abasic site in the DNA (Lindahl and Wood, 1999). In *E. coli*, two glycosylases excise alkylated bases from DNA, the constitutively expressed Tag DNA glycosylase (3-meA DNA glycosylase I) (Riazuddin and Lindahl, 1978) and the inducibly expressed AlkA DNA glycosylase (3-meA DNA glycosylase II), whose expression is induced in response to DNA alkylation damage (Evensen and Seeberg, 1982; Karran *et al.*, 1982). Structural studies have revealed that the Tag and AlkA DNA glycosylases as well as the *Saccharomyces cerevisiae* AlkA homolog, MAG, belong to the helix-hairpin-helix (HhH) superfamily of DNA glycosylases (Nash *et al.*, 1996; Drohat *et al.*, 2002). The human 3-meA DNA glycosylase, AAG, is a member of a separate family consisting of several highly homologous mammalian DNA glycosylases (Lau *et al.*, 1998).

The *E. coli* Tag protein has a narrow substrate range removing 3-meA and at a lower efficiency, 3-meG from DNA (Bjelland and Seeberg, 1987; Bjelland *et*

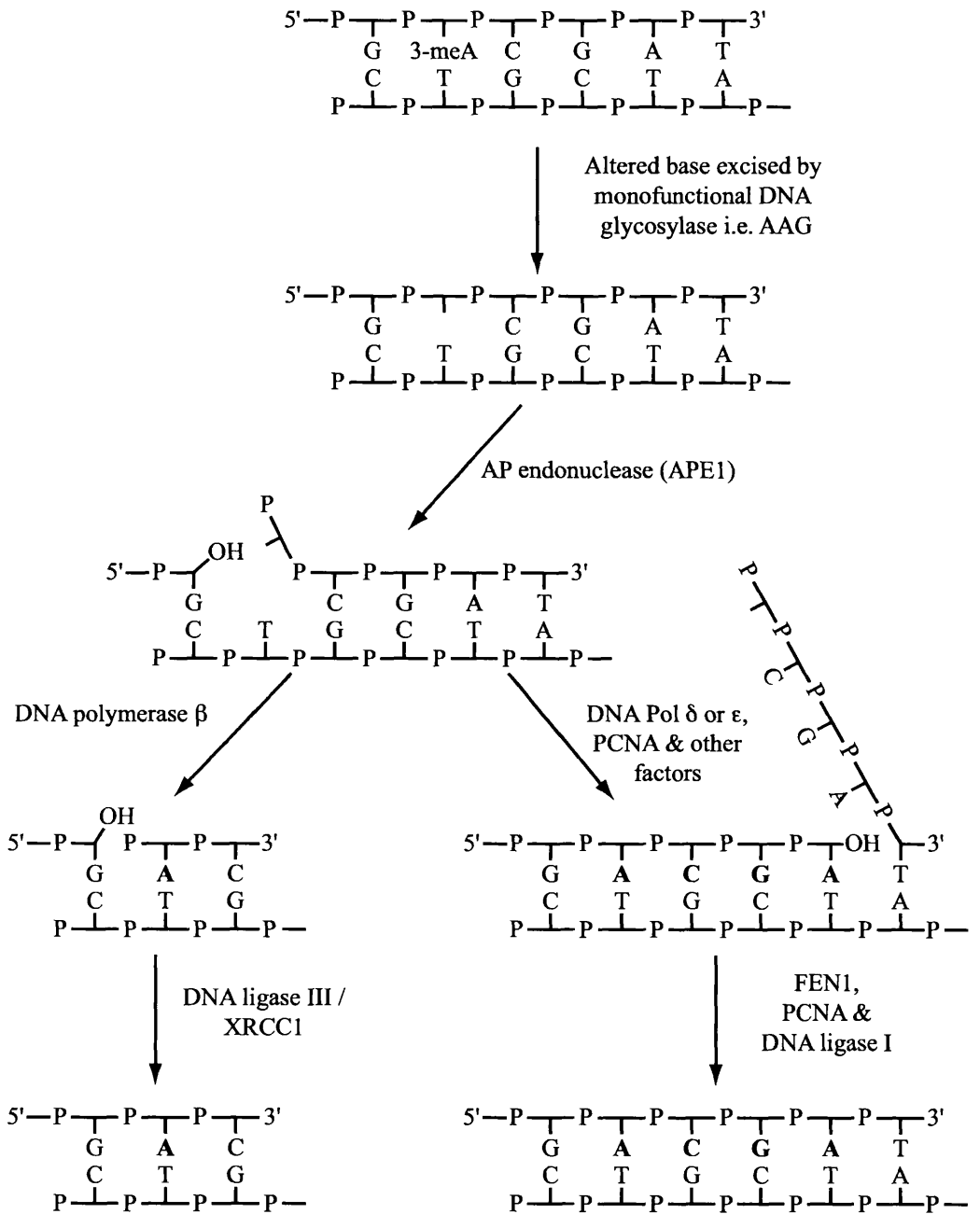


Figure 3. Base excision repair in mammalian cells, initiated by a monofunctional DNA glycosylase. [Adapted from: Klungland and Lindahl, 1997]

al., 1993). In contrast, both the *E. coli* AlkA and the human AAG glycosylases recognise chemically varied and structurally diverse lesions including the abundant methylated lesions 7-meG, 3-meA, 3-meG, O^2-meC, O^2-meT and 7-methyladenine (7-meA), as well as 1,N6-ethanoadenine and hypoxanthine (Saparbaev and Laval, 1994; Saparbaev *et al.*, 1995; Singer and Hang, 1997). Whereas, AlkA but not AAG can remove 8-methylguanine (8-meG)(Gasparutto *et al.*, 2002). In correlation with its broad substrate range, the *E. coli* AlkA glycosylase can inefficiently excise normal guanine and adenine bases from DNA (Berdal *et al.*, 1998; Posnick and Samson, 1999b). Therefore, *E. coli* cells repair low levels of damage using the relatively specific Tag DNA glycosylase. Cellular levels of the more promiscuous AlkA DNA glycosylase are kept low until the cells are challenged by DNA alkylation damage.

Crystallographic studies have provided insights into the substrate specificity of the *E. coli* AlkA, Tag and human AAG DNA glycosylases (Hollis *et al.*, 2000; Drohat *et al.*, 2002). It has been postulated that the broad substrate range of the *E. coli* AlkA DNA glycosylase can be at least partly attributed to the shape of AlkA's active site. Unlike the highly substrate specific DNA glycosylases which have deep, confined active sites, the active site of the AlkA glycosylase is a shallow groove, this presumably allows AlkA to accommodate a diverse range of substrates. As many of the substrates of AlkA are positively charged, it has been proposed that this enzyme recognises its various substrates by using non-specific interactions with active site electron-rich aromatic amino acid residues (Labahn *et al.*, 1996). However, an alternative model suggests that the AlkA DNA glycosylase is able to recognise and act on substrates with weakened glycosyl bonds (Berdal *et al.*, 1998). The substrate specificity of the constitutively expressed *E. coli* Tag protein may be limited to extremely labile lesions such as 3-

meA due to the absence of the invariant, catalytically important, aspartic acid residue present in all other HhH glycosylases (Drohat *et al.*, 2002).

In *E. coli,* formamidopyrimidine (FaPy) is excised by the formamidopyrimidine-DNA glycosylase, Fpg, which also excises 8-oxoguanine from DNA (Chetsanga and Lindahl, 1979; Tchou *et al.*, 1991). In eukaryotes, FaPy is excised by the 8-oxoguanine DNA glycosylase, Ogg1 (van der Kemp *et al.*, 1996).

Monofunctional DNA glycosylases, such as AlkA, Tag and AAG excise damaged bases leaving an abasic site that is recognised and cleaved by an AP endonuclease 5' to the abasic site (Lindahl, 1979). In *E. coli* the AP-endonucleoytic step is catalysed by endonuclease IV, which is damage inducible, or by exonuclease III, which is constitutively expressed (Doetsch and Cunningham, 1990). The mammalian AP endonuclease, APE1 is an exonuclease III homologue (Ramotar, 1997). The remnant 5'-deoxyribosephosphate (dRP) is released by the dRPase activity of the *E. coli* RecJ protein (Dianov *et al.*, 1994) or mammalian DNA polymerase β (Matsumoto and Kim, 1995). Bifunctional glycosylases such as Fpg, not only excise damaged bases, but also have an associated β-lyase activity that cleaves the phosphodiester bond 3' to the abasic site, the DNA backbone is then cleaved 5' to the abasic site by an AP-endonuclease. In both cases, the resulting gap is filled in by *E. coli* polymerase I or mammalian DNA polymerase β and sealed by a ligase, generating a "repair patch" only a single nucleotide in length.

An alternative "long patch" BER pathway that generates a repair patch of 2-13 nucleotides is sometimes employed in mammalian cells (Frosina *et al.*, 1996). The long patch pathway depends on enzymes normally involved in DNA

replication, such as, DNA polymerase δ / ε, PCNA, the structure-specific nuclease DNase IV (FEN-1) and DNA ligase I.

Nucleotide excision repair

The nucleotide excision repair (NER) pathway is the most versatile of all the DNA repair pathways and is able to repair various structurally unrelated lesions which distort the DNA structure and modify the chemistry of the DNA, for example, UV induced pyrimidine dimers (Batty and Wood, 2000). The main role of NER in the repair of alkylation damage is the removal of lesions with large alkyl groups, including O^6-ethylguanine and O^4-ethylthymine and bulkier alkyl derivatives.

The initial step in NER is recognition of the lesion. In non-transcribed DNA strands, distorting lesions are recognised by the human XPC-hHR23B complex (Figure 4) (Volker *et al.*, 2001) and in *E. coli* by the $UvrA_2B$ complex, which loads UvrB onto the site of the damage (Orren and Sancar, 1989). In transcribed DNA strands, RNA polymerases stall at DNA lesions and recruit proteins such as the *E. coli* Mfd protein (Selby and Sancar, 1993) or the human CSA and CSB proteins, which are thought to displace RNA and presumably recruit NER proteins (van den Boom *et al.*, 2002).

In *E. coli*, the UvrC protein, an endonuclease with two active sites, is recruited to the site of the damage and cuts at both sides of the lesion (Verhoeven *et al.*, 2000). The oligonucleotide (24-32 nucleotides long), which contains the lesion is then displaced by the UvrD protein (DNA helicase II). In human cells, the TFIIH complex is recruited to the site of damage and two of its subunits, the ATP-dependent helicases XPB and XPD together with XPA and RPA locally unwind the DNA around the lesion (Svejstrup *et al.*, 1996). The structure-specific

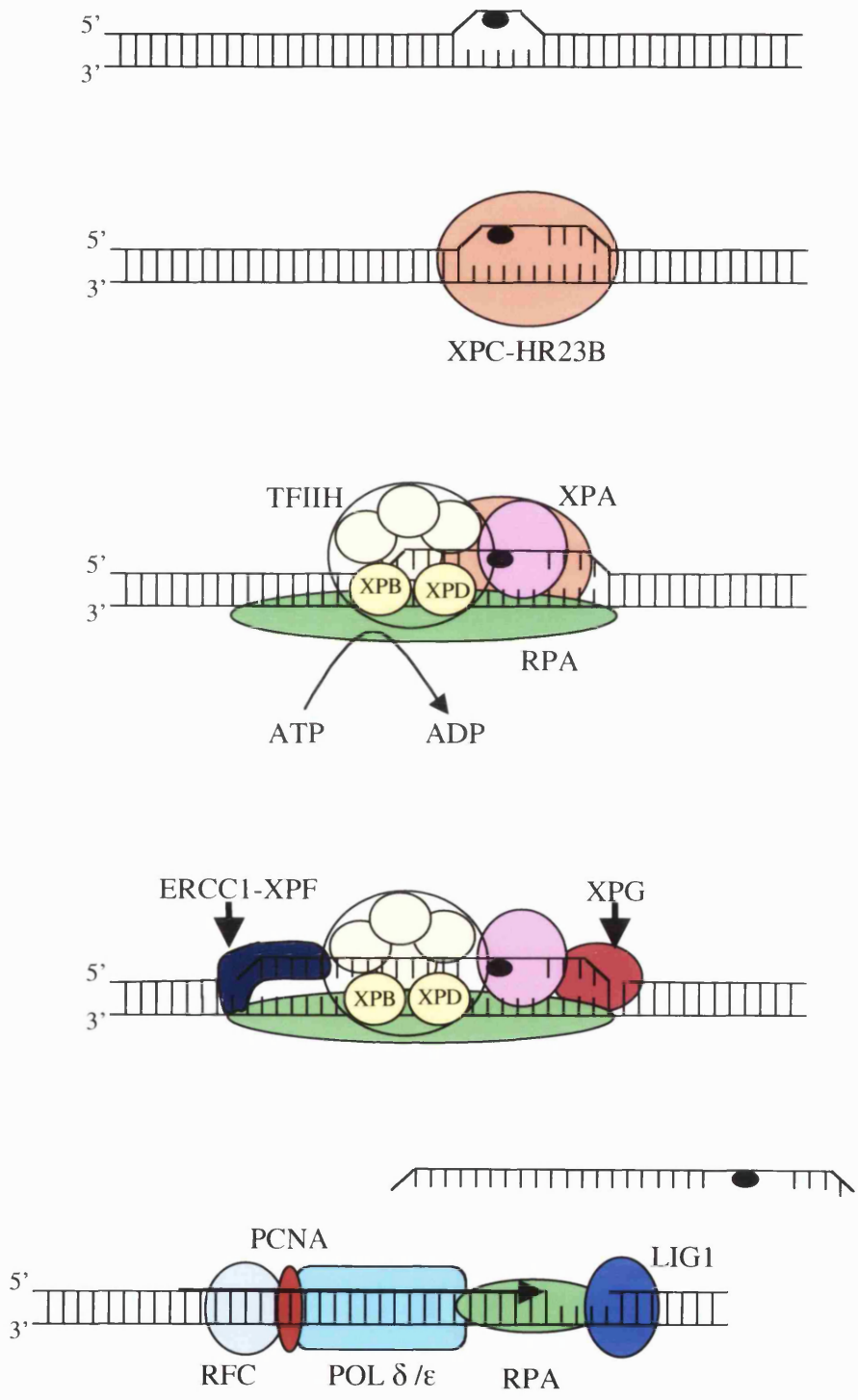


Figure 4. Nucleotide excision repair of non-transcribed DNA in mammalian cells [Adapted from Batty and Wood, 2000].

DNA endonucleases XPG and XPF/ERCC1 are recruited and make single-stranded DNA breaks 3' and 5' to the lesion respectively. It is not clear how human cells remove the oligonucleotide containing the DNA damage. Finally the *E. coli* DNA polymerase I, or the human DNA polymerase δ / ε, fills the gap and the nicks are sealed by a DNA ligase (Wood, 1997).

The *E. coli* UvrABC or the human XPF-ERCC1 proteins also have a role in the repair of interstrand crosslinks (ICL). These structure-specific DNA endonucleases cut both sides of the ICL, allowing for repair of the lesion by homologous recombination (McHugh *et al.*, 2001).

1.5 Tolerance of DNA alkylation damage

When the DNA repair mechanisms are overwhelmed, DNA alkylation damage can persist into S phase. Once initiated, it is imperative for cell survival that DNA replication is fully completed, therefore, cells have evolved DNA damage tolerance pathways that act to reinitiate DNA replication at stalled replication forks without removing the lesion. DNA damage tolerance activities in *E. coli* are encoded by the SOS regulon, which consists of over 30 unlinked genes whose expression is upregulated in response to regions of single stranded DNA produced at stalled replication forks (Chaudhury and Smith, 1985; Fernandez De Henestrosa *et al.*, 2000). The SOS response in *E. coli* upregulates an error-free and an error-prone mode of damage tolerance. The error-free pathway uses DNA recombination to synthesise past lesions. The error-prone pathway uses specialised DNA polymerases to read through sites of DNA damage, this translesion synthesis (TLS) prevents cell death and provides an option for environmental adaptation. Eukaryotic damage by-pass mechanisms have been studied most extensively in *S. cerevisiae*, where members of the Rad6

complementation group of genes regulate DNA damage tolerance, directing TLS and recombinational by-pass by ubiquitination of specific proteins in response to DNA damage (Broomfield *et al.*, 2001). Rad18 and Rad5 are E3 ubiquitin ligating enzymes which bind single stranded DNA at stalled replication forks and recruit the ubiquitin E2 conjugating enzymes, Rad6 and Mms2 / Ubc13 respectively. Recently, PCNA was the first Rad6 / Rad18 substrate to be identified (Hoege *et al.*, 2002).

Recombinational repair

Recombinational bypass of replication stalling DNA damage such as 3-meA has been proposed to occur by multiple, poorly defined, mechanisms (Kowalczykowski, 2000; Cox, 2001; McGlynn and Lloyd, 2002). The mechanism employed probably depends on the initial cause of the stalling and the resulting local DNA structure formed. If a lesion such as an alkylated base stalls leading strand DNA synthesis, lagging strand synthesis may continue beyond the lesion. The replication fork can be regressed to form a Holliday junction and the leading strand extended past the lesion using the lagging strand as a template. Alternatively, the newly synthesised strand can invade and use the homologous sister strand as a template to synthesise past the alkylated lesion.

Translesion synthesis

Translesion synthesis (TLS) is catalysed by specialised DNA polymerases, which synthesise past lesions that block the replicative DNA polymerase. In *E. coli,* TLS appears to be catalysed by the error prone DNA polymerases Pol V ($UmuD_2'C$), Pol IV (DinB) and Pol II (DinA) (Goodman, 2000). It is not clear which, if any of these DNA polymerases are involved in by-pass of alkylated

lesions but $UmuD_2'C$ (Pol V) is able to by-pass abasic sites, preferentially incorporating adenine opposite the lesion (Reuven *et al.*, 1999; Tang *et al.*, 1999). In *S. cerevisiae*, it is hypothesised that Rev1 initiates TLS of abasic sites by inserting a single cytosine residue opposite the abasic sites, this is extended by Pol ζ, which consists of two subunits Rev3 and Rev7 (Broomfield *et al.*, 2001).

There is evidence that the human error prone DNA polymerases Pol η, Pol ι and Pol κ are able to by-pass abasic sites. The Rad30 protein of yeast and its human homolog Pol η are able to catalyse accurate bypass of cyclobutane pyrimidine dimers, but show error prone bypass of O^6-meG (Lehmann, 2002).

Mismatch repair

Mismatch repair (MMR) is a postreplicative excision repair pathway that corrects misincorporated bases and base insertions / deletions produced by slippages of the replicative DNA polymerase, which have escaped the proofreading activity of these polymerases (Hsieh, 2001). However, the MMR repair system was found to be responsible for the cytotoxicity of O^6-meG (Karran and Bignami, 1994). It is proposed that, the MMR system acts on this mispair and enters a futile repair loop. The newly synthesised thymine is excised and the DNA resynthesised, potentially recreating the O^6-meG:T mismatch, which MMR again attempts to repair. As cell death does not occur until the second round of DNA synthesis (Tominaga *et al.*, 1997), replication of gapped MMR intermediates may result in the formation of lethal double strand breaks.

1.6 Inducible responses to DNA damage in *E. coli*

Bacteria induce several DNA repair mechanisms in response to DNA damage and this allows bacteria to survive in rapidly changing environments. In response to the presence of reactive oxygen species, the *soxRS* and *oxyR* systems of *E. coli* induce the expression of proteins that act to avoid or repair DNA oxidation damage (Pomposiello and Demple, 2001). The SOS response (see section 1.5) induces the expression of proteins involved in tolerance of DNA damage as well as in DNA recombination and DNA excision repair (Sutton *et al.*, 2000). Dissection of these inducible responses in *E. coli* have lead to the discovery of several DNA repair and tolerance activities, many of which have homologs in higher eukaryotes.

1.7 The adaptive response to alkylating agents

To defend against fluctuating levels of alkylating agents in the environment many divergent species of bacteria mount an inducible response to DNA alkylation damage. This adaptive response to alkylating agents enhances the resistance of the bacteria to the mutagenic and cytotoxic effects of DNA alkylating agents. The adaptive response was first observed when *E. coli* exposed to non-lethal doses of MNNG became more resistant to mutation and killing when later challenged with a higher dose of the agent (Samson and Cairns, 1977). This increased resistance to alkylating agents was dependent on active protein synthesis (Jeggo *et al.*, 1977; Samson and Cairns, 1977).

In *E. coli,* the adaptive response is triggered by formation of an innocuous type of DNA methylation damage, the Sp-diastereoisomer of DNA methyl phosphotriesters (Teo *et al.*, 1986). The methyltransferase domain at the amino-terminus of the Ada protein repairs this relatively innocuous form of damage. The

Ada protein transfers the methyl group from the DNA lesion onto an active site cysteine residue (Cys-38), converting Ada from a weak to a strong transcriptional activator of the *ada*, *alkA*, *alkB* and *aidB* genes (Figure 5) (Sedgwick and Lindahl, 2002). The adaptive response can also be activated to some extent by certain methylating agents such as MeI, which may act by directly methylating the active site cysteine of the Ada protein (Takahashi *et al.*, 1988).

The *alkA* and *aidB* genes map at 45 and 95 min on the *E. coli* genetic map, respectively, whereas the *ada* and *alkB* genes form a small operon that maps at 47 min. The initiation codon (ATG) of the *alkB* gene overlaps the termination codon (TAA) of the *ada* gene, however, as 10 fold more Ada protein is produced than AlkB protein, transcription may partially terminate at the end of the *ada* coding region (Kataoka and Sekiguchi, 1985; Kondo *et al.*, 1986).

The increased cellular levels of the *E. coli* Ada, AlkA, AlkB and AidB proteins convey resistance to the killing and mutagenic effects of alkylating agents. Two of these proteins, Ada and AlkA were found to be proteins that repair DNA alkylation damage. The Ada protein, apart from being the positive regulator of the adaptive response, also has a MTase domain at its carboxyl-terminus that repairs mutagenic alkylation damage such as O^6-meG by direct reversal (see section 1.4). The AlkA protein is a DNA glycosylase, which excises a wide range of alkylated DNA bases as part of the base excision repair pathway (see section 1.4). It has been proposed that AidB protein may be involved in the metabolic detoxification of endogenous and exogenous alkylating agents, thereby protecting cells against alkylation damage to DNA, but this hypothesis remains to be substantiated. The *aidB* gene appears to encode an isovaleryl-CoA dehydrogenase (IVD), which are required in higher eukaryotic cells for the degradation of leucine

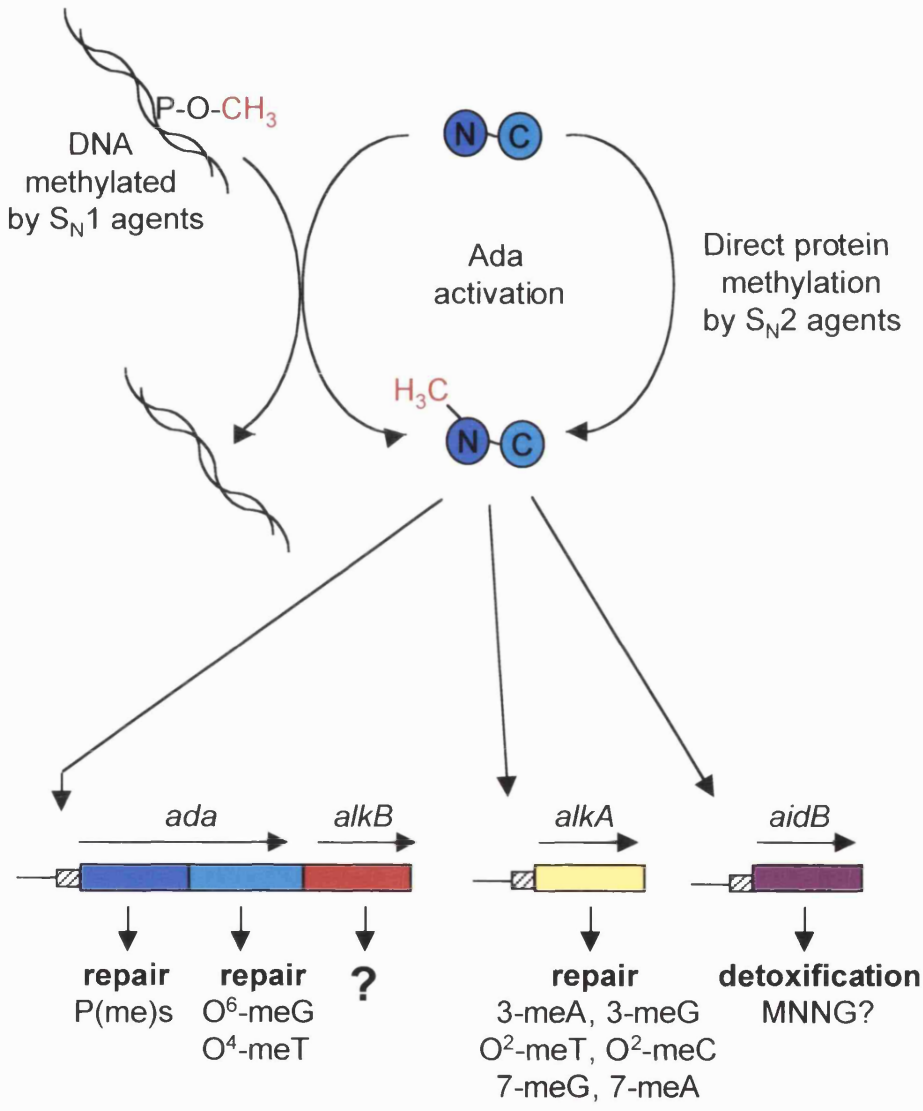


Figure 5. The *E. coli* adaptive response to alkylating agents. The Ada protein is the positive regulator of this response. The Ada protein is activated on repair of Sp-diastereoisomer of methylphosphotriesters in DNA and induces expression of the *ada*, *alkB*, *alkA* and *aidB* genes [Adapted from: Sedgwick and Lindahl, 2002].

during starvation conditions, however, *E. coli* has lost the other genes required for leucine catabolism (Landini *et al.*, 1994). Expression of *aidB* is less responsive to methylated Ada protein than the other genes of the adaptive response and requires ten fold higher concentrations of methylated Ada than the *ada* gene (Landini and Volkert, 1995).

1.8 The mystery of the AlkB protein

The *E. coli* AlkB protein has been implicated in the repair or tolerance of DNA alkylation damage (Kataoka *et al.*, 1983) and its expression was found to be induced as part of the adaptive response to alkylating agents (Kondo *et al.*, 1986). However, despite the early isolation of the *E. coli alkB* gene, no biochemical activity could be found and the function of the *E. coli* AlkB protein remained an enigma.

The *E. coli alkB* mutants are sensitive to S_N2 alkylating agents, and slightly sensitive to S_N1 agents (Kataoka *et al.*, 1983; Chen *et al.*, 1994). The AlkB protein was implicated in the repair or tolerance of DNA alkylation damage rather than the detoxification of alkylating agents, because *E. coli alkB* mutants exhibited a slight defect in the reactivation of MMS treated double stranded bacteriophage λ (Kataoka *et al.*, 1983).

The conservation of the *alkB* gene from bacteria to humans suggested that the AlkB protein plays an important cellular role. Two AlkB homologs have been reported in the literature. The *Caulobacter crescentus* AlkB homolog has 42 % identity and 78 % similarity to the *E. coli* AlkB protein; mutants of this *C. crescentus* protein exhibit sensitivity to MMS. Interestingly, expression of the *C. crescentus* AlkB homolog is cell cycle regulated with a pattern similar to activities required for DNA replication (Colombi and Gomes, 1997). A *Homo sapiens*

AlkB sequence homolog, hABH, was identified as an expressed sequence tag (EST). It was reported that expression of the hABH protein partially complemented the *E. coli alkB* mutant. The hABH gene mapped to chromosome 14(14q24) and was found to be expressed in most human tissues (Wei *et al.*, 1996).

Database searches reveal putative homologs of the *E. coli* AlkB protein in many other species of bacteria and eukaryotes including *Schizosaccharomyces pombe* (30 % identity, 45 % similarity) and *Arabidopsis thaliana* (39 % identity, 47 % similarity) (Figure 6). However, sequence homologs are not found in archaea or in *Saccharomyces cerevisiae*. Three genes, *YFW1, YFW12* and *YFW16,* were identified in a screen to identify *S. cerevisiae* AlkB homologs. These genes functionally complemented the *E. coli alkB* mutants but not the *alkA tag, recA* or *uvrB* mutants. However, these *S. cerevisiae* proteins have no sequence homology to the *E. coli* AlkB protein (Wei *et al.*, 1995). *YFW16* is allelic with the *STE11* gene encoding a serine / threonine-protein kinase that is involved in cell signalling in response to pheromone signals. Sequence analysis predicts that *YFW1* is a transmembrane glycoprotein, and that it shares a domain with the WSC family of cell wall integrity and stress response genes. *YFW12* is expected to belong to the Crisp family of cysteine-rich secretory proteins. The role of these proteins in protecting *E. coli* from alkylation damage is unclear.

Despite extensive testing no DNA repair activity could be assigned to the *E. coli* AlkB protein. *In vivo*, the *E. coli alkB* mutants were proficient in the repair of 3-meA, DNA strand breaks, abasic sites and other secondary lesions which may arise at abasic sites (Dinglay *et al.*, 1998). The AlkB protein exhibited no DNA glycosylase, DNA methyltransferase, AP endonuclease or DNA-dependent

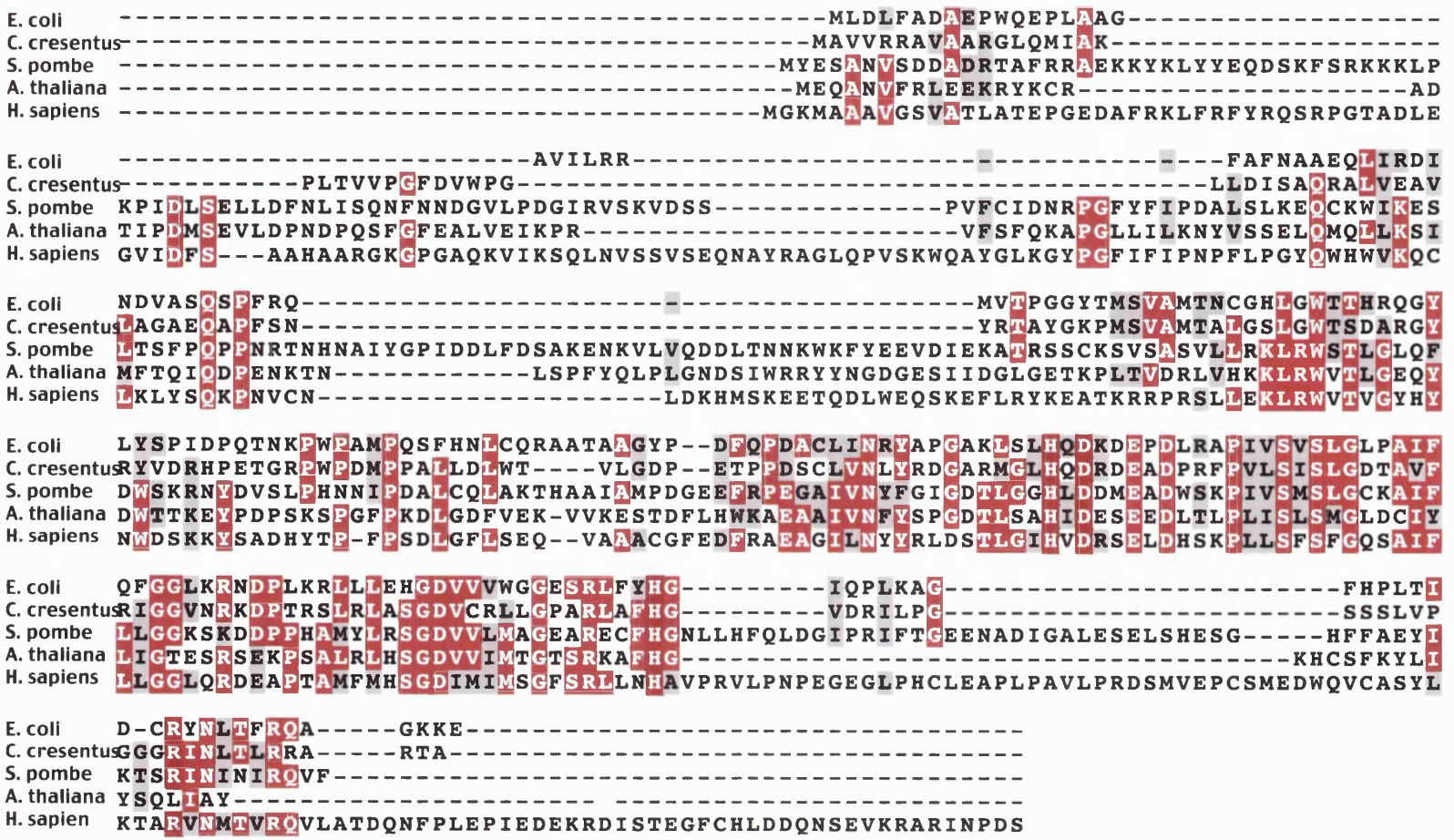


Figure 6. Multiple sequence alignment of several putative homologs of the *E. coli* AlkB protein. Accession numbers, *E. coli* P05050, *C. crescentus* O05725, *S. Pombe* O60066, *A. thaliana* Q9SA98, *H.sapiens* Q13686. **N** = Conserved residues **N** = Similar residues

ATPase activity *in vitro* (Kataoka and Sekiguchi, 1985; Kondo *et al.*, 1986). The failure to identify an *in vitro* activity for the AlkB protein would be explained if AlkB was to act as part of a multi-protein complex or if AlkB acted by regulating transcription of another DNA repair gene. However, as expression of *E. coli alkB* gene confers alkylation resistance to human cells (Chen *et al.*, 1994) and conversely, a human homolog was reported to convey MMS resistance to the *E. coli* mutant (Wei *et al.*, 1996), it was suggested that the AlkB protein works independently. The function of the AlkB protein did not appear to overlap with that of the *E. coli* AlkA, Tag, Ada, MutS, UmuC and UvrA proteins, thus the AlkB protein functions in a pathway independent of these activities. However, a slight functional backup by the RecA protein was observed (Volkert and Hajec, 1991; Dinglay *et al.*, 2000).

In *E. coli alkB* mutants, the modest defect in the reactivation of MMS-treated double stranded phage could be turned into a dramatic defect by using MMS-treated single stranded phage (Dinglay *et al.*, 2000). This suggests that either the AlkB protein is only able to act on single stranded DNA, or that the *E. coli* AlkB protein is able to repair or tolerate DNA alkylation damage that is produced specifically in single stranded DNA. Two lesions 1-methyladenine (1-meA) and 3-meC, are produced at higher levels in single stranded than double stranded DNA and are generated in greater amounts by S_N2 alkylating agents (Singer and Grunberger, 1983). However, *in vitro* the purified AlkB protein showed no activity against the lesions 1-meA or 3-meC in DNA (Dinglay *et al.*, 2000). Recombinant *E. coli* AlkB protein was found to bind approximately 10 fold more efficiently to single stranded DNA than double stranded DNA (Dinglay *et al.*, 2000).

E. coli alkB mutants are weakly sensitive to MMS-induced mutagenesis, showing small increases in the frequencies of G:C to A:T, G:C to T:A and A:T to

T:A base substitutions compared with wild-type strains (Dinglay *et al.*, 2000). The low frequency of mutations in the *E. coli alkB* mutants suggests that AlkB repairs or tolerates lesions that have a low capacity for mispairing during DNA synthesis. As the wild-type *E. coli* exhibit a lower MMS induced mutation frequency than the *alkB* mutant it is suggested that AlkB either repairs DNA damage accurately or is involved in the accurate bypass of replication stalling lesions.

Sequence analysis of the *E. coli alkB* gene had suggested a low level of homology to short regions of the oxidoreductase superfamily (Kondo *et al.*, 1986) or a novel predicted hydrolase domain (Aravind *et al.*, 1999). Neither of these predictions has been substantiated and the biochemical function of the AlkB protein remained a mystery

1.9 Aim

My aim was to develop an *in vitro* assay for the *E. coli* AlkB protein and to determine the biochemical function of the AlkB protein. As several alkylating agents are commonly used as cancer chemotherapy drugs it was also of interest to investigate human AlkB homologs.

CHAPTER 2

MATERIALS AND METHODS

CHAPTER 2. MATERIALS AND METHODS

2.1 Materials

Suppliers

Amersham Biosciences, BDH, Bio-Rad, Fermentas, Fisher Scientific International, Millipore, National Diagnositics, New England Biolabs (NEB), Novagen, PerkinElmer Life Sciences, Pierce, Qiagen, Roche, Sigma-Aldrich, Stratagene, TaKaRa Biosciences, Whatman.

Escherichia coli strains and plasmids

The *Escherichia coli* strains and plasmid used are listed in Table 2.

DNA oligonucleotides

The DNA oligonucleotides used are presented in Table 3.

Media

Luria - Bertani (LB) broth contained 10 g tryptone, 5 g yeast extract and 10 g of NaCl per litre adjusted to pH 7.0 using NaOH. LB agar contained the same ingredients as LB broth but also contained 15 g of agar per litre. NZY broth contained 5 g NaCl, 2 g $MgCl_2$, 5 g yeast extract and 10 g of NZ amine (casein hydrolysate) per litre adjusted to pH 7.5 with NaOH. SOC medium contained 20 g tryptone, 5 g yeast extract, 0.6 g NaCl, 0.2 g KCl, 2 g $MgCl_2$, 2.5 g $MgSO_4$ and 3.6 g of D-glucose per litre. 10 x M9 salts contained 60 g Na_2HPO_4, 30 g KH_2PO_4, 5 g NaCl and 10 g of NH_4Cl per litre adjusted to pH 7.4 with NaOH.

	Genotype	Source
***E. coli* Strain**		
AB1157	*leuB6 proA2 thi-1 hisG4 argE3 lacY1 galK2 ara-14 mtl-1 xyl-5 thr-1 tsx-33 rpsL31 supE44*	Laboratory stock
BS127	As BL21.DE3 but Δ(*ada-alkB*)::cm^r	B.Sedgwick
BS132	As AB1157 but *gyrA*.(DE3)	B.Sedgwick
BS133	As AB1157 but *gryA alkB22*.(DE3)	B.Sedgwick
BS150	As AB1157 but lac ΔU169/F'*proA*$^+$*B*$^+$ *lacI*^{q}ZΔ*M15*::Tn*10*(Tc^R)	B.Sedgwick
NovaBlue	*endA1 hsdR17*(r_{K12}^- m_{K12}^-) *supE44 thi-1 recA1 gyrA96 relA1 lac* F'*proA*$^+$*B*$^+$ *lacI*^{q}ZΔ*M15*::Tn*10*(Tc^R)	Novagen
BL21.DE3	*E. coli* B F$^-$ *ompT hsdS*(r_B^- m_B^-) *gal dcm* (DE3)	Novogen
BL21-CodonPlus(DE3)-RP	*E. coli* B F$^-$ *ompT hsdS*(r_B^- m_B^-) *dcm*$^+$ Tet^r gal λ(DE3) *endA* Hte [*argU proL* cam^r]	Stratagene
Plasmids		
pBAR67	ABH1 gene inserted into pET15b	B. Sedgwick
pBAR54	*alkB*$^+$ gene inserted into pET15b	B. Sedgwick
pSCT1	FLAG-tag inserted and His-tag removed from pBAR67	This study
pSCT2	FLAG-tag inserted and His-tag removed from pBAR54	This study

Table 2. *Escherichia coli* strains and plasmids.

DNA oligonucleotide	Sequence 5' to 3'	Length
116	AGACTACCATGGACTACAAGGACGACGATGACAAAGTCGACCATATGCTCGAG	53
117	CTCGAGCATATGGTCGACTTTGTCATCGTCGTCCTTGTAGTCCATGGTAGTCT	53
124	TTTTTTATTTTTTTATTTTTTTATTTTTTTATTTTTTTATTTTTT	41
133	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA AGAAAAAAGAAAAAAGAAAAAACAAGTCCCACGCTCACACACAATCC	94
134	GGATTGTGTGTGAGCGTGGGACTTG	25
Primer 1	ACACCCACACCCCAACAACCC	21
StA	TGTTTTTGTTTTTGTTTTTGTTTTTATTTTTGTTTTTGTTTTTGTTTTTGGGTTGTTGG GGTGTGGGTGT	70
StC	TGTTTTTGTTTTTGTTTTTGTTTTTCTTTTTGTTTTTGTTTTTGTTTTTGGGTTGTTGG GGTGTGGGTGT	70
Binding I	AACGCTACTACTATTAGTAGAATTGATGCCACCTTTTCAG	40
Poly(dA)	Poly(dA)	Average 310
Poly(dC)	Poly(dC)	Average 370

Table 3. DNA oligonucleotide sequences
The sequence highlighted in yellow encodes the FLAG-tag

2.2 Handling of hazardous materials

Handling of radioactive materials

Disposable gloves, protective glasses and a laboratory coat were worn when handling radioactive materials. Plexiglas shielding was used when handling strong beta emitting radioisotopes such as [^{32}P]. All tubes containing radioactive material were labelled and stored appropriately. Liquid, solid and scintillation waste was disposed of in the containers provided. The laboratory was regularly monitored for radioactive contamination. All radioactivity was used, stored and disposed of in accordance with Cancer Research UK local rules.

Handling of methylating agents

Disposable gloves, protective glasses and a laboratory coat were worn when handling methylating agents. Methylating agents were stored and used in the designated area. All methylating agents were used in the fume hood and were disposed of in 1 M NaOH.

2.3 Sub-cloning to produce FLAG-tagged AlkB and ABH1

Production of plasmid DNA and annealing of DNA oligonucleotides

Small-scale preparations of the pBAR54 and pBAR67 plasmids were made using the QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's instructions.

Two complementary DNA oligonucleotides (116 and 117) were designed, which encoded a FLAG-tag (DYKDDDDK) flanked by NdeI and NcoI restriction sites. The complementary DNA oligonucleotides were annealed by mixing 50 μg

of each of oligonucleotides 116 and 117 in 10 mM Tris.HCl pH 7.5, 1 mM EDTA and 100 mM NaCl in a total volume of 20 µl. The oligonucleotides were heated to 95 °C for 5 min and then allowed to cool slowly to room temperature.

Restriction enzyme digestions

The annealed oligonucleotides and the pBAR54 and pBAR67 plasmids were digested with NdeI and NcoI by incubating 2.5 µg of DNA with 40 units of each of the two restriction endonucleases (NEB) in NEBuffer4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM dithiothreitol (DTT), pH 7.9) at 37 °C overnight. Vector digestion was monitored by gel electrophoresis using 0.6 % agarose gel and visualised by UV.

Digested plasmids and oligonucleotides were purified by addition of an equal volume of saturated phenol / chloroform / isoamyl alcohol (25:24:1). The mixture was vortexed for 30 seconds and the phases were separated by a brief centrifugation at room temperature. The upper phase was recovered and the linear DNA was isolated by the addition of a tenth of the volume of sodium acetate pH 5.2, and three volumes of 100 % ice cold ethanol. After twenty minutes at -80 °C, the DNA was centrifuged at 15,000 g and the DNA pellets were washed twice with 80 % ice cold ethanol, before being dissolved in 10mM Tris.HCl pH 8.0.

Ligation reactions

The 5' phosphate residues were removed from 2.0 µg of linearised plasmid DNA by the addition of 10 units of calf intestinal alkaline phosphatase (CIP) (NEB) in 100 µl of NEBuffer 3 (100 mM NaCl, 50 mM Tris.HCl pH 7.9, 10 mM $MgCl_2$, 1 mM DTT) and incubated at 37 °C for 1 hour. The CIP was inactivated by heating to 75 °C for 15 minutes and the DNA was purified by phenol /

chloroform / isoamyl extraction followed by ethanol precipitation as above. The DNA was dissolved in 10 mM Tris.HCl pH 8.0.

The digested oligonucleotides encoding the FLAG-tag were then ligated into the pBAR54 and pBAR67 plasmids. Ligation reactions, with ratios of 1 : 3 and 1 : 200 of linearised vector to annealed oligonucleotide, were carried out using the Takara (version 2) ligation kit according to the manufacturer's instructions.

E. coli transformations

The ligation reactions were transformed into competent *recA*$^-$ NovaBlue *E. coli* (Novagen) according to the manufacturer's instructions and the bacteria were plated onto agar plates containing 50 μg/ml carbenicillin and incubated overnight at 37 °C. Plasmid DNA was isolated from individual colonies grown up overnight in LB broth supplemented with 50 μg/ml carbenicillin. Plasmids were screened by restriction enzyme digestion and DNA sequencing. The plasmid encoding the FLAG-tagged *E. coli alkB* gene was named pSCT1; the plasmid encoding the FLAG-tagged human ABH1 gene was named pSCT2

The pSCT1 plasmid was transformed into the expression strain *E. coli* BL21.DE3 (Novagen) according to the manufacturer's instructions, bacteria containing the pSCT1 plasmid were selected for by plating on agar containing 50 μg/ml carbenicillin. The pSCT2 plasmid containing the FLAG-tagged human ABH1 gene was transformed into BL21-CodonPlus(DE3)-RP (Stratagene) according to the manufacture's instructions, bacteria containing the pSCT2 plasmid and the plasmid encoding the rare tRNAs were selected for by plating onto LB agar containing 50 μg/ml carbenicillin and 50 μg/ml chloramphenicol.

Complementation of *E. coli alkB* mutant with FLAG-tagged AlkB

The single stranded ϕK phage was used in the complementation assays to avoid the requirement of the M13 phage for F' strains. Phage ϕK lysate was diluted to 5 x 10^4 pfu/ml in M9 minimal salts. MMS was added to a final concentrations of 0 mM, 7.5 mM and 15 mM in a total volume of 300 µl. After incubation at 30 °C for 30 min the reactions were diluted into 3 ml of M9 salts.

The BS132 wild type *E. coli* or the BS133 *alkB* mutant strains were transformed with expression vectors encoding the proteins of interest. These strains were grown in LB broth with 50 µg/ml carbenicillin until they reached an A_{600} of 2.0. The phage were added at various dilutions to 0.4 ml of bacteria in 3 ml of melted soft LB agar supplemented with 5 mM $CaCl_2$. The bacteria were then plated onto LB agar plates containing 50 µg/ml carbenicillin and 5 mM $CaCl_2$ and incubated at 37 °C for 3 hours. Phage survival was monitored by counting the plaques formed in the bacterial lawn.

2.4 Protein purification

Purification of *E. coli* His-tagged AlkB protein

The *E. coli* His-tagged AlkB protein was purified by a protocol modified from that previously described (Dinglay *et al.*, 2000). *E. coli* BL21.DE3/pBAR54 were grown up overnight in a 25 ml of LB broth containing 50 µg/ml carbenicillin. The starter culture was used to inoculate 750 ml of NZY media plus 20 % casein amino acids with 50 µg/ml carbenicillin and incubated at 37 °C with shaking. When the bacterial culture reached an A_{600} of 0.5, expression of the recombinant protein was induced by the addition of 1 mM isopropyl β-D-thiogalactoside (IPTG). After three hours of induction, the cells were harvested by centrifuging at

6000 g for 30 min at 4 °C. The cells were resuspended in 32 ml of ice cold resuspension buffer (50 mM Hepes.KOH pH 8.0, 300 mM NaCl, 5 % glycerol, 2 mM β mercaptoethanol and 2 mM EDTA) supplemented with CompleteTM protease inhibitors (Roche). The cells were lysed by sonication on ice (4 x 30 sec with 30 sec intervals). The cellular debris was removed by centrifugation (14 000 *g*, 20 min, 4 °C) and by filtering through a 0.45 μm filter (Millipore). The EDTA was removed from the lysate by dialysis against 1.5 litres of resuspension buffer minus EDTA (50 mM Hepes.KOH pH 8.0, 300 mM NaCl, 5 % glycerol and 2 mM β-mercaptoethanol) for approximately 3 hours at 4 °C. A 1 ml Ni-NTA (nitrilotriacetic acid) agarose column (Qiagen) was prepared by equilibration in the wash buffer (50 mM Hepes.KOH pH 8.0, 300 mM NaCl, 5 % glycerol, 2 mM β-mercaptoethanol and 1 mM imidazole). The dialysed extract was supplemented with 1 mM imidazole and loaded onto the Ni-agarose column. The column was washed with 20 ml of wash buffer and then 30 ml of wash buffer containing 40 mM imidazole followed by 5 ml of wash buffer containing 60 mM imidazole. The AlkB protein was eluted in wash buffer containing 250 mM imidazole; 1 ml fractions were collected. The fractions containing the purified His-tagged AlkB protein were identified by A_{280} readings and by using SDS-polyacrylamide gel electrophoresis. The purified His-tagged AlkB protein was dialysed into 30 mM potassium phosphate pH 7.5, 2 mM β-mercaptoethanol, 50 % glycerol, 300 mM NaCl, aliquots were stored at -80 °C.

Purification of *E. coli* FLAG-tagged AlkB protein

E. coli BL21.DE3/pSCT1 were grown up overnight in 25 ml of LB broth containing 50 μg/ml carbenicillin. 750 ml of LB with 50 μg/ml carbenicillin was inoculated and incubated at 37 °C with shaking. Expression of the recombinant

protein was induced by the addition of 0.5 mM IPTG to the cell culture at A_{600} 0.5. After three hours of induction, the cells were harvested by centrifugation at 6000 g for 30 min at 4 °C. The cells were resuspended in TBS (50 mM Tris.HCl pH 7.4, 150 mM NaCl) supplemented with 2 mM EDTA and CompleteTM protease inhibitors (Roche). The cells were lysed by sonication on ice (4 x 30 seconds, with 30 second intervals) and cellular debris was removed by centrifugation (14,000 *g*, 20 min, 4 °C) and filtration through a 0.45 μm filter (Millipore). The supernatant was dialysed against 1.5 litres of TBS for 3 hours and then applied to a 1 ml column of anti-FLAG M2 affinity gel (Sigma-Aldrich) pre-equilibrated with TBS. The column was washed with 32 column volumes (4 x 8 ml) of TBS (50 mM Tris.HCl pH 7.5, 150 mM NaCl). The protein was eluted with 5 column volumes (5 x 1 ml) of FLAG peptide (Sigma-Aldrich) at a concentration of 100 μg/ml suspended in TBS. 0.5 ml fractions were collected and the fractions containing the purified FLAG-tagged AlkB protein were identified by SDS-polyacrylamide gel electrophoresis. Centricon YM-10 concentrators (Millipore) pre-blocked with 1 % dried milk were used to concentrate the purified protein, remove the FLAG peptide and to exchange the buffer to 30 mM potassium phosphate pH 7.0, 2 mM β-mercaptoethanol, 50 % glycerol and 300 mM NaCl. The purified FLAG-tagged AlkB was stored in aliquots at -80 °C.

Purification of human FLAG-tagged ABH1 protein

E. coli BL21.DE3CodonPlus/pSCT2 were grown up overnight in 25 ml of LB broth containing 50 μg/ml carbenicillin and 50 μg/ml chloramphenicol. The overnight culture was used to inoculate 750 ml of LB containing 50 μg/ml carbenicillin and 50 μg/ml chloramphenicol, which was incubated at 37 °C with shaking. Expression of the recombinant protein was induced when the culture

reached an A_{600} 0.6 by the addition of 100 μM IPTG and incubation for 3 hours at 25°C. FLAG-tagged ABH1 was purified by affinity chromatography as described above for the FLAG tagged *E. coli* AlkB protein.

SDS-polyacrylamide gel electrophoresis

Protein samples were mixed with an equal volume of SDS sample buffer (4 % SDS, 125 mM Tris.HCl pH 6.8, 10 % β-mercaptoethanol, 20 % glycerol and 0.05 % bromophenol blue), heated to 90-100 °C for 10 min and loaded onto 12 % SDS-polyacrylamide gels. Pre-stained broad range SDS-standard markers (NEB) were also loaded and gels were run using the Mini-Protean II gel apparatus (BioRad). Gels were stained either by Coomassie blue or by using the Silver Stain Plus Kit (BioRad) according to the manufacturer's instructions.

Measurement of protein concentrations

The protein concentrations were determined by the method of Bradford (Bradford, 1976), using bovine serum albumin (BSA) as a standard.

2.5 *In vitro* assays of AlkB, ABH2 and ABH3

Preparation of oligonucleotides and M13 DNA

DNA oligonucleotide 116 was synthesised by the Cancer Research UK DNA synthesis facility on a commercial DNA synthesiser using cyanoethyl chemistry. To remove unincorporated precursors, the DNA oligonucleotide was dissolved in 0.3 M sodium acetate pH 5.2, 10 mM $MgCl_2$ and precipitated with three volumes of 100 % ice cold ethanol, before being dissolved in 10 mM Tris.HCl pH 8.0. After twenty minutes at -80 °C, the DNA oligonucleotide were centrifuged at 15,000 g and washed twice with 80 % ice cold ethanol. Poly(dA),

average length 310 residues and poly(dC), average length 370 residues, were purchased from Amersham Biosciences. The M13mp18 single stranded DNA was prepared as described (Sambrook *et al.*, 1989). Briefly, 5 ml of LB broth was inoculated with the *E. coli* strain BS150 and grown to an A_{600} of 0.5. The *E. coli* were infected with M13 phage (10^{10} pfu) and left to stand at room temperature for 5 minutes. 250 ml of pre-warmed LB broth was added and the bacteria were incubated at 37 °C for 5 hours with shaking. The bacteria were removed by centrifugation at 4000 g for 15 min at 4 °C. The supernatant was collected and the M13 phage particles were precipitated by the slow addition of 10 g of PEG 8000 and 7.5 g of NaCl over 1 hour with stirring. The phage were pelleted by centrifugation (10,000 g, 20 min, 4 °C) and resuspended in 10 mM Tris.HCl pH 8.0. The phage DNA was isolated by phenol / chloroform extraction and ethanol precipitation.

The DNA oligonucleotides and M13 DNA were each dissolved in 10 mM Tris.HCl pH 8.0. The concentration of the DNA was estimated from A_{260} readings using a spectrophotometer (Ultrospec II, LKB, Biochrom). An A_{260} of 1.0 corresponds to a single stranded DNA concentration of approximately 37 μg/ml.

[^{14}C]- and [^{3}H]- methylated DNA substrates

[^{3}H]DMS was obtained by custom order from PerkinElmer Life Sciences. The hexane was evaporated from 1 mCi of [^{3}H]DMS (1.74 Ci/mmole) by blowing nitrogen gas over the surface. The [^{3}H]DMS was added to 1.2 mg of poly(dA), poly(dC) or M13 DNA suspended in 10 mM sodium cacodylate pH 7.0. The reactions were incubated at 30 °C for 1 hour. The DNA was recovered by ethanol precipitation with the addition of 0.3 M sodium acetate pH 7.0, 10 mM $MgCl_2$ and three volumes of 100 % ice cold ethanol, incubtion at -80 °C followed by

centrifugation at 15,000 g for 20 minutes at 4 °C. The DNA pellets were washed twice with ice cold 80 % ethanol and dissolved in 10 mM Tris.HCl pH 8.0.

[^{14}C]MeI was obtained from Amersham Biosciences. 1 mCi of [^{14}C]MeI (58mCi/mmol) was condensed on dry ice for 2 hours. 2 mg of DNA oligonucleotide 116 or 0.6 mg of poly(dA) were dissolved in 0.8 ml of 10 mM sodium cacodylate pH 7.0 and 1.2 mg of poly(dC) was dissolved in 1.3 ml 50 mM Hepes.KOH pH 8.0. The DNA oligonucleotides were added to the [^{14}C]MeI and incubated at 30 °C for 6 hours. The DNA was recovered by ethanol precipitation and the DNA pellets were washed three times with 80 % ice cold ethanol.

The [^{3}H]- and [^{14}C]- methylated poly(dA) were concentrated by centrifugation at 5000 g in a Centricon YM-10 concentrator (Millipore) and washed with 10 mM Tris.HCl pH 8.0 to remove any unincorporated radioactivity. Due to the relative insolubility of poly(dC), the [^{3}H]- and [^{14}C]- methylated poly(dC) was instead dialysed against 2 litres of 10 mM Tris.HCl pH 8.0 for 5 hours. The specific activities of the [^{3}H]- and [^{14}C]- methylated DNA oligonucleotides were determined by scintillation counting.

Double stranded [^{14}C]-methylated DNA oligonucleotide

The [^{14}C]-methylated DNA oligonucleotide 116 was added to a 1.2 fold excess of the unmethylated complementary DNA oligonucleotide 117. These oligonucleotides were annealed by incubation at 60 °C for 2 min in the presence of 25 mM KCl followed by a slow cooling to room temperature. The DNA was confirmed to be double stranded by exonuclease III digestion. 100 U of exonuclease III (NEB) was added to 7 μg of the annealed oligonucleotides and incubated at 37 °C in the presence of NEBuffer 1 (10 mM Bis Tris Propane.HCl pH 7.0, 10 mM $MgCl_2$, 1 mM DTT) for 1 hour. The reactions were stopped by the

addition of 20 mM EDTA, the DNA was ethanol precipitated and the supernatant was recovered. Scintillation counting monitored the release of radioactive material from the DNA.

Methylated DNA substrates used to monitor direct reversal of DNA damage

A single stranded substrate containing [^{3}H]-cytosine residues interspersed between thymine residues was synthesised by the polymerase chain reaction (PCR) using a single primer. The DNA template, oligonucleotide 133, was replicated using Pfu polymerase and a 10 fold excess of a single primer (DNA oligonucleotide 134), in the presence of 150 μM dNTPs and 2.5 μM [^{3}H]-dCTP (16 Ci/mmol) (Amersham Biosciences). The PCR reactions were performed in a DNA thermal cycler 480 (Perkin Elmer) using 60 PCR cycles of 96 °C for 30 sec, 50 °C for 15 sec and 72 °C for 2 min. The DNA was recovered either by using the Qiaquick PCR purification kit (Qiagen) according to the manufacturer's instructions, or if pooling several PCR reactions, by phenol / chloroform extraction followed by ethanol precipitation. To obtain a high level of methylation, the [^{3}H]-PCR product was treated eight times with 50 mM DMS in 75 mM sodium cacodylate pH 7.4 at 30 °C for 2 hours. Between each treatment the DMS was removed by centrifugation of the DNA through a G-25 Sephadex mini-column (Amersham Biosciences) equilibrated in 75 mM sodium cacodylate pH 7.4, except for the final treatment for which the mini column was equilibrated in 10 mM Tris.HCl pH 8.0.

DNA oligonucleotide 124, which contains mostly thymine residues interspersed with adenine residues, was treated with 50 mM DMS in 75 mM sodium cacodylate, pH 7.4 at 30 °C, 2 times for 2 hours and then 4 times for 1 hour. Between each treatment the DMS was removed by centrifugation through a

G-25 Sephadex mini-column (Amersham Biosciences) equilibrated in 75 mM sodium cacodylate pH 7.4, except for the final treatment for which the mini column was equilibrated in 10 mM Tris.HCl pH 8.0.

AlkB, ABH2 and ABH3 *in vitro* assay conditions

The purified *E. coli* AlkB protein and the human ABH2 and ABH3 proteins were incubated with the methylated DNA substrates in a variety of conditions. The optimal conditions for the *E. coli* AlkB protein were determined to be 50 mM Hepes.KOH pH 8.0, 75 μM $Fe(NH_4)_2(SO_4)_2.6H_2O$, 1 mM α-ketoglutarate, 2 mM ascorbic acid and 50 μg/ml BSA for 15 min at 37 °C. Unless otherwise stated the purified human ABH2 protein was incubated with the DNA substrates in 50 mM Hepes.KOH pH 7.5, 25 μM $Fe(NH_4)_2(SO_4)_2.6H_2O$, 1 mM α-ketoglutarate, 2 mM ascorbic acid and 50 μg/ml BSA and the purified human ABH3 protein was incubated with DNA in 50mM Hepes.KOH pH 8, 75 μM $Fe(NH_4)_2(SO_4)_2.6H_2O$, 1 mM α-ketoglutarate, 2 mM ascorbic acid and 50 μg/ml BSA.

The AlkB, ABH2 and ABH3 reactions were stopped by the addition of EDTA to 11 mM and the DNA was ethanol precipitated by the addition of 0.3 M sodium acetate pH 7.0 and three volumes of ice cold ethanol in the presence of carrier calf thymus DNA. After incubation at -80 °C for 20 minutes, the DNA was pelleted by centrifugation at 15,000 g for 20 minutes at 4 °C. Approximately two thirds of the volume of the supernatants were transferred into scintillation vials and 5 ml of EcoScintA (National Diagnostics) was added. The radioactivity present was monitored using a Beckman LS60000IC scintillation counter. The data is expressed as the average of at least two replicates and each experiment was

repeated twice. The DNA pellets were washed twice with ice cold 80% ethanol, vacuum dried and stored at -20 °C for further analysis.

As approximately 62 % of the volume of the supernatants were removed for scintillation counting, when the AlkB, ABH2 or ABH3 reactions went to completion; approximately 620 cpm were recovered from 1000 cpm of poly(dC) substrate. From 1000 cpm of the poly(dA) substrate, when the reaction went to completion, approximately 400 cpm were recovered.

2.6 High Performance Liquid Chromatography (HPLC)

DNA hydrolysis

Purine bases were released from DNA by weak acid hydrolysis. The DNA pellets were resuspended in 10 µl of 0.1 M HCl and heated to 95 °C for 1 hour. Purine and pyrimidine bases were released from DNA substrates by strong acid hydrolysis using 90 % formic acid. 90 µl of 100% formic acid was added to 10 µl of DNA, the mixture was sealed in a glass vial and heated to 180 °C for 20 min. After cooling for 15 minutes at room temperature the acid mix was evaporated to dryness then resuspended in 10 µl of 0.1 M HCl.

HPLC

Methylated bases were analysed by HPLC on a Beckman system Gold - Programmable Solvent Module 126 and using the Beckman System Gold - Programmable Detector Module 166 for UV detection.

After weak acid hydrolysis of the DNA substrates the purines released were separated by HPLC. The purine bases were injected onto a Whatman partisil 10 cation exchange column pre- equilibrated in 0.1 M ammonium formate pH 3.6 and 20 % methanol, a gradient of methanol from 20 to 40 % was applied.

After the strong acid hydrolysis of DNA substrates, 3-meC was separated from other methylated bases including N^4-methylcytosine (N^4-meC) and 5-methylcytosine (5-meC) on a Whatman Partisil 10 cation exchange column pre-equilibrated in 0.1 M ammonium formate pH 3.6 and 5 % methanol, a gradient of methanol from 5 to 40 % was applied.

An HPLC separation was developed to obtain baseline resolution between 3-meC, cytosine, thymine and uracil bases. The bases released from strong acid hydrolysis of the DNA substrates were separated using an isocratic separation of 100 mM ammonium formate pH 3.6 and 2.5 % methanol on a Whatman partisil 10 cation exchange column

Quantification of bases

The identification of DNA bases eluted from the HPLC separations was achieved by co-migration with known standards. Bases released from non-radioactive DNA substrates were quantified by UV absorption, for example, the absorption of 1-meA and adenine bases was measured at A_{260}. The relative absorption of adenine to 1-meA at this wavelength was determined to be 1.04 by monitoring known amounts of these purines in the same conditions, this ratio was taken into account in the quantification. DNA bases hydrolysed from radioactive DNA substrates were quantified by scintillation counting. The HPLC column eluate was collected in fractions every 0.31 min (18.6 seconds), 5 ml of EcoScintA (National Diagnostics) was added to each fraction and the radioactivity present in the samples was monitored using a Beckman LS60000IC scintillation counter.

2.7 Primer extension assay

5' terminal [^{32}P]-labelling of primer

15 pmoles of primer1 were incubated with 150 μCi of adenosine 5'-[γ-^{32}P]-triphosphate (ATP)(Amersham Biosciences) and 60 units of T4 polynucleotide kinase (PNK) (NEB) which catalyses the transfer and exchange of phosphate from the γ position of ATP to the 5´ hydroxyl terminus of polynucleotides. The reaction was carried out in T4 PNK reaction buffer (70 mM Tris.HCl pH 7.6, 10 mM $MgCl_2$, 5 mM DTT) and was incubated at 37 °C for 40 minutes. The PNK was inactivated by incubation at 65 °C for 5 minutes. The unreacted radiolabelled ATP was removed by centrifugation through a G-50 Sephadex mini-column (Amersham Biosciences) equilibrated in 10 mM Tris.HCl pH 8.0.

DMS treatment of template DNA oligonucleotides

The DNA oligonucleotides, StA and StC, were suspended in 50 mM sodium cacodylate pH 7.0, 1 mM EDTA and various amounts of DMS were added. The methylation reactions were incubated at 30 °C for 30 minutes, the DMS was removed by centrifugation through G-50 Sephadex mini-columns (Amersham Biosciences) equilibrated in 10 mM Tris.HCl pH 8.0.

Pre-incubation of template DNA with AlkB in the optimised assay conditions

Methylated or unmethylated StA and StC DNA oligonucleotides were incubated with or without purified His-tagged AlkB in the optimised assay conditions (50 mM Hepes.KOH pH 8.0, 75 μM $Fe(NH_4)_2(SO_4)_2.6H_2O$, 1 mM α-ketoglutarate, 2 mM ascorbic acid and 50 μg/ml BSA) for 15 min at 37 °C. The low molecular weight co-factors were removed by centrifugation through a G-25

Sephadex mini-column (Amersham Biosciences) equilibrated in 10 mM Tris.HCl pH 8.0.

Primer extension reactions

200 fmoles of [^{32}P] labelled primer1 was incubated with 200 fmoles of the methylated, unmethylated or AlkB pre-incubated oligonucleotide StA and StC in the presence of 100 μM dNTPs and 1 x buffer (50 mM Tris.HCl pH 7.5, 5 mM $MgCl_2$, 1 mM DTT, 50 mM NaCl) in a total volume of 9 μl. To allow annealing of the DNA oligonucleotides, the reaction mixture was incubated for 5 min at room temperature before addition of 2 Units of the Klenow exo$^-$ DNA polymerase (NEB). The primer extension reactions were incubated at 37 °C for 15 minutes before being stopped by the addition of 10 μl of gel loading buffer (98 % deionised formamide, 20 mM EDTA, 0.025 % xylene cyanol and 0.025% bromophenol blue). The DNA was denatured by boiling at 95 °C for 10 minutes, reannealing was prevented by placing the samples immediately on ice. 2 μl of the reaction was loaded onto a DNA sequencing gel.

DNA sequencing gels

The reaction products were separated on a 0.4 mm, 12 % sequencing gel (SequaGel, National Diagnosis) run for 2 hours at 1400 volts on a Life Technologies model SA-32 gel apparatus using 90 mM Tris-borate pH 8.3, 2 mM EDTA as a running buffer. The gels were dried onto 3MM paper by heating to 80 °C for one hour on a gel dryer (Model 583, Bio-Rad) and exposed overnight to a molecular dynamics phosphor screen, then scanned and quantitated using the Molecular Dynamics Storm 860 Phosphoimager machine.

2.8 Binding of His-tagged AlkB to DNA

A 40mer oligonucleotide (Table 3, Binding I) was 5' phosphorylated using [γ-^{32}P]ATP and T4 PNK (see section 2.7) The oligonucleotide was then treated with 300 mM MMS or 100 mM MNU at 30 °C for 30 min. The methylating agent was removed by centrifugation through a Sephadex G50 column equilibrated in 10 mM Tris.HCl and 1 mM EDTA, pH 8.0. Varying amounts of His-tagged AlkB protein were incubated with the [^{32}P]-5' end labelled DNA oligomers (30,000 cpm / reaction) in 20 µl of 20 mM Tris.HCl pH 7.5, 100 mM KCl, 0.1 mM DTT, 10 % glycerol, at 30 °C for 30 min. After the addition of 1 ml of ice-cold buffer, the reaction was immediately filtered through nitrocellulose disc filters (HAW P02500 Scheibenfilter, Millipore) using a vacuum filter apparatus (Millipore). The filters were then washed with 10 ml of buffer and dried. The DNA bound to AlkB protein was quantitated by scintillation counting.

2.9 Sequence analysis

Nucleotide sequences and alignments were derived from standard BLAST searches of the nucleotide / protein database at the National Centre for Biotechnology Information (NCBI). Multiple alignments were performed using ClustalX and residues that are similar or identical to the consensus in the multiple alignment were identified using BOXSHADE software.

CHAPTER 3

RESULTS

ASSAY FOR DNA REPLICATION BY-PASS OF DNA METHYLATION DAMAGE

CHAPTER 3. ASSAY FOR DNA REPLICATION BY-PASS OF DNA METHYLATION DAMAGE

The function of the *E. coli* AlkB protein has been a longstanding mystery. *In vivo*, *E. coli alkB* mutants are sensitive to S_N2 methylating agents (Kataoka *et al.*, 1983) and are defective in the reactivation of MMS-treated single stranded DNA phage. Therefore, it was proposed that AlkB may either repair or tolerate DNA alkylation damage by acting on single stranded DNA, or by processing lesions generated in single stranded DNA by S_N2 agents, such as 1-meA or 3-meC (Dinglay *et al.*, 2000). Recombinant *E. coli* AlkB protein was found to bind approximately 10 fold more efficiently to single stranded DNA than double stranded DNA. Methylation of single stranded and double stranded DNA substrates with high doses of the S_N2 agent MMS (300 mM) increased the AlkB binding affinity by approximately twofold in both cases (Dinglay *et al.*, 2000). However, a larger increase in binding affinity was observed on pretreatment of single stranded DNA with 100 mM of the S_N1 agent MNU (Appendix A). Therefore, the stimulation of binding observed by high doses of methylation may reflect altered structural properties of the heavily alkylated DNA rather than a binding to a specific lesion processed by AlkB.

3.1 Potential role for AlkB in the tolerance of DNA methylation damage

Attempts to pinpoint the biochemical function of the AlkB protein have been unsuccessful. *In vitro*, the AlkB protein has been assayed for several enzymatic activities including DNA glycosylase, DNA methyltransferase, AP

endonuclease or DNA-dependent ATPase activities (Kondo *et al.*, 1986). There are several possible explanations for this apparent lack of *in vitro* enzymatic activity. For, example, the purified AlkB protein used in these assays may not have been active. The AlkB protein may have been inactivated during the protein purification process; it may lack an important cofactor or need to interact with another protein for activity. However, the ability of the *E. coli* AlkB protein to confer alkylation resistance to human cells implies that AlkB protein can act independently of other proteins (Chen *et al.*, 1994). It is possible that the AlkB protein was active in these *in vitro* assays but that the wrong substrates were used. However, either *in vitro* or by analysis of the *E. coli alkB* mutant, AlkB was tested for activity against all of the major DNA lesions generated by S_N2 methylating agents, including 3-meA, 7-meG, 1-meA and 3-meC, as well as FaPy and abasic sites, but no activity was observed (Kondo *et al.*, 1986; Dinglay *et al.*, 1998; Dinglay *et al.*, 2000). It is possible that the AlkB protein may not have an enzymatic activity, for example, AlkB may bind to and modify the structure of DNA.

In vitro assays for AlkB activity have almost exclusively tested the ability of the AlkB protein to repair DNA damage. However, the AlkB protein may be involved in the tolerance of alkylation damage. As *E. coli alkB* mutants have a low mutation frequency (Dinglay *et al.*, 2000) the AlkB protein could allow accurate bypass of lesions that stall DNA replication rather than mutagenic lesions. S_N2 methylating agents produce two lesions at significant levels that are known to stall DNA replication, 3-meA and 3-meC. 1-meA may also be a DNA replication stalling lesion as it may be unable to form stable base-pairs (Abbott and Saffhill, 1977; Boiteux and Laval, 1982; Saffhill, 1984; Beard *et al.*, 1996; Eom *et al.*, 1996). An *in vitro* assay was developed to determine whether AlkB could relieve stalling of DNA replication at 3-meA, 3-meC or 1-meA.

3.2 Development of an *in vitro* DNA replication stalling assay

To assay DNA replication stalling at 3-meC, 1-meA and 3-meA, DNA templates were required that contained these lesions at defined positions. Two oligonucleotides, StC and StA, were designed which contained mainly thymine residues interspersed with guanine residues, a single cytosine or adenine residue was located 45 nucleotides from their 5' end respectively (see Table 3). These oligonucleotides were methylated with an S_N2 methylating agent. The thymine and guanine residues should not be methylated to form any lesions that substantially stall DNA replication. It is expected that methylation of the cytosine residues at position 45 of the StC DNA oligonucleotide will produce 3-meC lesions. Whereas, methylation of the adenine residues at position 45 of the StA DNA oligonucleotide should generate a mix of methylated lesions. Two thirds of the lesions generated should be 1-meA and a sixth should be 3-meA, a small amount of the innocuous lesion 7-meA is also produced (Singer and Grunberger, 1983). The dose of DMS with which the template DNA was methylated was optimised to 5 mM for 30 minutes. At this dose of methylating agent a significant number of the adenines or cytosine residues in the template DNA oligonucleotides were methylated, but the integrity of the DNA was maintained.

A [^{32}P]-labelled DNA oligonucleotide (Primer 1), 21 nucleotides in length, was annealed to the methylated template DNA. This labelled primer was extended by *E. coli* Klenow exo$^-$, an *E. coli* DNA polymerase I derivative with no exonuclease activity. Ideally, the *E. coli* replicative polymerase, DNA polymerase III would have been used, however this was not commercially available. The reaction products were separated on a DNA sequencing gel. Maximal stalling was

observed at the methylated lesions when the polymerase extension reactions were carried out with 5 mM $MgCl_2$ for 15 minutes at 37 °C.

Figure 7 shows that in the replication bypass assay the 3-meC residues and the mix of 1-meA and 3-meA residues stall DNA replication. Under the conditions used, approximately half of the stalling is at the base before the lesions and approximately half of the stalling is opposite the lesions.

3.3 AlkB does not relieve stalling of DNA replication

10 pmole and 100 pmole of AlkB protein were added to primer extension reactions. The His-tagged AlkB protein did not appear to relieve stalling at the 3-meC or 1-meA / 3-meA residues (Figure 8). When 100 pmoles of AlkB were added to the assay, the polymerase became unable to incorporate a nucleotide opposite the lesions, this is possibly due to the high salt concentration of the AlkB storage buffer. Under the conditions tested, His-tagged AlkB, did not relieve stalling of DNA replication at 3-meC, 1-meA or 3-meA. Therefore, AlkB did not appear to repair or allow DNA replication bypass of the main replication stalling lesions produced by S_N2 methylating agents under the assay conditions used.

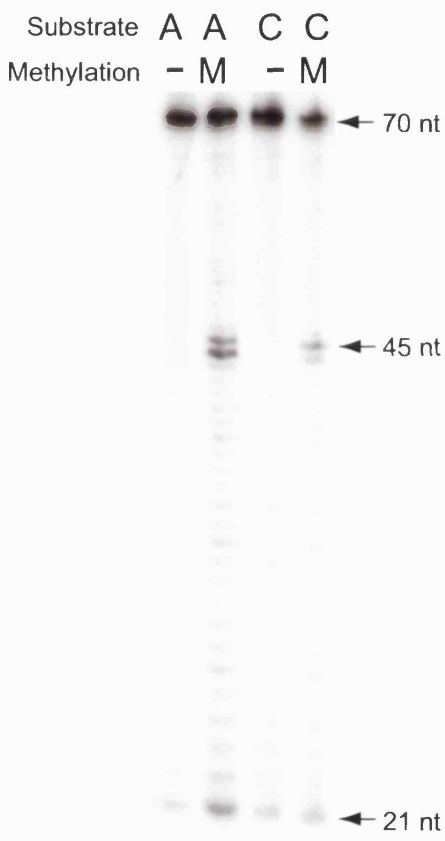


Figure 7. Stalling of DNA replication at 1-meA / 3-meA and 3-meC. A [^{32}P]-labelled 21mer oligonucleotide was annealed to a methylated (M) or unmethylated (-) 70mer DNA template that contained an A or C residue 45 nucleotides from its 5' end. The 21mer was extended by 40 pmoles of Klenow exo$^-$. The reactions were analysed on a 12 % sequencing gel.

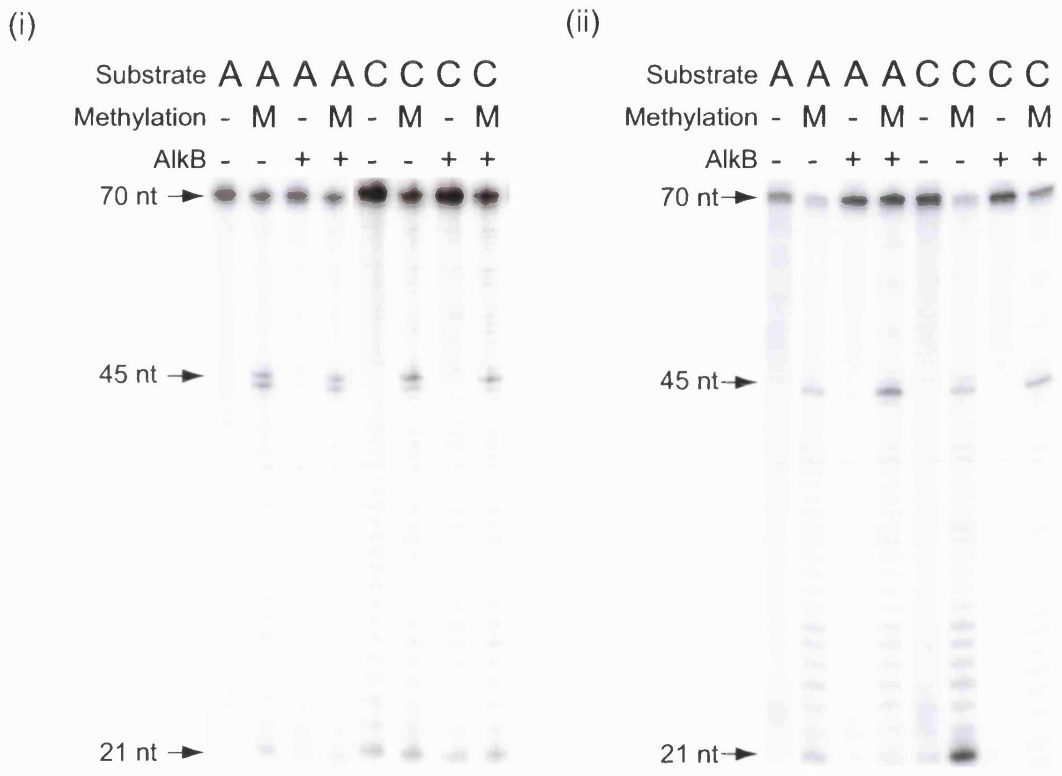


Figure 8. AlkB did not relieve stalling of DNA replication at 1-meA / 3-meA or 3-meC. The [^{32}P]-labelled 21mer was extended by Klenow exo$^-$ using DNA templates which contained either an A or C residue 45 nucleotides from their 5' end. These templates were methylated (M) or unmethylated (-) and the extensions were carried out in the presence or absence of (i) 10 pmoles of AlkB, or (ii) 100 pmoles of AlkB. The reactions were analysed on a 12% sequencing gel.

CHAPTER 4

RESULTS

THE *E. coli* ALKB PROTEIN REPAIRS DNA DAMAGE BY A DIRECT REVERSAL MECHANISM

CHAPTER 4. THE *E. coli* ALKB PROTEIN REPAIRS DNA DAMAGE BY A DIRECT REVERSAL MECHANISM

Soon after the discovery of the *alkB* gene, it was suggested that the AlkB protein had weak sequence homology to oxidoreductases (Kondo *et al.*, 1986). In 1999 it was proposed that the AlkB protein was a member of a novel family of hydrolases (Aravind *et al.*, 1999). However, no evidence was forthcoming to support either of these proposals and it appeared that more powerful sequence analysis would be required to identify any true homology between AlkB and proteins of known function.

Many functionally and evolutionarily important protein similarities are recognisable only through comparison of three-dimensional protein structures. As the number of protein folds is finite and representative structures exist for the most widespread folds, structural and functional information can be obtained if a reliable alignment can be established between the protein of interest and the amino acid sequence of a domain of known structure. However, this task is not trivial as when new protein families evolve from old ones, although the structures tend to remain basically conserved, the sequence rapidly diverges. The recent development of highly sensitive sequence comparison tools, such as PSI (Position Specific Iterating)-BLAST, significantly increases the ability to detect subtle sequence similarities and to make non-trivial structural predictions. PSI-BLAST is very effective as it constructs a multiple alignment from BLAST hits, then creates a position-specific weight matrix and uses it to iterate the search (Altschul and Koonin, 1998).

In 2001, Aravind and Koonin used sequence profile searches to suggest that AlkB had similarity to members of the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamiliy (Aravind and Koonin, 2001). All members of this superfamily have a double-stranded β-helix (DSBH) region that includes seven conserved strands arranged into two sheets in a jelly-roll topology (Roach *et al.*, 1997; Valegard *et al.*, 1998; Zhang *et al.*, 2000). Secondary structure prediction suggested that AlkB may contain this fold (Aravind and Koonin, 2001). Different members of the superfamily have specific inserts between the conserved strands of the DSBH fold, this accounts for the difficulty of identifying members of this superfamily by sequence based methods.

4.1 The α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily

The α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily is the largest subgroup of the mononuclear nonheme-Fe(II)-dependent oxygenases. Members of this superfamily use iron-oxo intermediates to catalyse a large variety of two electron oxidations. The general mechanism for reactions catalysed by this superfamily is shown in Figure 9. Members of this superfamily are distinguished by their requirement for Fe(II) and an α-keto acid co-substrate, typically α-ketoglutarate, for activity (Que and Ho, 1996). An atom of Fe(II) is co-ordinated in the enzyme's active site, α-ketoglutarate binds in a bidentate fashion to the Fe(II) ion, this triggers binding of the substrate and primes the active site for the activation of a molecule of dioxygen (Holme, 1975; Kivirikko and Pihlajaniemi, 1998). During the reaction, α-ketoglutarate is oxidatively decarboxylated to

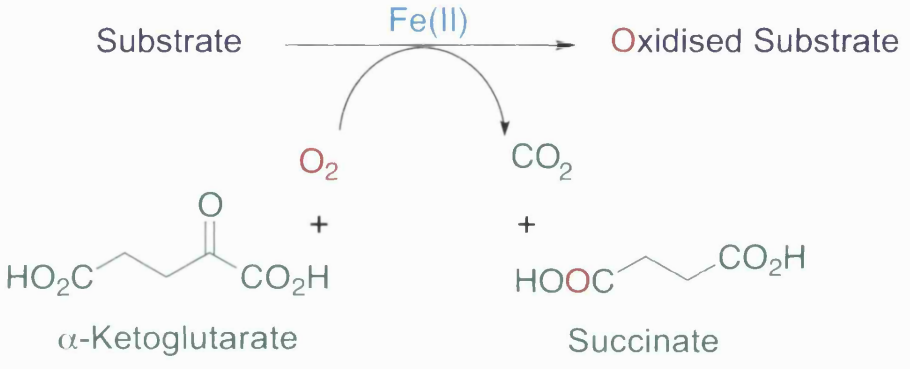


Figure 9. Generalised reaction mechanism for members of the α-ketoglutarate dependent-Fe(II)-dependent dioxygenase superfamily.

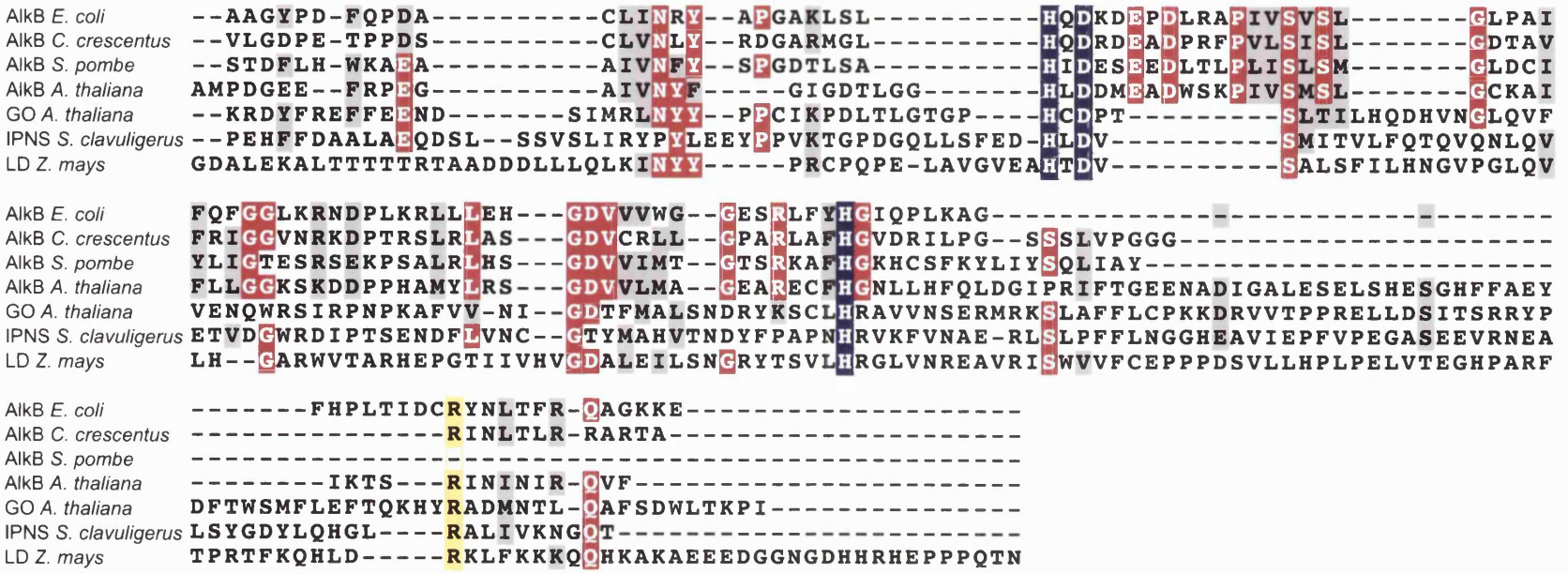


Figure 10. Putative AlkB homologs from several species contain the HXDX$_N$H iron binding motif. A multiple alignment of AlkB proteins with members of the α-ketoglutarate dependent-Fe(II) dioxygenase superfamily, GO = Gibberellin 20-oxidase (Q38844), IPNS = Isopenicillin N synthetase (P10621), LD = Leucoanthocyanidin dioxygenase (P41213).

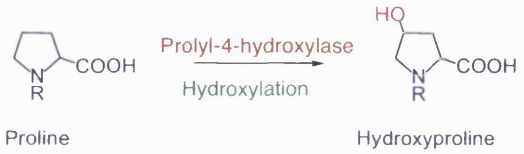
N = HXDX$_N$H iron binding motif N = Proposed interaction with α-ketoglutarate N = Other Identical residues N = Similar residues

produce succinate and CO_2 and a highly reactive ferryl (Fe(IV)=O) species is formed, which oxidises the prime substrate (Rohde *et al.*, 2003). The reactions catalysed by this superfamily do not require an external reducing agent for turnover, however ascorbic acid has been observed to stimulate activity of several superfamily members *in vitro*. It is not clear how this stimulation is achieved, however, it has been proposed that ascorbic acid may reduce inactive Fe(III) to Fe(II) and protect the enzyme from oxidative self-inactivation (Que and Ho, 1996).

AlkB appears to contain two motifs characteristic of members of this superfamily, an $HXDX_NH$ motif that is important for Fe(II) co-ordination and an arginine further downstream which may be involved in interactions with α-ketoglutarate (Figure 10).

The $HXDX_NH$ amino acid sequence, which is highly conserved amongst the α-ketoglutarate-Fe(II)-dependent dioxygenases forms a 2-His-1-carboxylate facial triad of Fe(II) co-ordinating residues. The 2-His-1-carboxylate facial triad is present in all mononuclear non-heme Fe(II) enzymes (Que, 2000), it anchors the iron in the active site and occupies one face of the Fe(II) co-ordination sphere, leaving the opposite face available to bind up to three ligands, such as other endogenous amino acid groups or exogenous ligands. This arrangement provides great variability in iron co-ordination environment and allows the mononuclear non-heme Fe(II) enzymes to use diverse mechanistic strategies to perform a variety of chemical transformations (Que, 2000). Different members of α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily catalyse various two electron oxidations, including hydroxylations, desaturation of carbon-carbon bonds and ring expansion (Figure 11). For example, deacetoxycephalosporin C synthase (DAOCS) catalyses the ring expansion of isopenicillin N to

(a) Hydroxylation by prolyl-4-hydroxylase



(b) Ring expansion by DAOCS

(c) The trifunctional role of CAS

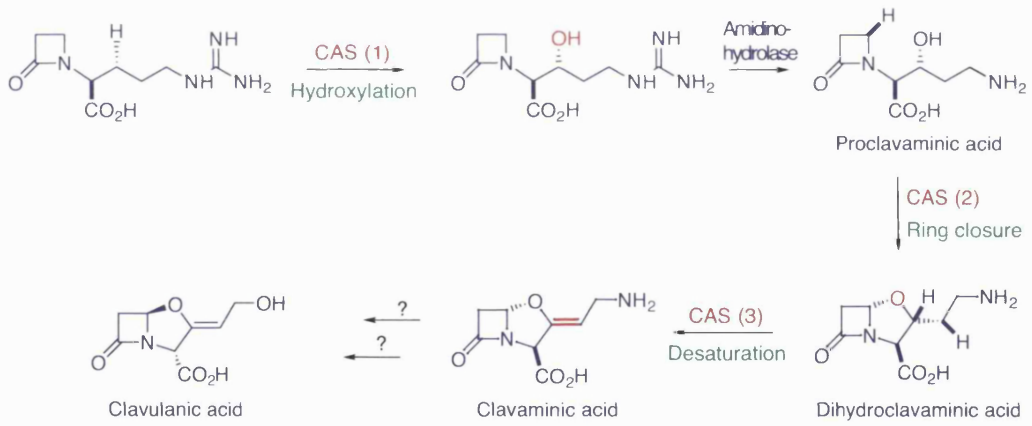


Figure 11. Members of the α-ketoglutarate-Fe(II)-dependent dioxygenases superfamily catalyse a variety of two electron oxidations. (a) Prolyl-4-hydroxylase catalyses the post-translational hydroxylation of proline residues in collagen. (b) DAOCS catalyses the ring expansion of Penicillin N. (c) The trifunctional role of CAS, hydroxylation, oxidative ring closure and desaturation in clavulanic acid synthesis.

deacetoxycephalosporin (DAOC) (Baldwin *et al.*, 1988). Prolyl-4-hydroxylase hydroxylates specific proline and lysine residues in collagen and other animal glycoproteins (Kivirikko *et al.*, 1989). Clavaminate synthase (CAS), a key enzyme in the synthesis of the β-lactamase inhibitor, clavulanic acid, catalyses three separate reactions, the first is hydroxylation, after a step catalysed by an amidinohydrolase, CAS then catalyses an oxidative ring closure followed by a desaturation reaction; Figure 11; (Salowe *et al.*, 1990).

If AlkB were a member of this superfamily it may explain why no *in vitro* activity has been found for the AlkB protein, as α-ketoglutarate and iron have never been added to any of the AlkB *in vitro* assays. However as the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily catalyse a wide range of different reactions, including hydroxylations, desaturations and oxidative ring closures, if the AlkB protein were a member of this superfamily it is not clear what reaction the AlkB protein would catalyse. Although no members of this superfamily are presently known to act on DNA, an *in vitro* assay was devised which could detect removal of methyl groups from DNA lesions as well as any modification of the methyl groups.

4.2 Assay development and optimisation

To test whether AlkB is a member of this superfamily and whether AlkB was able to modify or remove methylated bases from DNA a new *in vitro* assay was developed. DNA substrates were methylated with [^{14}C]-labelled methyl iodide (MeI) or [^{3}H]-labelled dimethyl sulphate (DMS) and incubated with purified His-tagged AlkB protein in the presence of the cofactors typically required by members of the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily. After incubation with the AlkB protein, the DNA substrates were

ethanol precipitated and the supernatant examined for release of ethanol soluble radioactive material. The methylated lesions remaining in the precipitated DNA were analysed by HPLC.

Initially, although the AlkB protein was incubated with the co-factors required by the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily, under the conditions used, AlkB did not remove or modify 1-meA, 3-meC, 3-meA, 7-meA or 7-meG in DNA (data not shown). However, after several variations in the assay conditions, a weak AlkB activity was identified and the conditions were optimised to maximise this activity.

Many of the α-ketoglutarate-Fe(II)-dependent dioxygenases have a narrow pH optimum. The optimal pH for AlkB activity was determined by assaying AlkB activity across the overlapping buffering ranges of Hepes (pH 6.8 to pH 8.2) and Bicine (pH 7.75 to pH 9.0). Phosphate buffers were avoided as they form insoluble salts with iron, resulting in low concentration of free Fe(II). The AlkB activity exhibited a relatively narrow pH optimum of between pH 8.0 and 8.5 depending on the buffer used, this variation is likely to be caused by differences in ionic strength of the buffers (Figure 12). The optimal pH for AlkB activity was taken to be pH 8.0 in 50 mM Hepes.KOH.

The method for purifying the His-tagged AlkB protein (Dinglay *et al.*, 2000) was altered to prevent the self-inactivation to which α-ketoglutarate-Fe(II)-dependent dioxygenases are particularly prone. At the start of the purification the weakly bound iron was removed from the active site of the AlkB protein, thus, temporarily disabling its catalytic activity. This was achieved by lysing the AlkB overexpressing bacteria in the presence of the metal chelator EDTA. The EDTA

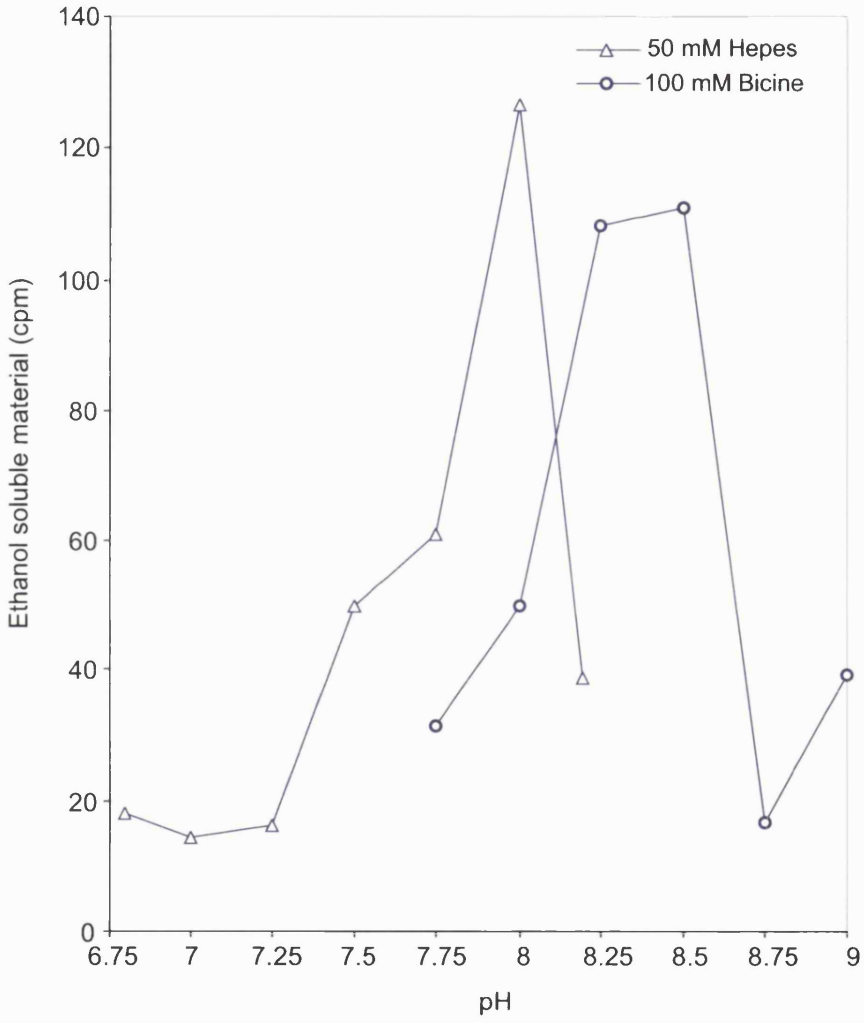


Figure 12. Effect of pH on *E. coli* AlkB activity.
0.5 pmoles of His-tagged AlkB were incubated with 2000 cpm of [^{14}C]-methylated poly(dA) in the presence of 1.4 mM α-ketoglutarate, 1 mM ascorbate, 100 μM $Fe(NH_4)_2(SO_4)_2$ and either 50 mM Hepes or 100 mM Bicine at various pHs at 37 °C for 15 min. The release of ethanol soluble radioactive material was monitored. The data is expressed as the average of three replicates.

was dialysed out of the crude extract before the His-tagged AlkB apoenzyme was purified on a nickel agarose column. AlkB activity could be restored by reconstituting Fe(II) back into the AlkB apoenzyme by the addition of $Fe(NH_4)_2(SO_4)_2$ to the assay mix.

The optimum iron concentration for maximal AlkB activity was determined by the addition of various concentrations of $Fe(NH_4)_2(SO_4)_2$ to the assay. A $Fe(NH_4)_2(SO_4)_2$ concentration of 75 μM was optimal (Figure 13a). It should be noted that the histidine tag attached to the AlkB protein possibly co-ordinates a small amount of the iron, slightly reducing the free iron concentration in the assay.

α-ketoglutarate is a co-substrate of reactions catalysed by the α ketoglutarate-dependent dioxygenases. The optimal α-ketoglutarate concentration for AlkB activity was determined. Several members of the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily suffer from substrate inhibition at high concentrations of α-ketoglutarate. Therefore, 1 mM α-ketoglutarate was chosen for the standard AlkB reaction conditions because at this concentration, near maximal AlkB activity was observed and there was no risk of substrate inhibition (Figure 13b).

The activities of many α-ketoglutarate-Fe(II)-dependent dioxygenases are stimulated by ascorbic acid. Maximal AlkB activity was observed in the presence of 2 mM ascorbic acid. In reactions where the amount of AlkB is limiting (such as in Figure 14a), the AlkB activity seems to be dependent on ascorbic acid, but at 100-fold higher AlkB concentrations where AlkB is in excess the effect of ascorbic acid was reduced (Appendix B). Thus, ascorbic acid was not essential but stimulated AlkB activity.

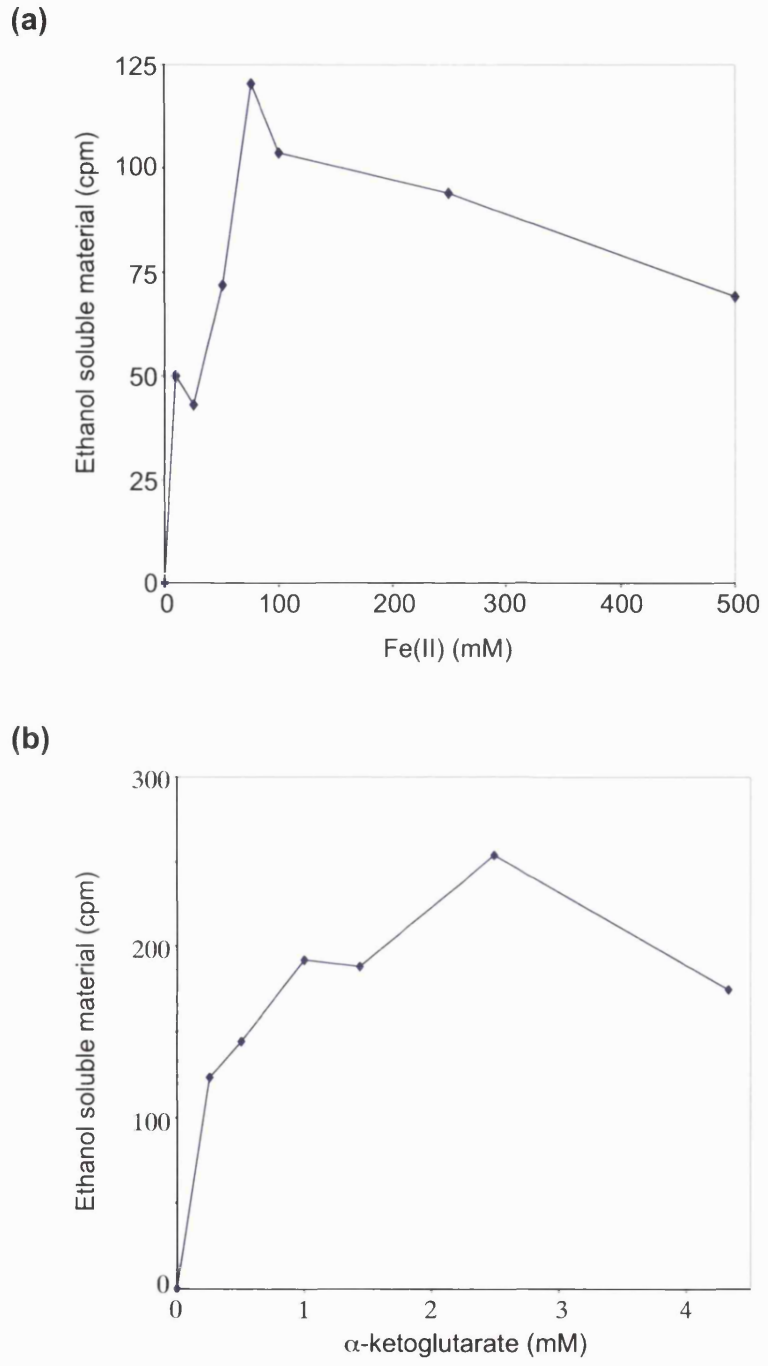


Figure 13. Effect of (a) iron and (b) α-ketoglutarate concentration on *E. coli* AlkB activity. 0.5 pmoles of His-tagged AlkB were incubated with 1200 cpm of [^{14}C]-methylated poly(dA) in the presence of (a) 50 mM Hepes.KOH pH 8.0, 1.5 mM α-ketoglutarate, 2 mM ascorbic acid and various concentrations of $Fe(NH_4)_2(SO_4)_2$ (b) 50 mM Hepes.KOH pH 8.0, 1 mM ascorbic acid, 100 μM $Fe(NH_4)_2(SO_4)_2$ and various concentrations of α-ketoglutarate at 37 °C for 15 min. The release of ethanol soluble radioactive material was monitored. The data is expressed as the average of two replicates.

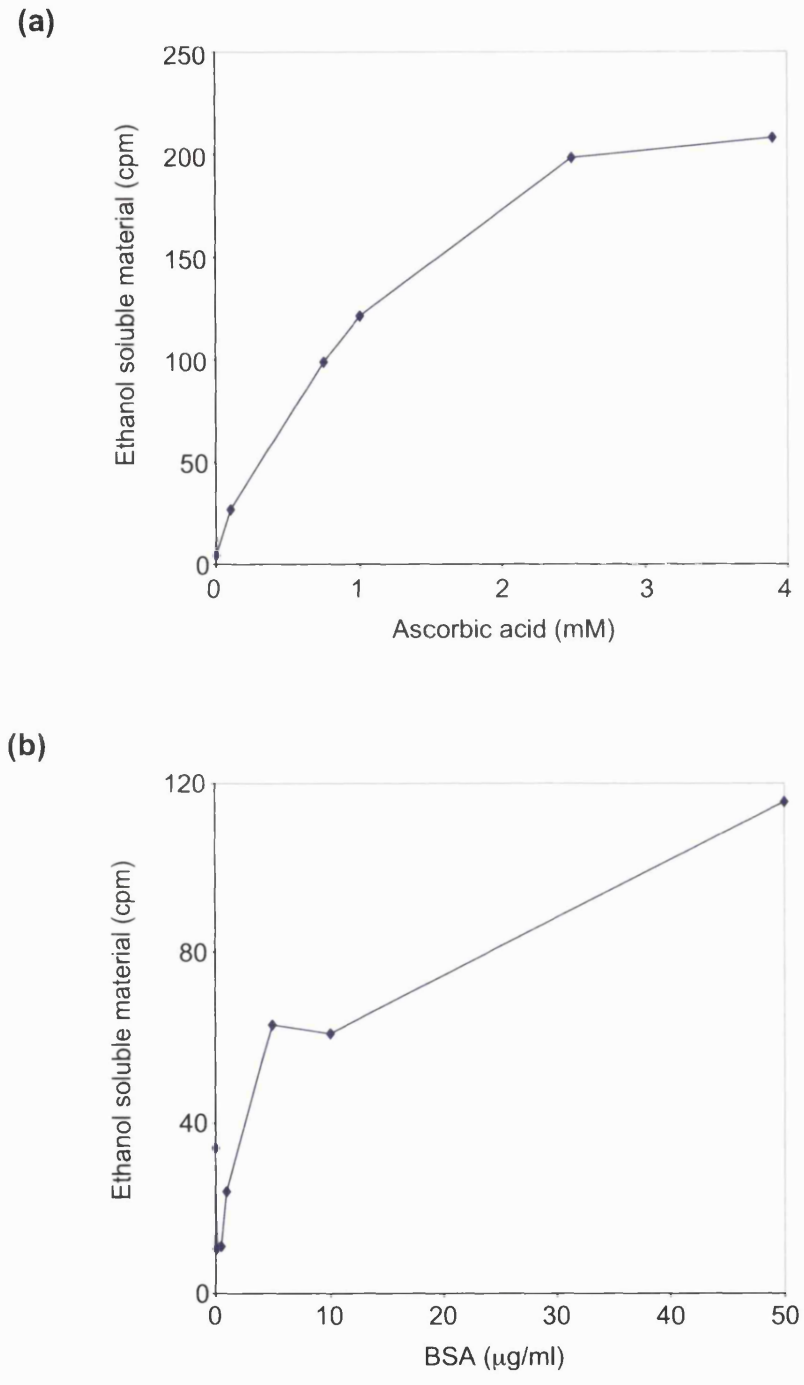


Figure 14. Effect of (a) ascorbic acid concentration (b) BSA concentration on *E. coli* AlkB activity. (a) 0.5 pmoles of His-tagged AlkB were incubated with 1200 cpm of [^{14}C]-methylated poly(dA) in the presence of 50 mM Hepes.KOH pH 8.0, 1.5 mM α-ketoglutarate, 100 μM $Fe(NH_4)_2(SO_4)_2$ and various concentrations of ascorbic acid (b) 0.125 pmoles of His-tagged AlkB were incubated with 1000 cpm of [^{14}C]-methylated poly(dA) in the presence of 50 mM Hepes.KOH pH 8.0, 2 mM α-ketoglutarate, 1 mM ascorbic acid, 75 μM $Fe(NH_4)_2(SO_4)_2$ and various concentrations of BSA at 37 °C for 15 min. The release of ethanol soluble radioactive material was monitored. The data is expressed as the average of two replicates.

Additional stimulation of the AlkB activity was found with the addition of bovine serum albumin (BSA) to the assay. BSA may increase AlkB activity by interacting with the enzymes sulphydryl groups or may increase the stability of the protein by increasing the protein concentration. The addition of 50 μg/ml of BSA to the reaction mix gave good activity (Figure 14b).

The optimal conditions for the AlkB assay were determined to be 50 mM Hepes.KOH pH 8.0, 75 μM $Fe(NH_4)_2(SO_4)_2$, 1 mM α-ketoglutarate, 2 mM ascorbic acid and 50 μg/ml bovine serum albumin.

4.3 AlkB is an α-ketoglutarate-Fe(II)-dependent dioxygenase

In the optimal conditions, the purified AlkB protein released ethanol soluble radioactive material from both [^{14}C]-methylated poly(dA) and poly(dC) substrates (Figure 15) as well as [^{3}H]-methylated substrates (Figure 16b). The amount of radioactive material released increased in relation to the amount of AlkB present.

To ensure that the observed activity was not being catalysed by Fe(II) co-ordinated by the recombinant AlkB protein's histidine-tag, a vector encoding a FLAG-tagged AlkB protein was constructed. The FLAG-tagged AlkB was shown to be active *in vivo*. The vector encoding the FLAG-tagged AlkB protein complemented the defect of *E. coli alkB* mutants to reactivate MMS treated ϕK phage (Figure 16a). The purified FLAG-tagged AlkB protein released ethanol soluble radioactive material from [^{3}H]-methylated poly(dA) (Figure 16b). Therefore, the observed activity is associated with the AlkB protein and not the histidine tag.

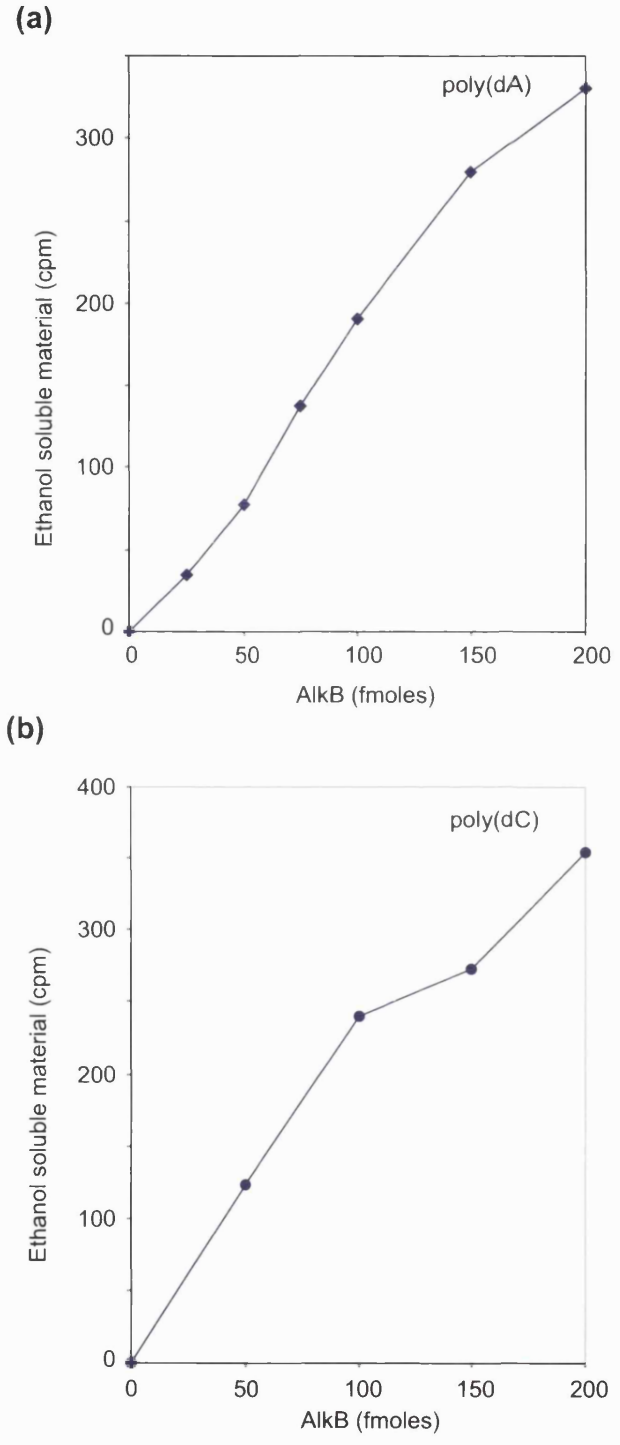


Figure 15. Release of ethanol soluble radioactive material from [^{14}C]-methylated poly(dA) and poly(dC) by *E. coli* AlkB in the presence of Fe(II) and α-ketoglutarate. (a) 1200 cpm of [^{14}C]-Methylated poly(dA) or (b) 800 cpm of [^{14}C]-Methylated poly(dC) was incubated with His-tagged AlkB in the optimised reaction mixture at 37 °C for 15 min and the release of ethanol soluble radioactive material was monitored. The data is expressed as the average of two replicates.

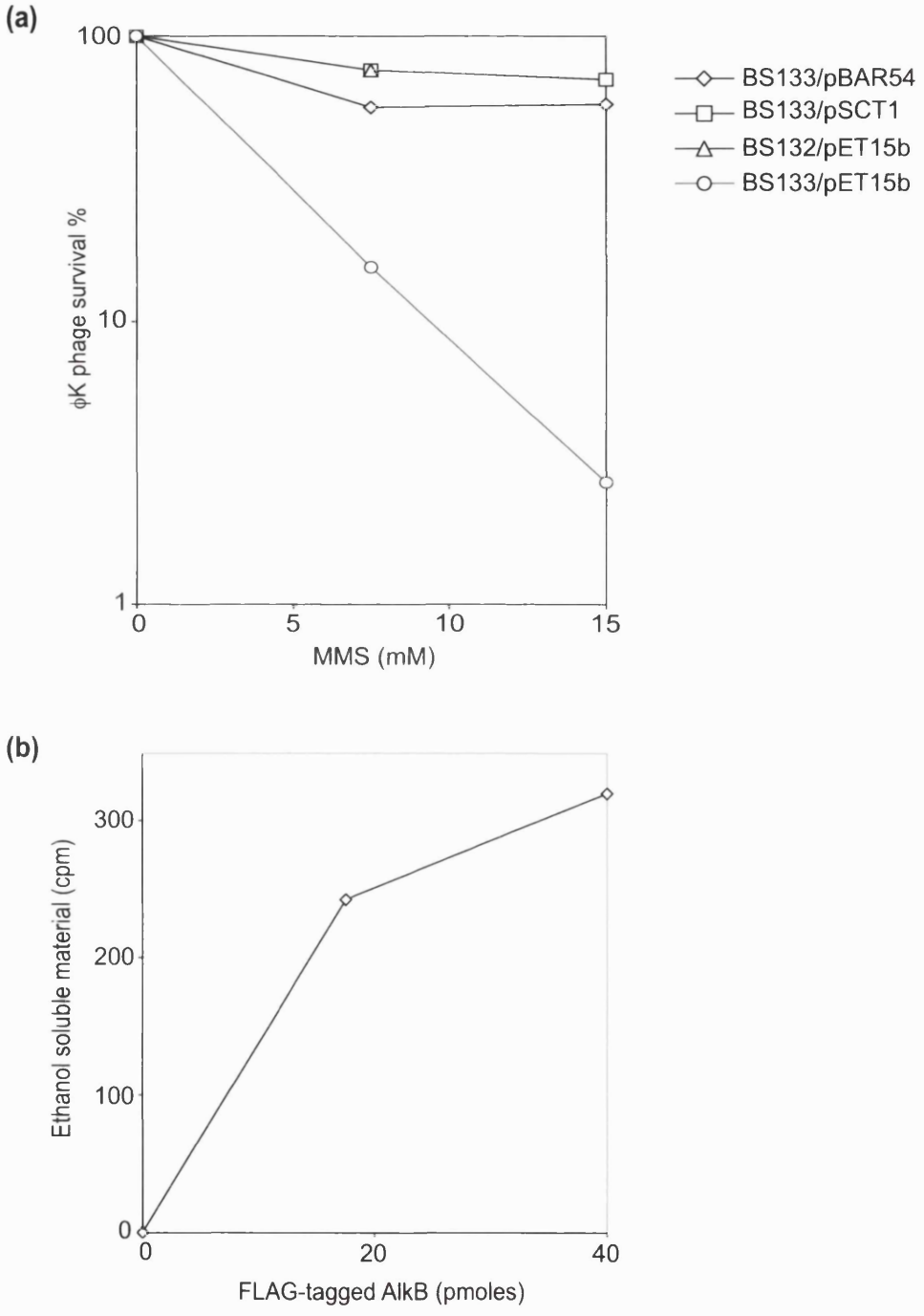


Figure 16. FLAG-tagged AlkB complements the *E. coli alkB* mutant phenotype and is active *in vitro*. (a) Reactivation of MMS treated φK single stranded DNA phage was monitored in various *E. coli* strains. Host strains were, BS132 wild type and BS133 *alkB* mutant. Plasmid pBAR54 encodes His-tagged AlkB protein and pSCT1 encodes FLAG-tagged AlkB protein. (b) Purified FLAG-tagged AlkB was incubated with 1200 cpm of [^{3}H]-methylated poly(dA) in 50 mM Hepes.KOH pH 8.0, 50 mM KCl, 5% glycerol, 2 mM β-mercaptoethanol, 1.43 mM α-ketoglutarate, 100 μM ascorbic acid and 100 μM $Fe(NH_4)_2(SO_4)_2$ at 37 °C for 15 min.

The activity of the AlkB protein on both the methylated poly(dA) and poly(dC) substrates is dependent on the presence of α-ketoglutarate and is stimulated by ascorbic acid. The AlkB activity is also dependent on the presence of Fe(II) as the reaction was inhibited by EDTA and this inhibition could be overcome by adding an excess of $Fe(NH_4)_2(SO_4)_2$ (Table 4a and b). The AlkB activity could not be restored by adding an excess of several other divalent cations including Zn(II), Mn(II), Co(II), Mg(II) and Ni(II) (Table 4c). A slight and variable activity was observed in the presence of Cu(II).

The requirement of AlkB for α-ketoglutarate, Fe(II) and stimulation by ascorbic acid is highly characteristic of the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily. The use of such distinctive cofactors shows that the observed activity is not catalysed by a contaminating exonuclease and demonstrates that AlkB is an α-ketoglutarate-Fe(II)-dependent dioxygenase.

4.4 AlkB is active on double stranded DNA

It had been suggested that AlkB repairs lesions that are only generated in single stranded DNA or that AlkB only acts on DNA when it is single stranded. To test the ability of AlkB to act on double stranded DNA, oligonucleotide 116 (Table 3), which contains equal quantities of all four DNA bases was methylated with [^{14}C]-MeI. The [^{14}C]-methylated oligonucleotide was annealed to the complementary DNA oligonucleotide 117 (Table 3) and was confirmed to be > 90% double stranded by digestion with the double strand specific exonuclease III (Appendix C). AlkB released ethanol soluble radioactive material from the double stranded substrate (Figure 17). Therefore, AlkB is able to act on double stranded and single stranded DNA.

(a)

[^{14}C]-methylated poly(dA)	% Activity
complete	100
no AlkB	<0.5
no α-ketoglutarate	<0.5
no ascorbic acid	<0.5
+ 0.2 mM EDTA	1.1
+ 0.2 mM EDTA + 0.7 mM $Fe(NH_4)_2(SO_4)_2$	99

(b)

[^{14}C]-methylated poly(dC)	% Activity
complete	100
no AlkB	<0.5
no α-ketoglutarate	0.9
no ascorbic acid	<0.5
+ 0.2 mM EDTA	<0.5
+ 0.2 mM EDTA + 0.7 mM $Fe(NH_4)_2(SO_4)_2$	83

(c)

	% Activity
complete	100
+ 0.2 mM EDTA	0
+ 0.2 mM EDTA + 0.7 mM $Fe(NH_4)_2(SO_4)_2$	100
+ 0.2 mM EDTA + 0.7 mM $CoSO_4$	<0.5
+ 0.2 mM EDTA + 0.7 mM $CuSO_4$	4
+ 0.2 mM EDTA + 0.7 mM $MgSO_4$	<0.5
+ 0.2 mM EDTA + 0.7 mM $MnSO_4$	<0.5
+ 0.2 mM EDTA + 0.7 mM $NiSO_4$	<0.5
+ 0.2 mM EDTA + 0.7 mM $ZnSO_4$	<0.5

Table 4. Requirements of AlkB activity. 2 pmoles of His-tagged AlkB were incubated with (a) and (c) 1200 cpm of [^{14}C]-methylated poly(dA) or (b) 800 cpm of poly(dC), in the complete reaction mix or in the specified conditions at 37 °C for 15 min. The release of ethanol soluble radioactive material was monitored.

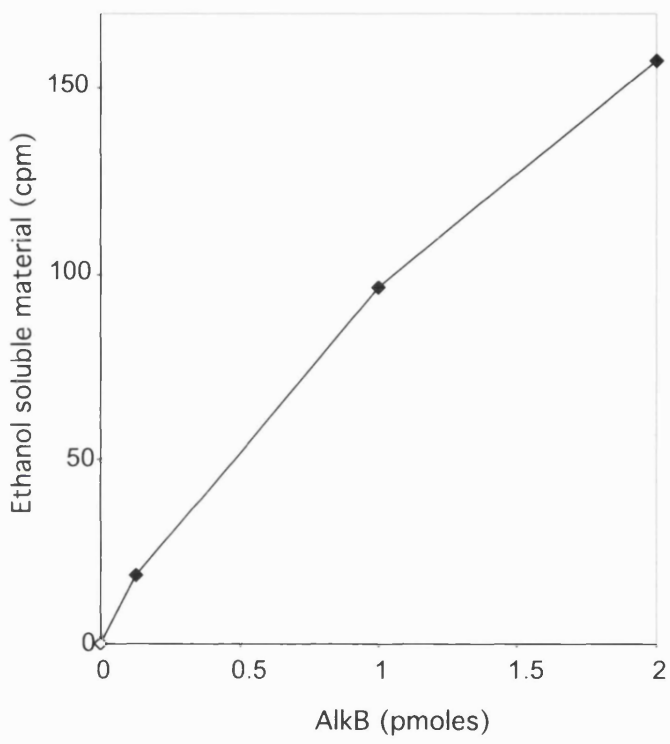


Figure 17. *E. coli* AlkB acts on double stranded DNA. [^{14}C]-methylated oligonucleotide 116 was annealed to oligonucleotide 117. 3000 cpm of the annealed DNA was incubated with His-tagged AlkB in the optimised conditions at 37 °C for 15 min. The release of ethanol soluble material was monitored. The data is expressed as the average of two replicates.

4.5 AlkB acts on 1-meA and 3-meC

AlkB releases ethanol soluble radioactive material from both the poly(dA) and poly(dC) [^{14}C]-methylated substrates. To determine which methylated bases the AlkB protein acts upon, after incubation with the AlkB protein the [^{14}C]-methylated bases remaining in the DNA were released by acid hydrolysis, separated by HPLC chromatography and quantified by scintillation counting.

Due to a weak N-glycosidic bond, adenine bases are released from DNA using a mild acid hydrolysis (0.1 M HCl, 95 °C, 1 hour). A standard HPLC protocol was used to separate the methylated adenines (Beranek *et al.*, 1980). However, to analyse the methylated cytosine bases present in the poly(dC) substrate, new methods were developed. The N-glycosidic bond of a cytosine residue is considerably less acid labile than that of an adenine residue, therefore, treatment with strong acid is required to release cytosine bases from the poly(dC) substrate. Perchloric acid hydrolysis of the DNA was considered inappropriate as it causes deamination of 3-meC to 3-methyluracil (3-meU). Therefore, hydrolysis of the poly(dC) substrate was carried out using 90 % formic acid at 180 °C for 20 minutes as it was shown that, providing that the length of the hydrolysis was short, formic acid caused very little deamination of 3-meC (Appendix D). An HPLC separation was developed which could differentiate 3-meC from other methylated cytosine bases such as N^4-methylcytosine (N^4-meC) and 5-methylcytosine (5-meC). This separation required a gradient of methanol from 5 to 40 %, applied to a Whatman Partisil 10 cation exchange column equilibrated in 0.1 M ammonium formate pH 3.6.

After incubation in the optimised conditions, with or without AlkB, the methylated lesions remaining in the [^{14}C]-methylated DNA substrates were analysed. The poly(dA) substrate contained three types of methylated lesions, 1-

meA was the most abundant lesion making up 64 % of the methylated adenines, of the remainder 31 % were 3-meA and 5 % were 7-meA. In the presence of Fe(II) and α-ketoglutarate, in the optimised conditions, during a 15 minute incubation 2.5 pmoles of AlkB was able to reduce the amount of 1-meA present in the substrate DNA by more than 95%. The AlkB activity did not affect the levels of 3-meA and 7-meA present in the substrate (Figure 18).

Analysis of the [^{14}C]-methylated poly(dC) substrate showed that 3-meC was the only detectable methylated base present. In the presence of Fe(II) and α-ketoglutarate in the optimised conditions, 2.5 pmoles of AlkB reduced the amount of 3-meC in these substrates by over 85 % in 15 minutes (Figure 19).

The poly(dA) and poly(dC) AlkB substrates may not be good models for DNA *in vivo*, as they are homopolymers and only 200-300 nucleotides in length. To ensure that AlkB is active on "real" DNA, [^{3}H]-DMS treated single stranded M13 DNA was used as a substrate. The ability of AlkB to repair the methylated purines in the 7 kb bacteriophage DNA was analysed. In this "real" M13 DNA substrate the AlkB protein was able to reduce the level of 1-meA present. However, AlkB did not act on the innocuous but extremely prevalent lesion 7-meG, nor was the AlkB protein active on 3-meA or 7-meA (Figure 20). After incubation of [^{3}H]-methylated substrates with AlkB, a small new HPLC peak is often observed (Figure 20). This peak is not seen when [^{14}C]-methylated DNA substrates are used in the AlkB reaction (Figure 18). The identity of the peak is unclear, however, the peak appears to be a reaction product that contains hydrogen but not carbon atoms from the methyl groups on the damaged DNA substrate.

Therefore, two substrates for the AlkB protein have been identified, 1-meA and 3-meC in DNA. It was also demonstrated that the AlkB protein does not effect

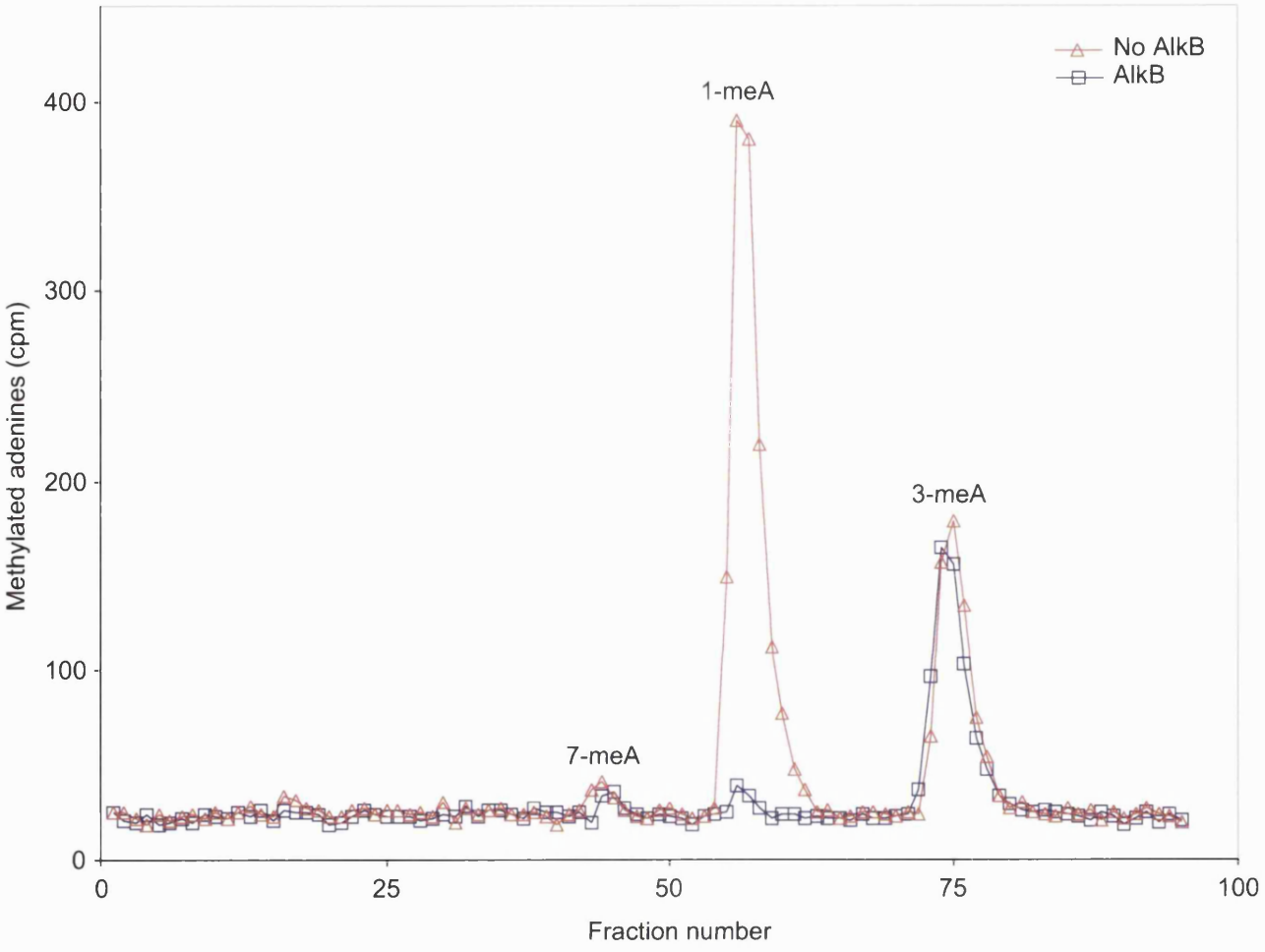


Figure 18. *E. coli* AlkB acts on 1-meA. 3000 cpm of [^{14}C]-MeI treated poly(dA) was incubated (Δ) without AlkB or (❐) with 2.5 pmoles His-tagged AlkB in the optimised conditions at 37 °C for 30 min. The [^{14}C]-labelled methylated bases remaining in the DNA substrate were analysed by HPLC and scintillation counting.

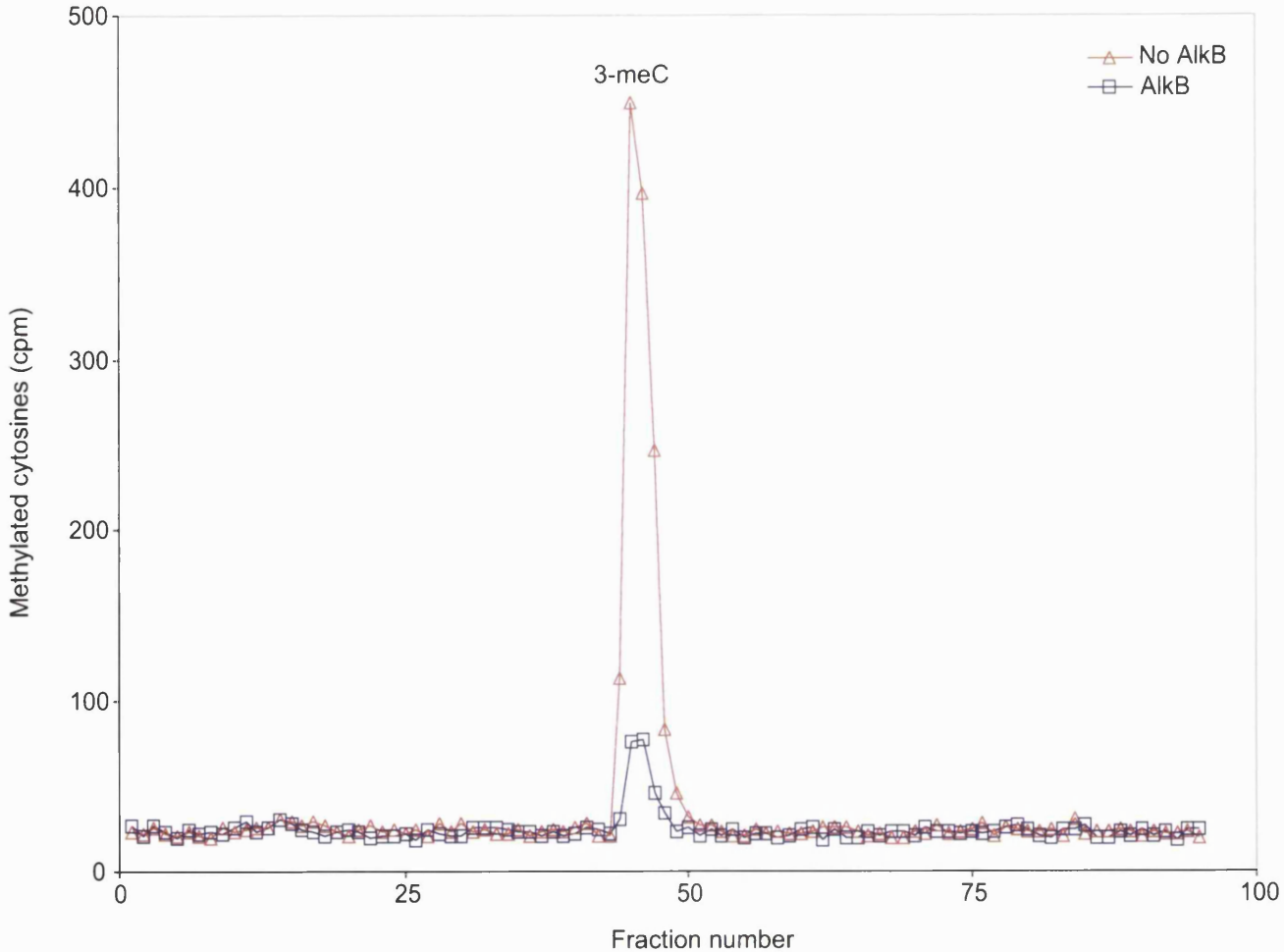


Figure 19. ***E. coli*** **AlkB acts on 3-meC.** 2000 cpm of [^{14}C]-MeI treated poly(dC) was incubated (Δ) without AlkB or (❐) with 2.5 pmoles His-tagged AlkB in the optimised conditions at 37 °C for 30 min. The [^{14}C]-labelled methylated bases remaining in the DNA substrate were analysed by HPLC and scintillation counting.

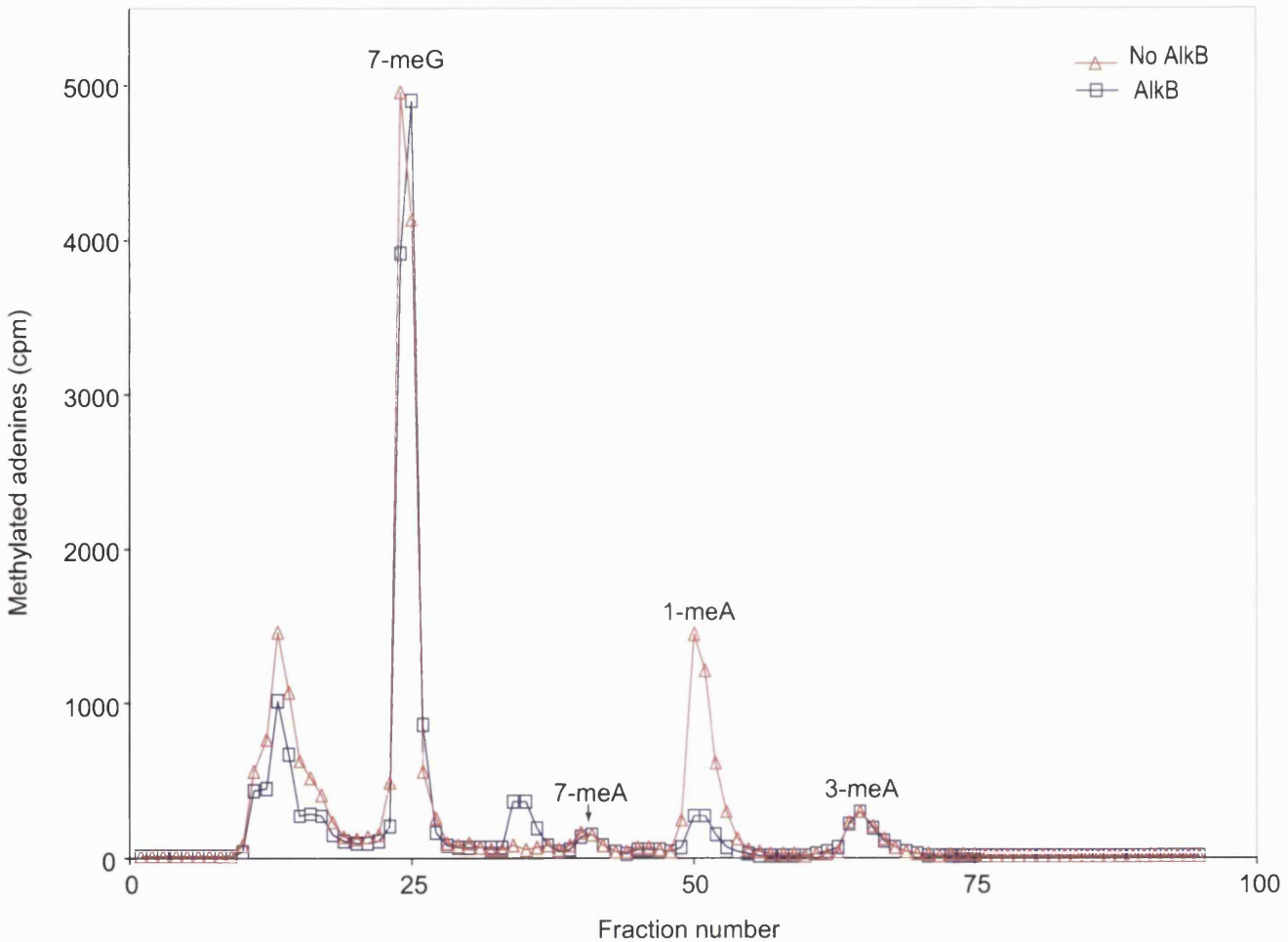


Figure 20. ***E. coli*** **AlkB acts on methylated single stranded M13 phage DNA.** 20,000 cpm of [^{3}H]-DMS treated single stranded M13 DNA was incubated (Δ) without AlkB or (❐) with 460 pmoles His-tagged AlkB in the optimised conditions at 37 °C for 30 min. The [^{3}H]-labelled methylated bases remaining in the DNA substrate were analysed by HPLC and scintillation counting.

all the other major lesions generated in DNA by SN2 methylating agents, 3-meA, 7-meA and 7-meG. The specificity of the AlkB protein for the lesions 1-meA and 3-meC rules out any possibility that AlkB is an exonuclease.

4.6 AlkB directly reverts 1-meA to adenine in DNA

AlkB releases radioactive material from DNA substrates containing both [^{14}C]- and [^{3}H]-methylated 1-meA and 3-meC, therefore, AlkB catalyses the loss of all of the atoms in the methyl group from the DNA substrates. There are several possible AlkB reaction mechanisms that could account for these results. For example, AlkB could excise the methylated base from the DNA or AlkB could remove the methyl group from the base. To further elucidate the reaction mechanism of the AlkB protein, the fate of the DNA base residues was followed during the AlkB reaction.

The base products of the AlkB reaction can be observed and quantified by UV absorption. If the AlkB protein were to excise the methylated base, when monitored by UV absorption, the 1-meA HPLC peak will simply disappear. However, if AlkB were to revert 1-meA to adenine, the 1-meA peak will disappear and an increase in the size of the adenine peak should be observed. Standard DNA methylation reactions methylate only a small proportion of the adenine residues, leaving the majority of adenines residues unmodified, therefore, any AlkB catalysed increase in the adenine HPLC peak would be very small. To ensure that it is possible to observe a large increase in the adenine peak, a method was developed to produce a DNA oligonucleotide substrate in which the majority of the adenine base residues were methylated.

A DNA oligonucleotide was designed that contained a small number of adenine residues interspersed between abundant thymine residues (DNA

oligonucleotide 124). The oligonucleotide was heavily methylated by DMS so as to methylate the majority of the adenine residues. The thymine residues are very inefficiently methylated and therefore act as a scaffold during the heavy methylation. To methylate as many of the adenine residues as possible, a high DMS concentration and a long exposure time were required. To prevent an increase in pH during the methylation reaction, which can cause the Dimroth rearrangement of 1-meA to N^6-methyladenine (N^6-meA) (Engel, 1975) the treatment buffer was refreshed after each 1 or 2 hour DMS treatment. During the treatment samples of the oligonucleotide were taken to monitor the extent to which the adenine residues had become methylated.

The adenine bases were released from the DNA substrate by weak acid hydrolysis, separated using the standard HPLC separation and quantified by UV absorption. The absorbance was measured at 260 nm as it is close to the absorption maximum of both 1-meA and adenine. To account for any intrinsic differences in the ability of 1-meA and adenine to absorb at this wavelength, the relative absorbance of these bases was calculated by measuring the A_{260} of known amounts of the purines under the HPLC separation conditions. The A_{260} ratio of adenine to 1-meA was found to be 1.04, this was taken into account during the quantification.

The amount of 1-meA present in the DNA substrate increased during the DMS treatment: after 2 hours approximately 24 % of adenine residues in the DNA substrate were methylated to form 1-meA, after 8 hours 76 % of the adenines in the DNA were 1-meA. The amount of 3-meA and 7-meA did not increase dramatically during the incubation with DMS, possibly due to instability of the glycosyl bond and loss of these modifications during the extensive DMS treatments (Appendix E).

The heavily methylated substrate was incubated with or without AlkB in the optimised assay conditions, the DNA was hydrolysed and individual bases quantified by HPLC and UV absorption. Approximately 76 % of the adenine residues in the DNA substrate were methylated to form 1-meA, 4 % of the adenines were recovered as 3-meA, 2 % were 7-meA and only 19 % of the adenines were unmodified. After incubation with AlkB, a decrease in the amount of 1-meA correlated with a stoichiometric increase in the amount of adenine recovered. AlkB increased the percentage of unmodified adenine present from 19 % to 76 % (Figure 21). Therefore, AlkB acts by a direct reversal mechanism and reverts 1-meA to adenine in DNA.

4.7 AlkB directly reverts 3-meC to cytosine in DNA

Attempts to use a similar DNA oligonucleotide substrate and UV absorption to examine whether AlkB reverts 3-meC to cytosine in DNA were complicated by the presence of thymine bases in the HPLC separation, which along with the cytosine bases were released from the DNA substrate by strong acid hydrolysis. To avoid this problem, a DNA substrate was used which contained radiolabelled cytosine residues, therefore, only cytosine and cytosine derived bases were detected by scintillation counting of the fractions obtained from the HPLC separation. A single stranded oligonucleotide substrate containing [^{3}H]-cytosine residues interspersed between abundant thymine residues was synthesised by PCR using DNA oligonucleotide 133 as a template and DNA oligonucleotide 134 as a single primer (Table 3), in the presence of 150 μM dNTPs and 2.5 μM [^{3}H]-dCTP.

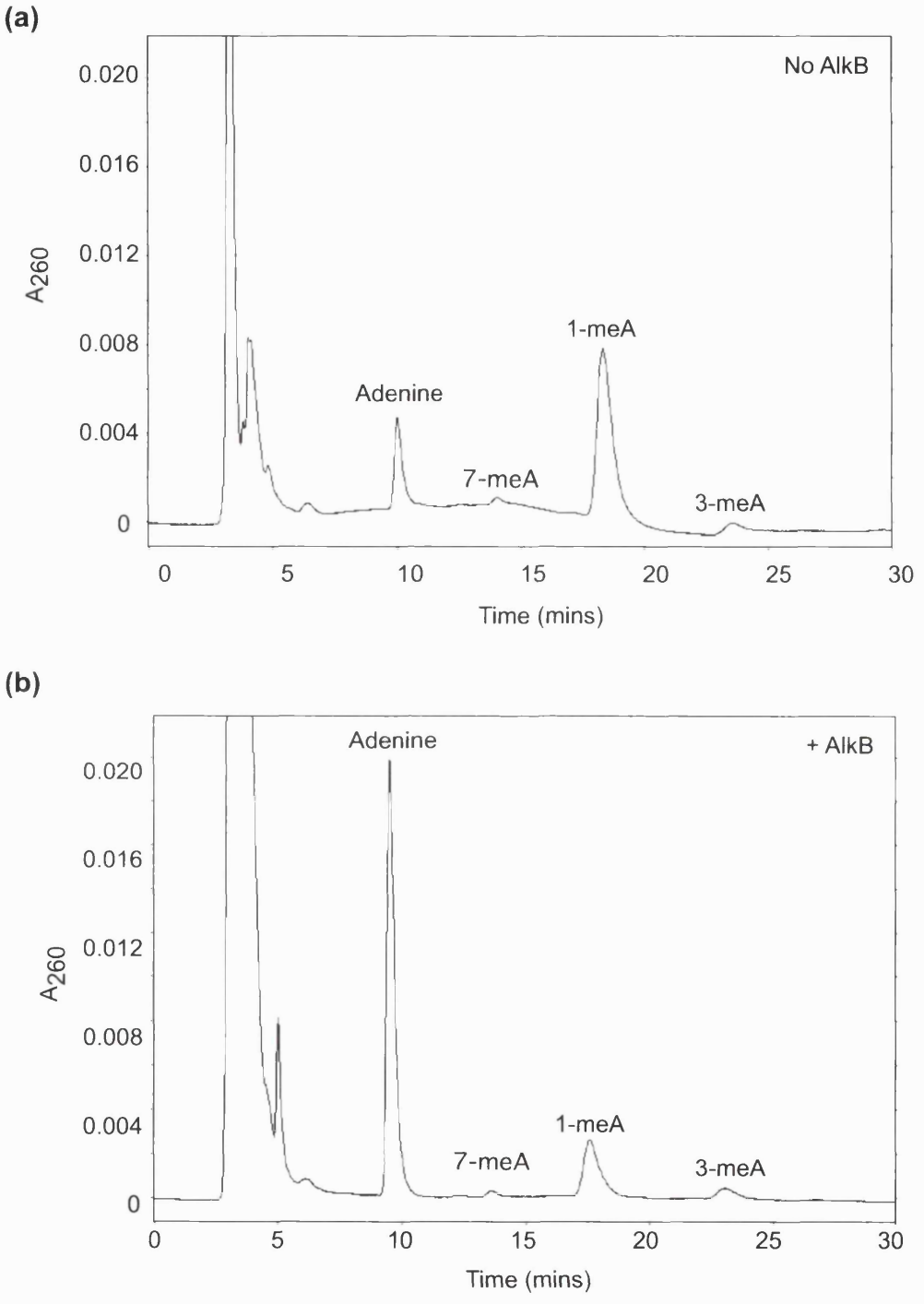


Figure 21. *E. coli* AlkB directly reverts 1-meA to adenine residues in DNA. The thymine-rich DNA oligodeoxynucleotide 124, containing five adenine residues was heavily methylated and then incubated (a) without or (b) with 0.92 nmoles of AlkB at 37 °C for 30 min. The bases present in the oligonucleotide were analysed by HPLC and A_{260} measurements. The early eluting peaks were oligo(dT) fragments.

Heavy methylation of this radioactively labelled oligonucleotide substrate was achieved by treating the DNA eight times for two hours with unlabelled DMS, the buffer was refreshed between each treatment.

A new HPLC separation was developed to obtain baseline resolution between 3-meC, cytosine and uracil bases; a Whatman partisil 10 cation exchange column was run isocratically in 100 mM ammonium formate pH 3.6, 2.5 % methanol.

AlkB was incubated with the heavily methylated [^{3}H]-cytosine containing oligonucleotide substrate. The bases were released from the DNA by formic acid hydrolysis followed by HPLC chromatography and scintillation counting. The DNA substrate contained approximately 70 % of the cytosines as 3-meC, and 30 % as unmodified cytosine. After incubation with AlkB, the amount of 3-meC decreased and a stoichiometric increase in cytosine was observed. AlkB increased the percentage of unmodified cytosine present from 30 % to 84 % (Figure 22). Therefore, AlkB also directly reverts 3-meC to cytosine in DNA.

4.8 1-meA is a DNA replication stalling lesion

There is no direct evidence that 1-meA is a replication stalling lesion. Previous experiments used methylated poly(dA) as a template for DNA replication and therefore were unable to distinguish replication stalling caused by 1-meA from that caused by 3-meA (Larson *et al.*, 1985). Using the DNA replication bypass assay (Chapter 3) and the newly optimised AlkB reaction to remove 1-meA from the template, it should be possible to determine whether 1-meA stalls DNA replication catalysed by the *E. coli* DNA polymerase I derivative, Klenow exo$^-$.

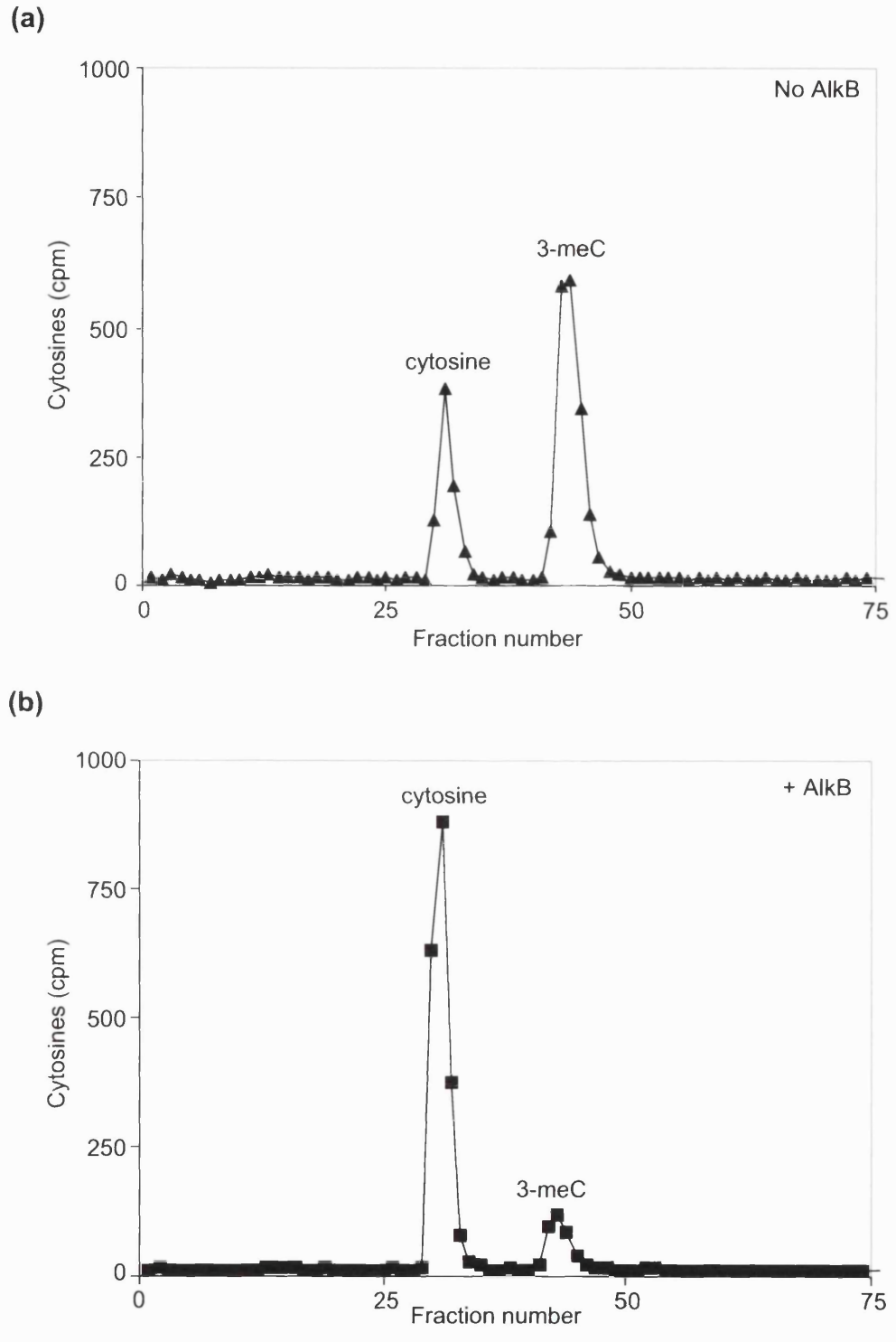


Figure 22. *E. coli* AlkB directly reverts 3-meC to cytosine residues in DNA. A thymine-rich oligodeoxynucleotide containing several [^{3}H]-cytosine residues was heavily methylated with DMS. 3300 cpm of this DNA substrate was incubated (a) without or (b) with 0.23 nmoles of AlkB in the optimised conditions at 37 °C for 30 min. The [^{3}H]-cytosine residues remaining in the oligonucleotide were analysed by HPLC and scintillation counting.

A DNA oligonucleotide containing a single adenine residue (StA, Table 3) was methylated to produce a mix of 1-meA and 3-meA lesions. This methylated oligonucleotide was incubated with or without the AlkB protein in the optimised conditions. The AlkB protein catalyses the direct reversal of 1-meA residues in the DNA to adenine residues, and the 3-meA residues in the DNA will be unaffected. A primer was then annealed to the template DNA and extended with the Klenow exo$^-$ DNA polymerase.

After pre-incubation of the template DNA with AlkB in the optimised conditions, which removed 1-meA from the DNA template, a reduction of replication stalling at the methylated adenine was observed (Figure 23). This indicates that 1-meA is a replication stalling lesion. The 3-meA that remains in the template DNA after the AlkB incubation still causes stalling of DNA replication, however, due to the limitations of this technique it is not possible to quantify the amount of replication stalling caused by 1-meA compared to 3-meA.

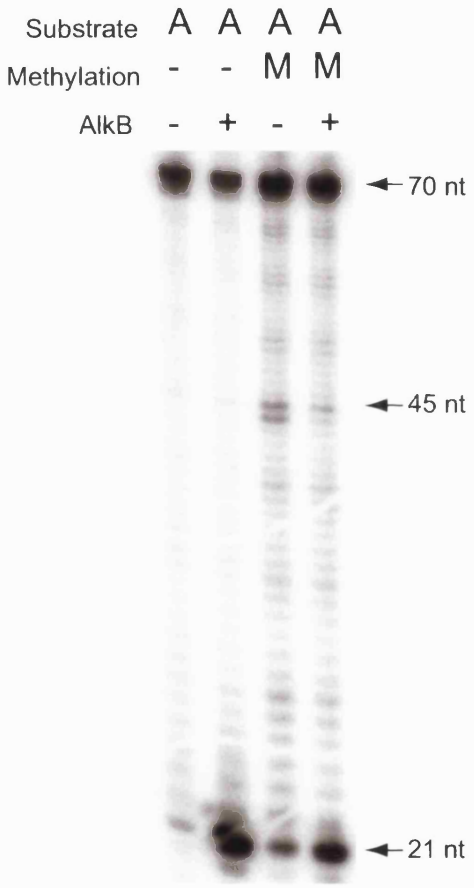


Figure 23. 1-meA is a DNA replication stalling lesion. A methylated (M) or unmethylated (-) 70mer DNA template that contains an A residue 45 nucleotides from its 5' end, was pre-incubated with or without AlkB in the optimal conditions. Following centrifugation through a G25 Sephadex mini-column, a [^{32}P]-labelled 21mer oligonucleotide was annealed and extended by Klenow exo$^-$. The reactions were stopped by the addition of formamide and run on a 12 % sequencing gel.

CHAPTER 5

RESULTS

HUMAN ALKB HOMOLOGS

CHAPTER 5. HUMAN ALKB HOMOLOGS

Certain alkylating agents are commonly used as cancer chemotherapy drugs. Identification of human homologs of AlkB protein could lead to a better understanding of the cytotoxicity of chemotherapy drugs and possibly to better treatments for cancer. Therefore, it was of interest to investigate human homologs of the *E. coli* AlkB protein.

5.1 Human AlkB sequence homologs

In 1996, a putative human homolog of the *E. coli* AlkB protein, hABH, was reported to partially complement the MMS sensitivity of an *E. coli alkB* mutant (Wei *et al.*, 1996). However, B. Sedgwick in our laboratory demonstrated that the sequences of cDNA clones of the hABH gene obtained from the IMAGE consortium differed from the reported sequence, but agreed with the predicted conceptual splice of the genomic DNA on chromosome 14q24 (accession AC008044 / XP_007409). Database searches by B. Sedgwick and P. Bates of the Biomolecular Modelling Laboratory also revealed two different potential human AlkB homologs, ABH2 (accession XP_058581) and ABH3 (accession NP_631917).

The ABH1, ABH2 and ABH3 proteins were of unknown function, but exhibited sequence similarity to the α-ketoglutarate-dependent-Fe(II)-dependent dioxygenase superfamily in their carboxyl-terminal regions (Figure 24). All three of these proteins contain the Fe(II) binding motif, $HXDX_NH$, that is characteristic of this superfamily and a conserved arginine that may be involved in α-ketoglutarate interactions.

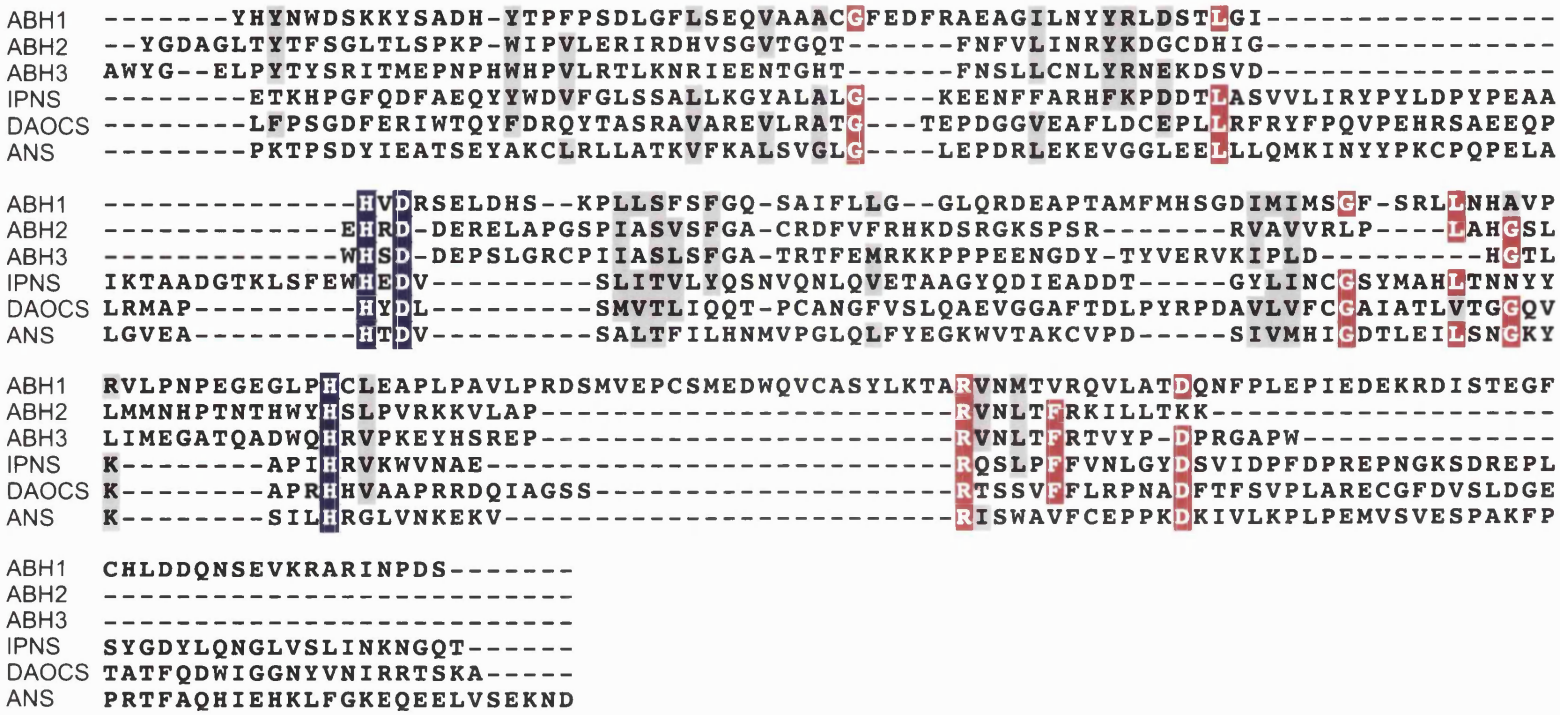


Figure 24. ABH2 and ABH3 have sequence homology to the α-ketoglutarate-Fe(II) dioxygenase superfamily. Multiple alignment the human AlkB homologs with IPNS (isopenicillin N synthase, *Aspergillus nidulans*), DAOCS (deacetoxycephalosporin C synthase, *Streptomyces clavuligerus*), and ANS (anthocyanidin synthase, *Arabidopsis thaliana*). N = $HXDX_NH$ iron binding motif
N = Other Identical residues N = Similar residues

5.2 ABH1, ABH2 and ABH3 *in vitro* assays

The function of the *E. coli* AlkB protein has been elucidated and AlkB has been shown to directly revert DNA methylation damage. The *in vitro* assay developed for the *E. coli* AlkB protein can be exploited to determine whether these AlkB sequence homologs, ABH1, ABH2 and ABH3 are α-ketoglutarate-dependent-Fe(II)-dependent dioxygenases. The assay can also be used to test whether these proteins, in the presence of the distinctive co-factors required by this superfamily, repair one or both of the methylated lesions, 1-meA and 3-meC in DNA.

The ABH1 gene was sub-cloned to produce a vector encoding a FLAG-tagged ABH1 protein. Purified FLAG-tagged ABH1 protein was incubated with [^{14}C]-methylated poly(dA) and poly(dC) substrates. The DNA substrates were ethanol precipitated and scintillation counting monitored the release of ethanol soluble radioactive material from the DNA. Purified FLAG-tagged ABH1 did not release radioactive material from a [^{14}C]-methyl iodide treated poly(dA) or poly(dC) substrate in the assay conditions optimised for the *E. coli* AlkB protein or after modification of the assay.

His- and GST-tagged ABH2 and ABH3 proteins were incubated with the [^{14}C]-methylated poly(dA) and poly(dC) substrates in the assay conditions optimised for the *E. coli* AlkB protein. Under these conditions the human ABH2 and ABH3 proteins exhibited a weak activity, resulting in the release of a small amount of radioactive material from the [^{14}C]-MeI treated poly(dC) and poly(dA) substrates. Therefore, the activities of the ABH2 and ABH3 proteins were investigated further.

5.3 Optimisation of ABH2 and ABH3 *in vitro* assays

In an attempt to increase the activity of the ABH2 and ABH3 proteins, the *in vitro* assay conditions were optimised for each of the human proteins. Members of the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily of enzymes often exhibit a narrow pH optima, therefore, the activity of the human ABH2 and ABH3 proteins were measured across a pH range. To determine the pH optimum of the ABH3 protein, the ability of the protein to release ethanol soluble radioactive material from the [^{14}C]-methylated poly(dA) substrate was measured in Hepes buffer from pH 6.8 to 8.0 and in Bicine buffer from pH 8.5 to 9.0. The pH optimum of the ABH3 protein appeared to be approximately pH 8.0 in 50 mM Hepes.KOH (Figure 25a). ABH3 appeared to exhibit a bell shaped pH curve, consistent with the involvement of two ionisable groups, one that must be protonated and one that must be unprotonated for catalysis or secondary structure formation. The human ABH3 protein and the *E. coli* AlkB protein both have bell shaped pH optima, peaking at approximately pH 8.0, this suggests that the ionisable amino acid groups that are important for catalysis by the AlkB protein may also be important for the ABH3 catalysed reaction.

The pH optimum of the human ABH2 protein was measured in Hepes buffer between pH 7.0 and 8.0 and at pH 8.5 in Bicine buffer. The maximum activity observed for ABH2 was at pH 7.0, although the actual optimum may be slightly lower than pH 7.0 (Figure 25b). The pH optimum for ABH2 appears to be different from that of the ABH3 protein and the *E. coli* AlkB protein, suggesting that different ionisable amino acid side chains may be important for catalysis by the ABH2 than by the ABH3 and AlkB proteins.

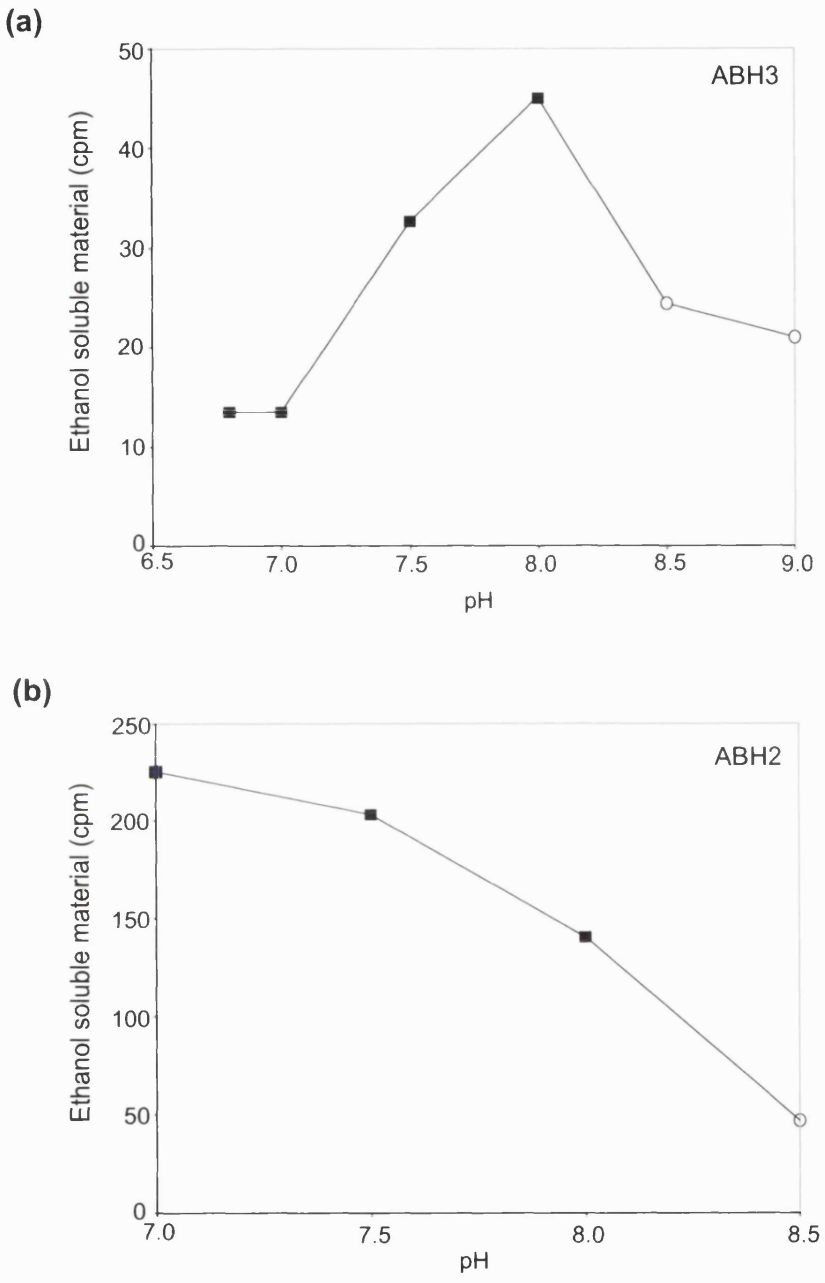


Figure 25. Effect of pH on ABH2 and ABH3 activities. 150 pmoles of (a) GST-tagged ABH3 was incubated with 800 cpm of [^{14}C]-methylated poly(dC) and (b) GST-tagged ABH2 was incubated with 1400 cpm of [^{14}C]-methylated poly(dA) in the presence of 2 mM ascorbic acid, 75 μM $Fe(NH_4)_2(SO_4)_2$ 1 mM α-ketoglutarate and 50 μg/ml BSA, with either 50 mM Hepes buffer (❑) or 100 mM Bicine buffer (❍) at various pHs at 37 °C for 30 min. The release of ethanol soluble radioactive material was monitored. The data is expressed as the average of two replicates.

α-ketoglutarate is a co-substrate for reactions catalysed by members of the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily. Therefore, the effect of α-ketoglutarate concentration on ABH2 and ABH3 activity was monitored. The ABH2 and ABH3 proteins required lower concentrations of α-ketoglutarate than the AlkB protein for which optimal activity was observed using 1 mM α-ketoglutarate (Figure 26). Addition of α-ketoglutarate above 0.25 mM did not significantly stimulate or inhibit the activity of either the ABH2 or the ABH3 protein. As only low ABH2 and ABH3 activities were observed, higher α-ketoglutarate concentrations may be required if more active proteins were obtained. Therefore, an α-ketoglutarate concentration of 1 mM was used for both the ABH2 and ABH3 *in vitro* assays.

Members of the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily use iron-oxo intermediates to catalyse two electron oxidation reactions. Therefore, as Fe(II) is required for catalysis, the $Fe(NH_4)_2(SO_4)_2$ concentration needed for optimal ABH2 and ABH3 activity was determined. The activity of both ABH2 and ABH3 peaked at 25 μM $Fe(NH_4)_2(SO_4)_2$ (Figure 27). The activities of the human proteins, especially ABH2, are low, therefore, higher concentrations of Fe(II) may be required in the future if more active protein were to be obtained. As only a slight inhibition of ABH3 activity is observed at iron concentrations over 25 μM, an iron concentration of 75 μM was used in ABH3 assays. Whereas, due to the inhibition of AHB2 activity at higher iron concentrations, an $Fe(NH_4)_2(SO_4)_2$ concentration of 25 μM was used in ABH2 assays.

The *in vitro* assay conditions which seemed most suitable for the ABH2 protein were 50 mM Hepes.KOH pH 7.5, 25 μM $Fe(NH_4)_2(SO_4)_2$, 1 mM α-ketoglutarate, 2 mM ascorbic acid and 50 μg/ml BSA and for the ABH3 protein

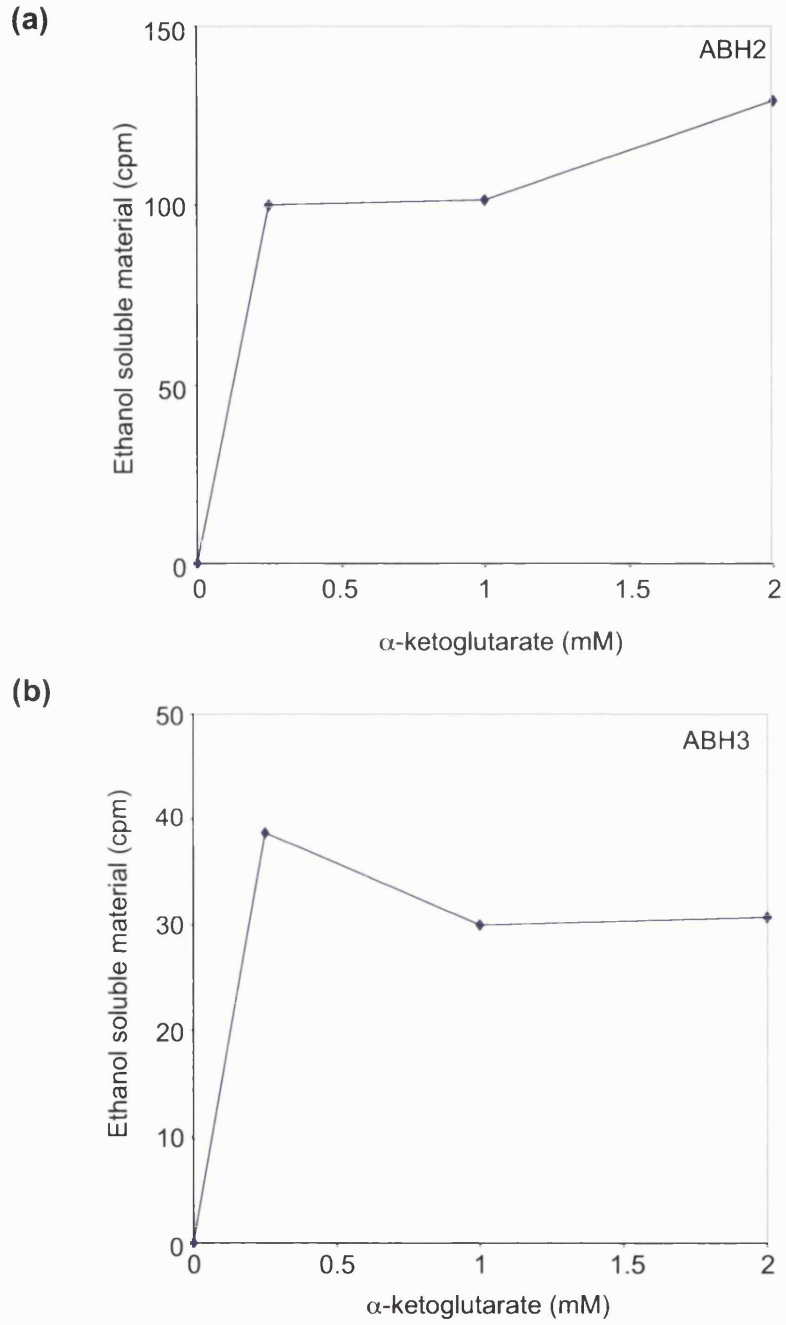


Figure 26. Effect of α-ketoglutarate concentration on ABH2 and ABH3 activities. 150 pmoles (a) GST-tagged ABH2 was incubated with 1400 cpm of [^{14}C]-methylated poly(dA) and (b) GST-tagged ABH3 was incubated with 400 cpm of [^{14}C]-methylated poly(dC) in the presence of 50 mM Hepes.KOH pH 8.0, 2 mM ascorbic acid, 75 μM $Fe(NH_4)_2(SO_4)_2$, 50 μg/ml BSA and various concentrations of α-ketoglutarate at 37 °C for 30 min. The release of ethanol soluble radioactive material was monitored. The data is expressed as the average of two replicates.

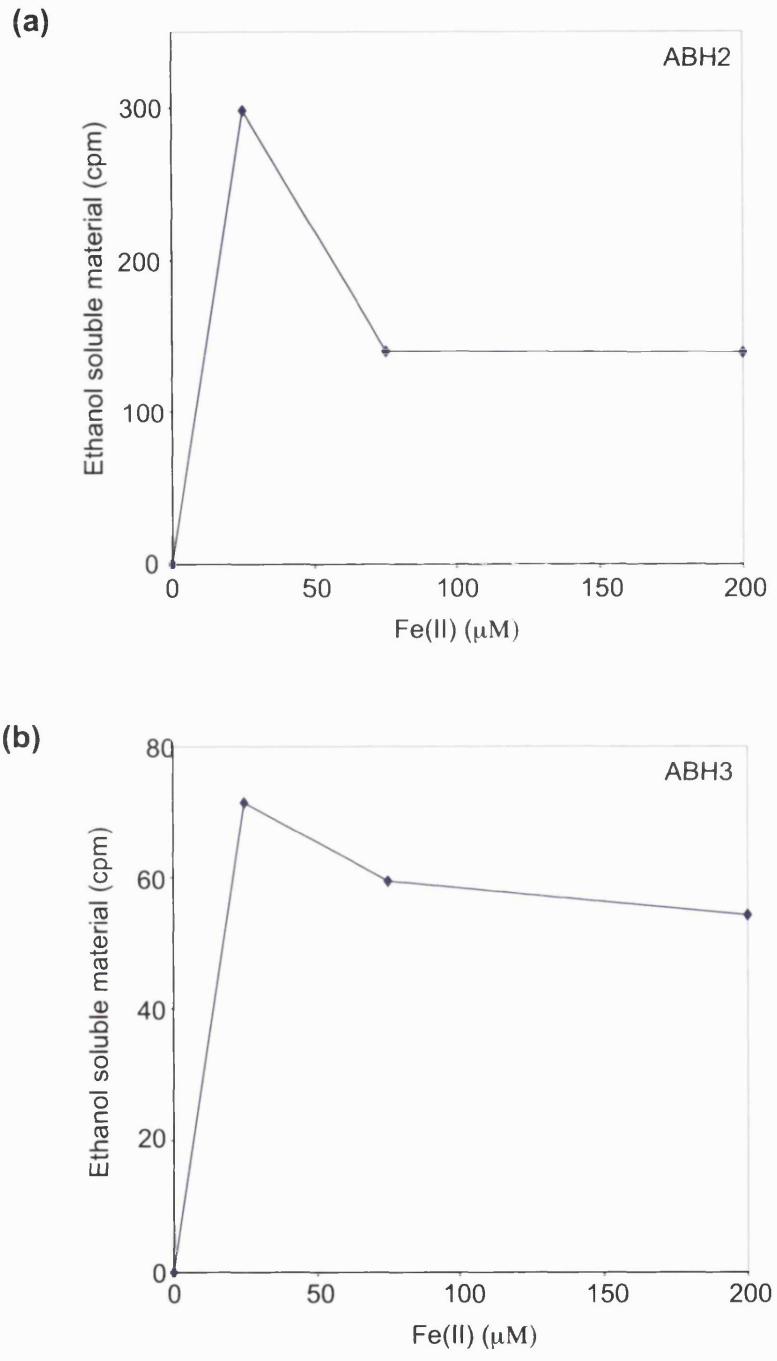


Figure 27. Effect of Fe(II) concentration on ABH2 and ABH3 activities. 150 pmoles of (a) GST-tagged ABH2 was incubated with 1400 cpm of [^{14}C]-methylated poly(dA) and (b) GST-tagged ABH3 was incubated with 400 cpm of [^{14}C]-methylated poly(dC) in the presence of 50 mM Hepes.KOH pH 8.0, 2 mM ascorbic acid, 1 mM α-ketoglutarate, 50 μg/ml BSA and various concentrations of $Fe(NH_4)_2(SO_4)_2$ at 37 °C for 30 min. The release of ethanol soluble radioactive material was monitored. The data is expressed as the average of two replicates.

the most appropriate assay conditions were 50 mM Hepes.KOH pH 8.0, 75 μM $Fe(NH_4)_2(SO_4)_2$, 1 mM α-ketoglutarate, 2 mM ascorbic acid and 50 μg/ml BSA.

It was observed that a white precipitate appeared in assays containing high concentrations of the ABH3 protein. The ABH3 protein was soluble in its storage buffer (50 mM Tris.HCl pH 8.5, 1 mM DTT, 150 mM NaCl, 10 % glycerol) but appeared to precipitate under the conditions used for the assay. After investigation, it became apparent that the NaCl in the storage buffer was the key to ABH3 solubility. The NaCl could be substituted for the more physiological salt, KCl, and the addition of 150 mM KCl to ABH3 assays prevented the formation of the white precipitate. However, the addition of 150 mM KCl to the assays did not significantly increase the activity of the ABH3 or ABH2 proteins (data not shown).

Despite the optimisation of the assay conditions, the activities of the human ABH2 and ABH3 proteins remain relatively low. It is not clear if this low activity is due to the predicted instability of the ABH2 and ABH3 proteins, or a problem arising during protein purification, such as misfolding of the proteins during overexpression. However, it is possible that to obtain full activity the ABH2 and ABH3 proteins need another co-factor, a post-translational modification or may need to interact with a protein partner.

In the newly optimised assay conditions titrations of the ABH2 and ABH3 proteins were incubated with [^{14}C]-methylated poly(dA) and poly(dC) substrates. The human ABH2 protein released ethanol soluble material from both the [^{14}C]-methylated poly(dA) and poly(dC) substrates (Figure 28). In a similar manner, the ABH3 protein was active on both the methylated poly(dA) and poly(dC) substrates (Figure 29). The ABH2 and ABH3 activities were both observed with either GST-

or His-tagged proteins and therefore are not catalysed by Fe(II) co-ordinated in the His-tag.

5.4 ABH2 and ABH3 are α-ketoglutarate-Fe(II)-dependent dioxygenases

To determine whether the ABH2 and ABH3 proteins are members of the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily, the co-factor requirements of the ABH2 and ABH3 activities were investigated. The activities of both the human proteins, ABH2 and ABH3, were found to be dependent on the presence of Fe(II), no ABH2 or ABH3 activity was observed in the absence of iron and the activities of both the ABH2 and ABH3 proteins were inhibited by the metal chelator EDTA (Table 5). α-ketoglutarate, a co-substrate of α-ketoglutarate-Fe(II)-dependent dioxygenases, was a requirement for the activity of both the ABH2 and ABH3 proteins (Table 5). ABH2 and ABH3 were both stimulated by the reducing agent ascorbic acid. The dependence of the activity of both ABH2 and ABH3 on α-ketoglutarate and Fe(II) and stimulation by ascorbic acid (Table 5), indicates that ABH2 and ABH3 are members of the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily.

5.5 ABH2 and ABH3 act on 1-meA and 3-meC

The *E. coli* AlkB protein has been demonstrated to directly revert 1-meA and 3-meC in DNA. To show that ABH2 and ABH3 are human functional homologs of the AlkB protein, it is important to demonstrate that the ABH2 and ABH3 protein act upon these lesions in DNA.

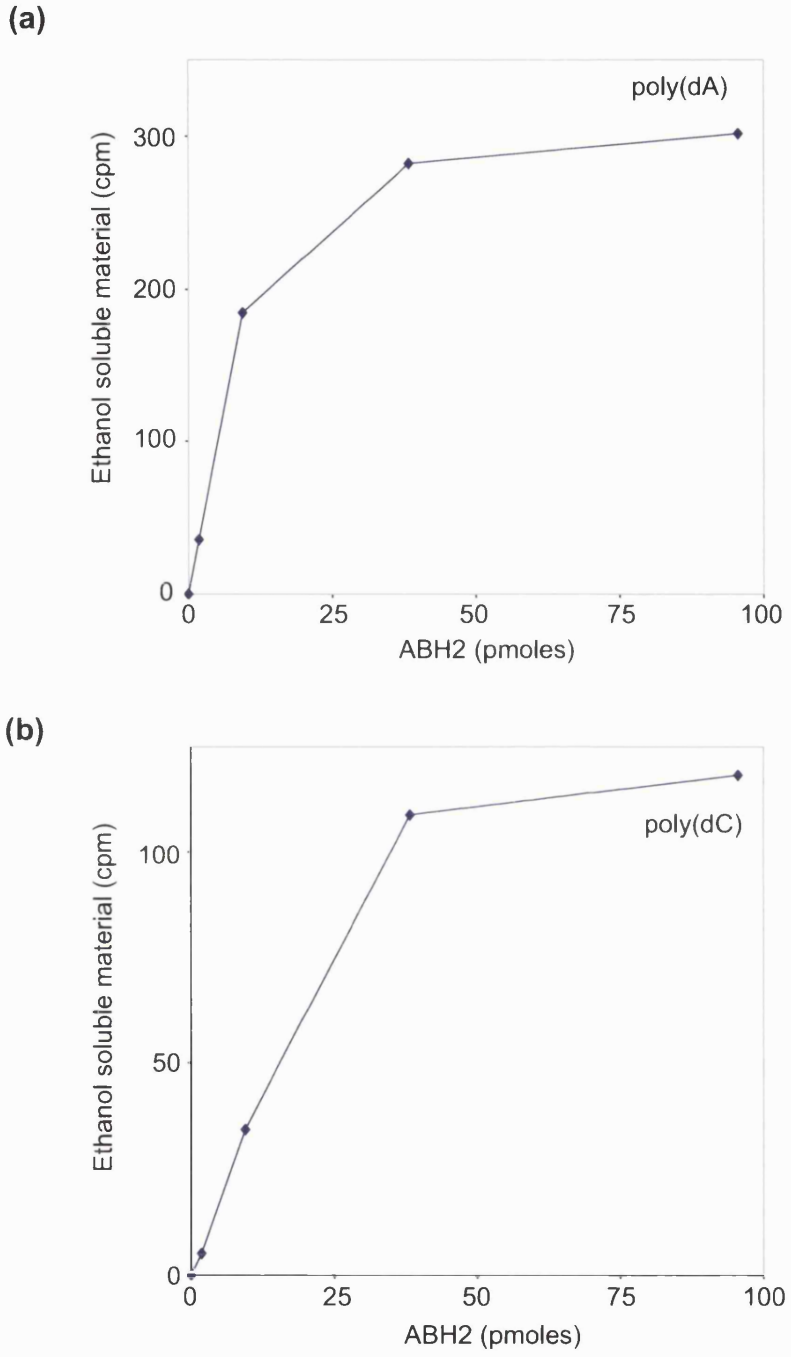


Figure 28. Release of ethanol soluble material from [^{14}C]-methylated poly(dA) and poly(dC) by ABH2 in the presence of Fe(II) and α-ketoglutarate. Various amounts of GST-tagged ABH2 were incubated with (a) 700 cpm of [^{14}C]-methylated poly(dA) and (b) 700 cpm of [^{14}C]-methylated poly(dC) in the presence of 50 mM Hepes pH 7.5, 2 mM ascorbic acid, 1 mM α-ketoglutarate, 25 μM $Fe(NH_4)_2(SO_4)_2$ and 50 μg/ml BSA at 37 °C for 30 min. The release of ethanol soluble radioactive material was monitored. The data is expressed as the average of two replicates.

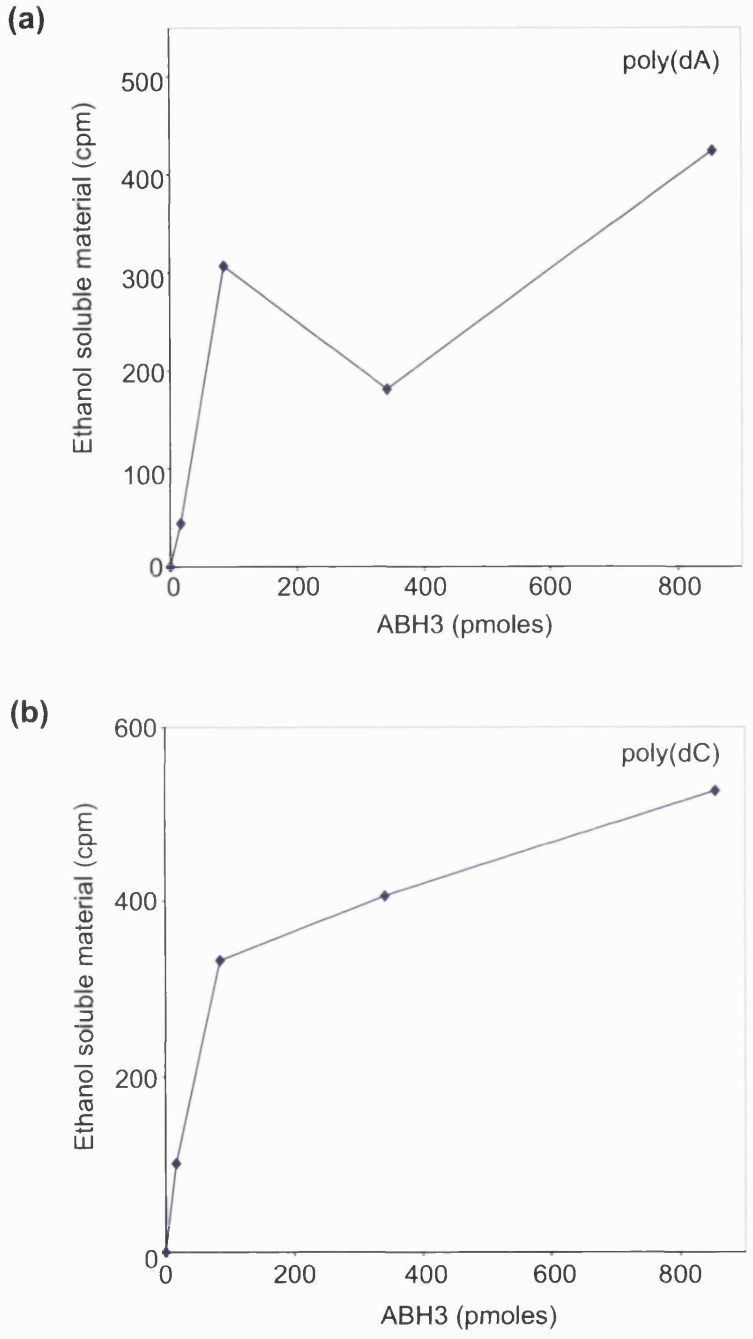


Figure 29. Release of ethanol soluble material from [^{14}C]-methylated poly(dA) and poly(dC) by ABH3 in the presence of Fe(II) and α-ketoglutarate. Various amounts of GST-tagged ABH3 were incubated with (a) 700 cpm of [^{14}C]-methylated poly(dA) and (b) 700 cpm of [^{14}C]-methylated poly(dC) in the presence of 50 mM Hepes.KOH pH 8.0, 2 mM ascorbic acid, 1 mM α-ketoglutarate, 75 μM $Fe(NH_4)_2(SO_4)_2$ and 50 μg/ml BSA at 37 °C for 30 min. The release of ethanol soluble radioactive material was monitored. The data is expressed as the average of two replicates.

(a)

	Activity (%)
Complete	100
No ABH2	<0.5
No α-ketoglutarate	<0.5
No ascorbic acid	<0.5
No $Fe(NH_4)_2(SO_4)_2$	<0.5
+ 10 mM EDTA	<0.5

(b)

	Activity (%)
Complete	100
No ABH3	<0.5
No α-ketoglutarate	<0.5
No ascorbic acid	2
No $Fe(NH_4)_2(SO_4)_2$	<0.5
+ 10 mM EDTA	<0.5

Table 5. Requirements of ABH2 and ABH3 activities. (a) 200 pmoles of His-tagged ABH2 or (b) 500 pmoles of His-tagged ABH3 were incubated with 2200 cpm of [^{14}C]-methylated poly(dA) in the optimised reaction conditions or in the conditions specified at 37 °C for 30 min. The release of ethanol soluble radioactive material was monitored.

To determine which lesions the human proteins were acting upon, after incubation with ABH2 and ABH3 in the optimised conditions, the methylated bases remaining in the [^{14}C]-methylated poly(dA) or poly(dC) substrates were released by acid hydrolysis, separated by HPLC and quantified by scintillation counting. Incubation with the ABH2 protein reduced the level of 1-meA present in the [^{14}C]-methylated poly(dA) substrate by 52 %, the amount of 3-meA and 7-meA present was unaffected (Figure 30). ABH2 also repaired 3-meC, completely removing 3-meC from the [^{14}C]-methylated poly(dC) substrate (Figure 31). The other human AlkB sequence homolog, ABH3, acted on 1-meA, reducing the amount of 1-meA present in the [^{14}C]-methylated poly(dA) substrate by 54 %, however, ABH3 did not act on 3-meA or 7-meA (Figure 32). ABH3 also totally removed 3-meC from the [^{14}C]-methylated poly(dC) substrate (Figure 33). Therefore, both of the human proteins ABH2 and ABH3 specifically act on 1-meA and 3-meC, the same lesions as are repaired by the *E. coli* AlkB protein.

As there are at least two human AlkB sequence homologs, it is of interest to determine whether both human proteins tackle both lesions with a similar efficiency. The relative ability of the ABH2 and ABH3 proteins to release of ethanol soluble radioactive material from [^{14}C]-methylated poly(dA) or poly(dC) substrates was assayed. It was found that the human ABH2 protein was approximately 4-fold more active on 1-meA than on 3-meC, whereas the human ABH3 protein was slightly more active, approximately 2-fold, on 3-meC than 1-meA (Figure 34).

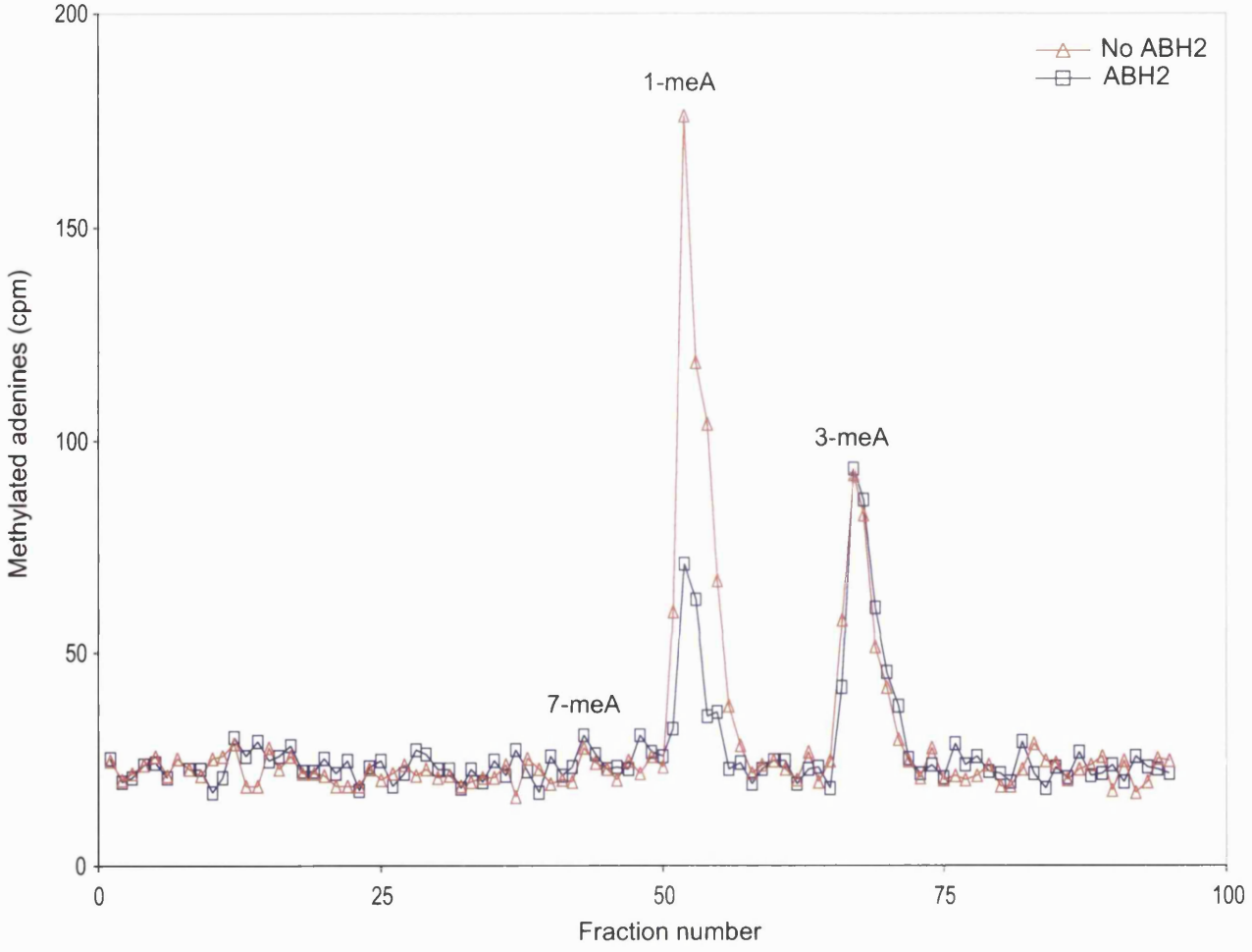


Figure 30. ABH2 acts on 1-meA. 1500 cpm of [^{14}C]-MeI treated poly(dA) was incubated (Δ) without ABH2 or (❐) with ABH2 in the optimised conditions at 37 °C for 30 min. The [^{14}C]-labeled methylated bases remaining in the DNA substrate were analysed by HPLC and scintillation counting.

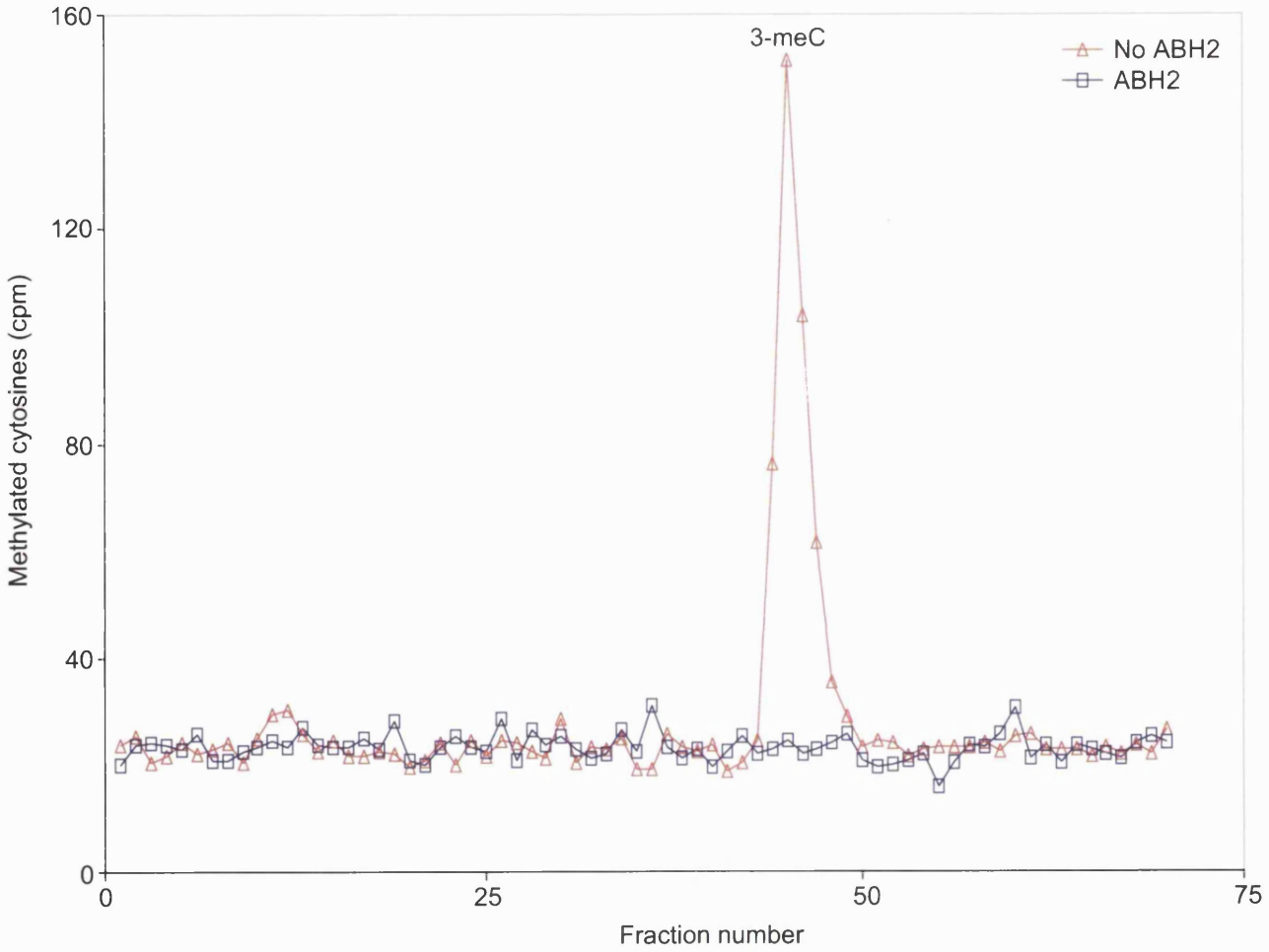


Figure 31. ABH2 acts on 3-meC. 1000 cpm of [^{14}C]-MeI treated poly(dC) was incubated (Δ) without ABH2 or (❐) with ABH2 in the optimised conditions at 37 °C for 30 min. The [^{14}C]-labeled methylated bases remaining in the DNA substrate were analysed by HPLC and scintillation counting.

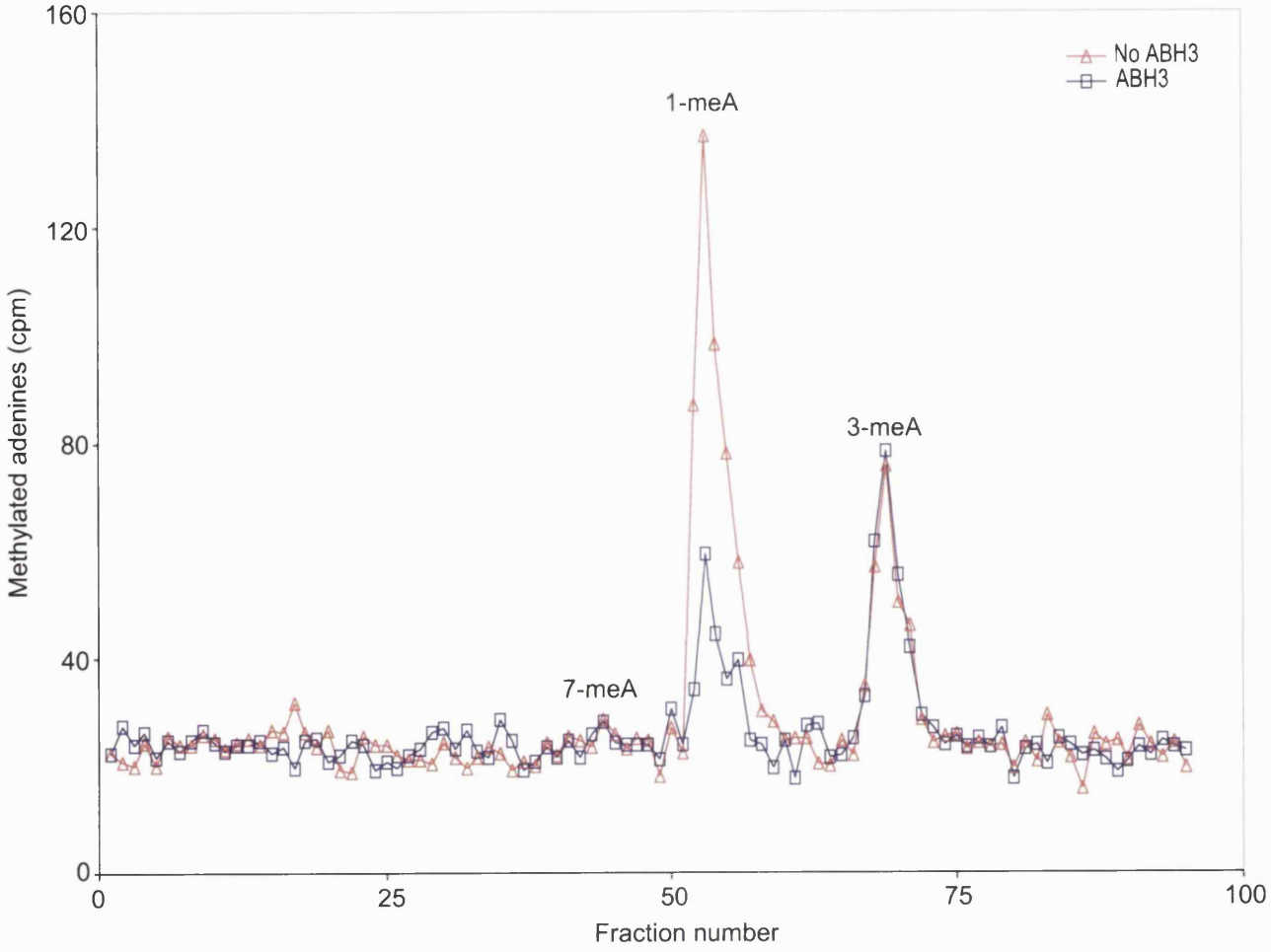


Figure 32. ABH3 acts on 1-meA. 1500 cpm of [^{14}C]-MeI treated poly(dA) was incubated (Δ) without ABH3 or (❐) with ABH3 in the optimised conditions at 37 °C for 30 min. The [^{14}C]-labeled methylated bases remaining in the DNA substrate were analysed by HPLC and scintillation counting.

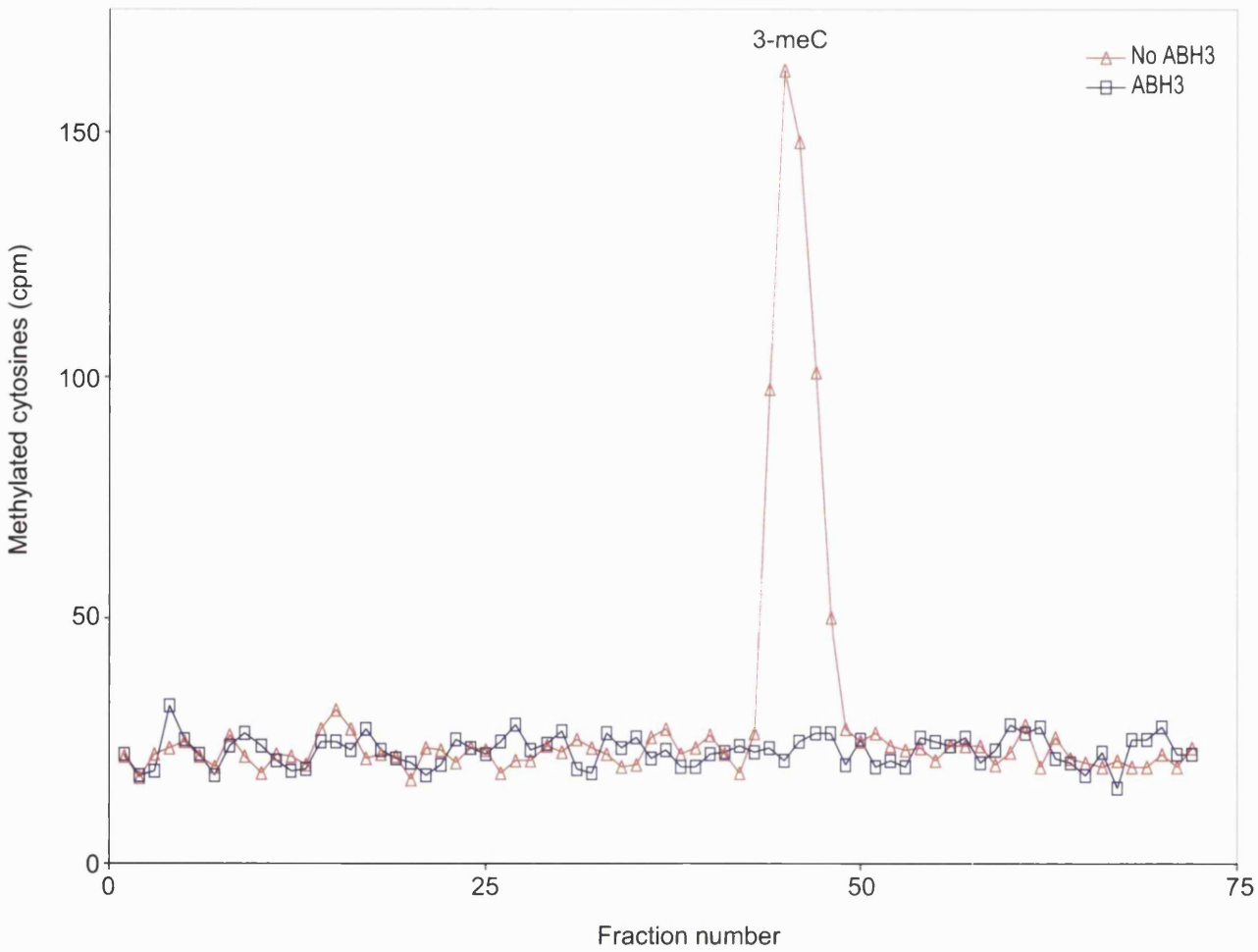


Figure 33. ABH3 acts on 3-meC. 1000 cpm of [^{14}C]-MeI treated poly(dC) was incubated (Δ) without ABH3 or (❐) with ABH3 in the optimised conditions at 37 °C for 30 min. The [^{14}C]-labeled methylated bases remaining in the DNA substrate were analysed by HPLC and scintillation counting.

5.6 ABH2 and ABH3 act on double stranded DNA

Both of the human AlkB homologs, ABH2 and ABH3 have been demonstrated to act on single stranded DNA substrates. The ability of ABH2 and ABH3 to act on double stranded DNA was tested by assaying the ability of these proteins to release ethanol soluble radioactive material from a [^{14}C]-MeI treated DNA oligonucleotide 116 annealed to a complementary strand (oligonucleotide 117). The DNA substrate was verified to be > 90 % double stranded by *E. coli* exonuclease III digestion. Both the human ABH2 and ABH3 enzymes were active on double stranded DNA (Figure 35).

5.7 ABH3 directly reverts 3-meC to cytosine

Here it has been demonstrated that the *E. coli* AlkB protein directly reverts 1-meA to adenine and 3-meC to cytosine in DNA. To establish whether ABH2 and ABH3 act by the same direct reversal mechanism, the ability of the human enzymes to revert 3-meC to cytosine in DNA was examined.

A single stranded oligonucleotide substrate containing [^{3}H]-cytosine residues interspersed between abundant thymine residues was synthesised by PCR using DNA oligonucleotide 133 as a template and DNA oligonucleotide 134 as a single primer (Table 3) in the presence of [^{3}H]-dCTP. The oligonucleotide was then repeatedly methylated with DMS until 62 % of the cytosine residues present were converted to 3-meC. The human ABH2 or ABH3 proteins were incubated with this heavily methylated substrate. The reversion of [^{3}H]-3-meC residues to [^{3}H]-cytosine residues in the DNA substrate was examined; the bases were released by strong acid hydrolysis, separated by HPLC chromatography and quantified by scintillation counting.

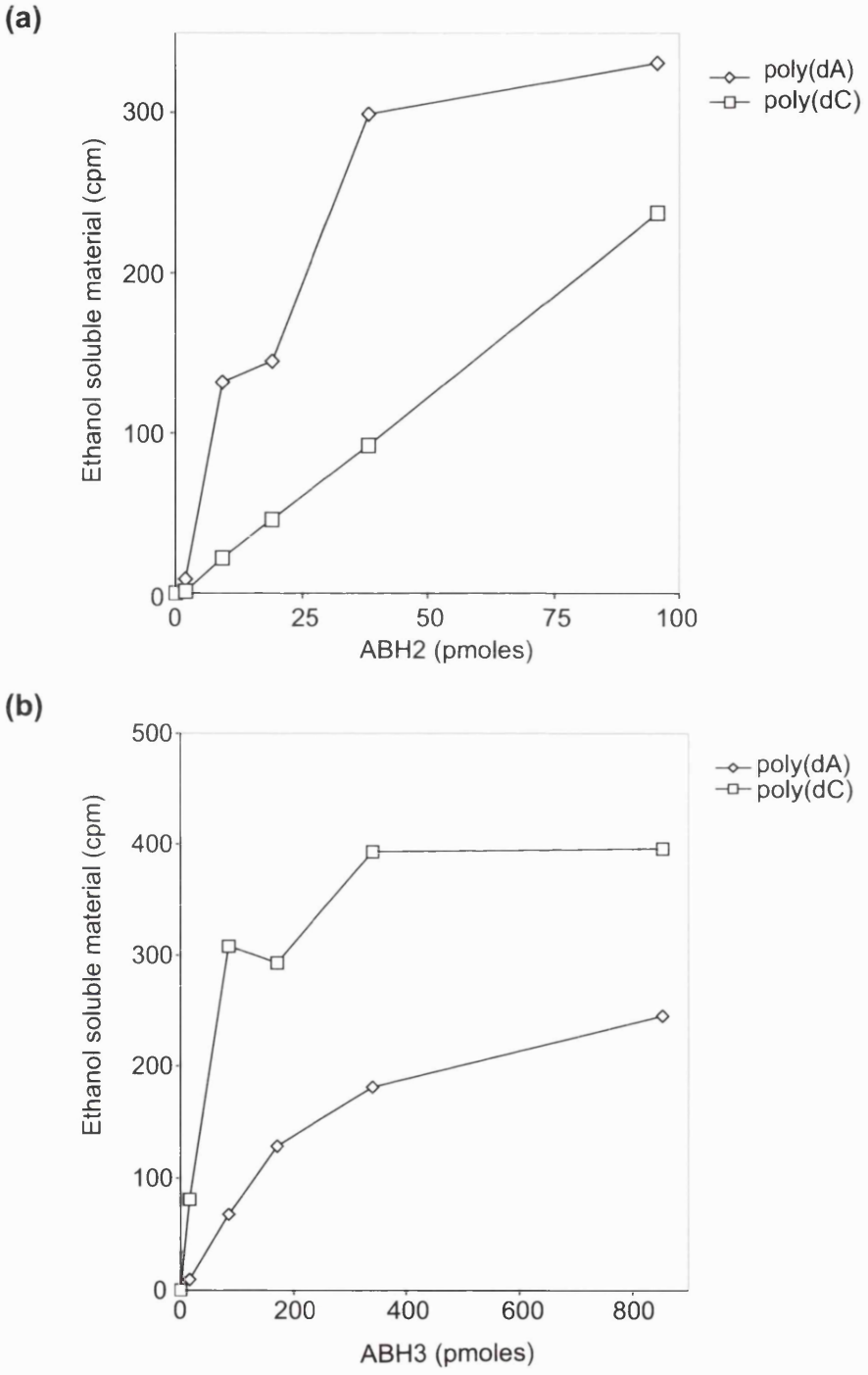


Figure 34. Efficiency of ABH2 and ABH3 proteins to act on 1-meA and 3-meC. Various amounts of (a) ABH2 and (b) ABH3 were incubated with 800 cpm of [^{14}C]-methylated poly(dA) and 500 cpm of [^{14}C]-methylated poly(dC) in the optimised reaction conditions at 37 °C for 30 min. The release of ethanol soluble radioactive material was monitored. The data is expressed as the average of three replicates.

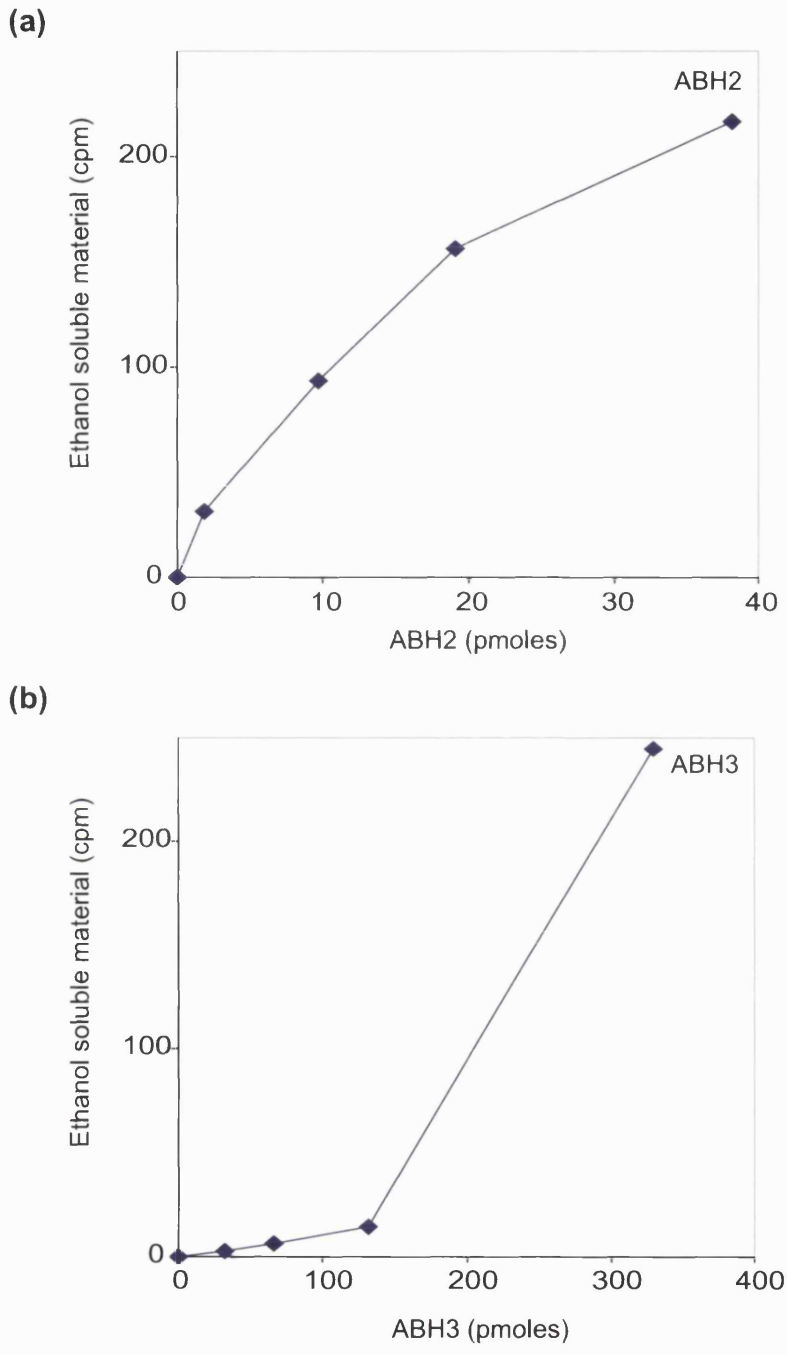


Figure 35. ABH2 and ABH3 act on double stranded DNA. [^{14}C]-methylated oligonucleotide 116 was annealed to oligonucleotide 117. 2660 cpm of this DNA substrate was incubated with various amounts of (a) ABH2 and (b) ABH3 in the optimised conditions at 37 °C for 30 min. The release of ethanol soluble material was monitored. The data is expressed as the average of two replicates.

The lower protein concentration of the ABH2 protein and its intrinsically low activity on 3-meC made it difficult to analyse of the ability of ABH2 to directly revert 3-meC. However, a small but reproducible (approximately 10 %), conversion of 3-meC to cytosine in DNA was catalysed by the ABH2 protein (data not shown). Therefore, ABH2 directly reverts 3-meC to cytosine residues in DNA, although inefficiently.

Incubation with the ABH3 protein decreased the proportion of 3-meC present in the substrate. Approximately 62 % of the cytosine residues in the DNA substrate were present as 3-meC, however after incubation with ABH3, only 32 % of the cytosine residues were present as 3-meC. The amount of unmodified cytosine present in the DNA substrate increased stoichiometrically, from 38 % in the substrate to 68 % after incubation with ABH3 (Figure 36). ABH3 therefore catalysed a decrease in 3-meC, linked to a stoichiometric increase in the amount of cytosine present. Therefore, ABH3 directly reverts the 3-meC residues to cytosine residues in DNA.

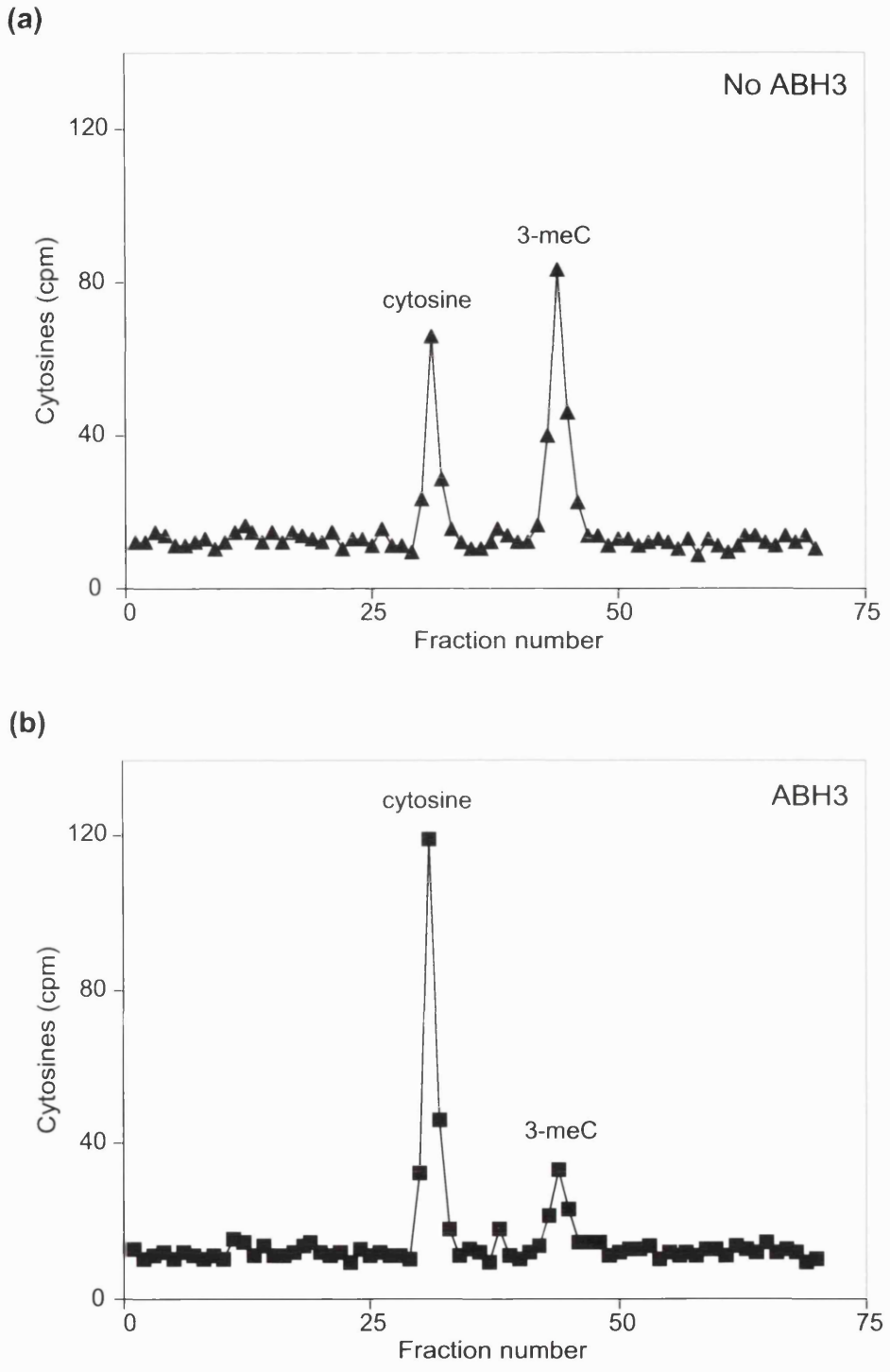


Figure 36. ABH3 directly reverts 3-meC to cytosine residues in DNA. A thymine-rich oligodeoxynucleotide containing several [^{3}H]-cytosine residues was heavily methylated with DMS. 1100 cpm of this substrate was incubated (a) without or (b) with 0.84 nmoles ABH3 in the optimised conditions at 37 °C for 30 min. The [^{3}H]-cytosine residues remaining in the oligonucleotide were analysed by HPLC and scintillation counting.

CHAPTER 6

DISCUSSION

CHAPTER 6. DISCUSSION

The function of the *E. coli* AlkB protein has eluded investigators for twenty years. Here the mystery is solved. The biochemical function of the *E. coli* AlkB protein and its human homologs ABH2 and ABH3 are elucidated. It is demonstrated that AlkB, ABH2 and ABH3 proteins are involved in DNA repair rather than tolerance of DNA methylation damage. It is shown that the purified AlkB, ABH2 and ABH3 proteins repair two lesions 1-meA and 3-meC, in both single and double stranded DNA. These proteins are shown to act by a direct reversal mechanism converting 1-meA and 3-meC to their unsubstituted parent residues in DNA. It is proposed that these proteins act by oxidative demethylation, an unprecedented mechanism of DNA repair.

6.1 AlkB, ABH2 and ABH3 are α-ketoglutarate-Fe(II)-dependent dioxygenases.

It was suggested, using sequence profile searches and protein fold recognition that the *E. coli* AlkB protein resembles members of the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily (Aravind and Koonin, 2001). Members of this superfamily catalyse a wide variety of two electron oxidation reactions including hydroxylation, ring expansion, desaturation of carbon-carbon bonds and oxidative ring closure, for a generalised reaction mechanism (see Figure 11). No members of this superfamily were known to act on DNA, however, an *in vitro* assay was developed to test whether in the presence of the distinctive co-factors required by this superfamily, the AlkB protein could modify or remove methylated DNA lesions.

My data demonstrates that the *E. coli* AlkB protein is an α-ketoglutarate-Fe(II)-dependent dioxygenase. The AlkB protein released radioactive material from [^{14}C]- and [^{3}H]-methylated DNA substrates (Figure 15 and 16) and this activity was dependent on the cofactors characteristic of the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily (Table 4). Enzymes of this superfamily use reactive iron-oxo species to catalyse their reactions; the observed AlkB activity exhibits an absolute requirement for Fe(II) (Table 4). The activity of the AlkB protein is dependent on the presence of α-ketoglutarate, a co-substrate for reactions catalysed by this superfamily, which during the reaction is oxidatively decarboxylated to produce succinate and CO_2. The activity of the AlkB protein is stimulated by ascorbic acid, a characteristic that has been observed in several other superfamily members. In collaboration, T. Henshaw and R. Hausinger (University of Michigan) demonstrated that succinate is generated and dioxygen is consumed during the AlkB reaction. They also showed that the spectroscopic signature of the AlkB protein is similar to that of other members of the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily.

Two human AlkB sequence homologs, the ABH2 and ABH3 proteins were active in the newly developed *in vitro* assay (Figure 28 and 29). The activities of both the ABH2 and ABH3 proteins are dependent on the presence of Fe(II) ions, α-ketoglutarate and are stimulated by ascorbic acid (Table 5). Therefore, the ABH2 and ABH3 proteins are also members of the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily.

6.2 AlkB, ABH2 and ABH3 act on 1-meA and 3-meC in DNA

Analysis of the methylated bases remaining in the DNA substrates after incubation with the *E. coli* AlkB protein, showed that AlkB acted specifically on two lesions, 1-meA and 3-meC (Figure 18 and 19). The AlkB protein did not show any activity against any other of the major lesions produced by S_N2 methylating agents such as 3-meA, 7-meA or 7-meG.

ABH2 and ABH3 are human functional homologs of the *E. coli* AlkB protein. ABH2 and ABH3 repair 1-meA and 3-meC in DNA (Figure 30, 31, 32 and 33) and were shown by B. Sedgwick in our laboratory to functionally complement the *E. coli alkB* mutant. These AlkB homologs were also identified independently (Aas *et al.*, 2003).

We were unable to confirm that the previously reported human homolog of AlkB (Wei *et al.*, 1996), referred to here as ABH1, is an AlkB functional homolog. *In vitro* assays with the FLAG-tagged ABH1 and by B. Sedgwick with the His-tagged ABH1 protein could detect no activity. B. Sedgwick demonstrated that the ABH1 gene did not complement the *E. coli alkB* mutant. The ABH1 protein has the closest sequence homology to the *E. coli* AlkB protein, so the possibility cannot be ruled out that ABH1 is an AlkB functional homolog that was inactive in the assays, or that ABH1 has a related activity that was not detected.

There are several other examples of duplication of DNA repair activities in human cells, for example, the uracil-DNA glycosylases, UNG and SMUG1, the thymine-DNA glycosylases, TDG and MBD4 and the apurinic endonucleases, APE1 and APE2 (Wood *et al.*, 2001). The occurrence of alternative or back-up DNA repair activities emphasises the complex task of counteracting DNA damage in the large genomes of higher eukaryotes.

6.3 AlkB, ABH2 and ABH3 act by direct reversal of base damage

As members of the α-ketoglutarate-Fe(II)-dependent dioxygenase superfamily catalyse a variety of reactions, it was not clear by what mechanism AlkB, ABH2 and ABH3 were acting. The specificity of the AlkB, ABH2 and ABH3 proteins for 1-meA and 3-meC, and the requirement for α-ketoglutarate ruled out the possibility that the observed activity is that of an exonuclease. The mode of action of AlkB is distinguished from that of MTases as it is enzymatic; from the data in Figure 15, a limiting amount of 0.1 pmole of AlkB protein removed an estimated 1.7 pmoles of 1-meA from methylated poly(dA) in 15 minutes.

It has been demonstrated that the AlkB, ABH2 and ABH3 proteins act by a direct reversal mechanism. Using DNA substrates in which the majority of the adenine or cytosine residues were methylated, the AlkB, ABH2 and ABH3 proteins were shown to stoichiometrically convert 3-meC residues in DNA to unmodified cytosine residues (Figure 22 and 36). In a similar manner the AlkB protein was also found to directly revert 1-meA residues to adenine residues in DNA (Figure 21).

6.4 Oxidative demethylation, an unprecedented mechanism of DNA repair

It is proposed that the AlkB, ABH2 and ABH3 proteins act by oxidative demethylation, an unprecedented mechanism of DNA repair. The α-ketoglutarate-Fe(II)-dependent dioxygenases, AlkB, ABH2 and ABH3, are proposed to hydroxylate 1-meA and 3-meC residues in DNA. The intermediates, which are expected to be 1-hydroxymethyladenine and 3-hydroxymethylcytosine, are

unstable and are expected to decompose to release formaldehyde, resulting in the direct reversal of 1-meA and 3-meC to the unmodified parent residues in DNA (Figure 37). My data shows that unmodified adenine and cytosine residues are products of the AlkB reaction, and in collaboration, T. Henshaw and R. Hausinger, confirmed that formaldehyde is also an AlkB reaction product. Therefore, the *E. coli* AlkB protein and its human homologs ABH2 and ABH3, directly revert DNA base damage by oxidative demethylation, a novel direct reversal mechanism of DNA repair. Our data are supported by a report of similar results (Falnes *et al.*, 2002).

Oxidative demethylation is a completely new mechanism of DNA repair. The only other direct reversal mechanism of DNA repair known to act in human cells is the reaction catalysed by the MGMT protein, which transfers the DNA methyl group from the DNA lesion onto a cysteine residue in the protein (Olsson and Lindahl, 1980). The SP lyase and photolyases also repair DNA damage by direct reversal mechanisms, however these enzymes have not been found in human cells. The SP lyase of *B. subtilis* cleaves the unique spore photoproduct (5-thyminyl-5,6-dihydrothymine dimer) by a free-radical mechanism to produce two thymine residues (Rebeil and Nicholson, 2001) and the photolyases use light energy to monomerise cyclobutyl pyrimidine dimers and pyrimidine-pyrimidone 6-4 photoproducts (Thompson and Sancar, 2002).

None of the α-ketoglutarate-Fe(II)-dependent dioxygenases were previously known to act on DNA, however, thymine hydroxylase is involved in the pyrimidine salvage pathway in *Neurospora crassa* and *Rhodotorula glutinis* and catalyses the oxidation of the methyl group of free thymine to its alcohol, aldehyde and carboxylic acid (5-carboxyuracil) in three successive reactions. A carboxylase

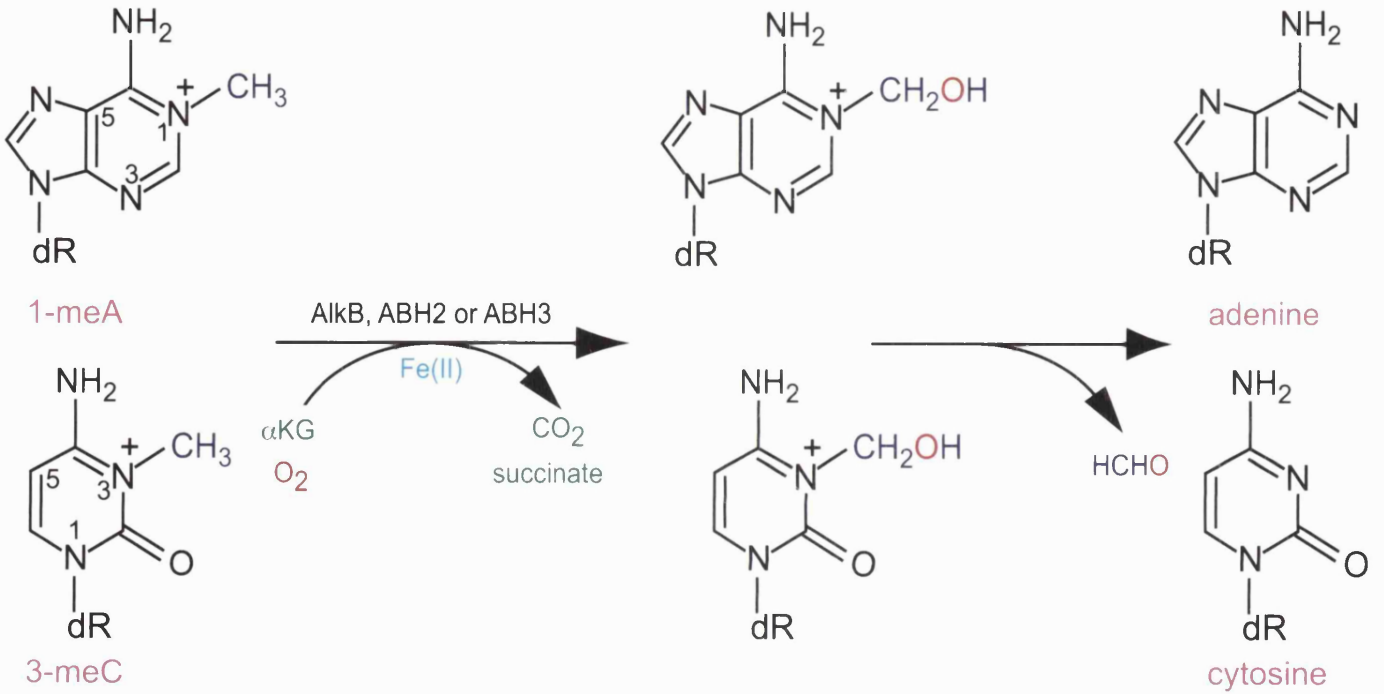


Figure 37. Oxidative demethylation, a novel direct reversal mechanism of DNA repair catalysed by AlkB, ABH2 and ABH3.

is then required to convert 5-carboxyuracil to uracil. Thymine hydroxylase is a promiscuous enzyme and *in vitro* can catalyse several other reactions including the demethylation of free 1-methylthymine bases to thymine (Bankel *et al.*, 1977; Thornburg *et al.*, 1993). Enzymes from other protein superfamilies catalyse demethylation reactions, for example, the heme containing cytochrome P450 superfamily of proteins can demethylate N-methylamines (Karki and Dinnocenzo, 1995) and amine oxidases, which are 'protein radical enzymes', use a modified tyrosine residue to deaminate methylamine (Lizcano *et al.*, 2000).

The lesions 1-meA and 3-meC are generated predominantly in single stranded DNA by S_N2 methylating agents. Therefore, this mechanism of direct reversal may have evolved to repair lesions in single stranded DNA where it is impossible to perform excision repair due to the lack of a complementary DNA strand. The MTases, which repair methylation damage by direct reversal are unable to act on 1-meA and 3-meC as these lesions are methylated on nitrogen atoms and the nitrogen-carbon bond is too strong. Therefore, a different direct reversal mechanism, oxidative demethylation, is required which is able to remove the tightly bound methyl groups from these lesions in single stranded DNA. However, it has been shown that AlkB, ABH2 and ABH3 are all able to repair DNA damage in double stranded DNA (Figure 17 and 35); this is advantageous to the cell as DNA is unfolded only transiently.

6.5 Oxidative demethylation, an accurate DNA repair mechanism

Direct reversal of 1-meA and 3-meC to unsubstituted parent residues in DNA by oxidative demethylation should be a highly accurate form of DNA repair, and this agrees with *in vivo* observations that repair by AlkB is non-mutagenic (Dinglay *et al.*, 2000). The formaldehyde generated by the oxidative

demethylation reaction could cause DNA-DNA or DNA-protein crosslinks. However, because formaldehyde is generated in several metabolic processes, cells have highly efficient mechanisms for its inactivation and metabolism (Ma and Harris, 1988). Therefore, the extremely low level of formaldehyde generated are unlikely to be a threat to the genome, especially compared to the danger of unrepaired 1-meA / 3-meC in the DNA.

6.6 1-meA and 3-meC are cytotoxic lesions

The AlkB protein protects against the cytotoxic rather than the mutagenic effects of DNA alkylation damage (Kataoka *et al.*, 1983). As the AlkB protein has been demonstrated to repair 1-meA and 3-meC, these must be primarily cytotoxic lesions. The 1 position of adenine and the 3 position of cytosine are involved in DNA base-pairing, therefore alkylation at these sites will interfere with the formation of base-pairs. 3-meC is known to inhibit DNA replication *in vitro* and under high fidelity conditions does not mispair during DNA synthesis, therefore, 3-meC is considered to be a cytotoxic lesion (Boiteux and Laval, 1982; Saffhill, 1984). 1-meA has been less well studied, but, 1-meA has been found to be one of the elements that contribute to the stalling of reverse transcriptase during the replication of retroviruses such as the human immunodeficiency virus (HIV); (Renda *et al.*, 2001). Direct evidence is presented here that 1-meA is a DNA replication stalling lesion. Pre-incubation of methylated DNA with the purified AlkB protein, in the optimised conditions, reduces the level of DNA replication stalling observed at methylated adenine residues (Figure 23).

The *E. coli alkB* mutants are weakly sensitive to MMS-induced mutagenesis, showing a small increase in the frequencies of G:C to A:T, G:C to T:A and A:T to T:A base substitutions (Dinglay *et al.*, 2000). Although, 1-meA

and 3-meC are expected to be cytotoxic lesions, *in vivo* these lesions cause a very low level of base substitution mutagenesis. Therefore, in the absence of the AlkB protein, DNA polymerases may occasionally bypass these 1-meA and 3-meC in an error prone manner.

6.7 The specificity of AlkB, ABH2 and ABH3 for 1-meA and 3-meC in DNA

The *E. coli* AlkB protein and its human homologs, ABH2 and ABH3, work specifically on 1-meA and 3-meC in DNA, but not on the other major lesions generated by S_N2 alkylating agents. It could be suggested that the AlkB, ABH2 and ABH3 proteins recognise 1-meA and 3-meC in DNA by their inability to form normal base-pairs, distorting the DNA helix. However, as AlkB and its human homologs are able to repair these lesions in single stranded DNA, helix distortion appears not to be a prime determinant for recognition of these lesions. 1-meA and 3-meC are both positively charged and this may help the AlkB protein locate these lesions, however, the positively charged lesion 3-meA is not an AlkB substrate. The methyl groups of 1-meA and 3-meC are located in similar chemical and structural environments; both these lesions are methylated on nitrogen atoms that are double bonded to carbon atoms with attached amino groups, this may aid AlkB in distinguishing these lesions from other methylated bases. It is likely that the AlkB, ABH2 and ABH3 proteins use a combination of these factors in the recognition of 1-meA and 3-meC in DNA. More extensive testing for AlkB activity against minor alkylated lesions, such as 3-methylthymine, may help to elucidate the substrate specificity of the AlkB, ABH2 and ABH3 proteins.

Although both human AlkB homologs act on 1-meA and 3-meC, they act with different efficiencies. The ABH2 protein is observed to be slightly more active on 1-meA and the ABH3 protein on 3-meC in DNA (Figure 34), however, it has been reported that these preferences are much greater (Aas *et al.*, 2003). This discrepancy may be due to differences in assay conditions. The *E. coli* AlkB protein is active on 1-ethyladenine, however the human proteins ABH2 and ABH3 act only very inefficiently on this substrate (Duncan *et al.*, 2002). The differences observed in the relative activities of the AlkB, ABH2 and ABH3 proteins for different lesions may be due, at least in part, to differences in side chains in and around the substrate binding pocket of these proteins. The differences in the pH optima of the proteins (Figure 12 and 25) is an indication that the ABH2 protein may use different amino acids in catalysis or in secondary structure formation from those used by the AlkB and ABH3 protein.

6.8 AlkB and ABH3 repair RNA methylation damage

AlkB homologues have been identified in plant RNA viruses replicating without a DNA intermediate, which lead to the suggestion that AlkB homologs may act on RNA (Aravind and Koonin, 2001). The *E. coli* AlkB protein and the human ABH3 protein, but not the ABH2 protein have been found to act on 1-meA and 3-meC in RNA oligonucleotides (Aas *et al.*, 2003). Therefore, it is proposed that the AlkB and ABH3 proteins are able to repair methylation damage of mRNA, tRNA and rRNA. Repair of RNA makes sense energetically as transcription of genes is energy consuming, as is building ribosome particles. However, tRNA and rRNA both contain 1-meA and 3-meC as natural enzyme-mediated modifications. In fact, the occurrence 1-meA at position 58 of the TψC loop has been reported in tRNAs from all three domains of life, suggesting an evolutionarily conserved role

in tRNA structure or function (Bjork *et al.*, 1987). It is not clear how or if AlkB and ABH3 distinguishes between these essential modifications and RNA methylation damage.

6.9 Cell cycle regulation

The lesions repaired by AlkB and its functional homologs are produced predominantly in single stranded DNA, therefore it was suggested that these lesions would be generated at replication forks and transcription bubbles (Dinglay *et al.*, 2000). Expression of the *Caulobacter crescentus alkB* sequence homolog was found to be cell cycle regulated with a pattern similar to activities required for DNA replication (Colombi and Gomes, 1997). Also the *E. coli alkB* mutants are more sensitive to alkylating agents during active growth than in stationary phase when only minimal DNA replication and transcription occurs (Dinglay *et al.*, 2000). The human ABH2 protein relocates to replication foci during S phase (Aas *et al.*, 2003) and is therefore implicated in the repair of DNA at replication forks, whereas ABH3 appears to have a role in maintenance of non-replicating DNA and of RNA possibly at sites of transcription.

6.10 Implications for cancer treatment

The human AlkB homologues, ABH2 and ABH3 both repair the DNA lesions 1-meA and 3-meC in DNA. These lesions are expected to be overwhelmingly cytotoxic lesions, they are unlikely to contribute to carcinogenesis. Therefore, humans with mutations in the ABH2 and / or ABH3 genes would not be expected to be cancer prone.

S_N1 alkylating agents, such as streptozotocin, temozolomide, dacarbazine and procarbazine are commonly used in chemotherapy. The main cytotoxic lesion produced by these S_N1 agents is O^6-meG, which via the action of the MMR system is a lethal lesion (Karran and Bignami, 1994). Many tumours up-regulate MGMT and become resistant to the killing effects of O^6-meG (Gerson, 2002; Margison and Santibanez-Koref, 2002). Therefore, the main determinant of susceptibility of some tumours to S_N1 alkylating agents may be the ability to repair 3-meA and minor alkylated lesions. As 1-meA and 3-meC are generated much more efficiently by S_N2 than S_N1 methylating agents it is not clear how significant a role 1-meA and 3-meC play in the cytotoxicity of S_N1 agents. However, as 1-meA and 3-meC are produced at higher levels in single stranded DNA such as that found at replication forks, these lesions may be formed and therefore play a role in cytotoxicity in rapidly dividing cells such as tumour cells. However, expression of the ABH2 and ABH3 genes does not appear to correlate with the proliferative status of the tissue (Duncan *et al.*, 2002). The ABH2 gene was rather intriguingly named Prostate cancer antigen 1 when originally submitted to the GenBank database, however the reason for this name is unclear.

Small drug molecules which antagonise the activity of the ABH2 and / or ABH3 proteins may be useful adjuncts to cancer therapy that uses simple alkylating agents, as these antagonists may sensitise rapidly dividing cells to the chemotherapy drugs. High resolution crystal structures of the human AlkB homologues may allow detailed atomic level screening and analysis to design small molecule antagonists of these enzymes. A useful starting place for drug design may come from analysis of the induction of starfish oocyte maturation, where the free base of 1-meA is used as a hormone. Several molecules, such as 8-dimethylamino-1-methyladenine have been found to inhibit the interaction between the 1-meA free base and its receptors (Monsees *et al.*, 1993)..

As myelosuppression is the acute dose-limiting toxicity of alkylating agents, it has been suggested that *ex vivo* gene therapy using retroviruses to express MGMT in bone marrow may help combat the myelosuppressive effects of the alkylating agents (Kleibl and Margison, 1998). If 1-meA and 3-meC are shown to contribute to the cytotoxicity of the S_N1 agents used in chemotherapy, the inclusion of viruses encoding ABH2 and / or ABH3 in this treatment may enhance the protection of the rapidly dividing bone marrow.

6.11 Other cellular demethylations

It is hypothesised that a similar oxidative demethylation mechanism could participate in removal of other forms of chemically stable alkylations, such as 5-methylcytosine (5-meC) or methyl-lysine residues in histones. Demethylation of 5-meC residues in mammalian DNA is considered to be an important mechanism of epigenetic control (Mayer *et al.*, 2000) but evidence for 5-meC demethylase activity is controversial (Smith, 2000). Our results suggest that these activities should now be reinvestigated using the conditions necessary for oxidative demethylation.

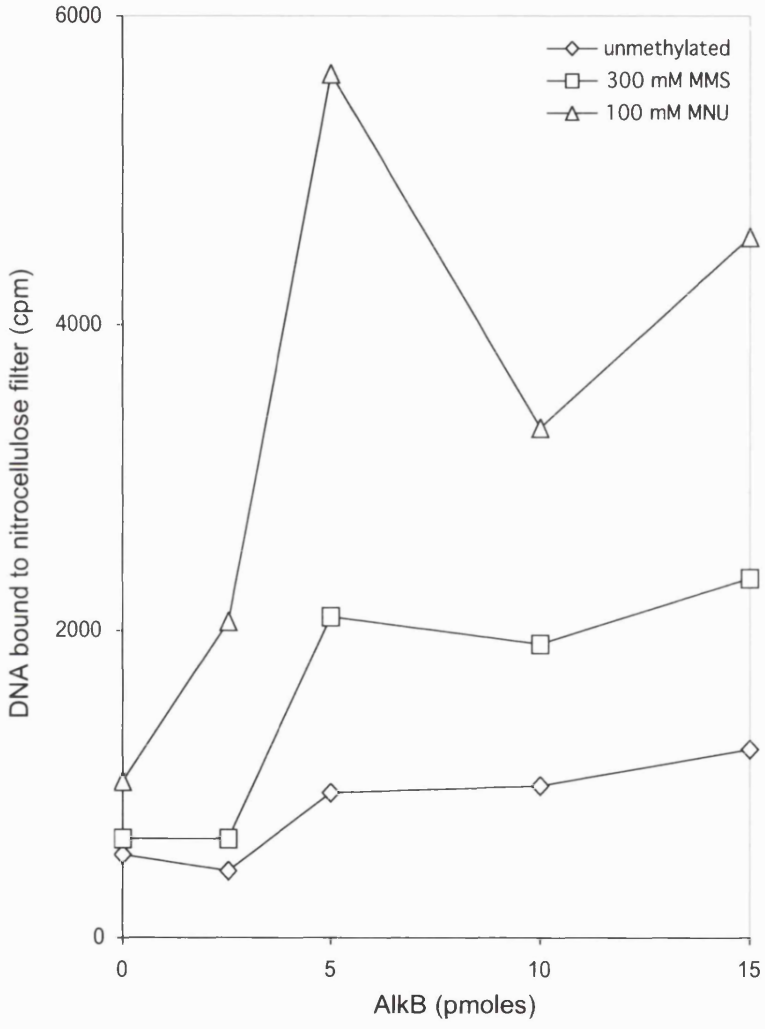
Proteins have been described which have an amino-terminal AlkB-like domain fused to a carboxyl-terminal domain that is predicted to transfer methyl groups from SAM to DNA, RNA or small molecules (Aravind and Koonin, 2001), these proteins could be involved in the regulation of cellular processes via controlled methylation and demethylation.

CONCLUSION

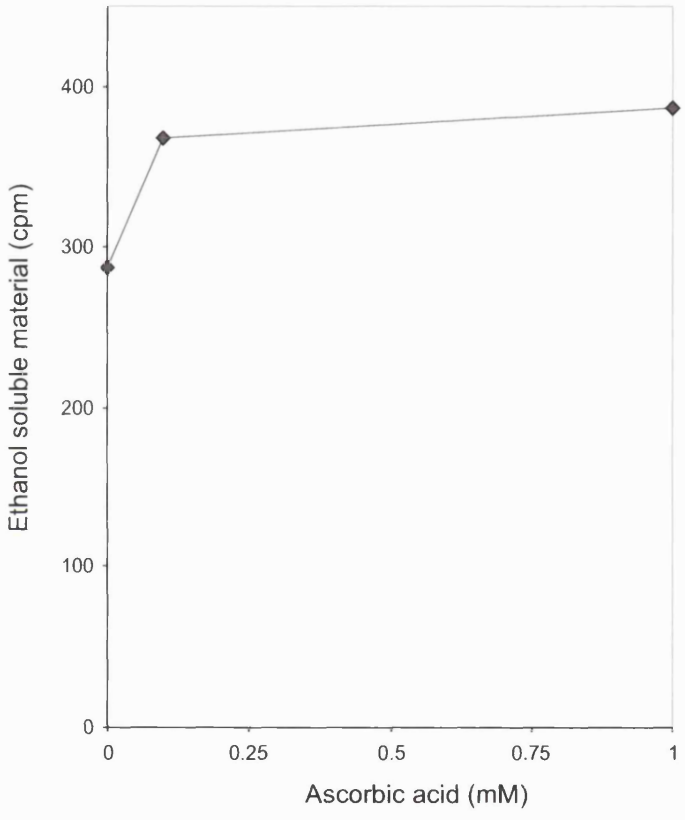
Here, the function of the enigmatic *E. coli* AlkB protein and its two human homologs ABH2 and ABH3 have been elucidated. *In vitro* assays were used to

demonstrate that the AlkB, ABH2 and ABH3 are α-ketoglutarate-Fe(II)-dependent dioxygenases. The AlkB, ABH2 and ABH3 proteins were shown to directly revert two lesions, 1-meA and 3-meC in DNA, to their unsubstituted parent residues. Therefore, it is proposed that AlkB, ABH2 and ABH3 act by oxidative demethylation, a novel and unprecedented mechanism of DNA repair.

APPENDIXES



Appendix A. Binding of AlkB to S_N1 methylated DNA. A ^{32}P end labeled 40mer oligonucleotide was methylated with 300 mM MMS or 100 mM MNU. Various amounts of His-tagged AlkB were incubated with these sustrates (30,000 cpm / reaction) at 30 °C for 30 min. Reaction mixtures were passed through nitrocellulose filters and the DNA bound to the retained AlkB was quantified by scintillation counting. The data is expressed as the average of two replicates.



Appendix B. The effect of ascorbic acid on reactions containing a high AlkB concentration. 320 pmoles of His-tagged AlkB were incubated with 1000 cpm of [^{14}C]-methylated poly(dA) in the presence of 50 mM Hepes.KOH pH 8.0, 1.5 mM α-ketoglutarate, 100 μM $Fe(NH_4)_2(SO_4)_2$ and various concentrations of ascorbic acid for 15 min at 37 °C. The release of ethanol soluble material was monitored. The data is expressed as the average of two replicates.

	% of cpm released
no exonuclease III	0
exonuclease III	93.3

Appendix C. Annealed DNA is double stranded. 100 units of exonuclease III (NEB) was added to 7 µg of the annealed oligonucleotides 116 and 117 and incubated at 37 °C for 1 hour. The reactions were stopped by the addition of EDTA, the DNA was ethanol precipitated and the supernatant was recovered. Scintillation counting monitored the release of radioactive material from the DNA.

Time (min)	% A_{260}	
0	82.4	17.6
25	82.6	17.4
50	61.7	38.3

Appendix D. 3-meC is not deaminated by a 25 min in 90 % formic acid at 180 °C. 30 µg of 3-methyl-dC were incubated with 90 % formic acid at 180 °C for various times. The formic acid was evaporated off and the DNA bases resuspended in 0.1 M HCl. The bases were separated by HPLC and quantitated by A_{260} measurements

Time (hours)	% of A_{260}			
	A	1-meA	3-meA	7-meA
2	71.6	24.0	2.9	1.5
4	54.2	39.2	4.4	2.1
5	38.2	53.5	5.8	2.4
6	31.3	59.7	6.3	2.6
7	21.8	71.4	4.4	2.4
8	17.1	76.3	4.3	2.2

Appendix E. Quantitation of methylated purines produced by heavy methylation of oligonucleotide 124. DNA oligonucleotide 124, which contains mostly thymine residues interspersed with adenine residues, was treated repeatedly with 50 mM DMS. Between each treatment a sample of the DNA was taken and the methylated bases present examined by HPLC and A_{260} measurements.

REFERENCES

Aas, P.A., Otterlei, M., Falnes, P.O., Vagbo, C.B., Skorpen, F., Akbari, M., Sundheim, O., Bjoras, M., Slupphaug, G., Seeberg, E. and Krokan, H.E. (2003) Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature*, **421**, 859-863.

Abbott, P.J. and Saffhill, R. (1977) DNA-synthesis with methylated poly(dA-dT) templates: possible role of O^4-methylthymine as a pro-mutagenic base. *Nucleic Acids Res*, **4**, 761-769.

Altschul, S.F. and Koonin, E.V. (1998) Iterated profile searches with PSI-BLAST-a tool for discovery in protein databases. *Trends Biochem Sci*, **23**, 444-447.

Aravind, L. and Koonin, E. (2001) The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. *Genome Biol*, **2**, RESEARCH0007.

Aravind, L., Walker, D.R. and Koonin, E.V. (1999) Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic Acids Res*, **27**, 1223-1242.

Baldwin, J.E., Adlington, R.M., Crouch, N.P., Coates, J.B., Keeping, J.W., Schofield, C.J., Shuttleworth, W.A. and Sutherland, J.D. (1988) Substrate specificity of cloned deacetoxycephalosporin C/deacetylcephalosporin C synthetase. *J Antibiot*, **41**, 1694-1695.

Bankel, L., Lindstedt, G. and Lindstedt, S. (1977) Thymine 7-hydroxylase from *Neurospora crassa*. Substrate specificity studies. *Biochim Biophys Acta*, **481**, 431-437.

Barrows, L. and Magee, P. (1982) Nonenzymatic methylation of DNA by S-adenosylmethionine in vitro. *Carcinogenesis*, **3**, 349-351.

Batty, D.P. and Wood, R.D. (2000) Damage recognition in nucleotide excision repair of DNA. *Gene*, **241**, 193-204.

Beard, W.A., Osheroff, W.P., Prasad, R., Sawaya, M.R., Jaju, M., Wood, T.G., Kraut, J., Kunkel, T.A. and Wilson, S.H. (1996) Enzyme-DNA interactions

required for efficient nucleotide incorporation and discrimination in human DNA polymerase beta. *J Biol Chem*, **271**, 12141-12144.

Beranek, D.T., Weis, C.C. and Swenson, D.H. (1980) A comprehensive quantitative analysis of methylated and ethylated DNA using high pressure liquid chromatography. *Carcinogenesis*, **1**, 595-606.

Berdal, K.G., Johansen, R.F. and Seeberg, E. (1998) Release of normal bases from intact DNA by a native DNA repair enzyme. *EMBO J*, **17**, 363-367.

Biswas, T., Ramana, C.V., Srinivasan, G., Boldogh, I., Hazra, T.K., Chen, Z., Tano, K., Thompson, E.B. and Mitra, S. (1999) Activation of human O^6-methylguanine-DNA methyltransferase gene by glucocorticoid hormone. *Oncogene*, **18**, 525-532.

Bjelland, S., Bjoras, M. and Seeberg, E. (1993) Excision of 3-methylguanine from alkylated DNA by 3-methyladenine DNA glycosylase I of *Escherichia coli*. *Nucleic Acids Res*, **21**, 2045-2049.

Bjelland, S. and Seeberg, E. (1987) Purification and characterization of 3-methyladenine DNA glycosylase I from *Escherichia coli*. *Nucleic Acids Res*, **15**, 2787-2801.

Bjork, G.R., Ericson, J.U., Gustafsson, C.E., Hagervall, T.G., Jonsson, Y.H. and Wikstrom, P.M. (1987) Transfer RNA modification. *Annu Rev Biochem*, **56**, 263-287.

Bodell, W. and Singer, B. (1979) Influence of Hydrogen bonding in DNA and polynucleotides on reaction of Nitrogens and Oxygens toward Ethylnitrosourea. *Biochemistry*, **18**, 2860-2863.

Boiteux, S. and Laval, J. (1982) Mutagenesis by alkylating agents: coding properties for DNA polymerase of poly(dC) template containing 3-methylcytosine. *Biochimie*, **64**, 637-641.

Boldogh, I., Ramana, C.V., Chen, Z., Biswas, T., Hazra, T.K., Grosch, S., Grombacher, T., Mitra, S. and Kaina, B. (1998) Regulation of expression

of the DNA repair gene O^6-methylguanine-DNA methyltransferase via protein kinase C-mediated signaling. *Cancer Res*, **58**, 3950-3956.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*, **72**, 248-254.

Broomfield, S., Hryciw, T. and Xiao, W. (2001) DNA postreplication repair and mutagenesis in *Saccharomyces cerevisiae*. *Mutat Res*, **486**, 167-184.

Chaudhury, A.M. and Smith, G.R. (1985) Role of *Escherichia coli* RecBC enzyme in SOS induction. *Mol Gen Genet*, **201**, 525-528.

Chen, B., Carroll, P. and Samson, L. (1994) The *Escherichia coli* AlkB protein protects human cells against alkylation-induced toxicity. *J Bacteriol*, **176**, 6255-6261.

Chetsanga, C.J. and Lindahl, T. (1979) Release of 7-methylguanine residues whose imidazole rings have been opened from damaged DNA by a DNA glycosylase from *Escherichia coli*. *Nucleic Acids Res*, **6**, 3673-3684.

Colombi, D. and Gomes, S. (1997) An *alkB* gene homolog is differentially transcribed during the *Caulobacter crescentus* cell cycle. *J Bacteriol*, **179**, 3139-3145.

Cox, M.M. (2001) Recombinational DNA repair of damaged replication forks in *Escherichia coli*: questions. *Annu Rev Genet*, **35**, 53-82.

Daniels, D.S. and Tainer, J.A. (2000) Conserved structural motifs governing the stoichiometric repair of alkylated DNA by O^6-alkylguanine-DNA alkyltransferase. *Mutat Res*, **460**, 151-163.

Dianov, G., Sedgwick, B., Daly, G., Olsson, M., Lovett, S. and Lindahl, T. (1994) Release of 5'-terminal deoxyribose-phosphate residues from incised abasic sites in DNA by the *Escherichia coli* RecJ protein. *Nucleic Acids Res*, **22**, 993-998.

Dinglay, S., Gold, B. and Sedgwick, B. (1998) Repair in *Escherichia coli alkB* mutants of abasic sites and 3-methyladenine residues in DNA. *Mutat Res*, **407**, 109-116.

Dinglay, S., Trewick, S.C., Lindahl, T. and Sedgwick, B. (2000) Defective processing of methylated single-stranded DNA by *E. coli alkB* mutants. *Genes Dev*, **14**, 2097-2105.

Doetsch, P.W. and Cunningham, R.P. (1990) The enzymology of apurinic/apyrimidinic endonucleases. *Mutat Res*, **236**, 173-201.

Drohat, A.C., Kwon, K., Krosky, D.J. and Stivers, J.T. (2002) 3-Methyladenine DNA glycosylase I is an unexpected helix-hairpin-helix superfamily member. *Nat Struct Biol*, **9**, 659-664.

Dronkert, M.L. and Kanaar, R. (2001) Repair of DNA interstrand cross-links. *Mutat Res*, **486**, 217-247.

Duncan, T., Trewick, S.C., Koivisto, P., Bates, P.A., Lindahl, T. and Sedgwick, B. (2002) Reversal of DNA alkylation damage by two human dioxygenases. *Proc Natl Acad Sci USA*, **99**, 16660-16665.

Engel, J.D. (1975) Mechanism of the Dimroth rearrangement in adenosine. *Biochem Biophys Res Commun*, **64**, 581-586.

Eom, S.H., Wang, J. and Steitz, T.A. (1996) Structure of Taq polymerase with DNA at the polymerase active site. *Nature*, **382**, 278-281.

Evensen, G. and Seeberg, E. (1982) Adaptation to alkylation resistance involves the induction of a DNA glycosylase. *Nature*, **296**, 773-775.

Falnes, P., Johansen, R. and Seeberg, E. (2002) AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli. Nature*, **419**, 178-182.

Fernandez De Henestrosa, A.R., Ogi, T., Aoyagi, S., Chafin, D., Hayes, J.J., Ohmori, H. and Woodgate, R. (2000) Identification of additional genes

belonging to the LexA regulon in *Escherichia coli*. *Mol Microbiol*, **35**, 1560-1572.

Frosina, G., Fortini, P., Rossi, O., Carrozzino, F., Raspaglio, G., Cox, L.S., Lane, D.P., Abbondandolo, A. and Dogliotti, E. (1996) Two pathways for base excision repair in mammalian cells. *J Biol Chem*, **271**, 9573-9578.

Garcia-Santos Mdel, P., Calle, E. and Casado, J. (2001) Amino acid nitrosation products as alkylating agents. *J Am Chem Soc*, **123**, 7506-7510.

Gasparutto, D., Dherin, C., Boiteux, S. and Cadet, J. (2002) Excision of 8-methylguanine site-specifically incorporated into oligonucleotide substrates by the AlkA protein of *Escherichia coli*. *DNA Repair*, **1**, 437-447.

Gerson, S.L. (2002) Clinical relevance of MGMT in the treatment of cancer. *J Clin Oncol*, **20**, 2388-2399.

Goffin, C. and Verly, W.G. (1983) Interstrand DNA crosslinks due to AP (apurinic/apyrimidinic) sites. *FEBS Lett*, **161**, 140-144.

Goodman, M.F. (2000) Coping with replication 'train wrecks' in *Escherichia coli* using Pol V, Pol II and RecA proteins. *Trends Biochem Sci*, **25**, 189-195.

Goodtzova, K., Kanugula, S., Edara, S., Pauly, G.T., Moschel, R.C. and Pegg, A.E. (1997) Repair of O^6-benzylguanine by the *Escherichia coli* Ada and Ogt and the human O^6-alkylguanine-DNA alkyltransferases. *J Biol Chem*, **272**, 8332-8339.

Hecht, S.S., Abbaspour, A. and Hoffman, D. (1988) A study of tobacco carcinogenesis. XLII. Bioassay in A/J mice of some structural analogues of tobacco-specific nitrosamines. *Cancer Lett*, **42**, 141-145.

Hix, S., Morais Mda, S. and Augusto, O. (1995) DNA methylation by tert-butyl hydroperoxide-iron (II). *Free Radic Biol Med*, **19**, 293-301.

Hoege, C., Pfander, B., Moldovan, G.L., Pyrowolakis, G. and Jentsch, S. (2002) RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature*, **419**, 135-141.

Hollis, T., Ichikawa, Y. and Ellenberger, T. (2000) DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, *Escherichia coli* AlkA. *EMBO J*, **19**, 758-766.

Holme, E. (1975) A kinetic study of thymine 7-hydroxylase from *Neurospora crassa*. *Biochemistry*, **14**, 4999-5003.

Hsieh, P. (2001) Molecular mechanisms of DNA mismatch repair. *Mutat Res*, **486**, 71-87.

Jeggo, P., Defais, T.M., Samson, L. and Schendel, P. (1977) An adaptive response of *E. coli* to low levels of alkylating agent: comparison with previously characterised DNA repair pathways. *Mol Gen Genet*, **157**, 1-9.

Karki, S.B. and Dinnocenzo, J.P. (1995) On the mechanism of amine oxidations by P450. *Xenobiotica*, **25**, 711-724.

Karran, P. and Bignami, M. (1994) DNA damage tolerance, mismatch repair and genome instability. *Bioessays*, **16**, 833-839.

Karran, P., Hjelmgren, T. and Lindahl, T. (1982) Induction of a DNA glycosylase for N-methylated purines is part of the adaptive response to alkylating agents. *Nature*, **296**, 770-773.

Kataoka, H. and Sekiguchi, M. (1985) Molecular cloning and characterization of the *alkB* gene of *Escherichia coli*. *Mol Gen Genet*, **198**, 263-269.

Kataoka, H., Yamamoto, Y. and Sekiguchi, M. (1983) A new gene (*alkB*) of *Escherichia coli* that controls sensitivity to methyl methane sulfonate. *J Bacteriol*, **153**, 1301-1307.

Kivirikko, K.I., Myllyla, R. and Pihlajaniemi, T. (1989) Protein hydroxylation: prolyl 4-hydroxylase, an enzyme with four cosubstrates and a multifunctional subunit. *FASEB J*, **3**, 1609-1617.

Kivirikko, K.I. and Pihlajaniemi, T. (1998) Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases. *Adv Enzymol Relat Areas Mol Biol*, **72**, 325-398.

Kleibl, K. and Margison, G.P. (1998) Increasing DNA repair capacity in bone marrow by gene transfer as a prospective tool in cancer therapy. *Neoplasma*, **45**, 181-186.

Kondo, H., Nakabeppu, Y., Kataoka, H., Kuhara, S., Kawabata, S. and Sekiguchi, M. (1986) Structure and expression of the *alkB* gene of *Escherichia coli* related to the repair of alkylated DNA. *J Biol Chem*, **261**, 15772-15777.

Kowalczykowski, S.C. (2000) Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem Sci*, **25**, 156-165.

Labahn, J., Scharer, O.D., Long, A., Ezaz-Nikpay, K., Verdine, G.L. and Ellenberger, T.E. (1996) Structural basis for the excision repair of alkylation-damaged DNA. *Cell*, **86**, 321-329.

Lai, D.Y. and Arcos, J.C. (1980) Minireview: dialkylnitrosamine bioactivation and carcinogenesis. *Life Sci*, **27**, 2149-2165.

Landini, P., Hajec, L.I. and Volkert, M.R. (1994) Structure and transcriptional regulation of the *Escherichia coli* adaptive response gene *aidB*. *J Bacteriol*, **176**, 6583-6589.

Landini, P. and Volkert, M.R. (1995) RNA polymerase alpha subunit binding site in positively controlled promoters: a new model for RNA polymerase-promoter interaction and transcriptional activation in the *Escherichia coli ada* and *aidB* genes. *EMBO J*, **14**, 4329-4335.

Larson, K., Sahm, J., Shenkar, R. and Strauss, B. (1985) Methylation-induced blocks to in vitro DNA replication. *Mutat Res*, **150**, 77-84.

Lau, A.Y., Scharer, O.D., Samson, L., Verdine, G.L. and Ellenberger, T. (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. *Cell*, **95**, 249-258.

Lawley, P. (1984) Carcinogenesis by Alkylating Agents. In Searle, C. (ed.), *Chemical Carcinogenesis*. American Chemical Society, Washington, D.C., Vol. 1.

Lehmann, A.R. (2002) Replication of damaged DNA in mammalian cells: new solutions to an old problem. *Mutat Res*, **509**, 23-34.

Lindahl, T. (1979) DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and base excision-repair. *Prog Nucleic Acid Res Mol Biol*, **22**, 135-192.

Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709-715.

Lindahl, T. (2001) Keynote: past, present, and future aspects of base excision repair. *Prog Nucleic Acid Res Mol Biol*, **68**, xvii-xxx.

Lindahl, T., Demple, B. and Robins, P. (1982) Suicide inactivation of the *E. coli* O6-methylguanine-DNA methyltransferase. *EMBO J*, **1**, 1359-1363.

Lindahl, T. and Wood, R.D. (1999) Quality control by DNA repair. *Science*, **286**, 1897-1905.

Lizcano, J.M., Unzeta, M. and Tipton, K.F. (2000) A spectrophotometric method for determining the oxidative deamination of methylamine by the amine oxidases. *Anal Biochem*, **286**, 75-79.

Loechler, E.L., Green, C.L. and Essigmann, J.M. (1984) In vivo mutagenesis by O^6-methylguanine built into a unique site in a viral genome. *Proc Natl Acad Sci USA*, **81**, 6271-6275.

Ma, T.H. and Harris, M.M. (1988) Review of the genotoxicity of formaldehyde. *Mutat Res*, **196**, 37-59.

Margison, G.P., Cooper, D.P. and Potter, P.M. (1990) The *E. coli* ogt gene. *Mutat Res*, **233**, 15-21.

Margison, G.P. and Santibanez-Koref, M.F. (2002) O^6-alkylguanine-DNA alkyltransferase: role in carcinogenesis and chemotherapy. *Bioessays*, **24**, 255-266.

Matsumoto, Y. and Kim, K. (1995) Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair. *Science*, **269**, 699-702.

Mayer, W., Niveleau, A., Walter, J., Fundele, R. and Haaf, T. (2000) Demethylation of the zygotic paternal genome. *Nature*, **403**, 501-502.

McCarthy, T.V. and Lindahl, T. (1985) Methyl phosphotriesters in alkylated DNA are repaired by the Ada regulatory protein of *E. coli*. *Nucleic Acids Res*, **13**, 2683-2698.

McGlynn, P. and Lloyd, R.G. (2002) Recombinational repair and restart of damaged replication forks. *Nat Rev Mol Cell Biol*, **3**, 859-870.

McHugh, P.J., Spanswick, V.J. and Hartley, J.A. (2001) Repair of DNA interstrand crosslinks: molecular mechanisms and clinical relevance. *Lancet Oncol*, **2**, 483-490.

Monsees, T., Meijer, L. and Jastorff, B. (1993) 8-Dimethylamino-1-methyladenine, a novel potent antagonist of the 1-methyladenine receptor in starfish oocytes. *Eur J Biochem*, **213**, 155-165.

Myers, L.C., Cushing, T.D., Wagner, G. and Verdine, G.L. (1994) Metal-coordination sphere in the methylated Ada protein-DNA co-complex. *Chem Biol*, **1**, 91-97.

Nash, H.M., Bruner, S.D., Scharer, O.D., Kawate, T., Addona, T.A., Spooner, E., Lane, W.S. and Verdine, G.L. (1996) Cloning of a yeast 8-oxoguanine DNA glycosylase reveals the existence of a base-excision DNA-repair protein superfamily. *Curr Biol*, **6**, 968-980.

O'Connor, T.R., Boiteux, S. and Laval, J. (1988) Ring-opened 7-methylguanine residues in DNA are a block to in vitro DNA synthesis. *Nucleic Acids Res*, **16**, 5879-5894.

Olsson, M. and Lindahl, T. (1980) Repair of alkylated DNA in *Escherichia coli.* Methyl group transfer from O^6-methylguanine to a protein cysteine residue. *J Biol Chem*, **255**, 10569-10571.

Orren, D.K. and Sancar, A. (1989) The (A)BC excinuclease of *Escherichia coli* has only the UvrB and UvrC subunits in the incision complex. *Proc Natl Acad Sci USA*, **86**, 5237-5241.

Pomposiello, P.J. and Demple, B. (2001) Redox-operated genetic switches: the SoxR and OxyR transcription factors. *Trends Biotechnol*, **19**, 109-114.

Posnick, L. and Samson, L. (1999a) Influence of S-Adenosylmethionine pool size on spontaneous mutation, Dam methylation, and cell growth of *Escherichia coli. J Bacteriol*, **181**, 6756-6762.

Posnick, L.M. and Samson, L.D. (1999b) Imbalanced base excision repair increases spontaneous mutation and alkylation sensitivity in *Escherichia coli. J Bacteriol*, **181**, 6763-6771.

Que, L., Jr. (2000) One motif--many different reactions. *Nat Struct Biol*, **7**, 182-184.

Que, L., Jr. and Ho, R.Y. (1996) Dioxygen Activation by Enzymes with Mononuclear Non-Heme Iron Active Sites. *Chem Rev*, **96**, 2607-2624.

Ramotar, D. (1997) The apurinic-apyrimidinic endonuclease IV family of DNA repair enzymes. *Biochem Cell Biol*, **75**, 327-336.

Rebeck, G.W. and Samson, L. (1991) Increased spontaneous mutation and alkylation sensitivity of *Escherichia coli* strains lacking the ogt O6-methylguanine DNA repair methyltransferase. *J Bacteriol*, **173**, 2068-2076.

Rebeil, R. and Nicholson, W.L. (2001) The subunit structure and catalytic mechanism of the *Bacillus subtilis* DNA repair enzyme spore photoproduct lyase. *Proc Natl Acad Sci USA*, **98**, 9038-9043.

Renda, M.J., Rosenblatt, J.D., Klimatcheva, E., Demeter, L.M., Bambara, R.A. and Planelles, V. (2001) Mutation of the methylated tRNA(Lys)(3) residue A58 disrupts reverse transcription and inhibits replication of human immunodeficiency virus type 1. *J Virol*, **75**, 9671-9678.

Reuven, N.B., Arad, G., Maor-Shoshani, A. and Livneh, Z. (1999) The mutagenesis protein UmuC is a DNA polymerase activated by UmuD', RecA, and SSB and is specialized for translesion replication. *J Biol Chem*, **274**, 31763-31766.

Riazuddin, S. and Lindahl, T. (1978) Properties of 3-methyladenine-DNA glycosylase from *Escherichia coli*. *Biochemistry*, **17**, 2110-2118.

Roach, P.L., Clifton, I.J., Hensgens, C.M., Shibata, N., Schofield, C.J., Hajdu, J. and Baldwin, J.E. (1997) Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation. *Nature*, **387**, 827-830.

Rohde, J.U., In, J.H., Lim, M.H., Brennessel, W.W., Bukowski, M.R., Stubna, A., Munck, E., Nam, W. and Que, L., Jr. (2003) Crystallographic and spectroscopic characterization of a nonheme Fe(IV)-O complex. *Science*, **299**, 1037-1039.

Rydberg, B. and Lindahl, T. (1982) Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J*, **1**, 211-216.

Saffhill, R. (1984) Differences in the promutagenic nature of 3-methylcytosine as revealed by DNA and RNA polymerising enzymes. *Carcinogenesis*, **5**, 691-693.

Saffhill, R. and Hall, J.A. (1985) The incorporation of O^6-methyldeoxyguanosine monophosphate and O^4-methyldeoxythymidine monophosphate into polynucleotide templates leads to errors during subsequent in vitro DNA synthesis. *Chem Biol Interact*, **56**, 363-370.

Salowe, S.P., Marsh, E.N. and Townsend, C.A. (1990) Purification and characterization of clavaminate synthase from *Streptomyces clavuligerus*:

an unusual oxidative enzyme in natural product biosynthesis. *Biochemistry*, **29**, 6499-6508.

Sambrook, J., Fritsch, E. and Maniatis, T. (1989) *Molecular cloning: A laboratory manual.* Cold Spring Harbour Laboratory Press, New York.

Samson, L. and Cairns, J. (1977) A new pathway for DNA repair in *Escherichia coli*. *Nature*, **267**, 281-283.

Samson, L., Han, S., Marquis, J.C. and Rasmussen, L.J. (1997) Mammalian DNA repair methyltransferases shield O^4-MeT from nucleotide excision repair. *Carcinogenesis*, **18**, 919-924.

Saparbaev, M., Kleibl, K. and Laval, J. (1995) *Escherichia coli*, *Saccharomyces cerevisiae*, rat and human 3-methyladenine DNA glycosylases repair 1,N6-ethenoadenine when present in DNA. *Nucleic Acids Res*, **23**, 3750-3755.

Saparbaev, M. and Laval, J. (1994) Excision of hypoxanthine from DNA containing dIMP residues by the *Escherichia coli*, yeast, rat, and human alkylpurine DNA glycosylases. *Proc Natl Acad Sci USA*, **91**, 5873-5877.

Sedgwick, B. (1982) Genetic mapping of *ada* and *adc* mutations affecting the adaptive response of *Escherichia coli* to alkylating agents. *J Bacteriol*, **150**, 984-988.

Sedgwick, B. (1997) Nitrosated peptides and polyamines as endogenous mutagens in O^6-alkylguanine-DNA alkyltransferase deficient cells. *Carcinogenesis*, **18**, 1561-1567.

Sedgwick, B. and Lindahl, T. (2002) Recent progress on the Ada response for inducible repair of DNA alkylation damage. *Oncogene*, **21**, 8886-8894.

Selby, C.P. and Sancar, A. (1993) Molecular mechanism of transcription-repair coupling. *Science*, **260**, 53-58.

Shuker, D.E. and Margison, G.P. (1997) Nitrosated glycine derivatives as a potential source of O^6-methylguanine in DNA. *Cancer Res*, **57**, 366-369.

Singer, B. (1975) The chemical effects of nucleic acid alkylation and their relation to mutagenesis and carcinogenesis. *Prog Nucleic Acid Res Mol Biol*, **15**, 219-284.

Singer, B. and Grunberger, D. (1983) *Molecular Biology of Mutagens and Carcinogens*. Plenum Press, New York.

Singer, B. and Hang, B. (1997) What structural features determine repair enzyme specificity and mechanism in chemically modified DNA? *Chem Res Toxicol*, **10**, 713-732.

Singer, B., Pergolizzi, R.G. and Grunberger, D. (1979) Synthesis and coding properties of dinucleoside diphosphates containing alky pyrimidines which are formed by the action of carcinogens on nucleic acids. *Nucleic Acids Res*, **6**, 1709-1719.

Smith, S.S. (2000) Gilbert's conjecture: the search for DNA (cytosine-5) demethylases and the emergence of new functions for eukaryotic DNA (cytosine-5) methyltransferases. *J Mol Biol*, **302**, 1-7.

Srivenugopal, K.S., Yuan, X.H., Friedman, H.S. and Ali-Osman, F. (1996) Ubiquitination-dependent proteolysis of O^6-methylguanine-DNA methyltransferase in human and murine tumor cells following inactivation with O^6-benzylguanine or 1,3-bis(2-chloroethyl)-1-nitrosourea. *Biochemistry*, **35**, 1328-1334.

Sutton, M.D., Smith, B.T., Godoy, V.G. and Walker, G.C. (2000) The SOS response: recent insights into umuDC-dependent mutagenesis and DNA damage tolerance. *Annu Rev Genet*, **34**, 479-497.

Svejstrup, J.Q., Vichi, P. and Egly, J.M. (1996) The multiple roles of transcription/repair factor TFIIH. *Trends Biochem Sci*, **21**, 346-350.

Swain, C. and Scott, C. (1953) Quantative correlation of relative rates. Comparison of hydroxide ion with other nucleophilic reagents towards alkyl halides, esters, epoxides and acyl halides. *J. Am. Chem. Soc*, **75**, 141-147.

Takahashi, K., Kawazoe, Y., Sakumi, K., Nakabeppu, Y. and Sekiguchi, M. (1988) Activation of Ada protein as a transcriptional regulator by direct alkylation with methylating agents. *J Biol Chem*, **263**, 13490-13492.

Tang, M., Shen, X., Frank, E.G., O'Donnell, M., Woodgate, R. and Goodman, M.F. (1999) UmuD'(2)C is an error-prone DNA polymerase, *Escherichia coli* pol V. *Proc Natl Acad Sci USA*, **96**, 8919-8924.

Taverna, P. and Sedgwick, B. (1996) Generation of an endogenous DNA-methylating agent by nitrosation in *Escherichia coli. J Bacteriol*, **178**, 5105-5111.

Tchou, J., Kasai, H., Shibutani, S., Chung, M.H., Laval, J., Grollman, A.P. and Nishimura, S. (1991) 8-oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity. *Proc Natl Acad Sci USA*, **88**, 4690-4694.

Teo, I., Sedgwick, B., Kilpatrick, M., McCarthy, T. and Lindahl, T. (1986) The intracellular signal for induction of resistance to alkylating agents in *E. coli*. *Cell*, **45**, 315-324.

Thompson, C.L. and Sancar, A. (2002) Photolyase/cryptochrome blue-light photoreceptors use photon energy to repair DNA and reset the circadian clock. *Oncogene*, **21**, 9043-9056.

Thornburg, L.D., Lai, M.T., Wishnok, J.S. and Stubbe, J. (1993) A non-heme iron protein with heme tendencies: an investigation of the substrate specificity of thymine hydroxylase. *Biochemistry*, **32**, 14023-14033.

Tominaga, Y., Tsuzuki, T., Shiraishi, A., Kawate, H. and Sekiguchi, M. (1997) Alkylation-induced apoptosis of embryonic stem cells in which the gene for DNA-repair, methyltransferase, had been disrupted by gene targeting. *Carcinogenesis*, **18**, 889-896.

Valegard, K., van Scheltinga, A.C., Lloyd, M.D., Hara, T., Ramaswamy, S., Perrakis, A., Thompson, A., Lee, H.J., Baldwin, J.E., Schofield, C.J., Hajdu, J. and Andersson, I. (1998) Structure of a cephalosporin synthase. *Nature*, **394**, 805-809.

van den Boom, V., Jaspers, N.G. and Vermeulen, W. (2002) When machines get stuck--obstructed RNA polymerase II: displacement, degradation or suicide. *Bioessays*, **24**, 780-784.

van der Kemp, P.A., Thomas, D., Barbey, R., de Oliveira, R. and Boiteux, S. (1996) Cloning and expression in *Escherichia coli* of the OGG1 gene of *Saccharomyces cerevisiae*, which codes for a DNA glycosylase that excises 7,8-dihydro-8-oxoguanine and 2,6-diamino-4-hydroxy-5-N-methylformamidopyrimidine. *Proc Natl Acad Sci USA*, **93**, 5197-5202.

Vaughan, P., Sedgwick, B., Hall, J., Gannon, J. and Lindahl, T. (1991) Environmental mutagens that induce the adaptive response to alkylating agents in *Escherichia coli. Carcinogenesis*, **12**, 263-268.

Verhoeven, E.E., van Kesteren, M., Moolenaar, G.F., Visse, R. and Goosen, N. (2000) Catalytic sites for 3' and 5' incision of *Escherichia coli* nucleotide excision repair are both located in UvrC. *J Biol Chem*, **275**, 5120-5123.

Volker, M., Mone, M.J., Karmakar, P., van Hoffen, A., Schul, W., Vermeulen, W., Hoeijmakers, J.H., van Driel, R., van Zeeland, A.A. and Mullenders, L.H. (2001) Sequential assembly of the nucleotide excision repair factors *in vivo*. *Mol Cell*, **8**, 213-224.

Volkert, M. and Hajec, L. (1991) Molecular analysis of the *aidD6::Mu dl* (*bla lac*) fusion mutation of *Escherichia coli* K12. *Mol Gen Genet*, **229**, 319-323.

Vora, R.A., Pegg, A.E. and Ealick, S.E. (1998) A new model for how O^6-methylguanine-DNA methyltransferase binds DNA. *Proteins*, **32**, 3-6.

Wei, Y., Carter, K., Wang, R. and Shell, B. (1996) Molecular cloning and functional analysis of a human cDNA encoding an *Escherichia coli* AlkB homolog, a protein involved in DNA alkylation damage repair. *Nucleic Acids Res*, **24**, 931-937.

Wei, Y., Chen, B. and Samson, L. (1995) Suppression of *Escherichia coli alkB* mutants by *Saccharomyces cerevisiae* genes. *J Bacteriol*, **177**, 5009-5015.

Wood, R.D. (1997) Nucleotide excision repair in mammalian cells. *J Biol Chem*, **272**, 23465-23468.

Wood, R.D., Mitchell, M., Sgouros, J. and Lindahl, T. (2001) Human DNA repair genes. *Science*, **291**, 1284-1289.

Wuosmaa, A.M. and Hager, L.P. (1990) Methyl chloride transferase: a carbocation route for biosynthesis of halometabolites. *Science*, **249**, 160-162.

Xiao, W. and Samson, L. (1993) *In vivo* evidence for endogenous DNA alkylation damage as a source of spontaneous mutation in eukaryotic cells. *Proc Natl Acad Sci USA*, **90**, 2117-2121.

Zhang, Z., Ren, J., Stammers, D.K., Baldwin, J.E., Harlos, K. and Schofield, C.J. (2000) Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase. *Nat Struct Biol*, **7**, 127-133.

Defective processing of methylated single-stranded DNA by *E. coli alkB* mutants

Suneet Dinglay, Sarah C. Trewick, Tomas Lindahl, and Barbara Sedgwick[1]

Imperial Cancer Research Fund, Clare Hall Laboratories, South Mimms, Hertfordshire EN6 3LD, UK

Escherichia coli alkB **mutants are very sensitive to DNA methylating agents. Despite these mutants being the subject of many studies, no DNA repair or other function has been assigned to the AlkB protein or to its human homolog. Here, we report that reactivation of methylmethanesulfonate (MMS)-treated single-stranded DNA phages, M13, f1, and G4, was decreased dramatically in *alkB* mutants. No such decrease occurred when using methylated λ phage or M13 duplex DNA. These data show that *alkB* mutants have a marked defect in processing methylation damage in single-stranded DNA. Recombinant AlkB protein bound more efficiently to single- than double-stranded DNA. The single-strand damage processed by AlkB was primarily cytotoxic and not mutagenic and was induced by SN2 methylating agents, MMS, DMS, and MeI but not by SN1 agent *N*-methyl-*N*-nitrosourea or by γ irradiation. Strains lacking other DNA repair activities, *alkA tag, xth nfo, uvrA, mutS,* and *umuC,* were not defective in reactivation of methylated M13 phage and did not enhance the defect of an *alkB* mutant. A *recA* mutation caused a small but additive defect. Thus, AlkB functions in a novel pathway independent of these activities. We propose that AlkB acts on alkylated single-stranded DNA in replication forks or at transcribed regions. Consistent with this theory, stationary phase *alkB* cells were less MMS sensitive than rapidly growing cells.**

[*Key Words*: DNA repair; DNA alkylation; AlkB]

Received March 27, 2000; revised version accepted June 15, 2000.

Alkylating agents arise endogenously in cells and also occur widely in the environment (Rebeck and Samson 1991; Vaughan et al. 1991; Taverna and Sedgwick 1996). As a consequence, cells need protection against such compounds, which is provided by activities that specifically remove alkylation lesions from DNA. Inducible resistance of *Escherichia coli* to the cytotoxic and mutagenic effects of simple alkylating agents involves the increased expression of the *ada, alkA,* and *alkB* genes (Lindahl et al. 1988). The functions of the Ada and AlkA proteins have been studied in detail, whereas that of AlkB remains unclear. Ada, a multifunctional protein, directly demethylates O^6-methylguanine and methylphosphotriesters in DNA by transferring methyl groups onto two of its own cysteine residues. It also positively regulates the adaptive response using *S*-diastereoisomers of methylphosphotriesters as the inducing signal (Lindahl et al. 1988). AlkA is a 3-methyladenine-DNA glycosylase and excises the toxic lesion 3-methyladenine from DNA. It can also excise other altered bases, such as hypoxanthine and N^6-ethenoadenine (Matijasevic et al. 1992; Saparbaev and Laval 1994). The resulting abasic sites are repaired by the base excision repair pathway (Lindahl et al. 1997). O^6-methylguanine-DNA methyltransferases and 3-methyladenine-DNA glycosylases are conserved in prokaryotes and eukaryotes (Pegg et al. 1995). An additional *E. coli* function, AidB, is induced by high concentrations of alkylating agents and is possibly involved in inactivation of certain alkylating agents (Landini et al. 1994).

Conservation of AlkB protein from bacteria to humans indicates its importance for cellular defence against alkylating agents (Wei et al. 1996), but its function remains elusive despite its identification in 1983 (Kataoka et al. 1983). The *alkB* gene forms a small operon with *ada* and is regulated from the *ada* promoter (Lindahl et al. 1988). The AlkB protein prevents death from cells' exposure to methylmethanesulfonate (MMS) and dimethylsulphate (DMS) but is less effective in protection against N-methyl-N'-nitro-N-nitroguanidine (MNNG) and *N*-methyl-*N*-nitrosourea (MNU; Kataoka et al. 1983; Chen et al. 1994). A small defect in the reactivation of MMS-treated λ bacteriophage in an *alkB* mutant suggests a role for AlkB in DNA repair (Kataoka et al. 1983), but the mechanism is unknown. *AlkB* mutants are not defective in the repair of several different types of potentially toxic lesions that may be generated by methylating agents in duplex DNA. These lesions include 3-methyladenine, DNA strand breaks, abasic sites, and secondary lesions that may arise at abasic sites such as DNA–protein cross-links and DNA interstrand cross-links (Dinglay et al. 1998). Purified AlkB protein is devoid of detectable DNA glycosylase, DNA methyltransferase,

[1]Corresponding author.
E-MAIL b.sedgwick@icrf.icnet.uk; FAX 171-269-3801

nuclease, or DNA-dependent ATPase activity in standard enzyme assays (Kondo et al. 1986) and has no sequence similarity to other proteins of known function in the databases. Homologs of AlkB have been identified in *Homo sapiens* and *Caulobacter crescentus* (Wei et al. 1996; Colombi and Gomes 1997), and recent database searches reveal a wide distribution of other putative AlkB homolog through evolution (data not shown). Overexpression of the *E. coli* AlkB protein confers MMS resistance to human cells (Chen et al. 1994), and conversely, the human protein confers alkylation resistance to *E. coli alkB* mutants (Wei et al. 1996), suggesting that AlkB proteins act independently and not via formation of multiprotein complexes. Expression of the *C. crescentus alkB* gene is not induced by alkylation damage but is cell-cycle regulated with a pattern similar to activities required for DNA replication (Colombi and Gomes 1997).

In this article, we describe a substantial defect in the reactivation of MMS-treated single-stranded DNA phages in *alkB* mutants and show that AlkB protein is required to process toxic DNA damage induced in single-stranded DNA by SN2 methylating agents.

Results

AlkB processes methylated single-stranded DNA

AlkB mutants are sensitive to killing by MMS but only marginally sensitive to MNNG. They have a small defect in the reactivation of MMS-treated λ phage, indicating a defect in DNA repair (Kataoka et al. 1983). Differences in the known spectra of methylated bases induced by MMS and MNNG were considered as a possible explanation for the *alkB* phenotype. The sites methylated by MMS in duplex DNA are also modified by MNNG, whereas in single-stranded DNA some sites are more reactive with MMS than with MNNG (Singer and Grunberger 1983). To examine the possibility that the AlkB protein processes damage induced in single-stranded DNA, reactivation of MMS-treated M13 phage was monitored in an *alkB117::Tn3* mutant. Survival of the methylated phage was strikingly low in the *alkB* mutant. The lethal MMS dose resulting in 10% M13 survival (LD10) was fourfold lower for the *alkB* mutant than for the wild type (Fig. 1A). The survival of untreated phage was the same in both strains. Similar observations were made using two other single-stranded DNA phages, f1 and G4, when they were treated with MMS and transfected into *alkB117::Tn3* mutants (Fig. 1B,C), whereas no similar defect was apparent in the reactivation of MMS-treated λ, a double-stranded DNA phage (Fig. 1D). The pronounced defect in reactivation of MMS-treated M13 was also observed in a second *alkB* mutant, HK82 (*alkB22*; data not shown). These observations indicate that the AlkB protein is required specifically to process damaged single-stranded DNA or lesions formed more frequently in single strands but recognized in both single or duplex DNA.

Instead of using intact phage, purified M13 DNA in its duplex or single-stranded form was treated with MMS, transformed by heat shock into wild type and *alkB117::Tn3* strains and plaque-forming units were monitored. The transformation efficiency of MMS-treated single-stranded DNA was markedly less in the *alkB* mutant than in the wild type, the LD50 being fivefold less in the *alkB* mutant (Fig. 2B). In contrast, double-stranded M13 DNA treated with up to 100 mM MMS transformed wild type and *alkB* strains with equal frequencies and decreased by less than twofold in both strains (Fig. 2A). These observations confirmed that AlkB is required to process methylation lesions in single-stranded DNA.

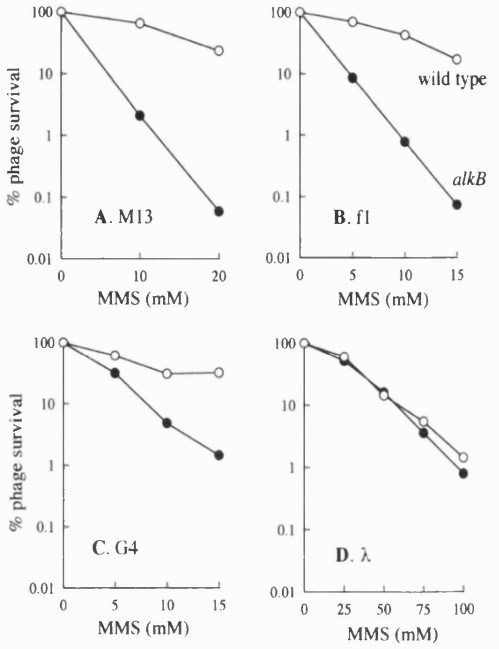


Figure 1. Defective reactivation of MMS-treated single-stranded DNA phages in an *alkB* mutant. Phages M13, f1, and G4 were treated with various doses of MMS at 30°C for 30 min and immediately plated to estimate survival in wild type (○) and *alkB117::Tn3* (●) strains. Double-stranded DNA phage λ was similarly treated but at 37°C. (*A*) λ phage were plated on AB1157 (wild type) and BS87 (*alkB117::Tn3*); (*B*) M13 and (*C*) f1 phage, on AB1157/F′ and BS87/F′; (*D*) G4 phage on *Escherichia coli* C-1 and BS159 (*E. coli* C-1 but *alkB117::Tn3*).

AlkB preferentially binds to single-stranded DNA

To tag the AlkB protein at its amino terminus with six histidines, the *alkB* gene was subcloned into a pET15b vector (Studier et al. 1990). Expression of the subcloned gene was IPTG (isopropyl β-D-thiogalactoside) inducible. The new plasmid construct, pBAR54, complemented MMS sensitivity of an *alkB* mutant, demonstrating that the his-tagged AlkB protein was active in vivo (data not shown). The his-tagged protein was purified by Ni-NTA-agarose column chromatography (Fig. 3A), and its binding affinities to single-stranded and duplex DNA in nonmethylated and methylated forms were compared.

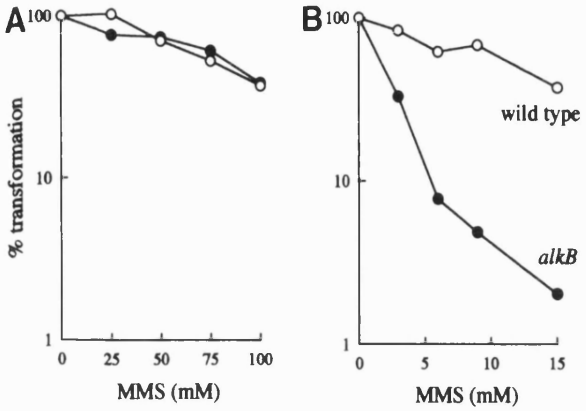


Figure 2. Decreased frequency of transformation by MMS treated single-stranded M13 DNA in an *alkB* mutant. M13 DNA, 20 ng double stranded or 100 ng single stranded, treated with MMS at various concentrations at 30°C for 30 min was transformed into AB1157 (wild type; ○) and BS87 (*alkB117::Tn3*; ●) strains. (*A*) Double-stranded M13 DNA; (*B*) Single-stranded M13 DNA.

The purified protein was incubated with 5′-^{32}P end-labeled 40-mer oligonucleotides, and binding was monitored by nitrocellulose filter binding assays. AlkB protein bound to both single- and double-stranded DNA but showed a much greater affinity for single-stranded DNA. Preferential binding of AlkB to single-stranded DNA was also confirmed using a gel-shift assay (Ausubel et al. 1999; data not shown). Pretreatment of the single- and double-stranded substrates with a high dose of MMS (300 mM) increased the AlkB binding affinity by approximately twofold in both cases (Fig. 3B). However, a similar increase of approximately 2.5-fold was also observed on pretreatment of the single-stranded DNA with 300 mM MNU (data not shown). AlkB mutants are not especially sensitive to MNU (Kataoka et al. 1983), so the stimulation by high doses of these two methylating agents may reflect altered structural properties of the heavily alkylated DNA rather than a binding to a specific lesion processed by AlkB.

AlkB processes DNA damage induced by SN2 methylating agents

SN1 and SN2 alkylating agents react through unimolecular and bimolecular pathways of nucleophilic substitution, respectively. AlkB mutants are sensitive to SN2 methylating agents, MMS and DMS, but much less sensitive to SN1 agents, MNNG and MNU (Kataoka et al. 1983; Chen et al. 1994). To ascertain whether this characteristic also applies to the survival of single-stranded DNA phage in an *alkB* mutant, reactivation of M13 after treatment with DMS, methyl iodide (MeI, also an SN2 agent), MNU, or γ rays was examined in AB1157/F′ (wild type) and BS87/F′ (*alkB117::Tn3*) strains. After exposure to DMS or MeI, M13 survival was much lower in the *alkB* mutant compared with the wild type strain, whereas after treatment with MNU or γ rays, survival decreased similarly in both strains (Fig. 4). LD10 of DMS was fivefold lower and LD50 of MeI sevenfold lower in the *alkB* mutant. Thus, damage in single-stranded DNA processed by the AlkB protein is induced specifically by the SN2 agents MMS, DMS, and MeI but not by MNU or γ rays.

AlkB function is independent of other DNA repair pathways

AlkA and Tag are 3-methyladenine-DNA glycosylases that repair the toxic lesion 3-methyladenine. To determine whether these activities influence survival of damaged single-stranded DNA, M13 phage were treated with MMS and their survival was assayed in an *alkA tag* mutant. This mutant was not defective in reactivating methylated M13 phage, and an *alkA tag* Δ*(ada-alkB)* mutant was no more deficient than the single *alkB* mutant (Fig. 5A). In contrast, the *alkA tag* mutant had a striking defect in reactivation of MMS-treated λ phage,

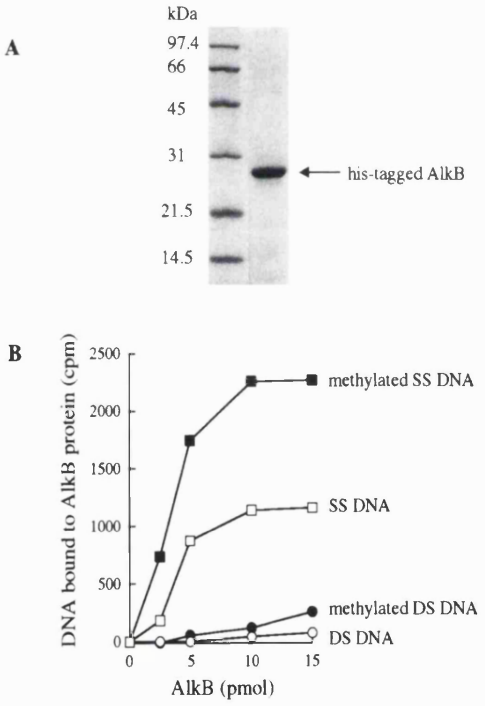


Figure 3. Binding of AlkB to DNA. (*A*) His-tagged AlkB protein was purified by Ni-NTA-agarose column chromatography and visualized by SDS–polyacrylamide gel electrophoresis and Coomassie blue staining. Sizes of molecular weight markers (kD) are indicated. (*B*) ^{32}P-5′-end labeled single- or double-stranded 40-mer oligonucleotides were methylated by treatment with 300 mM MMS. Various amounts of his-tagged AlkB protein were incubated with these substrates (30,000 cpm/reaction) at 30°C for 30 min. Reaction mixtures were passed through nitrocellulose filters and DNA bound to retained AlkB protein quantitated by scintillation counting. The substrates were (□) single-stranded DNA; (■) methylated single-stranded DNA; (○) double-stranded DNA; (●) methylated double-stranded DNA.

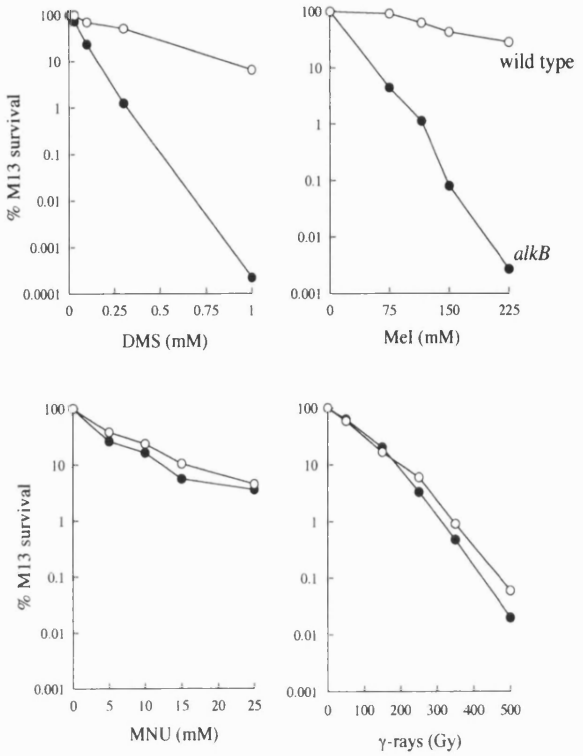


Figure 4. Defective reactivation of DMS and MeI (but not MNU and γ ray) treated M13 phage in an *alkB* mutant. M13 phage were treated with various doses of DMS, MeI, or MNU at 30°C for 30 min or with γ rays (2.71 Gy/min) for various times and immediately plated to estimate survival in AB1157/F′ (wild type; ○) and BS87/F′ (*alkB117::Tn3*; ●) strains.

whereas an *alkB* mutant showed no defect (Fig. 5B). Reactivation of MMS-treated M13 phage was also not defective in *xth nfo* double mutants lacking apurinic endonucleases or in *umuC*, *uvrA*, or *mutS* mutants defective in error-prone replication, nucleotide excision repair, or mismatch repair (data not shown). A *recA* mutant showed a small reproducible defect in reactivation of methylated M13 phage, and a *recA alkB* double mutant had a slightly greater defect than an *alkB* single mutant. The *recA* and *alkB* mutant defects were therefore additive, indicating that the two activities work independently (Fig. 5C).

Processing of mutagenic DNA damage by AlkB

The effect of AlkB activity on the spectrum of base substitutions induced by MMS was examined. Initially, the frequency of *lacZ* mutations arising in MMS-treated M13mp18 was analyzed after transfection of F′/wild-type and F′/*alkB* strains. The mutation frequencies were low (in the range of 10^{-4}–10^{-5}) but slightly higher in the *alkB* mutant than in the wild type (data not shown). With the aim of increasing the frequency of base substitution mutations, the SOS response and error-prone replication were induced by direct treatment of cells with MMS (Schendel and Defais 1980; Banerjee et al. 1990). Six F′*lacZ*$^-$/Δ*lac* strains (CC101–CC106) that revert to F′*lacZ*$^+$/Δ*lac*, each by different targeted base substitution mutations, were used (Cupples and Miller 1989). Small but reproducible increased frequencies of G:C to A:T, G:C to T:A, and A:T to T:A base substitutions were observed in *alkB117::Tn3* derivatives of CC102, CC104, and CC105, respectively, compared with the relevant wild-type strains (Fig. 6). Other types of base substitutions in *alkB* derivatives of CC101, CC103, and CC106 were not detected (data not shown). *Ada ogt* mutants are sensitive to induction of GC to AT transition mutations by DNA methylating agents (Mackay et al. 1994). The *alkB* mutants were only weakly sensitive to MMS mutagenesis compared with CC102 Δ(*ada-alkB*) *ogt* (Fig. 5).

AlkB mutants in stationary phase are less sensitive to MMS

Stationary phase cells have fewer DNA replication forks (Kornberg and Baker 1992) and are less active in transcription than rapidly growing cells and may, therefore, contain fewer regions of single-stranded DNA. Consequently, *alkB* cells deficient in processing damaged single-stranded DNA may be less sensitive to MMS in

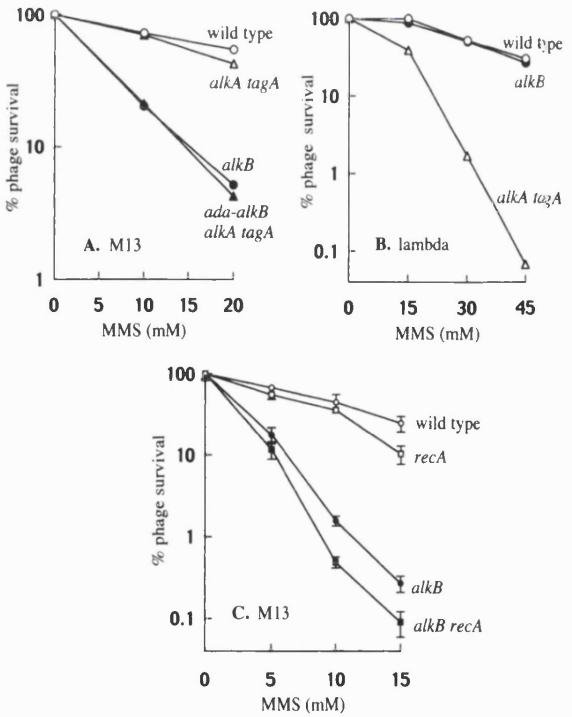


Figure 5. Survival of MMS treated M13 and λ phages in *alkA tag* and *recA* mutants. The phage were treated with increasing doses of MMS for 30 min at 23°C (*A*) or 30°C (*B*,*C*) and immediately plated on various strains. (*A*) M13 transfection of: (○) AB1157/F′ (wild type); (●) BS87/F′ (*alkB117::Tn3*); (△) GC4803/F′ (*alkA1 tagA1*); (▲) BS122/F′ (Δ*ada-alkB25*::Cam^r *alkA1 tagA1*). (*B*) λ Transfection of: (○) AB1157 (wild type); (●) BS87 (*alkB117::Tn3*); (△) GC4803 (*alkA1 tagA1*). (*C*) M13 transfection of: (○) AB1157/F′ (wild type); (●) BS87/F′ (*alkB117::Tn3*); (□) SD4/F′ (Δ*recA*); (■) SD5/F′ (Δ*recA alkB117::Tn3*).

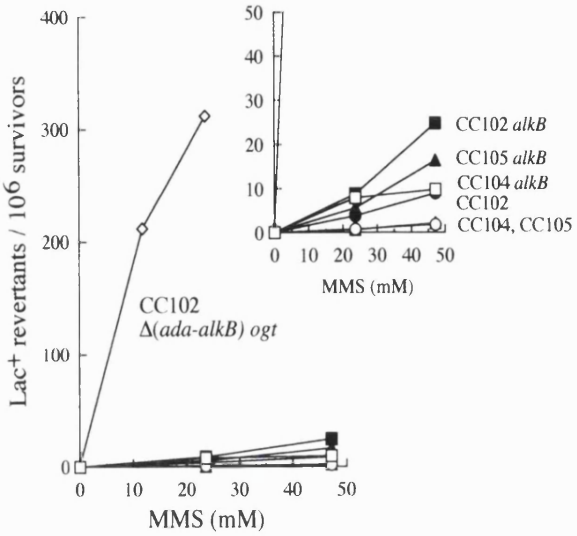


Figure 6. MMS mutagenesis of *E. coli alkB* mutants. CC101–CC106 (F'*lacZ*⁻/Δ*lacZ*) and their *alkB* derivatives were treated with increasing concentrations of MMS at 37°C for 20 min and immediately plated to monitor Lac⁺ revertants and survivors. (□) CC102; (■) CC102 *alkB117::Tn3*; (○) CC104; (●) CC104 *alkB117::Tn3*; (△) CC105; (▲) CC105 *alkB117::Tn3*; (◇) CC102 Δ*(ada-alkB25)*::Cam^r *ogt*-1::Kan^r. The same data are displayed on two different scales.

stationary phase than during exponential growth. As expected, exponentially proliferating *alkB* cells were much more sensitive to MMS than wild-type cells growing at a similar rate. The MMS sensitivity of *alkB* cells was significantly reduced when in stationary phase, whereas wild type stationary and exponential cells had only a small difference in sensitivity (Fig. 7A). This latter observation indicated that uptake or reactivity of MMS was not dramatically reduced in stationary phase and so was not the reason for decreased sensitivity of the stationary *alkB* cells. A difference between exponential and stationary *alkB* cells was not observed in the reactivation of MMS-treated M13 phage in agreement with the concept that the reduced sensitivity of *alkB* stationary cells to direct MMS treatment is due to a low content of single-stranded DNA sequences (Fig. 7B).

Discussion

Homologs of the *alkB* gene have been identified in several bacterial genomes, *Schizosaccharomyces pombe*, *Drosophila melanogaster*, *Arabidopsis thaliana*, and *Homo sapiens*, but not in *Saccharomyces cerevisiae* (data not shown; Wei et al. 1996; Colombi and Gomes 1997). Persistence of the AlkB protein through evolution indicates an important functional role in cellular responses to alkylating agents that make up the largest group of environmental genotoxic compounds. No significant homology of AlkB to other known DNA-processing activities has been found by database searches, although a novel hydrolase domain has been suggested (Aravind et al. 1999). Early observations indicated a possible minor role for AlkB in processing damage in methylated duplex DNA (Kataoka et al. 1983). Here, by phage reactivation experiments and cellular transformation with isolated DNA, we observed an extreme deficiency in the ability of *alkB* mutants to process methylated single-stranded DNA but little if any defect in processing double-stranded DNA. These observations provide conclusive evidence that AlkB protein processes DNA damage and deals with lesions produced in single-stranded DNA. In addition, we have shown that AlkB binds preferentially to single-stranded DNA. These findings provide crucial steps forward in elucidating the function of the AlkB protein.

The *E. coli* Tag and AlkA 3-methyladenine-DNA glycosylases excise toxic 3-methyladenine residues from duplex DNA. AlkA protein in vitro can also act on single-stranded DNA but with a low efficiency (Bjelland and Seeberg 1996). By phage reactivation experiments, we found that an *alkA tag* strain was not defective in processing methylated single-stranded DNA in vivo.

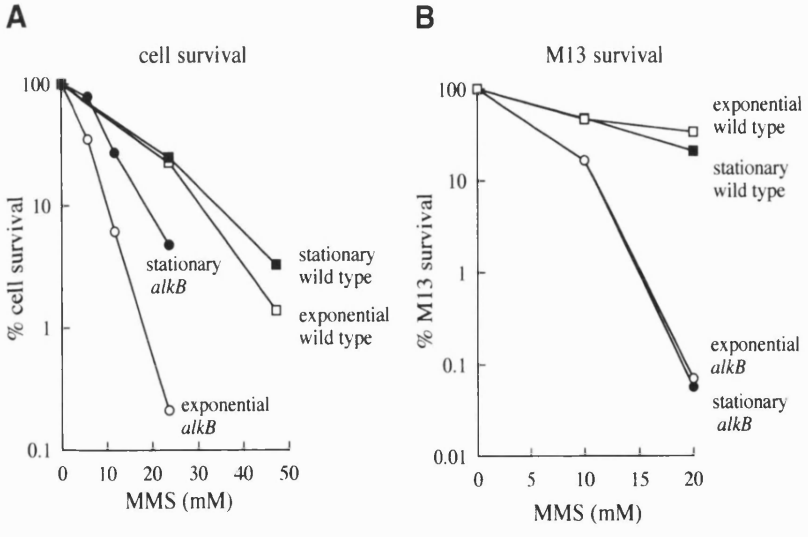


Figure 7. Sensitivity of exponential and stationary phase cells to MMS and their ability to reactivate MMS-treated M13 phage. (*A*) Exponential cultures (A_{450} 0.5) and overnight cultures (A_{450} 1.3) that had been in stationary phase for 16 hr were exposed to various concentrations of MMS for 20 min and then assayed for survival. (*B*) M13 phage were treated with various concentrations of MMS at 30°C for 30 min and assayed for survival in the exponential and stationary phase cultures. The titer of untreated phage (100% survival) was the same in stationary and exponential cells. (□) Exponential AB1157 (wild type); (■) stationary AB1157; (○) exponential BS87 (*alkB117::Tn3*); (●) stationary BS87.

This observation suggests that AlkA is either not active on DNA single strands in vivo or that the apurinic sites resulting from its activity on single-stranded DNA have a similar toxicity to 3-methyladenine. The *alkA tag* Δ(*ada-alkB*) mutant was no more defective in processing single-stranded DNA than the *alkB* single mutant. Processing of methylated lesions in DNA single strands by AlkB therefore does not involve cooperation with 3-methyladenine-DNA glycosylases. Additive sensitivity of an *alkA alkB* double mutant to MMS has been noted previously (Volkert and Hajec 1991).

The *alkB* mutants investigated were only weakly susceptible to MMS-induced base substitution mutagenesis. Thus, the lesions processed by AlkB in DNA single strands have a low capacity for mispairing during DNA replication. Also, processing of DNA damage by AlkB protein in wild-type strains reduced mutagenesis rather than causing it and, so, is unlikely to involve inaccurate replication past blocking lesions. In addition to this, survival of MMS-treated M13 phage was not reduced in a *umuC* mutant, indicating that AlkB protein does not cooperate with UmuC to allow replication past the damage. Considering the possibility that AlkB may be involved in accurate lesion bypass, it is of note that the survival of MMS-treated M13 phage was not reduced in *xth nfo* or *uvrA* mutants. Base excision or nucleotide excision repair therefore do not excise the damage from double-stranded DNA after lesion bypass events. A *recA* mutant had a small defect in processing methylated single-stranded DNA. Our evidence indicated that AlkB and RecA proteins act in different processes and, therefore, RecA may provide a minor alternative pathway for dealing with the damage in single-stranded DNA.

A unique characteristic of *alkB* mutants is their extreme sensitivity to SN2 but not SN1 methylating agents (Kataoka et al. 1983). Here, the cytotoxic lesions processed by AlkB in single-stranded DNA were similarly induced by several SN2 methylating agents, DMS, MMS, and MeI, but not by the SN1 agent MNU or by γ irradiation. Both SN1 and SN2 methylating agents induce N^7-methylguanine and N^3-methyladenine in single-stranded DNA (Singer and Grunberger 1983). Modification at these sites destabilizes the glycosyl bond, and any base loss results in toxic apurinic sites. Since MNU does not induce the lesions that are processed by AlkB protein but does induce N^7-methylguanine, N^3-methyladenine, and apurinic sites, these lesions were excluded as substrates of AlkB. The observation that AlkB protein processes damaged single-stranded DNA also eliminates DNA interstrand cross-links as its substrate. Our attention was drawn to sites that are normally protected from methylation by hydrogen bonding in duplex DNA but that are more reactive in single-stranded DNA. Thus, N^1-methyladenine and N^3-methylcytosine are induced by MMS more readily in single than double strands, and this effect is less pronounced for MNU (Singer and Grunberger 1983). N^3-methylcytosine residues block DNA replication in vitro, and this may also be the case for N^1-methyladenine because of disruption of base pairing and inability to form stable base pairs (Abbott and Saffhill 1977; Boiteux and Laval 1982; Saffhill 1984; Larson et al. 1985). Because of their potential cytotoxicity, we propose these lesions as candidate substrates for the AlkB protein. However, active removal of radiolabeled N^1-methyladenine or N^3-methylcytosine promoted by AlkB from cellular DNA in vivo or from DNA substrates by purified AlkB protein has not been detected (data not shown). Also, the spectrum of base substitution mutations in an MMS-treated *alkB* mutant did not point to a particular modified base as the substrate of AlkB. The mutation frequencies for three out of six possible substitutions showed a small increase, but the mutations occurred in both GC and AT base pairs.

The specificity of AlkB protein in processing damage in DNA single strands suggests that AlkB acts at DNA replication forks or at sites of transcription. This model is supported by the observation that rapidly growing AlkB cells are more sensitive to MMS than those in stationary phase, whereas the growth stage of the cells did not affect survival of MMS-treated M13 phage. Lesions that arise in the replication fork and block DNA synthesis will require rapid repair or bypass replication. We propose that AlkB is involved in either of these processes functioning in an apparently accurate manner and playing a similar critical role in the cellular defence against methylating agents both in *E. coli* and mammalian cells.

Materials and methods

Materials

MMS, DMS, and MeI were purchased from Aldrich; M13mp18 RF1 DNA from Pharmacia Biotech; and MNU was a kind gift from P. Swann, University College London.

Bacterial strains

E. coli strains are listed in Table 1. New *E. coli K12* strains were constructed by transduction using P1 *cml clr* 100 bacteriophage (Sedgwick 1982). The *alkB117::Tn3*, Δ(*ada-alkB25*::Cam^r), and Δ(*srlR-recA*)*306::Tn10* transductants were selected on LB agar containing 50 μg/ml carbenicillin, 20 μg/ml chloramphenicol, or 15 μg/ml tetracyline, respectively. Enhanced MMS sensitivity of *alkB* transductants compared with the parent strains was verified by streaking 10 μl of cultures (A_{450} 0.4) across a gradient of 0–11.8 mM MMS in a 10-cm square Luria-Bertani (LB) agar plate and incubating at 37°C. F′*proAB*$^+$ *lacI*^{Q} *lacZΔM15* Tn10 was transferred from XL1-Blue (Stratagene) into several strains and selected by plating on LB agar containing 15 μg/ml tetracycline and 200 μg/ml streptomycin for counterselection. Most F′ strains used in M13 and f1 phage survival and mutagenesis experiments contained this F′ factor. The exceptions were Δ(srlR-recA)306::Tn10 strains that carried F′*proAB*$^+$ *lacI*^{Q} *lacZΔM15* Tn5 (Stratagene) selected on 40 μg/ml kanamycin. F′148 (his$^+$-aroD$^+$) was transferred from KLF48/KL159 (Coli Genetic Stock Center) into BS87 (*alkB117::Tn3*) and selected by plating on M9 minimal agar supplemented with 20 μg/ml required amino acids except histidine and 50 μg/ml carbenicillin. F′148/BS87 was then used to transfer the *alkB117::Tn3* mutation into *E. coli C-1* by F′-mediated transfer (Miller 1972), and BS159 (*E. coli C-1*

Table 1. E. coli *K12 and* E. coli *C strains*

Strain	Genotype	Source or derivation
AB1157	*argE3 hisG4 leuB6 Δ(gpt-proA)62 thr-1 ara-14 galK2 lacY1 mtl-1 xylA5 thi-1 rpsL31 glnV44 tsx-33 rfbD1 mgl-51 kdgK51*	Laboratory stock
BS87	as AB1157 but *alkB117*::Tn3	Sedgwick 1992
GW7107	as AB1157 but Δ(*ada-alkB25*)::Cam^r	G.C. Walker
GC4803	as AB1157 but X::Tn5 *tagA1 alkA1*	S. Boiteux
BS122	as GC4803 but Δ(*ada-alkB25*)::Cam^r	P1(GW7101)xGC4803
RPC501	as AB1157 but *nfo-1*::Kan^r Δ*xth*	B. Weiss
BS121	RPC501 but Δ(*ada-alkB25*)::Cam^r	P1(GW7101)xRPC501
RW202	*trpE65 lon-11 sulA1* Δ(*srlR-recA*)306::Tn10	B. Bridges
SD4	AB1157 but Δ(*srlR-recA*)*306*::Tn10	P1(RW202)xAB1157
SD5	BS87 but Δ(*srlR-recA*)*306*::Tn10	P1(RW202)xBS87
E. coli C-1	wild type	CGSC
BS159	*E. coli C-1* but *alkB117::Tn3*	F148/BS87x*E.coli* C-1
CC101-CC106	*ara* Δ(*lac proB*)$_{XIII}$/F′*lacI*$^-$*Z*$^-$ *proB*$^+$	Cupples and Miller 1989
SD11-SD16	CC101-CC106 but *alkB117*::Tn3	P1(BS87)xCC102-CC106
PT11	CC102 but Δ(*ada-alkB25*)::Cam^r *ogt-1*::Kan^r	Taverna and Sedgwick 1996

All strains are *E. coli* K12 unless specified to be *E. coli* C.

alkB117::Tn3) was selected on M9 minimal agar containing carbenicillin without amino acid supplements.

Preparation and titration of bacteriophage lysates

Bacteria were grown in LB broth. Tetracycline was added for strains carrying the F′ *proAB*$^+$ *lacI*^{Q} *lacZΔM15* Tn10 factor. M13mp18 and f1 phage lysates of strain AB1157/F′ and λ_{gv} lysates of AB1157 were prepared as described (Sambrook et al. 1989; Dinglay et al. 1998). G4 phage lysates were prepared using *E. coli* C-1. G4 phage (2×10^4 pfu) and 5 mM $CaCl_2$ were added to 1 ml *E. coli* C-1 culture (A_{600} 0.25) and incubated without shaking at 37°C for 10 min. Thirty ml LB broth containing 5 mM $CaCl_2$ were then added and incubated for 6 hr. After adding 100 µl chloroform, the lysate was centrifuged at 7600*g* for 10 min, and the supernatant retained. To titer, M13, f1, and G4 phage were serially diluted and 100-µl aliquots were plated with 100 µl of late exponential cultures of host bacteria (A_{600} 0.8) in 3 ml melted soft LB agar on LB agar plates and incubated overnight at 37°C. Phage survival was monitored by plaque formation. Phage λ were titered as described previously (Dinglay et al. 1998).

Survival of bacteriophage after treatment with DNA-damaging agents

MNU was dissolved in 10 mM potassium acetate (pH 4.5) and aliquots stored at -20°C. Phage lysates were diluted to 8×10^9 pfu/ml in M9 minimal salts and 10 mM $MgSO_4$ and mixed with an equal volume of methylating agent (MMS, DMS, MeI, or MNU) freshly diluted to various concentrations in the same medium. After incubation at 30°C for 30 min (unless otherwise indicated), the phage suspensions were diluted immediately in M9 salts and 1 mM $MgSO_4$ and titered for survival. M13 phage (4×10^8 pfu/ml) exposed for various times to γ irradiation emitted by a CSL 15–137 Cs source at 2.71 Gy/min were similarly titered for survival.

Transformation with MMS-treated M13 single-stranded or double-stranded DNA

Isolation of M13mp18 single-stranded DNA, preparation of competent cells by treatment with $CaCl_2$, and transformation of these cells with M13 DNA were as described (Sambrook et al. 1989). To assay for pfu, AB1157 or BS87 (*alkB117::Tn3*) cells transformed with M13 DNA were plated in LB soft agar together with AB1157/F′ or BS87/F′, respectively. The frequency of transformation was assayed over several concentrations of single-stranded or double-stranded M13 DNA in order to define the linear range. In this range, 20 ng double-stranded DNA gave approximately 6000 transfectants and 100 ng single-stranded DNA gave approximately 2000 transfectants. When treating with MMS, 1 µl DNA (100 ng double stranded or 500 ng single stranded) was incubated with 1 µl MMS at various concentrations in M9 minimal salts and 10 mM $MgSO_4$ at 30°C for 30 min. The MMS was diluted immediately by adding 18 µl 10 mM Tris-HCl and 1 mM EDTA (pH 8). Four microliters of the treated DNA was added to 50 µl competent cells to monitor the transformation frequency.

Sensitivity of alkB mutants to MMS mutagenesis

Strains CC101–CC106 (Miller 1992) and their *alkB117::Tn3* derivatives were grown in M9 minimal salts media to A_{450} 0.5. Aliquots were treated with various concentrations of MMS at 37°C for 20 min, washed in M9 salts containing 1 mM $MgSO_4$, and then serially diluted in the same buffer. Cells were plated on LB agar to estimate survival and on minimal media plates containing 0.2% lactose to monitor Lac$^+$ mutant colonies. The plates were incubated at 37°C.

Sensitivity of exponential and stationary phase cells to MMS

Cells were cultured in M9 minimal media supplemented with 0.2% casein amino acid hydrolysate (Sigma-Aldrich) and thiamine hydrochloride (Miller 1992). Cultures were exposed to MMS either during exponential growth at A_{450} 0.5 or 16 hr after entering stationary phase at A_{450} 1.3. The MMS treatments were at 37°C for 20 min, and the cells were immediately diluted and plated on LB agar plates to monitor cell survival.

Subcloning of the alkB gene and purification of his-tagged AlkB protein

Oligonucleotide primers were synthesized on an Applied Biosystems 394 DNA Synthesizer. The *alkB* gene in plasmid pCS70

(Teo et al. 1984) was amplified by PCR, using Pfu polymerase (Stratagene) and two primers 5′-GGAGAGCATATGTTGGATCTGTTTGCCGAT-3′ and 5′-ATTCGGATCCTTATTCTTTTTTACCTGCCT-3′, to engineer *Nde*I and *Bam*HI restriction sites at the 5′ and 3′ ends of the gene, respectively. The PCR product was digested with *Nde*I and *Bam*HI and inserted into the vector pET15b (Novagen). The DNA sequence of the insert was verified to be correct by sequencing both DNA strands. The new construct, pBAR54, encoded the AlkB protein with a tag of six histidines attached to its amino terminus. This plasmid was transformed into BL21.DE3, in which expression of the cloned gene was induced by IPTG (Studier et al. 1990). SDS-PAGE and Western blotting using anti-AlkB polyclonal antibodies monitored induction of the AlkB protein.

BL21.DE3/pBAR54 was cultured in 270 ml LB broth and 50 μg/ml carbenicillin to A_{600} 0.5 at 37°C. IPTG 1 mM was added and the incubation continued for 3 hr. The cells were harvested, washed in PBSA, and resuspended in 8.5 ml 50 mM Hepes-KOH (pH 8) 2 mM β-mercaptoethanol, 5% glycerol, and 300 mM NaCl. After sonication, the extract was clarified by centrifugation. The extract (55 mg total protein) was supplemented with 1 mM imidazole and loaded onto a 1-ml Ni-NTA (nitrilotriacetic acid)-agarose column (Qiagen) previously equilibrated in buffer (50 mM Hepes-KOH at pH 8, 2 mM β-mercaptoethanol, 5% glycerol, 100 mM NaCl, 1 mM imidazole). The column was washed with 20 ml buffer and then 30 ml buffer containing 40 mM imidazole followed by 5 ml buffer containing 60 mM imidazole. The AlkB protein was eluted in buffer containing 250 mM imidazole. A_{280} readings and visualization by SDS–polyacrylamide gel electrophoresis located the fractions containing pure AlkB protein. The purified his-tagged AlkB protein (1.9 mg) was dialysed into 30 mM potassium phosphate (pH 7.5), 2 mM DTT, 3 mM EDTA, 300 mM NaCl, and 50% glycerol and stored at -80°C.

Binding of his-tagged AlkB protein to DNA

A 40-mer oligonucleotide, 5′-AACGCTACTACTATTAGTAGAATTGATGCCACCTTTTCAG-3′, was 5′ phosphorylated using [γ-^{32}P]ATP and T4 polynucleotide kinase (New England Biolabs). To prepare double-stranded DNA, the end-labeled oligonucleotide was annealed to a twofold excess of complementary strand by heating at 95°C for 2 min and cooling slowly to room temperature (~4 hr). Single- and double-stranded oligonucleotides were treated with 300 mM MMS at 30°C for 30 min and the MMS removed by centrifugation through a Sephadex G50 column equilibrated in 10 mM Tris-HCl and 1 mM EDTA (pH 8). Varying amounts of his-tagged AlkB protein were incubated with [^{32}P]-5′ end-labeled DNA oligomers (30,000 cpm/reaction) in 20 μl buffer (20 mM Tris-HCl at pH 7.5, 100 mM KCl, 0.1 mM DTT, 10% glycerol) at 30°C for 30 min. After addition of 1 ml ice-cold buffer, the reaction mixture was immediately filtered through nitrocellulose disc filters (HAW P02500 Scheibenfilter, Millipore) using a vacuum filtration apparatus (Millipore). The filters were washed with 10 ml of buffer and dried. Scintillation counting quantitated labeled DNA bound to AlkB protein.

Acknowledgments

We thank Lauren Posnick, Peter Karran, and Richard Wood for discussions and John Sguoros and Michael Mitchell for help with homology searches. This work was supported by the Imperial Cancer Research Fund.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Abbott, P.J. and Saffhill, R. 1977. DNA synthesis with methylated (dA-dT) templates: Possible role of O^4-methylthymine as a pro-mutagenic base. *Nucleic Acids Res.* **4:** 761–769.

Aravind, L., Walker, D.R., and Koonin, E.V. 1999. Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic Acids Res.* **27:** 1223–1242.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. 1999. *Current protocols in molecular biology*. Wiley Interscience, New York, NY.

Banerjee, S.K., Borden, A., Christensen, R.B., LeClerc, J.E., and Lawrence, C.W. 1990. SOS-dependent replication past a single *trans- syn* T-T cyclobutane dimer gives a different mutation spectrum and increased error rate compared with replication past this lesion in uninduced cells. *J. Bacteriol.* **172:** 2105–2112.

Bjelland, S. and Seeberg, E. 1996. Different efficiencies of the Tag and AlkA DNA glycosylases from *Escherichia coli* in the removal of 3-methyladenine from single-stranded DNA. *FEBS Lett.* **397:** 127–129.

Boiteux, S. and Laval, J. 1982. Mutagenesis by alkylating agents: Coding properties for DNA polymerase of poly(dC) template containing 3-methylcytosine. *Biochimie* **64:** 637–641.

Chen, B.J., Carroll, P., and Samson, L. 1994. The *Escherichia coli* AlkB protein protects human cells against alkylation-induced toxicity. *J. Bacteriol.* **176:** 6255–6261.

Colombi, D. and Gomes, S.L. 1997. An *alkB* homolog is differentially transcribed during the *Caulobacter crescentus* cell cycle. *J. Bacteriol.* **179:** 3139–3145.

Cupples, C.G. and Miller, J.H. 1989. A set of *lacZ* mutations in *Escherichia coli* that allow rapid detection of each of the six base substitutions. *Proc. Natl. Acad. Sci.* **86:** 5345–5349.

Dinglay, S., Gold, B., and Sedgwick, B. 1998. Repair in *Escherichia coli alkB* mutants of abasic sites and 3-methyladenine residues in DNA. *Mutat. Res.* **407:** 109–116.

Kataoka, H., Yamamoto, Y., and Sekiguchi, M. 1983. A new gene (*alkB*) of *Escherichia coli* that controls sensitivity to methyl methane sulfonate. *J. Bacteriol.* **153:** 1301–1307.

Kondo, H., Nakabeppu, Y., Kataoka, H., Kuhara, S., Kawabata, S., and Sekiguchi, M. 1986. Structure and expression of the *alkB* gene of *Escherichia coli* related to the repair of alkylated DNA. *J. Biol. Chem.* **261:** 15772–15777.

Kornberg, A. and Baker, T.A. 1992. *DNA replication*. W.H. Freeman, New York, NY.

Landini, P., Hajec, L.I., and Volkert, M.T. 1994. Structure and transcriptional regulation of the *Escherichia coli* adaptive response gene *aid B. J. Bacteriol.* **176:** 6583–6589.

Larson, K., Sahm, J., Shenkar, R., and Strauss, B. 1985. Methylation-induced blocks to *in vitro* DNA replication. *Mutat. Res.* **150:** 77–84.

Lindahl, T., Karran, P., and Wood, R.D. 1997. DNA excision repair pathways. *Curr. Opin. Genet. Dev.* **7:** 158–169.

Lindahl, T., Sedgwick, H., Sekiguchi, M., and Nakabeppu, Y. 1988. Regulation and expression of the adaptive response to alkylating agents. *Annu. Rev. Biochem.* **57:** 133–157.

Mackay, W.J., Han, S., and Samson, L.D. 1994. DNA alkylation repair limits spontaneous base substitution mutations in *Escherichia coli*. *J. Bacteriol.* **176:** 3224–3230.

Matijasevic, Z., Sekiguchi, M., and Ludlum, D.B. 1992. Release of N^2, 3-ethenoguanine from chloroacetaldehyde-treated DNA by *Escherichia coli* 3-methyladenine-DNA glycosylase

II. *Proc. Natl. Acad. Sci.* **89:** 9331–9334.
Miller, J.H. 1972. *Experiments in molecular genetics.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
———. 1992. *A short course in bacterial genetics.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Pegg, A.E., Dolan, M.E., and Moschel, R.C. 1995. Structure, function, and inhibition of O^6–alkylguanine-DNA alkyltransferase. *Prog. Nucleic Acid Res. Mol. Biol.* **51:** 167–223.
Rebeck, G.W. and Samson, L. 1991. Increased spontaneous mutation and alkylation sensitivity of *Escherichia coli* strains lacking the *ogt* O^6-methylguanine-DNA methyltransferase. *J. Bacteriol.* **173:** 2068–2076.
Saffhill, R. 1984. Differences in the promutagenic nature of 3-methylcytosine as revealed by DNA and RNA polymerising enzymes. *Carcinogenesis* **5:** 691–693.
Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. *Molecular cloning: A laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Saparbaev, M. and Laval, J. 1994. Excision of hypoxanthine from DNA containing dIMP residues by the *Escherichia coli,* yeast, rat, and human alkylpurine DNA glycosylases. *Proc. Natl. Acad. Sci.* **91:** 5873–5877.
Schendel, P.F. and Defais, D. 1980. The role of *umuC* gene product in mutagenesis by simple alkylating agents. *Mol. Gen. Genet.* **177:** 661–665.
Sedgwick, B. 1982. Genetic mapping of *ada* and *adc* mutations affecting the adaptive response to *Escherichia coli* to alkylating agents. *J. Bacteriol.* **150:** 984–988.
———. 1992. Oxidation of methylhydrazines to mutagenic methylating derivatives and inducers of the adaptive response of *Escherichia coli* to alkylation damage. *Cancer Res.* **52:** 3693–3697.
Singer, B. and Grunberger, D. 1983. *Molecular biology of mutagens and carcinogens.* Plenum Press, New York, NY.
Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185:** 60–89.
Taverna, P. and Sedgwick, B. 1996. Generation of an endogenous DNA-methylating agent by nitrosation in *Escherichia coli. J. Bacteriol.* **178:** 5105–5111.
Teo, I., Sedgwick, B., Demple, B., Li, B., and Lindahl, T. 1984. Induction of resistance to alkylating agents in *E.coli*: The *ada*$^+$ gene product serves both as a regulatory protein and as an enzyme for repair of mutagenic damage. *Embo J.* **3:** 2151–2157.
Vaughan, P., Sedgwick, B., Hall, J., Gannon, J., and Lindahl, T. 1991. Environmental mutagens that induce the adaptive response to alkylating agents in *Escherichia coli. Carcinogenesis* **12:** 263–268.
Volkert, M.R. and Hajec, L.I. 1991. Molecular analysis of the *aidD6*::Mud1(*bla lac*) fusion mutation of *Escherichia coli* K12. *Mol. Gen. Genet.* **229:** 319–323.
Wei, Y., Carter, K.C., Wang, R., and Shell, B.K. 1996. Molecular cloning and functional analysis of a human cDNA encoding an *Escherichia coli* AlkB homolog, a protein involved in DNA alkylation damage repair. *Nucleic Acids Res.* **24:** 931–937.

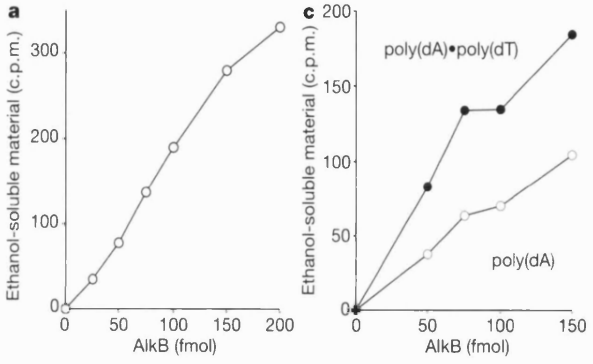
Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage

Sarah C. Trewick*†, Timothy F. Henshaw†‡, Robert P. Hausinger‡, Tomas Lindahl* & Barbara Sedgwick*

* *Cancer Research UK London Research Institute, Clare Hall Laboratories, South Mimms, Hertfordshire EN6 3LD, UK*
‡ *Departments of Chemistry or Microbiology & Molecular Genetics, Michigan State University, East Lansing, Michigan 48824, USA*
† *These authors contributed equally to this work*

Methylating agents generate cytotoxic and mutagenic DNA damage. Cells use 3-methyladenine-DNA glycosylases to excise some methylated bases from DNA, and suicidal O^6-methylguanine-DNA methyltransferases to transfer alkyl groups from other lesions onto a cysteine residue[1,2]. Here we report that the highly conserved AlkB protein repairs DNA alkylation damage by means of an unprecedented mechanism. AlkB has no detectable nuclease, DNA glycosylase or methyltransferase activity; however, *Escherichia coli alkB* mutants are defective in processing methylation damage generated in single-stranded DNA[3–5]. Theoretical protein fold recognition had suggested that AlkB resembles the Fe(II)- and α-ketoglutarate-dependent dioxygenases[6], which use iron-oxo intermediates to oxidize chemically inert compounds[7,8]. We show here that purified AlkB repairs the cytotoxic lesions 1-methyladenine and 3-methylcytosine in single- and double-stranded DNA in a reaction that is dependent on oxygen, α-ketoglutarate and Fe(II). The AlkB enzyme couples oxidative decarboxylation of α-ketoglutarate to the hydroxylation of these methylated bases in DNA, resulting in direct reversion to the unmodified base and the release of formaldehyde.

DNA alkylating agents occur endogenously, are present in the environment and are used in chemotherapy[9,10]. *Escherichia coli* exposed to these compounds respond by inducing the expression of four genes, *ada*, *alkA*, *aidB* and *alkB*. Ada protein is an O^6-methylguanine-DNA methyltransferase and also regulates this adaptive response. AlkA is a 3-methyladenine-DNA glycosylase, and AidB is proposed to destroy certain alkylating agents[1,2]. AlkB is conserved from bacteria to mammals, but its role has not been resolved despite the early isolation of an *E. coli alkB* mutant[3]. Expression of *E. coli alkB* confers alkylation resistance to human cells[11], and conversely, a human homologue has been reported to convey methyl methanesulphonate (MMS) resistance to the *E. coli* mutant[12]. AlkB processes the cytotoxic DNA damage generated in single-stranded DNA by S_N2 methylating agents, such as MMS, dimethylsulphate (DMS) and methyl iodide[5]. 1-Methyladenine and 3-methylcytosine are predominant forms of base damage only in single-stranded DNA because the modification sites are normally protected by base pairing[13,14]. These lesions are cytotoxic because they stall DNA replication, and are not removed by known DNA repair pathways. We previously proposed 1-methyladenine and 3-methylcytosine in DNA as candidate substrates of the AlkB protein[5].



b

	Activity (%)
Complete	100
No AlkB	<0.5
No α-ketoglutarate	<0.5
No ascorbate	<0.5
+ 0.2 mM EDTA	1.1
+ 0.2 mM EDTA + Fe^{2+} 0.7 mM	99

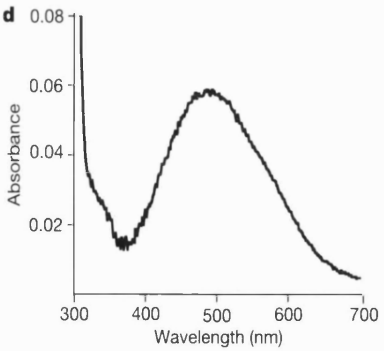


Figure 1 Release of ethanol-soluble material from methylated poly(dA) by AlkB in the presence of Fe(II) and α-ketoglutarate. **a**, [^{14}C]Methyl iodide-treated poly(dA) (1,200 c.p.m.) was incubated with AlkB in the complete reaction mixture, and the release of radioactive material was monitored. **b**, Requirements for AlkB activity (2 pmol AlkB). **c**, Comparison of AlkB activity on [^{14}C]-labelled methylated poly(dA) (1,200 c.p.m.) (open circles) and after annealing to poly(dT) (filled circles), assayed at 20 °C for 30 min to maintain stable duplexes. **d**, Difference absorption spectrum of anaerobic AlkB (0.35 mM) in the presence of α-ketoglutarate (1 mM) and Fe(II) (0.3 mM) minus the spectrum without Fe(II).

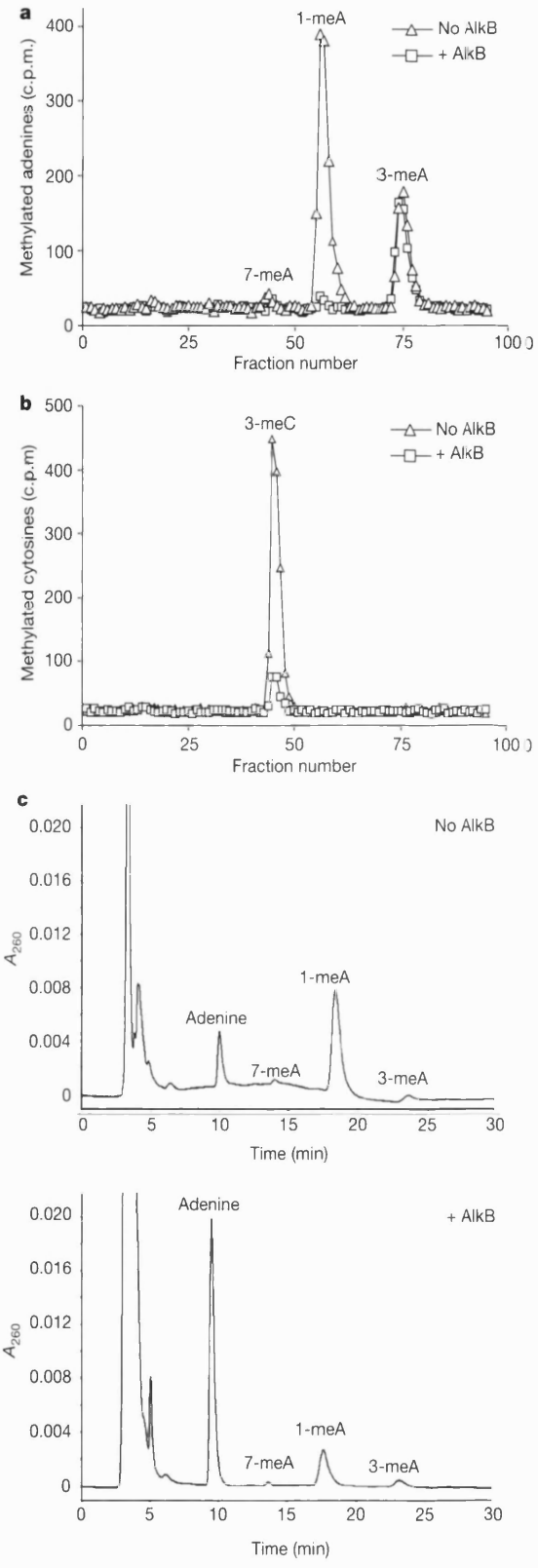


Figure 2 Repair of 1-methyladenine (1-meA) and 3-methylcytosine (3-meC) by AlkB. **a**, **b**, [^{14}C]Methyl iodide-treated poly(dA) (**a**) or poly(dC) (**b**) were incubated without AlkB (triangles) or with 2.5 pmol AlkB (squares). The [^{14}C]-labelled methylated bases remaining in the substrates were analysed by HPLC and scintillation counting. 3-meA and 7-meA, 3- and 7-methlyadenine, respectively. **c**, Direct reversion of 1-methyladenine to adenine by AlkB. A thymine-rich oligodeoxynucleotide containing five adenine residues was heavily methylated and then incubated without (top) or with (bottom) 920 pmol AlkB at 37 °C for 30 min. The bases present in the oligonucleotide were analysed by HPLC and A_{260} measurements. The early eluting peaks were oligo(dT) fragments.

Nevertheless, many attempts to develop assays for this enzyme were unsuccessful. Theoretical sequence profile and fold recognition searches suggested that AlkB may be either a hydrolase or an α-ketoglutarate- and Fe(II)-dependent dioxygenase[6,15]. Here we demonstrate DNA repair activity for AlkB using DNA substrates containing 1-methyladenine or 3-methylcytosine, α-ketoglutarate as a co-substrate, and Fe(II) as a cofactor.

A substrate containing ^{14}C-labelled methylated adenine residues was prepared by treating poly(dA) with [^{14}C]methyl iodide. In the presence of α-ketoglutarate and Fe(II), purified His-tagged AlkB protein released ethanol-soluble radioactive material from this methylated substrate (Fig. 1a). Similar activity was also observed with FFlag-tagged AlkB protein (data not shown). The activity was dependent on α-ketoglutarate, inhibited by EDTA and stimulated by ascorbate (Fig. 1b). Inhibition by EDTA was overcome by adding an excess of $Fe(NH_4)_2(SO_4)_2{\cdot}6H_2O$, thus demonstrating a requirement for Fe(II). At 100-fold higher AlkB concentrations, the effect of ascorbate was reduced. Thus, ascorbate was not essential but stimulated AlkB activity, probably by regenerating Fe(II) from Fe(III)) (data not shown). The conditions for this assay were optimized to 50 mM HEPES-KOH, pH 8, 75 μM $Fe(NH_4)_2(SO_4)_2{\cdot}6H_2O$, 1 mM α-ketoglutarate, 2 mM ascorbate, and 50 μg ml^{-1} bovine serum albumin (BSA). To determine whether AlkB could also act on double-stranded DNA, the [^{14}C]-labelled methylated poly(dA) was annealed to poly(dT). AlkB was approximately threefold more active on the double- compared with the single-stranded substrate (Fig. 1c).

To examine which methylated adenines were processed by AlkB, the poly(dA) substrate was incubated with AlkB, acid hydrolysed, and analysed by high-performance liquid chromatography (HPLC). In the poly(dA) polymer treated with [^{14}C]methyl iodide, 1-methyladenine was the principal methylated base, 3-methyladenine was also abundant, but 7-methyladenine was a minor product[14]. After incubation with AlkB, the levels of 1-methyladenine were reduced whereas the amounts of 3-methyladenine and 7-methyladenine remained unchanged (Fig. 2a). Thus, AlkB specifically catalysed removal of 1-methyladenine from methylated poly(dA). From the data in Fig. 1a, 0.1 pmol AlkB protein removed 1.7 pmol 1-methyladenine from methylated poly(dA) in 15 min, indicating that AlkB acts enzymatically and is not consumed during the reaction. This distinguishes its mode of action from O^6-methylguanine-DNA methyltransferase.

A second modification that is formed to a greater extent in single- compared with double-stranded DNA is 3-methylcytosine. This lesion was also considered as a candidate substrate of AlkB. AlkB protein released radioactive material from poly(dC) treated with [^{14}C]methyl iodide (data not shown). HPLC analysis showed that 3-methylcytosine was the only detectable modified base in this substrate[14], and that it disappeared on incubation with AlkB protein (Fig. 2b). We have therefore identified two substrates of AlkB, 1-methyladenine and 3-methylcytosine, that are both generated in single-stranded DNA on treatment with S_N2 methylating agents.

The essential requirement for both α-ketoglutarate and Fe(II), inhibition by EDTA and stimulation by ascorbate strongly supports the proposal that AlkB is a dioxygenase that is dependent on α-ketoglutarate and Fe(II). Direct evidence that AlkB binds Fe(II) and α-ketoglutarate was obtained by examining the absorption spectrum of the anaerobic protein in the presence or absence of Fe(II) or α-ketoglutarate. The protein with bound metal and cofactor had an absorption peak at 500 nm (Fig. 1d). This chromophore is attributed to a weak charge transfer from Fe(II) to α-ketoglutarate, and is a spectroscopic signature of Fe(II)- and α-ketoglutarate-dependent dioxygenases[16–18]. The AlkB absorption is slightly shifted compared with the 530-nm transition observed in other family members, and was not perturbed further on addition of methylated DNA.

We propose that AlkB repairs 1-methyladenine and 3-methylcytosine in DNA by oxidative demethylation. In such a mechanism, the lesions would be reverted to adenine and cytosine, formaldehyde would be generated and O_2 consumed. To demonstrate that AlkB directly reverts 1-methyladenine in DNA to adenine residues, a non-radioactive substrate was prepared in which 76% of the adenine residues were methylated to form 1-methyladenine. This was achieved by repeated treatments with DMS of an oligonucleotide containing adenine residues interspersed between inefficiently methylated thymine residues[14]. Only 4% of the adenines were recovered as 3-methyladenine, and 2% were 7-methyladenine (Fig. 2c). The low amount of 3-methyladenine might be caused by instability of the glycosyl bond and loss of this modification during

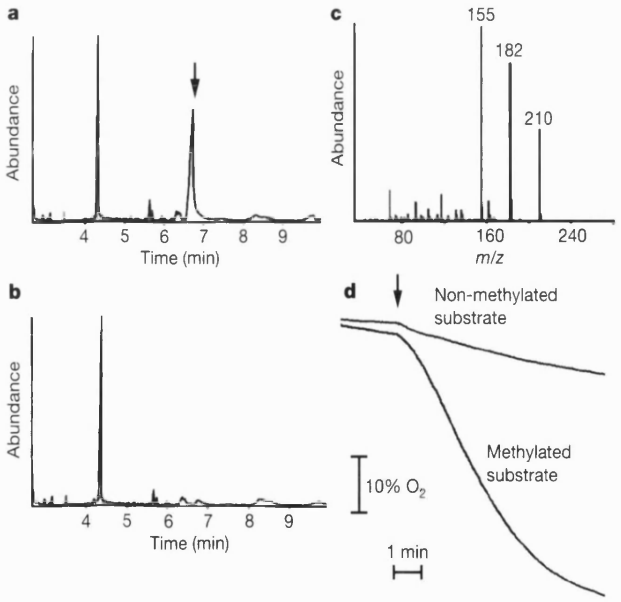


Figure 3 Production of formaldehyde and consumption of O_2 by AlkB. Methylated oligo(dA)$_{34}$•oligo(dT) was incubated with AlkB. The products were derivatized with PFPH and analysed by gas chromatography mass spectrometry. Gas chromatography traces are shown of derivatized products generated when the DNA was methylated (**a**) or not methylated (**b**). **c**, Mass spectrum of the product indicated by the arrow in **a**. **d**, Consumption of O_2 during incubation of AlkB with either methylated or non-methylated oligo(dA)$_{34}$•oligo(dT) was determined by using an oxygen electrode. AlkB (9 μM) was added at the point indicated by the arrow.

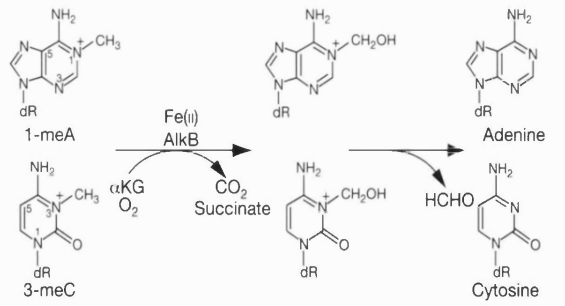


Figure 4 Schematic representation of the repair of 1-methyladenine (1-meA) and 3-methylcytosine (3-meC) residues in DNA by oxidative demethylation catalysed by AlkB. Oxidation of 1-methyladenine and 3-methylcytosine by AlkB requires O_2, α-ketoglutarate (αKG) and Fe(II), and generates CO_2 and succinate. Oxidized methyl groups are released as formaldehyde resulting in direct reversal of the lesions to the unmodified base residues. Note that the alkylation positions in 1-methyladenine and 3-methylcytosine are equivalent, although these sites in the pyrimidine rings of A and C are numbered differently by standard nomenclature. dR, deoxyribose.

the extensive DMS treatments. The heavily methylated substrate was incubated with AlkB in the optimized assay conditions, the DNA hydrolysed and individual bases quantified by HPLC and measurements of absorbance at 260 nm (A_{260}). In the presence of AlkB, a decrease in the amount of 1-methyladenine correlated with a stoichiometric increase in the amount of adenine recovered (Fig. 2c). We conclude that AlkB converts 1-methyladenine directly to adenine in DNA.

To determine whether formaldehyde (HCHO) was a reaction product, AlkB was incubated with MMS-treated poly(dA) annealed to poly(dT), the products were derivatized with pentafluorophenylhydrazine (PFPH), and analysed by gas chromatography. One derivatized product (arrow in Fig. 3a) arose only when the DNA substrate was methylated and Fe(II) and α-ketoglutarate were present (Fig. 3a, b), and had a mass spectrum identical to that for the HCHO-PFPH adduct[19] (Fig. 3c). The release of HCHO was also monitored by a coupled spectrophotometric assay using formaldehyde dehydrogenase[20], and again was detected only in complete assay conditions with the heavily methylated DNA substrate (turnover number 1.5 s^{-1}, data not shown). A yield of 140 μM HCHO correlated stoichiometrically with the amount of 1-methyladenine in the DNA substrate (7–10% of the adenines), the amount of O_2 consumed (Fig. 3d) and the succinate generated (data not shown). This stoichiometric relationship further verifies the proposal that AlkB is an α-ketoglutarate-dependent dioxygenase. In the absence of methylated DNA, an observed slow consumption of both oxygen (Fig. 3d) and α-ketoglutarate was consistent with a partial uncoupling of α-ketoglutarate decomposition and hydroxylation of the methylated DNA bases. Such uncoupling is a well known property of this family of enzymes[21,22].

Oxidative demethylation is an unprecedented mechanism of DNA repair. The proposed reaction mechanism by which AlkB repairs 1-methyladenine and 3-methylcytosine in DNA is shown in Fig. 4. Owing to the stability of the N–C bond in 1-methyladenine and 3-methylcytosine, demethylation by hydrolysis would be energetically unfavourable; consequently oxidative demethylation by reactive iron-oxygen species is required. Direct reversal of this base damage to unsubstituted parent residues would be a highly accurate form of DNA repair and agrees with *in vivo* observations that repair by AlkB is non-mutagenic[5]. *Escherichia coli alkB* mutants are more sensitive to alkylating agents during active growth than in stationary phase probably because 1-methyladenine and 3-methylcytosine are produced in single-stranded regions of DNA in replication forks and transcription bubbles[5]. DNA unfolds only transiently during replication and transcription, so it is beneficial that AlkB repairs its substrates not only in single strands but also, and even more efficiently, after DNA re-annealing (Fig. 1c). Of note, *Caulobacter crescentus alkB* expression is cell-cycle regulated with a pattern similar to activities required for DNA replication[23].

We have extended the family of α-ketoglutarate- and Fe(II)-dependent dioxygenases to include AlkB, as recently proposed on theoretical grounds[6]. These enzymes catalyse a variety of reactions including hydroxylations, desaturations and oxidative ring closures, and account for oxidation of proline in collagen, steps in the biosynthesis of several antibiotics and cellular metabolites, as well as biodegradation of selected compounds[7,8]. No other members of this family are presently known to act on DNA, but it is notable that a fungal enzyme, thymine-7-hydroxylase, oxidizes the methyl group of free thymine[24]. We suggest that oxidative demethylation requiring free-radical chemistry could be involved in other mechanisms of active removal of chemically stable adducts from DNA or histones, for example, the 5-methylcytosine demethylase activity[25,26] that is apparently important in epigenetic control. □

Methods

AlkB protein assay and substrates

A total of 0.6 mg poly(dA) (average length 310 residues; Amersham Biosciences) was treated with 1 mCi [^{14}C]methyl iodide (58 mCi mmol^{-1}) (Amersham Biosciences) in 0.8 ml^{-1} of 10 mM sodium cacodylate, pH 7, at 30 °C for 6 h. The methylated polymer was precipitated in ethanol, dissolved in 10 mM Tris-HCl, pH 8, and had a specific actitivity of 1,860 counts per min (c.p.m.) per μg. To prepare a double-stranded substrate, [^{14}C]-labelled methylated poly(dA) was annealed to poly(dT) at 20 °C. A total of 1.2 mgg poly(dC) (average length 370 residues) in 1.3 ml of 50 mM HEPES-KOH, pH 8, wwas similarly treated with [^{14}C]methyl iodide to yield a specific activity of 700 c.p.m. μg^{-1}. His-tagged AlkB was overexpressed and purified as previously described by nickel-aagarose affinity chromatography[5], except that EDTA was added to the sonication buffer, annd the extract was dialysed against the same buffer without EDTA. A plasmid coding for r Flag-tagged AlkB was constructed by digestion of pBAR54 (ref. 5) with *Nde*I and *Nco*I I to remove DNA coding for the His tag. This DNA fragment was replaced by a double-stranded oligonucleotide coding for the Flag tag. The Flag-tagged protein was purified by immnunoaffinity chromatography (Sigma). Purified AlkB was incubated with thee [^{14}C]methyl iodide-treated substrates in 50 mM HEPES-KOH, pH 8, 75 μM $Fe(NH_4)_2(SO_4)_2{\cdot}6H_2O$, 1 mM α-ketoglutarate, 2 mM ascorbate, and 50 μg ml^{-1} BBSA for 15 min at 37 °C. The reaction was stopped by adding 11 mM EDTA, and the substrate was precipitated in ethanol in the presence of carrier calf thymus DNA. We monitored ethanol-soluble radioactive material by scintillation counting.

HPLC analysis

Methylated adenine residues were released from [^{14}C]-labelled methylated poly(dA) by hydrolysis in 0.1 M HCl at 95 °C for 1 h; 3-methylcytosine was released from methhylated poly(dC) by treatment with 90% formic acid at 180 °C for 20 min. Methylated baseses were analysed by HPLC on a Whatman Partisil 10 cation exchange column in 0.1 M ammonium formate, pH 3.6. A gradient of MeOH from 20 to 40% was applied to separate thee methylated adenine derivatives, and a gradient from 5 to 40% MeOH was used to a analyse 3-methylcytosine.

High-level methylation of an oligonucleotide

A 41-base oligonucleotide (TTTTTT(ATTTTTT)$_5$) was treated with 50 mM DMSS in 75 mM sodium cacodylate, pH 7.4, at 30 °C, twice for 2 h, and then four times fonr 1 h. Between each treatment the DMS was removed by centrifugation through a G25 Sepphadex mini-column equilibrated in the same buffer. The level of methylation of adenine rresidues was examined by acid hydrolysis, HPLC and A_{260} measurements. The A_{260} ratio oof adenine to 1-methyladenine was 1.04, which was determined by monitoring knowwn amounts of these purines in the same conditions.

Gas chromatography mass spectrometry

Oligo(dA) (25 residues) was treated with 500 mM MMS at 30 °C for 30 min and, s after removal of the MMS, annealed to oligo(dT). AlkB (38 μM) was incubated with thhis substrate (50 μM) in 50 mM HEPES-KOH, pH 8, containing 1 mM α-ketoglutaraate, 75 μM $Fe(NH_4)_2(SO_4)_2{\cdot}6H_2O$, and 2 mM ascorbate in a final volume of 200 μl. Afafter 10 min at 30 °C, 200 μl 15 mM PFPH in 1.2 M phosphoric acid was added, and incuubation continued at 50 °C for 2 h. The reaction mixtures were extracted with 100 μl of 855:15 hexane:dichloromethane, and 20 μl was analysed by gas chromatography mass spectrometry as described[19].

O_2 consumption

We mixed 50 mM HEPES, pH 8, with 1 mM α-ketoglutarate, 75 μM $Fe(NH_4)_2(SO_4)_2{\cdot}6H_2O$, 100 μM ascorbate, and 50 μM oligonucleotide substrate, prrepared as described for gas chromatography mass spectrometry, for 3 min in a YSI 5300 c oxygen electrode at 30 °C to allow equilibration with atmospheric oxygen. The electrode v was precalibrated with air-saturated water (236 μM O_2). AlkB was added through a gaas-tight syringe to a final concentration of 9 μM.

Received 7 March; accepted 6 June 2002; doi:10.1038/nature00908.

1. Lindahl, T., Sedgwick, B., Sekiguchi, M. & Nakabeppu, Y. Regulation and expression of the a adaptive response to alkylating agents. *Annu. Rev. Biochem.* **57,** 133–157 (1988).
2. Seeberg, E. & Berdal, K. G. in *Base Excision Repair of DNA Damage: Repair of Alkylatiom Daamage to DNA* (ed. Hickson, I. D.) 151–168 (Landes Bioscience, Austin, 1999).
3. Kataoka, H., Yamamoto, Y. & Sekiguchi, M. A new gene (*alkB*) of *Escherichia coli* that contrrols sensitivity to methyl methane sulfonate. *J. Bacteriol.* **153,** 1301–1307 (1983).
4. Kondo, H. *et al.* Structure and expression of the *alkB* gene of *Escherichia coli* related to the rrepair of alkylated DNA. *J. Biol. Chem.* **261,** 15772–15777 (1986).
5. Dinglay, S., Trewick, S. C., Lindahl, T. & Sedgwick, B. Defective processing of methylated sinngle-stranded DNA by *E. coli alkB* mutants. *Genes Dev.* **14,** 2097–2105 (2000).
6. Aravind, L. & Koonin, E. V. The DNA-repair protein AlkB, EGL-9, and leprecan define new fammilies of 2-oxoglutarate- and iron-dependent dioxygenases. *Genome Biol.* 2 RESEARCH0007 (2001).
7. Prescott, A. G. & Lloyd, M. D. The iron(II) and 2-oxoacid-dependent dioxygenases and theirir role in metabolism. *Nat. Prod. Rep.* **17,** 367–383 (2000).
8. Ryle, M. J. & Hausinger, R. P. Non-heme iron oxygenases. *Curr. Opin. Chem. Biol.* **6,** 1993–201 (2002).
9. Rebeck, G. W. & Samson, L. Increased spontaneous mutation and alkylation sensitivity of *Escscherichia coli* strains lacking the *ogt* O^6-methylguanine-DNA methyltransferase. *J. Bacteriol.* **173,** 20688–2076 (1991).
10. Vaughan, P., Sedgwick, B., Hall, J., Gannon, J. & Lindahl, T. Environmental mutagens that inoduce the adaptive response to alkylating agents in *Escherichia coli*. *Carcinogenesis* **12,** 263–268 (1991).
11. Chen, B. J., Carroll, P. & Samson, L. The *Escherichia coli* AlkB protein protects human ccells a against alkylation-induced toxicity. *J. Bacteriol.* **176,** 6255–6261 (1994).
12. Wei, Y., Carter, K. C., Wang, R. & Shell, B. K. Molecular cloning and functional analysis of a a human cDNA encoding an *Escherichia coli* AlkB homolog, a protein involved in DNA alkylation daamage repair. *Nucleic Acids Res.* **24,** 931–937 (1996).

13. Bodell, W. J. & Singer, B. Influence of hydrogen bonding in DNA and polynucleotides on reaction of nitrogens and oxygens toward ethylnitrosourea. *Biochemistry* **18,** 2860–2863 (1979).
14. Singer, B. & Grunberger, D. *Molecular Biology of Mutagens and Carcinogens: Reactions of Directly Acting Agents with Nucleic Acids* 45–96 (Plenum, New York, 1983).
15. Aravind,, L., Walker, D. R. & Koonin, E. V. Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic Acids Res.* **27,** 1223–1242 (1999).
16. Hegg, E.. L. *et al.* Herbicide-degrading α-keto acid-dependent enzyme: metal coordination environment and mechanistic insights. *Biochemistry* **38,** 16714–16726 (1999).
17. Pavel, E.. G. *et al.* Circular dichroism and magnetic circular dichroism spectroscopic studies of the non-heme ferrous active site in clavaminate synthase and its interaction with α-ketoglutarate co-substrate. *J. Am. Chem. Soc.* **120,** 743–753 (1998).
18. Ryle, M.. J., Padmakumar, R. & Hausinger, R. P. Stopped-flow kinetic analysis of *Escherichia coli* taurine/α-ketoglutarate dioxygenase: interactions with α-ketoglutarate, taurine, and oxygen. *Biochemistry* **38,** 15278–15286 (1999).
19. Heck, Hl. D. A., White, E. L. & Cassanova-Schmitz, M. Determination of formaldehyde in biological tissues by gas chromatography/mass spectrometry. *Biomed. Mass Spectrom.* **9,** 347–353 (1982).
20. Lizcano,, M. J., Unzeta, M. & Tipton, K. F. A spectrophotometric method for determining the oxidative deamination of methylamine by amine oxidases. *Anal. Biochem.* **286,** 75–79 (2000).
21. de Jong,. L., Albracht, S. P. & Kemp, A. Prolyl 4-hydroxylase activity in relation to the oxidation state of enzyme--bound iron. The role of ascorbate in peptidyl proline hydroxylation. *Biochim. Biophys. Acta* **704,** 326–332 (1982).
22. Myllyla, R., Majamaa, K., Gunzler, V., Hanauske-Abel, H. M. & Kivirikko, K. I. Ascorbate is consumed stoichiometrically in the uncoupled reactions catalysed by prolyl 4-hydroxylase and lysyl hydroxylase. *J. Biol. Chem.* **10,** 5403–5405 (1984).
23. Colombii, D. & Gomes, S. L. An *alkB* homologue is differentially transcribed during the *Caulobacter crescentus* cell cycle. *J. Bacteriol.* **179,** 3139–3145 (1997).
24. Thornburg, L. D., Lai, M.-T., Wishnok, J. S. & Stubbe, J. A non-heme iron protein with heme tendencies: an investigation of the substrate specificity of thymine hydroxylase. *Biochemistry* **32,** 14023–14033 (1993).
25. Mayer, W., Niveleau, A., Walter, J., Fundele, R. & Haaf, T. Demethylation of the zygotic paternal genome.. *Nature* **403,** 501–502 (2000).
26. Smith, S. S. Gilbert's conjecture: the search for DNA (cytosine-5) demethylases and the emergence of new functions for eukaryotic DNA (cytosine-5) methyltransferases. *J. Mol. Biol.* **302,** 1–7 (2000).

Acknowledgements

We thank J. McCracken for advice, and T. Duncan and P. Robins for technical help. This work was supported by Cancer Research UK, and by NIH grants to R.P.H. and J. McCracken.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to B.S. (e-mail: b.ssedgwick@cancer.org.uk) or R.P.H. (e-mail: hausinge@msu.edu).

Reversal of DNA alkylation damage by two human dioxygenases

Tod Duncan*†, Sarah C. Trewick*†, Pertti Koivisto*, Paul A. Bates‡, Tomas Lindahl*§, and Barbara Sedgwick*

*Cancer Research UK London Research Institute, Clare Hall Laboratories, South Mimms, Hertfordshire EN6 3LD, United Kingdom; and ‡Biomolecular Modelling Laboratory, Cancer Research UK London Research Institute, Lincoln's Inn Fields, London WC2A 3PX, United Kingdom

Edited by Martin Gellert, National Institutes of Health, Bethesda, MD, and approved November 12, 2002 (received for review September 30, 2002)

The *Escherichia coli* AlkB protein protects against the cytotoxicity of methylating agents by repair of the DNA lesions 1-methyladenine and 3-methylcytosine, which are generated in single-stranded stretches of DNA. AlkB is an α-ketoglutarate- and Fe(II)-dependent dioxygenase that oxidizes the relevant methyl groups and releases them as formaldehyde. Here, we identify two human AlkB homologs, ABH2 and ABH3, by sequence and fold similarity, functional assays, and complementation of the *E. coli alkB* mutant phenotype. The levels of their mRNAs do not appear to correlate with cell proliferation but tissue distributions are different. Both enzymes remove 1-methyladenine and 3-methylcytosine from methylated polynucleotides in an α-ketoglutarate-dependent reaction, and act by direct damage reversal with the regeneration of the unsubstituted bases. AlkB, ABH2, and ABH3 can also repair 1-ethyladenine residues in DNA with the release of acetaldehyde.

Although single-stranded regions of DNA occur *in vivo* within replication forks and transcription bubbles, the susceptibility of single-stranded DNA to alkylating agents has been little investigated. The major lesions generated in single-stranded DNA are 1-alkyladenine and 3-alkylcytosine; these modification sites are protected by the complementary strand in duplex DNA (1). The 3-methylcytosine (3-meC) lesions block replication and are potentially cytotoxic (2). The *Escherichia coli* AlkB function counteracts toxicity by alkylating agents and its expression is induced by exposure to such agents (3, 4). Expression of *E. coli* AlkB in mammalian cells also confers increased resistance to alkylating agents (5). We have shown that AlkB specifically repairs alkylation damage in single-stranded DNA *in vivo*, and binds preferentially to single-stranded DNA *in vitro* (6). These results indicated that AlkB repairs 1-methyladenine (1-meA) and/or 3-meC residues in DNA, but the reaction mechanism was unknown. In an important lead, protein fold analysis combined with weak sequence homology suggested that AlkB is a member of the family of α-ketoglutarate (αKG)- and Fe(II)-dependent dioxygenases (7). These enzymes are involved in a variety of metabolic reactions; however, a fungal member of the family can perform catabolic oxidative demethylation of the free base 1-methylthymine (8). Biochemical assays with purified AlkB protein recently demonstrated that AlkB is indeed an αKG- and Fe(II)-dependent dioxygenase that oxidatively demethylates 1-meA and 3-meC residues in single-stranded as well as double-stranded DNA. The methyl group is released from the lesion as free formaldehyde, with accompanying regeneration of the unsubstituted base residue in DNA (9, 10).

Because alkylating agents are environmental carcinogens, and also are used clinically as cytotoxic anticancer drugs, it was of interest to determine whether human cells have a counterpart to the *E. coli* AlkB protein. Here, we identify and characterize two human AlkB homologs encoded on different chromosomes.

Materials and Methods

Single-Stranded DNA Phage Reactivation. M13 phage was used to examine reactivation of methyl methanesulfonate (MMS)-treated single-stranded DNA in transfected *E. coli alkB* mutant cells expressing cloned AlkB protein or putative human homologs. In some experiments, the single-stranded φK phage was used to avoid the M13 requirement for F′ strains (11). MMS treatment of the phage and estimation of reactivation were as described (6). Cells were grown in the presence of 100 μM isopropyl β-D-thiogalactopyranoside for 3 h at 25°C to induce expression of cloned genes before transfection.

Molecular Cloning of Human ABH1, ABH2, and ABH3. ABH1, ABH2, and ABH3 cDNAs were obtained from the IMAGE consortium and amplified by PCR using *Pfu* Turbo polymerase. ABH1 was subcloned into the expression vector pET15b. ABH2 and ABH3 were subcloned by using Gateway technology (Invitrogen). In this case, the amplified ORFs flanked by attB1 and attB2 sites were cloned into the entry vector p221 by recombination, and subsequently shuttled into pDEST.GST or pDEST17 for expression of GST- or hexahistidine-tagged fusion proteins, respectively.

Purification of *E. coli* AlkB and Human ABH1, ABH2, and ABH3 Proteins. His-tagged *E. coli* AlkB protein was overexpressed and purified as described (9). Expression plasmids encoding GST- or hexahistidine-tagged ABH1, ABH2, or ABH3 were transformed into *E. coli* Codon plus. To induce expression, 0.1 mM β-D-thiogalactopyranoside was added to cultures at A_{600} = 0.6, and incubation continued for 3.5 h at 25°C. Harvested cells expressing hexahistidine-tagged proteins were disrupted in 25 mM Tris-HCl, pH 8.0/300 mM NaCl/10 mM imidazole/1 mM EDTA/1 mM β-mercaptoethanol by using a French press or by treatment with lysozyme and sonication. Lysates were clarified by centrifugation at 17,000 × *g* for 30 min at 4°C and dialyzed against lysis buffer lacking EDTA. These lysates were loaded in batch onto Ni^{2+} nitrilotriacetic acid agarose resin (Qiagen, Valencia, CA). The resin was washed with lysis buffer containing 40 mM imidazole but no EDTA. Bound proteins were then eluted in 25 mM Tris·HCl, pH 8.0/150 mM NaCl/250 mM imidazole. When purifying GST-tagged proteins, the lysis buffer contained 50 mM Tris·HCl, pH 8.0/300 mM NaCl/1 mM EDTA/1 mM DTT. Clarified lysates were loaded batchwise onto glutathione-Sepharose resin (Amersham Pharmacia), and washed with lysis buffer lacking EDTA. Specifically bound proteins were eluted in 50 mM Tris·HCl, pH 8.5/150 mM NaCl/1 mM DTT/10 mM reduced glutathione. Cultures (2.4 liters) of overexpressing strains yielded 2.5 mg of GST.ABH2 or 12 mg of GST.ABH3 in soluble and 90–95% pure form. However, the enzyme activities were unstable during isolation and on storage, with ABH2 being more labile than ABH3.

Substrates and Assays of *E. coli* AlkB and Human ABH2 and ABH3. Poly(dA) and poly(dC) (Amersham Pharmacia) were treated with [^{14}C]MeI (58 mCi/mmol; 1 Ci = 37 GBq) (Amersham

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: 3-meC, 3-methylcytosine; 1-meA, 1-methyladenine; αKG, α-ketoglutarate; MMS, methyl methanesulfonate; DMS, dimethyl sulfate; DNPH, 2,4-dinitrophenylhydrazone; 1-etA, 1-ethyladenine; EtI, ethyl iodide.

†T.D. and S.C.T. contributed equally to this work.

§To whom correspondence should be addressed. E-mail: tomas.lindahl@cancer.org.uk.

www.pnas.org/cgi/doi/10.1073/pnas.262589799

Pharmacia) as described (9). The specific activities of the methylated substrates were 1,860 and 700 cpm/μg, respectively. The same method was used to treat 0.6 mg poly(dA) with 250 μCi [^{14}C] ethyl iodide (EtI; 31.5 mCi/mmol) (ICN) except that the treatment was at 37°C for 18 h. The specific activity of the ethylated poly(dA) was 120 cpm/μg.

A single stranded substrate containing [^{3}H]cytosine residues interspersed between thymines was synthesized by PCR. The template, 5′-$(A_6G)_9A_6$CAAGTCCCACGCTCACACACA-ATCC-3′, was replicated by using *Pfu* polymerase and a single primer, 5′-GGATTGTGTGTGAGCGTGGGACTTG-3′, in the presence of 150 μM dNTPs and 2.5 μM [^{3}H]dCTP (16 Ci/mmol) (Amersham Pharmacia). To obtain a high level of methylation, the ^{3}H-labeled PCR product was treated 8 times with 50 mM dimethyl sulfate (DMS) in 75 mM sodium cacodylate (pH 7.4) at 30°C for 2 h. Between each treatment the DMS was removed by centrifugation through a Sephadex G-25 mini-column equilibrated in the same buffer. The level of methylation of [^{3}H]cytosine residues was determined by formic acid hydrolysis and HPLC using a Whatman Partisil 10 cation exchange column run isocratically in 100 mM formate (pH 3.6), 2.5% methanol; 70% of the [^{3}H]cytosines were methylated to 3-me[^{3}H]C.

Purified AlkB protein and putative human homologs were assayed as described (9). Briefly, the enzymes were incubated with [^{14}C]MeI- or [^{14}C]EtI-treated substrates in 50 mM Hepes·KOH, pH 8/75 μM $Fe(NH_4)_2(SO_4)_2$/1 mM αKG/2 mM ascorbate/50 μg/ml BSA for 15 min at 37°C. The reaction was stopped by adding EDTA, and the polynucleotide substrate was ethanol precipitated. Ethanol-soluble radiolabeled material was monitored by scintillation counting. Methylated adenine and cytosine residues remaining in the ^{14}C-methylated polynucleotides were released by acid hydrolysis and analyzed by HPLC and scintillation counting (9). ^{14}C-ethylated adenines were analyzed by HPLC on a Phenomenex Luna 2 reverse phase column in 50 mM ammonium acetate, pH 6.5. A linear gradient of 1–46% methanol was applied to separate the ethylated adenine bases.

Detection of [^{14}C]Aldehydes. To determine whether AlkB released formaldehyde or acetaldehyde from ^{14}C-ethylated poly(dA), the ethanol-soluble ^{14}C-labeled material was derivatised with 2,4-dinitrophenylhydrazone (DNPH) (12). The DNPH derivatives were analyzed by reverse phase HPLC on a Phenomenex Hypersil column using a linear 50–90% methanol gradient in water and quantitated by scintillation counting. DNPH derivatised formaldehyde and acetaldehyde markers were monitored at A_{254}.

Northern Hybridization. Multiple tissue Northern blots (Ambion 3142/3143) were probed, washed, and stripped according to the manufacturer's instructions. Each blot contained 2 μg of poly(A)$^+$ RNA from 10 different normal tissues or 10 different cell lines derived from various carcinomas. Briefly, blots were prehybridized in hybridization buffer (Ambion) for 30 min at 42°C. Full-length ORF probes were labeled with [^{32}P]dCTP by PCR and purified through Sephadex G-50 spin columns. The probe was added directly to the hybridization buffer to a specific activity of 10^6 cpm/ml. Blots were hybridized for 16 h at 42°C before being washed twice with 0.3 M NaCl/0.03 M citrate/0.1% SDS and twice with 0.15 M NaCl/0.015 M citrate/0.1% SDS at 42°C. The bound probe was visualized by fluorography. A β-actin probe was used as a control to verify integrity of the samples on the blots.

Results

Identification of Human AlkB Homologs. A putative human homolog, ABH, of the *E. coli* AlkB protein was previously reported to partially complement the MMS sensitivity of an *E. coli alkB* mutant (13). With the aim of further characterizing this potential

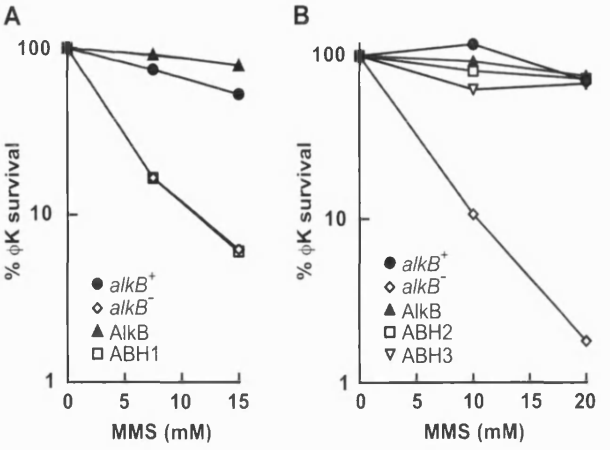


Fig. 1. Complementation of *E. coli alkB* mutant phenotype by overexpression of human ABH2 or ABH3 proteins, but not ABH1. Reactivation of MMS-treated ϕK single-stranded DNA phage was monitored in various *E. coli* strains. (*A*) Host strains were *alkB*$^+$ or *alkB22* derivatives of *E. coli* BL21.DE3. Plasmid pBAR54 encodes *E. coli* AlkB protein and pBAR65 encodes human ABH1 (not tagged). ●, *alkB*$^+$/pET15b; ◇, *alkB22*/pET15b; ▲, *alkB22*/pBAR54; □, *alkB22*/pBAR65. (*B*) Host strains were *alkB*$^+$ or *alkB22* derivatives of *E. coli* AB1157 *gyrA*. Plasmid pBAR32 encodes *E. coli* AlkB protein, pGST.ABH2 encodes GST tagged ABH2 and pGST.ABH3 encodes GST-tagged ABH3. ●, *alkB*$^+$/pGEX6P-1; ◇, *alkB22*/pGEX6P-1; ▲, *alkB22*/pBAR32; □, *alkB22*/pGST.ABH2; ▽, *alkB22*/pGST.ABH3. The proteins overexpressed in the *alkB* mutant are indicated with symbols in the corner of the figures.

homolog, we obtained two independent full-length cDNA clones of the ABH gene from the IMAGE consortium (IMAGE IDs 1337358 and 2284109). DNA sequencing showed that both cDNAs contained ORFs encoding identical proteins (data not shown). The N and C termini of these ORFs, however, differed from those reported for ABH by Wei *et al.* in 1996 (13), but agreed with the predicted conceptual splice of the genomic DNA sequence on chromosome 14q24 (GenBank accession no. AC008044) and also with a further cDNA (GenBank accession no. XM_007409). The ORF of the IMAGE 1337358 cDNA, which we refer to as ABH1, was subcloned into the pET15b vector and overexpressed in an *alkB22* derivative of *E. coli* BL21.DE3. Induction by 50 μM β-D-thiogalactopyranoside at 37°C for 3 h resulted in 15% solubility of the overexpressed ABH1 protein (data not shown). *E. coli alkB* mutants are defective in reactivation of MMS treated single stranded DNA phage (6). Overexpression of *E. coli* AlkB protein complemented this phenotype, but this was not the case for human ABH1 (Fig. 1*A*). Moreover, the purified hexahistidine-tagged ABH1 protein did not release soluble radiolabeled material from a [^{14}C]methyl iodide treated poly(dA) substrate in the optimised AlkB assay conditions (9) or after various modifications of this assay (data not shown). We were therefore unable to support the previous suggestion (13) that human ABH1 is a functional AlkB homolog.

BLAST searches of the NCBI human DNA databases revealed a different potential AlkB homolog (Fig. 2). This ORF, described as "similar to prostate cancer antigen-1," was predicted from the conceptual splice of DNA sequences on chromosome 12q23.3 (accession no. XP_058581). Searching the databases with XP_058581 revealed yet another possible AlkB homolog. This second protein known as prostate cancer antigen-1 is encoded by the DEPC-1 gene located on chromosome 11q11 (accession no. NP_631917). Both these proteins are of unknown function. We refer to the two proteins as ABH2 (XP_058581) and ABH3 (NP_631917), having 30.8 and 23.1% core region identity to *E. coli* AlkB, respectively (Fig. 2*A*). Analysis of ABH2 and ABH3 for known protein domains by using the program DOMAINFISHING (14) indicated that ABH2 and ABH3 were similar to αKG- and Fe(II)-

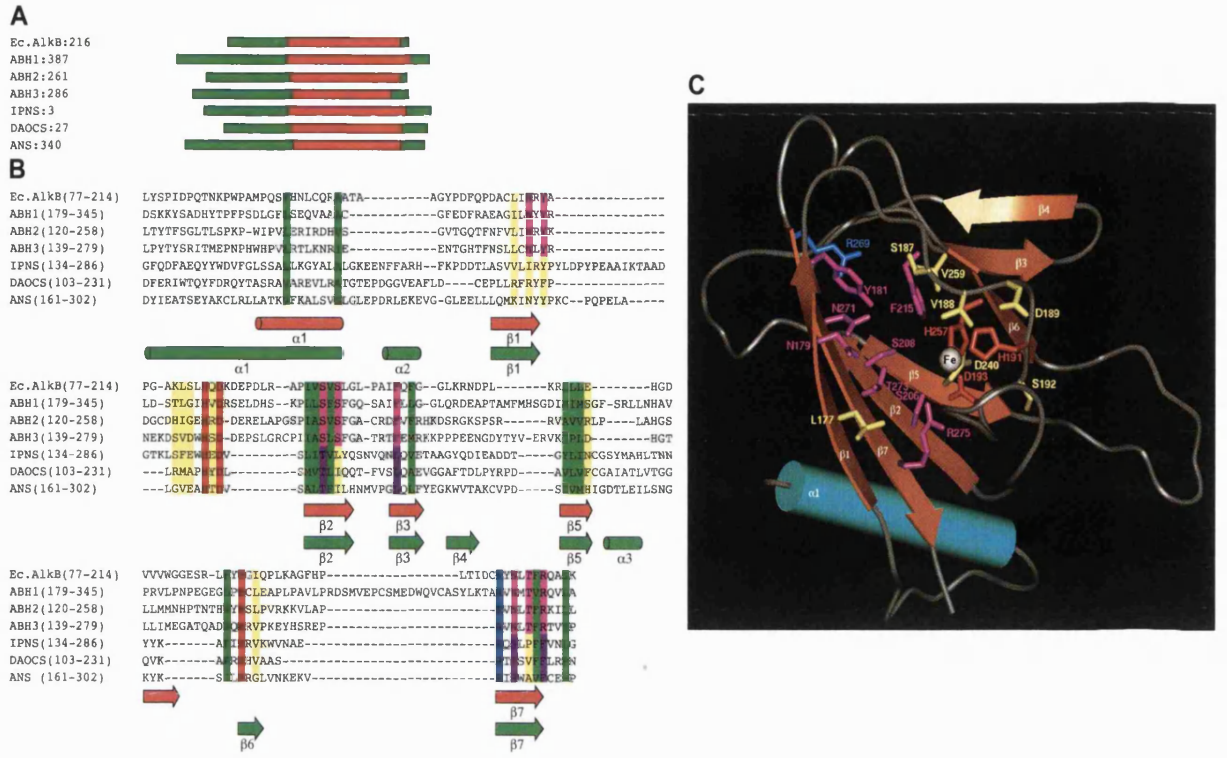


Fig. 2. (*A*) Approximate location of the core of the αKG–Fe(II) dioxygenase fold within AlkB, its putative human homologs, and three crystal structures from the αKG–Fe(II) dioxygenase superfamily: IPNS (isopenicillin N synthase from *Aspergillus nidulans*) (15), DAOCS (deacetoxycephalosporin C synthase from *Streptomyces clavuligerus*) (16), and ANS (anthocyanidin synthase from *Arabidopsis thaliana*) (17). (*B*) Multiple alignment within the core region of the AlkB homologs and the three crystal structures described above. Predicted secondary structures of ABH2 and ABH3 were calculated with the program PSI-PRED (19). The location of average secondary structure elements are shown as cylinders for α-helices and solid arrows for β-strands. These are colored green for known secondary structure locations (crystal structures) and red for predicted secondary structure locations (AlkB homologs). Residue background colors indicate the following properties: green, conserved hydrophobics; red, coordinate the Fe(II) ion; blue, conserved within the substrate binding pocket for all sequences; violet, conserved within the substrate binding pocket of the crystal structures only; magenta, conserved within the binding pocket for all AlkB homologs only; yellow, within the substrate binding pocket for all sequences but not conserved. (*C*) Three-dimensional model for the core region of ABH3. By using the known crystal structures of the three αKG-dependent DNA dioxygenases as templates and the alignment (*B*), a three-dimensional model was constructed for the core of ABH3 by using the program 3D-JIGSAW (24). The overall quality of side-chain packing and stereochemistry of the final model were checked by using program QUANTA 2000 (Accelrys, Orsay, France). The conserved α-helix is colored cyan. The β-strands, forming the jellyroll structure, are colored orange. The approximate location of the Fe(II) ion is represented as a white sphere. Side chains are colored according to the scheme described in *B*.

dependent dioxygenases in their C-terminal regions. By using the known crystal structures of three αKG- and Fe(II)-dependent dioxygenases (15–17) and the predicted secondary structures of AlkB (18), ABH2, and ABH3, a well defined alignment within the core region of these enzymes was obtained (Fig. 2 *A* and *B*). Both ABH2 and ABH3 contain the Fe(II)-binding motif, $HXDX_NH$, which is characteristic of αKG-dependent dioxygenases, and a conserved arginine that may be involved in αKG interactions (4). Considerable structural and sequence variation occurred outside the core regions. Based on the superimposed core regions of the three crystal structures, a 3D model was constructed for the core of ABH3 (19) (Fig. 2*C*), which exhibits a jellyroll topology containing one 3- and one 4- β-strand sheets.

Full length ABH2 and ABH3 cDNAs were obtained from the IMAGE consortium (IDs 5179189 and 5484474, respectively). The ORFs were subcloned to obtain plasmid constructs (pGST.ABH2 and pGST.ABH3) encoding GST-tagged fusion proteins. These plasmids were transformed into an *E. coli alkB22* mutant, and their ability to complement the defective reactivation of a MMS treated single-stranded DNA phage was examined. At the MMS doses used, both plasmids fully complemented the *alkB* phenotype so that phage reactivation was similar to that in an $alkB^+$ strain or in the *alkB22* mutant overexpressing *E. coli* AlkB protein (Fig. 1*B*). This complementation provided evidence that both ABH2 and ABH3 are functional human homologs of the *E. coli* AlkB protein.

Enzymatic Activities of ABH2 and ABH3. The GST- and hexahistidine- tagged ABH2 and ABH3 proteins were overexpressed in *E. coli* and purified by affinity chromatography. *E. coli* AlkB protein removes 1-meA and 3-meC from single stranded DNA (9). The ability of the two human proteins to repair these DNA modifications was therefore examined. The purified proteins were incubated with [^{14}C]MeI-treated poly(dA) or poly(dC) substrates, and the removal of 1-meA and 3-meC monitored by acid hydrolysis of the substrate and HPLC chromatography. Initially, assay conditions of pH 8 and 75 μM Fe(II) that are optimal for AlkB were used, but these were subsequently modified in the case of ABH2, which had greater activity at pH 7.5 and 25 μM Fe(II). Both GST- and hexahistidine-tagged ABH2 and ABH3 removed both 1-meA and 3-meC from the methylated polynucleotides (Fig. 3). These activities were totally dependent on αKG, stimulated by ascorbate and fully inhibited by EDTA (data not shown). These properties are characteristic of αKG- and Fe(II)-dependent dioxygenases. ABH2 and ABH3, like AlkB, are therefore members of this family of enzymes. AlkB oxidizes the methyl groups of 1-meA and 3-meC in DNA to release formaldehyde (9, 10). By assaying the release of ethanol-soluble [^{14}C]formaldehyde from ^{14}C-methylated poly(dA) or poly(dC), ABH2 was ≈4-fold more active on 1-meA than on 3-meC, whereas ABH3 was slightly more active (≈2-fold) on 3-meC than 1-meA (data not shown). Both enzymes

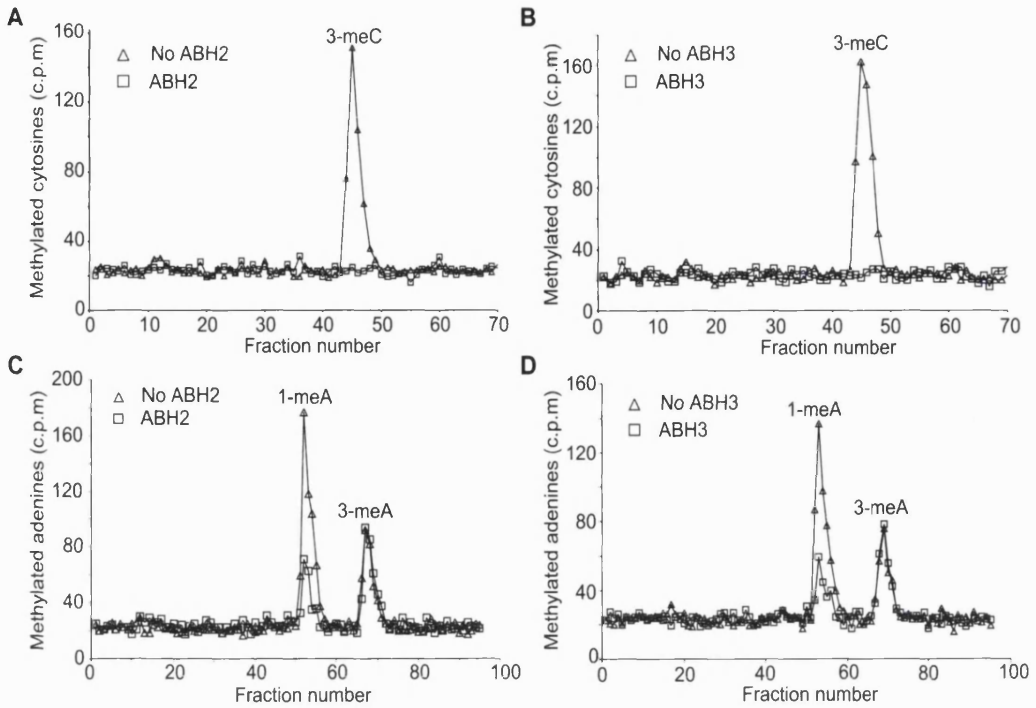


Fig. 3. Repair of 1-methyladenine and 3-meC residues by ABH2 and ABH3. ABH2 or ABH3 (1 nmol) were incubated with [^{14}C]MeI-treated poly(dC) (*A* and *B*) or poly(dA) (*C* and *D*) for 30 min under standard assay conditions. The ^{14}C-labeled methylated bases remaining in the substrates were analyzed by HPLC and scintillation counting. △, no enzyme; □, ABH2 or ABH3.

were also similarly active on a [^{14}C]MeI-treated oligonucleotide (50% GC) annealed to a complementary strand, which was verified to be >90% double stranded by *E. coli* exonuclease III digestion (data not shown).

E. coli AlkB directly reverts 1-meA in DNA to adenine by oxidative demethylation (9, 10). Here, we demonstrate that AlkB also directly reverts 3-meC to cytosine, and show that ABH2 and ABH3 have the same activity. A single stranded oligonucleotide

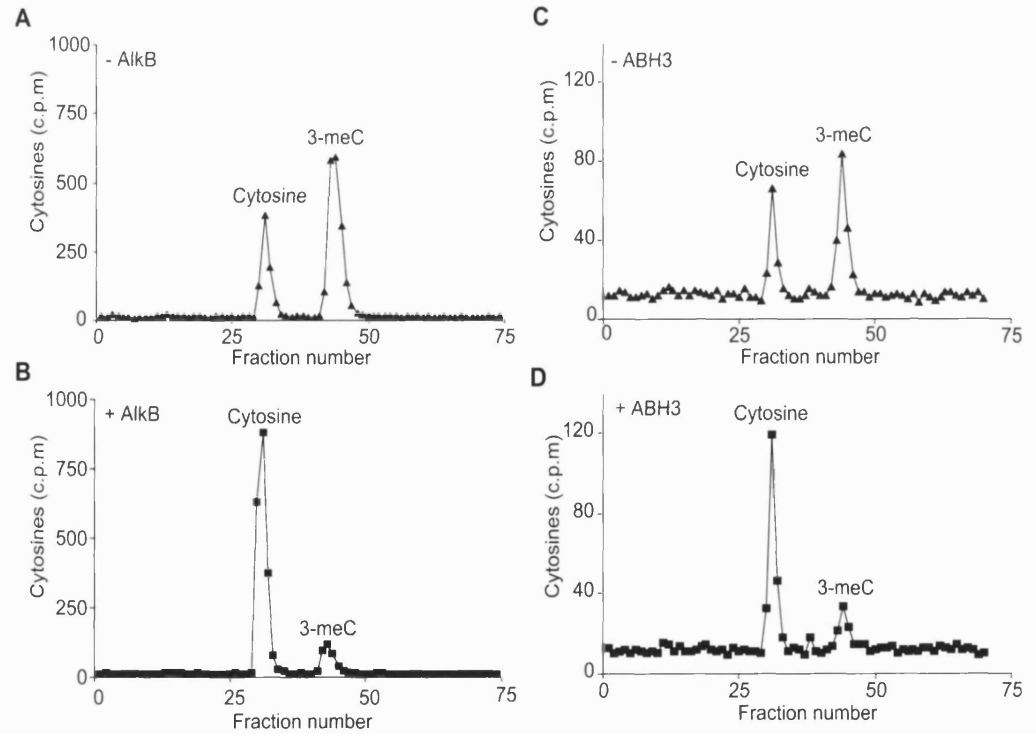


Fig. 4. Direct reversion of 3-meC to cytosine by *E. coli* AlkB and human ABH3. A thymine-rich oligodeoxynucleotide containing several [^{3}H]cytosine residues was heavily methylated with DMS and then incubated with or without 0.23 nmol AlkB (*A* and *B*) or 0.84 nmol ABH3 (*C* and *D*) at 37°C for 30 min. Purified ABH3 was less active than AlkB; therefore, a reduced amount of substrate was incubated with ABH3 so that reversal of 3-meC to cytosine was readily seen. The [^{3}H]cytosine residues remaining in the oligonucleotide were analyzed by HPLC and scintillation counting.

BIOCHEMISTRY

substrate containing [^{3}H]cytosine was synthesized by PCR using a single primer and then repeatedly methylated with DMS until 70% of the cytosines were converted to 3-meC. AlkB, ABH2, or ABH3 were incubated with this substrate and reversion of 3-me[^{3}H]C to [^{3}H]cytosine examined by acid hydrolysis and HPLC chromatography. Both AlkB and ABH3 directly reverted 3-meC to cytosine (Fig. 4), confirming that ABH3 acts by the same mechanism as AlkB; only partial reversion was achieved with ABH2 (data not shown) because of its lower activity on 3-meC and the lability of this enzyme.

Repair of Ethylated DNA Bases. Bacterial and human 3-methyladenine-DNA glycosylases and O^6-methylguanine-DNA methyltransferases also repair bulkier ethylated adducts, but usually with lesser efficiency (20). AlkB uses a completely different mechanism to these enzymes to remove aberrant methyl groups from DNA. It was therefore of interest to determine whether this activity can also repair ethyl adducts. To investigate whether *E. coli* AlkB can repair ethyl lesions generated in single stranded DNA, the ability of an *alkB* mutant to reactivate an EtI-treated single-stranded DNA phage was examined. Survival of the ethylated M13 phage was greatly decreased in the *alkB* mutant compared with the wild type, indicating that AlkB protein repairs DNA ethylation damage (Fig. 5*A*). A single-stranded substrate containing 1-ethyladenine (1-etA) was prepared by treating poly(dA) with [^{14}C]EtI. In standard assay conditions, *E. coli* AlkB protein released [^{14}C]ethanol-soluble material from this substrate. Analysis of the ethylated bases remaining in the polynucleotide showed that AlkB had completely removed the 1-etA lesions. A small amount of 3-etA was also present in the substrate, but was not removed by AlkB activity (Fig. 5*B*). AlkB therefore repaired 1-etA but not 3-etA, which parallels its previously observed activity on 1-meA but not 3-meA (9). By quantifying the release of [^{14}C]ethanol soluble material, AlkB was observed to repair 1-etA with a similar but slightly lower efficiency than 1-meA (data not shown). ABH2 and ABH3 proteins were also shown to remove 1-etA from ^{14}C-ethylated poly(dA). Both human proteins repaired 1-etA inefficiently, however, at ≈0.5% of the rate of 1-meA as followed by measuring the release of radioactive material from ^{14}C-ethylated poly(dA) and ^{14}C-methylated poly(dA) in parallel, and also by measuring the reduction of 1-etA remaining in the polynucleotide substrate (data not shown).

AlkB oxidizes the methyl group of 1-meA in DNA to release formaldehyde (9, 10). Depending on whether AlkB oxidizes 1-etA in DNA initially at carbon-1 or carbon-2 of the ethyl group, the product released could be either formaldehyde or acetaldehyde. To determine the nature of the aldehyde released, AlkB was incubated with [^{14}C]EtI treated poly(dA) and the ethanol soluble ^{14}C-product derivatized with DNPH and examined by HPLC. The radiolabeled product co-chromatographed with the DNPH derivative of acetaldehyde and not formaldehyde (Fig. 5*C*). AlkB therefore oxidizes at carbon-1 of the ethyl adduct generating acetaldehyde.

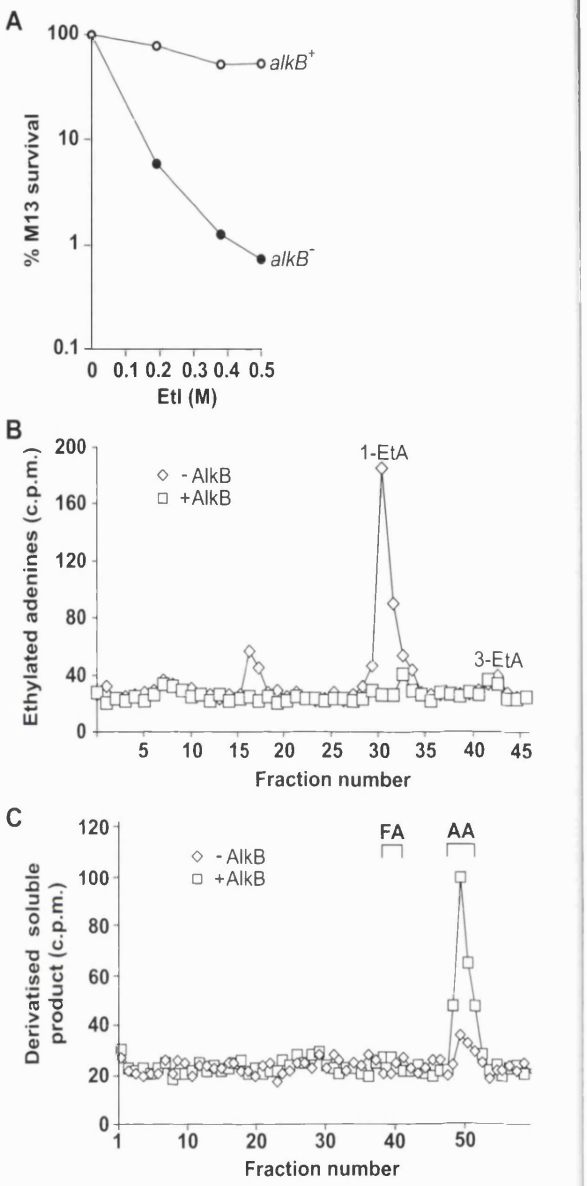


Fig. 5. Repair of 1-ethyladenine by *E. coli* AlkB. (*A*) Defective reactivation of EtI treated M13 single-stranded DNA phage in an *E. coli alkB* mutant. Host strains were *alkB*$^+$ and *alkB117*::*Tn3* derivatives of F′/ABII57. ○, *alkB*$^+$; ●, *alkB117*::*Tn3*. (*B*) Removal of 1-etA from [^{14}C]EtI-treated poly(dA) by 10 pmols AlkB protein. The ^{14}C-ethylated adenines remaining in the substrate were examined by HPLC. ◇, No AlkB; □, +AlkB. (*C*) AlkB oxidizes 1-etA to release acetaldehyde. AlkB (10 pmols) was incubated with [^{14}C]EtI-treated poly(dA), and the [^{14}C]ethanol soluble material derivatized with DNPH. ◇, No AlkB; □, +AlkB. DNPH derivatives of formaldehyde and acetaldehyde were run as markers and monitored at A_{254}.

Tissue Distribution of ABH2 and ABH3. The mRNAs of ABH2 and ABH3 were clearly detected in all of 10 different human tissues and several different cell lines derived from carcinomas by Northern hybridization, but were present at variable levels. ABH2 mRNA was present at relatively high levels in liver and bladder, whereas ABH3 was highest in spleen, prostate, bladder, and colon tissues. In the cell lines derived from carcinomas, ABH2 mRNA was relatively high in HeLa, whereas ABH3 mRNA was barely detectable in HeLa cells or in two of three cell lines derived from Burkitt's lymphomas (data not shown). There was no obvious correlation between expression of the ABH2 and ABH3 genes and the proliferative status of the tissue or cell type. Yellow fluorescent protein fusion constructs with the two human proteins showed that ABH2 localized to cell nuclei, where it was present diffusely throughout the nucleoplasm and accumulated in nucleoli, whereas ABH3 was present diffusely in nuclei and to a lesser extent in cytoplasm. Neither ABH2 nor ABH3 showed the punctate cytoplasmic staining typical of mitochondrial proteins (data not shown).

Discussion

We have identified two human homologs, ABH2 and ABH3, of the *E. coli* AlkB protein by complementation of the *alkB* mutant phenotype. Like AlkB, both proteins have conserved domains and biochemical requirements characteristic of αKG-dependent

dioxygenases, and are therefore new members of this family of enzymes. Both proteins had DNA repair activities similar to AlkB, and removed both 1-meA and 3-meC from single- and double-stranded DNA substrates. The relative efficiency of repair of the two aberrant bases differed between ABH2 and ABH3, but this effect was small and may not have biological significance. One clear difference between AlkB and the human enzymes was the rapid repair of both 1-meA and 1-etA by AlkB, whereas both human proteins were only weakly active on 1-etA. Our findings suggest that ABH2 and ABH3 have similar activities in DNA repair. Redundancy of human DNA repair enzymes is not unprecedented. Other examples include the uracil–DNA glycosylases, UNG and SMUG1, and the thymine–DNA glycosylases, TDG and MBD4 (21). The occurrence of alternative or backup DNA repair activities emphasizes the complex task of counteracting DNA damage in the large genomes of higher eukaryotes. Even though we were unable to detect an AlkB-like function for the previously reported ABH1 protein, highly conserved residues in AlkB, ABH2, and ABH3 are also present in ABH1 (Fig. 2*B*). It is therefore possible that ABH1 has a related activity, but with a different substrate specificity to AlkB, ABH2, and ABH3.

The molecular model of the ABH3 core (Fig. 2*C*) indicates residues involved in catalysis and substrate specificity of the AlkB family of enzymes. The conserved residues of AlkB, ABH2, and ABH3 (shown by the magenta and blue colored side chains in Fig. 2*C*) may be essential for the very similar or identical function of these enzymes. Subtle differences in their relative activities and specificities may be caused by variation in side chains in and around the substrate-binding pocket (shown by yellow colored side chains in Fig. 2*C*). The 3D model will help to identify relevant target residues for site-directed mutagenesis. However, high-resolution crystal structures of human AlkB homologues would be required before detailed atomic level screening and analysis can be used to design small molecule antagonists of these enzymes. Such inhibitors could be useful adjuncts to cancer therapy that uses simple alkylating agents.

AlkB and its human homologs enzymatically reverse 1-meA and 3-meC in DNA to adenine and cytosine by oxidative demethylation, indicating an accurate mode of DNA repair. Other activities that directly reverse DNA damage are the suicidal O^6-methylguanine–DNA methyltransferases that transfer the DNA methyl group to a cysteine residue in the protein (22–24), light requiring photolyases that monomerize cyclobutyl pyrimidine dimers and the *Bacillus subtilis* spore SP lyase, which cleaves the unique spore photoproduct (5-thyminyl-5,6-dihydrothymine dimer) to two thymine residues by a free-radical mechanism (25). Photolyase and SP lyase have not been found in human cells. Thus, the only two known mechanisms for direct reversal of DNA base damage in mammalian cells are caused by the O^6-methylguanine-DNA methyltransferase and the activities characterized here of the ABH2 and ABH3 DNA dioxygenases.

We thank Peter Robins for help in protein purification.

1. Bodell, W. J. & Singer, B. (1979) *Biochemistry* **18,** 2860–2863.
2. Boiteux, S. & Laval, J. (1982) *Biochimie* **64,** 637–641.
3. Kataoka, H., Yamamoto, Y. & Sekiguchi, M. (1983) *J. Bacteriol.* **153,** 1301–1307.
4. Sedgwick, B. & Lindahl, T. (2002) *Oncogene Rev.*, in press.
5. Chen, B. J., Carroll, P. & Samson, L. (1994) *J. Bacteriol.* **176,** 6255–6261.
6. Dinglay, S., Trewick, S. C., Lindahl, T. & Sedgwick, B. (2000) *Genes Dev.* **14,** 2097–2105.
7. Aravind, L. & Koonin, E. V. (2001) *Genome Biol.* **2,** research0007.1–0007.8.
8. Thornburg, L. D., Lai, M.-T., Wishnok, J. S. & Stubbe, J. (1993) *Biochemistry* **32,** 14023–14033.
9. Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T. & Sedgwick, B. (2002) *Nature* **419,** 174–178.
10. Falnes, P. O., Johansen, R. F. & Seeberg, E. (2002) *Nature* **419,** 178–182.
11. Kodaira, K.-I., Oki, M., Kakikawa, M., Kimoto, H. & Taketo, A. (1996) *J. Biochem.* **119,** 1062–1069.
12. Houlgate, P. R., Dhingra, K. S., Nash, S. J. & Evans, W. H. (1989) *Analyst* **114,** 355–360.
13. Wei, Y., Carter, K. C., Wang, R. & Shell, B. K. (1996) *Nucleic Acids Res.* **24,** 931–937.
14. Contreras-Moreira, B. & Bates, P. A. (2002) *Bioinformatics* **18,** 1141–1142.
15. Roach, P. L., Clifton, I. J., Fulop, V., Harlos, K., Barton, G. J., Hajdu, J., Andersson, I., Schofield, C. J. & Baldwin, J. E. (1995) *Nature* **375,** 700–704.
16. Lee, H.-J., Lloyd, M. D., Harlos, K., Clifton, I. J., Baldwin, J. E. & Schofield, C. J. (2001) *J. Mol. Biol.* **308,** 937–948.
17. Wilmouth, R. C., Turnbull, J. J., Welford, R. W. D., Clifton, I. J. & Prescott, A. G. (2002) *Structure (London)* **10,** 93–103.
18. Aravind, L., Walker, D. R. & Koonin, E. V. (1999) *Nucleic Acids Res.* **27,** 1223–1242.
19. Jones, D. T. (1999) *J. Mol. Biol.* **292,** 195–202.
20. Tudek, B., Van Zeeland, A. A., Kusmierek, J. T. & Laval, J. (1998) *Mutat. Res.* **407,** 169–176.
21. Wood, R. D., Mitchell, M., Sgouros, J. & Lindahl, T. (2001) *Science* **291,** 1284–1289.
22. Olsson, M. & Lindahl, T. (1980) *J. Biol. Chem.* **255,** 10569–10571.
23. Foote, R. S., Mitra, S. & Pal, B. C. (1980) *Biochem. Biophys. Res. Commun.* **97,** 654–659.
24. Bogden, J. M., Eastman, A. & Bresnick, E. (1981) *Nucleic Acids Res.* **9,** 3089–3103.
25. Rebeil, R. & Nicholson, W. L. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 9038–9043.
26. Bates, P. A. & Sternberg, M. J. E. (1999) *Proteins*, Suppl. 3, 47–54.