The Versatile *E. coli* Adaptive Response Protein AlkB Mitigates Toxicity and Mutagenicity of Etheno-, Ethano-, and Methyl-modified Bases *in vivo*

By

Lauren Elizabeth Frick

S.B., Materials Science and Engineering (2003)
Massachusetts Institute of Technology

Submitted to the Department of Biological Engineering in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Molecular Systems Pharmacology and Toxicology

at the

Massachusetts Institute of Technology

September 2007

Signature of Author: _____
Department of Biological Engineering
August 24, 2007

Certified by: _____
John M. Essigmann
William and Betsy Leitch Professor of Chemistry and Toxicology
Thesis Supervisor

Accepted by: _____
Alan J. Grodzinsky
Professor of Electrical, Mechanical, and Biological Engineering
Chair, Course XX Graduate Program Committee

This doctoral thesis has been examined by a committee of the Department of Biological Engineering as follows:

Peter C. Dedon
Professor of Toxicology and Biological Engineering
Chair

John M. Essigmann
William and Betsy Leitch Professor of Chemistry and Toxicology
Thesis Supervisor

Michael R. Volkert
Professor of Molecular Genetics and Microbiology
University of Massachusetts Medical School

The Versatile *E. coli* Adaptive Response Protein AlkB Mitigates Toxicity and Mutagenicity of Etheno-, Ethano-, and Methyl-modified Bases *in Vivo*

by
Lauren Elizabeth Frick

Submitted to the Department of Biological Engineering on August 24, 2007 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Molecular Systems Pharmacology and Toxicology

The *Escherichia coli* AlkB protein is an exceptionally versatile DNA repair enzyme. Its expression is induced upon exposure to alkylating agents as part of the Ada-mediated adaptive response. This member of the α-ketoglutarate/iron(II)-dependent dioxygenase family was originally discovered to reverse directly methylated lesions formed preferentially in single-stranded regions of DNA, such as 1-methyladenine and 3-methylcytosine. Repair proceeds via an oxidative demethylation pathway, in which the aberrant methyl group is hydroxylated and spontaneously lost as formaldehyde. Since these early studies, the list of lesions repaired by AlkB through this pathway has been extended to include 1-methylguanine, 3-methylthymine, 3-ethylcytosine, and 1-ethyladenine.

Furthermore, the protein possesses a second, distinct chemical mechanism through which it can repair another class of lesions, the etheno-adducts formed by the reaction of DNA with metabolites of the carcinogen vinyl chloride or with breakdown products generated by lipid oxidation. In this case, direct repair proceeds through epoxidation of the etheno bond, creating an intermediate that hydrolyzes to a glycol form and finally releases the two-carbon bridge as glyoxal, restoring the unadducted adenine or cytosine. Thus, the AlkB protein bridges the repair of alkylative lesions with those induced by oxidative stress and embodies the multi-faceted protection required to preserve genomic stability and coding information despite the constant threats to which organisms are exposed.

Herein, we exploit and characterize a pair of *E. coli* strains differing only in AlkB status to demonstrate the ability of AlkB to repair the etheno-lesions, the structural analog 1,N^6-ethanoadenine (EA), and 3-methyluracil *in vivo*. Additionally, we establish the ability of the EA "repair product" to form interstrand cross-links in certain sequence contexts of duplex DNA. We also show that although the adaptive response proteins repair lesions generated by oxidative stress, oxidative agents do not induce expression of the response. Finally, we establish that certain hypothesized substrates for AlkB are not in fact repaired by the enzyme, nor are they repaired by another adaptive response protein, AidB. This work extends the current knowledge regarding the amazing ability of AlkB to protect cellular nucleic acids from damage arising from a diverse array of both endogenous and exogenous sources.

Thesis Supervisor: John M. Essigmann
Title: William and Betsy Leitch Professor of Chemistry and Toxicology

ACKNOWLEDGEMENTS:

I would not have completed this journey without the help, guidance, and support of many. I owe countless thanks to those who have assisted me along my way, and I will try here to acknowledge as many as possible.

To my advisor, John, a fantastic teacher who was instrumental in my interest in, acceptance to, and now completion of graduate school in Biological Engineering. Thank you for taking the time to guide me through the PhD process and to teach me the things that your wisdom and experience have taught you. I have enjoyed the opportunity to work under and learn from you and I thank you for all the ways in which you have prepared me to move forward in my career and in life.

To my thesis committee, Professors Michael R. Volkert, Peter C. Dedon, and Leona D. Samson, who are always full of good ideas, criticisms and advice. I appreciate all of the time that each of you has put in on my behalf and all of the things that you have taught me about academia and the scientific process. A special thanks to Mike for the many miles he has logged traveling to and from Boston for my meetings.

To all those who took the time to read and offer valuable feedback and comments on the thesis during its preparation, including Kyle Proffitt, Eunsuk Kim, Hector Hernandez, Charles Morton, and Jim Delaney, thank you, I couldn't have completed it without you. Special thanks to Hector for his help with DNA structure analysis and for the creation of the beautiful figure in Chapter 5.

To the Essigmann lab, my colleagues and friends for the past four years. For all the times that I needed help, advice, or a place to vent, you were there. A special thanks to Jim Delaney, who took the time out of his own work to teach me critical things such as what a pipette looks like and how to run a gel. I very much appreciate all of the knowledge you've shared over the years and the assistance you've been so reliable in providing. A similar thanks goes to Peter Rye and Sarah Delaney, whose vast experience and willingness to share and to teach the unenlightened were crucial to the success of many experiments.

To Peter Slade, whose generous assistance with mass spectrometry was incredibly helpful and left me with a deeper understanding of the intricacies and uses of the technique.

To my friend, mentor, and sounding board Sarah Delaney, who taught me how to be a graduate student, how to do research, how to just get brave, dive in, and try something that I'd never done before, how to prepare to give an experiment the best shot at working, and how to figure out what went wrong if it failed anyway. Your successes have been an inspiration to me, and I would not have made it through this process without your friendship and advice (about both lab and life), and for those I am ever grateful.

To my fellow BE graduate students, thanks for the laughs, the memories, and the empathy. Knowing that I was not alone in this quest was hugely important to my successful completion of it. Special thanks to my printmaster, Josh Apgar.

To my volleyball girls, my most reliable source of stress relief. On and off the court, you guys are endless fun, and as both teammates and friends, you have been vitally important to me throughout my time at MIT.

To my super-supportive ex-roommates, Carolyn and JJ, you guys are the greatest. From Kiraz dates to camping to triathlons (almost!), you two have kept me grounded while expanding my horizons. We'll all be Doctors soon!

To James Peters, who exemplified the highest level of quality mentoring. The impending retirement of my beloved high school physics teacher, whose career dream was to send a student to MIT, inspired my application to and subsequent presence at the Institute. I thank him, and the few other supportive Norwinites, Patricia Coury, Gary Good, and Marsha Clark, for believing in me and pushing me to test the limits of what a clueless girl from western Pennsylvania could achieve, for without them, I certainly would not be here.

To my family, for always feeding me (and perhaps even believing in) the cliché that, if I worked hard enough, I could be whatever I wanted to be when I grew up. To my father, for passing along his passion for numbers and science (and for keeping his propensity for giving really long, drawn-out answers to simple questions to himself!). To my mother, for devoting herself to giving her children every possible opportunity to excel in life. Thank you for your hard work, your high standards, and your sacrifices. Your unwavering love and unpressuring support have been critical along the path to this milestone. To my brother Brian, who always gave me a hard time. About everything. Thank you for pushing me and teaching me not to sell myself short. To my sister Jessica, whose independence, tenacity, and self-sufficiency are inspirational. You will forge your own path in this world, and the rest of us will only follow in awe.

To my incredible boyfriend Charles, for all his love and support throughout the years and especially during this rather frazzling process. Your unselfish patience, advice, and attempts at humor were critical to the maintenance of my remaining sanity. From the day I set foot on campus, you have been there for me, keeping me in line. For all the nights you slept on the couch to ensure the quality of my rest and for the countless hours you've spent allowing me to vent my anxieties, concerns and frustrations, I am indebted. You are far and away the best thing MIT has given me. The first star I see...

TABLE OF CONTENTS:

ABBREVIATIONS:

εA	1,N^6-ethenoadenine
εC	3,N^4-ethenocytosine
αKG	α-ketoglutarate
2AP	2-aminopurine
A	adenine
aa	amino acids
amu	atomic mass units
AP	abasic (apurinic/apyrimidinic)
BCNU	1,3-bis(2-chloroethyl)-1-nitrosourea
BER	base excision repair
BME	β-mercaptoethanol
bp	base pairs
C	cytosine
CAA	chloroacetaldehyde
CoA	coenzyme A
CPD	cyclobutane pyrimidine dimer
CRAB	competitive replication of adduct bypass
DCPIP	2,6-dichlorophenolindophenol
DMS	dimethylsulfate
DMSO	dimethylsulfoxide
dsDNA	double-stranded DNA
e3A	3-ethyladenine
e3C	3-ethylcytosine
e3U	3-ethyluracil
EA	1,N^6-ethanoadenine
EDTA	ethylenediaminetetraacetic acid
EMS	ethyl methanesulfonate
ENU	*N*-nitroso-*N*-ethylurea
ESI-TOF	electrospray ionization – time of flight
FAD	flavin adenine dinucleotide
g	gravity
G	guanine
GC-MS	gas chromatography – mass spectrometry
h	hour
IPTG	isopropyl-beta-D-1-thiogalactopyranoside
IVD	isovaleryl coenzyme A dehydrogenase
kb	kilobase
kDa	kilodalton
Lrp	leucine responsive protein
m1A	1-methyladenine
m1G	1-methylguanine
m3A	3-methyladenine
m3C	3-methylcytosine
m3G	3-methylguanine

m3T	3-methylthymine
m3U	3-methyluracil
m7A	7-methyladenine
m7G	7-methylguanine
m8G	8-methylguanine
MALDI-TOF	matrix assisted laser desorption ionization - time of flight
MeCl	Methyl chloride
MeI	methyl iodide
MMS	methyl methanesulfonate
MNNG	*N*-methyl-*N*'-nitro-*N*-nitrosoguanidine
MNU	*N*-nitroso-*N*-methylurea
MPTE	methylphosphotriester
O2mC	O^2-methylcytosine
O2mT	O^2-methylthymine
O4mT	O^4-methylthymine
O6mG	O^6-methylguanine
ONPG	o-nitrophenyl-β-D-galactoside
PAGE	polyacrylamide gel electrophoresis
PCI	phenol/chloroform/isoamyl alcohol
PCR	polymerase chain reaction
PFBHA	*O*-(2,3,4,5,6-pentafluorobenzyl)hydroxylamine
REAP	restriction endonuclease and postlabeling analysis of mutation frequency
ROS	reactive oxygen species
SDS	sodium dodecyl sulfate
SIN-1	3-morpholinosydnonimine
ssDNA	single-stranded DNA
ssRNA	single-stranded RNA
T	thymine
TE	Tris - EDTA
TES	Tris – EDTA – sodium chloride
THF	tetrahydrofuran
TLC	thin layer chromatography
U	uracil
UDG	uracil DNA glycosylase
X-gal	5-bromo-4-chloro-3-indolyl-β-galactoside

LIST OF FIGURES:

LIST OF TABLES:

CHAPTER 1:

Introduction to the *Escherichia coli* Adaptive Response to Alkylating Agents

Alkylation Damage

DNA damaging agents are ubiquitous and hence unavoidable. Food, air, tobacco smoke, environmental agents, and chemotherapeutics are all exogenous sources of exposure to chemicals that react deleteriously with DNA, either directly or following metabolic activation (*1-3*). Additionally, such agents are created spontaneously by cells as a result of normal metabolism as well as at elevated levels in response to inflammation (*4*). In order to maintain replicative and transcriptional potential and to preserve genomic integrity, organisms have evolved several methods of combating the many types of DNA damage.

Alkylating agents, one major class of DNA damaging species, cause the addition of methyl or larger alkyl groups to the DNA bases at any of fifteen possible sites (Figure 1.1) in addition to the anionic oxygen of the phosphodiester backbone (*5-7*). The spectrum of lesions generated by a given alkylating agent is determined by the mechanism of modification as well as the secondary structure of the target DNA (Table 1.1). Alkylators fall into two chemical categories (Figure 1.2): S_N1 agents, which typically react through a "unimolecular" mechanism (meaning spontaneous departure of the leaving group) with ring nitrogens, exocyclic oxygens of the base, and oxygens of the sugar-phosphate backbone, and S_N2 agents, which act through a "bimolecular" reaction (leaving group departure occurs only upon reaction with another species) to target mainly ring nitrogens, including those exposed only in single stranded DNA-- i.e., N1 of adenine and N3 of cytosine. Additionally, methylated lesions can arise through the reaction of DNA with methyl radicals. These modified base lesions are both toxic and mutagenic (*8-10*), and their repair facilitates cellular survival.

Endogenous Alkylating Agents. Cells unexposed to external sources of alkylation accumulate methylated DNA bases spontaneously, indicating that compounds capable of modifying DNA exist or are generated *in vivo* (*7,11*). *S*-Adenosyl-methionine (SAM), the methyl donor molecule utilized in many cellular enzymatic methylation processes, has been shown to alkylate DNA (*12-15*) by an S_N2 mechanism at purine ring nitrogens, generating mostly 7-methylguanine (m7G) and 3-methyladenine (m3A) (*14,15*).

However, modulating SAM levels over a 100-fold range in *E. coli* does not significantly affect the levels of spontaneous mutational events (*16*), suggesting that another source of methyl damage is probably responsible for the observed spontaneous mutagenesis.

Nitrosation of amino acids and peptides by endogenously-formed nitric oxide in bacteria is also thought to produce compounds capable of alkylating DNA (*17*) and could be responsible for the spontaneous mutations observed in *E. coli* (*18-20*). Nitrosation of glycine and its derivatives generates compounds that alkylate guanine at the O^6 position, creating mutagenic adducts (*17,21*).

While bacterial generation of alkylating compounds may be particularly relevant to adduct levels in the intestine, endogenous damage in humans is not limited to that generated by microbial tenants. Human metabolic processes also generate nitrosamines, *N*-nitroso species, bisulfite and hydroxylamines in quantities large enough to be considered a threat to genomic integrity (*22,23*). Additionally, the reactive nitrogen species generated in high concentration at sites of inflammation can react with and damage DNA in this manner (*4*), perhaps with grim consequence, as the DNA hypermethylation associated with inflammation is also associated with increased mortality in chronic kidney disease (*24*). Hypermethylation has also been shown to contribute to uncontrolled growth in a tumor model (*25*) and aberrant DNA methylation is linked to carcinogenesis (*26*), demonstrating that repair of these lesions in DNA is critical to the maintenance of organismal health.

Exogenous Alkylating Agents. In addition to the internal generation of alkylating agents, humans are continuously exposed to reactive compounds in the environment. For example, methyl chloride (MeCl) is generated industrially as well as naturally through the detoxification of chlorine by some algae and fungi and is thought to be the most abundant environmental alkylating agent known (*27*). A significant portion is also produced by the decomposition of plant material (*28*). Other S_N2 methyl halides capable of modifying DNA, such as bromomethane and iodomethane (*29*), are generated in marine environments, and some of these compounds are found in measurable quantities in ambient air (*30*).

Food can also be a source of exposure to exogenous alkylating agents, as reactive, mutagenic nitroso-species are formed by the reaction of nitrite with amines and amides (*2,18,31*). Cured foods and fish in particular contain high levels of alkylating *N*-nitroso compounds (*31*).

Tobacco smoke contains several potently carcinogenic nitroso compounds, both tobacco-specific and more common (*32*), that have been shown to cause DNA alkylation at O^6G as well as other sites (*33*). Additionally, pollutants and pesticides are capable of alkylating DNA (*34-38*).

Beyond incidental environmental exposure, humans can be subjected to very high doses of alkylating agents via chemotherapy (*39,40*). Indeed, alkylating compounds are the oldest type of anti-cancer drug, and they are still relied upon today to combat certain forms of cancer (*39,40*). DNA damage seems to be a critical component of the efficacy of these drugs, and increased DNA repair is a common cause of developed resistance (*41*). Both mono- and bi-functional alkylators are employed in this capacity, and these drugs function by creating a range of adducts from simple methyl lesions to interstrand cross-links (*40,42-45*).

In order to study the effects of alkylation damage on living systems in a controlled environment, several chemicals have come into common use as alkylating agents in the laboratory. These include the S_N1 agents *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine (MNNG), *N*-nitroso-*N*-methylurea (MNU), and *N*-nitroso-*N*-ethylurea (ENU), the S_N2 agents methyl methanesulfonate (MMS), ethyl methanesulfonate (EMS), and dimethylsulfate (DMS), and the bifunctional 1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU).

In addition to alkylation damage induced by environmental factors, the cellular genome is also subject to other types of exogenous damage, such as that caused by ultraviolet (UV) light or ionizing radiation, that create a wide spectrum of lesions against which cells and organisms must defend themselves (*46-55*).

Alkylation Repair

Because alkylation damage is essentially universal, and its effects on genomic integrity so deleterious, virtually all organisms encode multiple repair activities, both redundant and complementary, to enable survival despite such heavy damage loads.

Direct Reversal. The most faithful method of repairing modified DNA is direct reversal of base damage, which offers the advantage of allowing repair to proceed on intact DNA, ensuring the conservation of coding information and avoiding the creation of repair intermediates that can be even more toxic to a cell than the original lesion. The *E. coli* genome encodes several proteins that restore native DNA through the various methods of direct reversal, some of which are involved in the adaptive response to alkylating agents, outlined below.

Methyltransferase. The mutagenic adduct O^6-methylguanine (O6mG) mispairs with thymine during replication, creating G:C to A:T transitions (*56*), the most common mutation seen following cellular exposure to S_N1 alkylating agents (*57,58*). This mutagenic and carcinogenic adduct (*59,60*) is repaired in a single step by alkyltransferase proteins (*61,62*) through a suicidal S_N2 transfer of the methyl group to an active site cysteine residue. *E. coli* possesses two versions of this repair activity, the constitutively-expressed Ogt and the inducible Ada. Both are capable of repairing O^4-methylthymine (O4mT) in addition to O6mG, and the Ada protein can remove the methyl group from the *S*-diastereomer of methylphosphotriesters (MPTEs) in the DNA backbone as well (*7*) (Figure 1.3). Furthermore, DNA methyltransferase proteins are capable of removing methyl groups from RNA bases (*63*), although whether this reaction takes place *in vivo* is not yet clear. Alkyltransferase proteins are conserved among many species (*5,64-67*), which suggests that the value of error-free damage repair warrants the high energetic cost of the continuous synthesis required by a suicidal protein.

Photolyase. Although not a mechanism employed in the repair of methylated bases, photolyases represent a noteworthy method of direct reversal of damage. Exposure of

cells to UV light causes the formation of cyclobutane pyrimidine dimers (CPDs) and 6-4 pyrimidine-pyrimidones, which interfere with the replication and transcriptional processes of the cell (*68-72*). A second method of direct reversal, termed photoreactivation, is employed in the repair of CPDs. *E. coli* DNA photolyase is a 49 kDa enzyme (*73,74*) that binds specifically to CPDs and, with the aid of two chromophores, reverses the cyclization of the two pyrimidines in a light-dependent electron-transfer reaction (*75-77*). This class of direct reversal enzymes is also highly conserved among species, with the notable exception of placental mammals, in which photolyase activity seems to be absent (*7,78*). Importantly, *E. coli* DNA photolyase is capable of repairing *cis-syn* U-U cyclobutane dimers in RNA in a light-dependent manner (*79*), and this reaction constitutes a second example of demonstrated RNA repair.

Dioxygenase. The most recently discovered way in which cells employ direct reversal in DNA repair is through the mechanism of oxidative demethylation exemplified by the AlkB family of proteins (Figure 1.4). These enzymes use non-heme iron (II) and a dioxygen molecule as co-factors and α-ketoglutarate as a co-substrate to hydroxylate erroneous methyl or larger alkyl groups from damaged bases, creating an unstable hydroxyalkyl intermediate that decomposes in water to release formaldehyde (or the corresponding aldehydic species for larger alkyl groups) and the repaired base (*80,81*). AlkB can also repair unsaturated cyclic adducts by exploiting the same reactive iron(IV)-oxo active site intermediate to epoxidize the double bond, creating a glycol species that is released as glyoxal (*82,83*). While many species encode proteins homologous to AlkB, many others do not (*84-91*), raising the possibility that proteins of unrelated sequence can perform this chemistry (*92*), that another system is capable of repairing the AlkB substrate adducts, or that failure to repair these cytotoxic and mutagenic lesions is somehow tolerated in certain species.

Base Excision Repair. In addition to repair by one of the aforementioned direct reversal mechanisms, DNA bases damaged by methylation can be subject to excision from the polynucleotide strand by one of several glycosylase enzymes, each of which is specific

for a particular set of substrates. *E. coli* utilize two methylation-specific glycosylases, the constitutively-expressed Tag (also called 3-methyladenine DNA glycosylase I or TagA) and the inducible AlkA (also called 3-methyladenine DNA glycosylase II or AidA), to increase the rate of *N*-glycosyl bond cleavage of alkylated bases by three orders of magnitude over that of unmodified bases (*7*). The Tag protein specifically removes m3A from DNA, and is also capable of excising the analogous 3-ethyladenine (e3A) and 3-methylguanine (m3G) lesions to some extent (*93*). In contrast, the inducible AlkA protein (Figure 1.5) can efficiently excise m3A, 7-methylguanine (m7G), 7-methyladenine (m7A), and 1,N^6-ethenoadenine (εA) (*7*), and possesses some activity on O^2-methylthymine (O2mT) and O^2-methylcytosine (O2mC) as well (*94*). Excision of alkylated bases by these enzymes represents the first step in the base excision repair (BER) pathway, and the resulting abasic (AP) site is then processed by one of the AP endonucleases, which nicks the DNA strand by hydrolyzing the phosphodiester bond 5' to the AP site. The nicked strand can then be acted upon by an enzyme with DNA-deoxyribophosphodiesterase activity, creating a single nucleotide gap. This gap is filled by DNA synthesis using the information encoded on the complementary strand. Following resynthesis, the strand being repaired is sealed by the action of DNA ligase, completing the process and restoring intact double-stranded DNA. Although base excision repair is the pathway most often employed (*7*), it carries with it certain risks, as it creates AP site and nicked strand intermediates, both of which are known to be extremely toxic or mutagenic to a cell if encountered during replication (*53,55,95-97*).

Nucleotide Excision Repair. A second repair process involving the incision and resynthesis of a damaged DNA strand is termed nucleotide excision repair. This pathway was discovered through the observation that irradiated *E. coli* possessed two methods of removing the radiation-induced dimers from the genome, one of which depended on light (photolyase, discussed above) and one of which did not (*98,99*). Importantly, the light-independent process caused dimers to disappear from the bulk DNA and to appear in the acid soluble phase, suggesting that they had not been reversed, but rather excised from the genome. Further characterization of this pathway revealed it to be a five-step process

requiring damage location and recognition, incision of the damaged DNA strand, excision of a stretch of nucleotides, resysnthesis of the removed region, and ligation to seal the newly-repaired strand (*7*). In *E. coli*, this process is carried out through the coordinated effort of three proteins, UvrA, UvrB, and UvrC. Bacterial strains defective in one or more of these proteins are sensitive to many types of damage, including alkylation (*100-103*), suggesting that nucleotide excision repair is an active pathway through which wild-type *E. coli* protect themselves against the detrimental effects of exposure to alkyating agents.

The *Escherichia coli* Adaptive Response to Alkylating Agents

The adaptive response to alkylating agents is one multi-faceted strategy employed by *E. coli* to preserve genomic integrity in the face of attack by certain DNA damaging agents, the best characterized of which form small alkyl-DNA adducts. This response to alkylation damage involves the up-regulation of expression of four genes, three of which produce proteins of established function (Figure 1.6). The gene products produced as part of the adaptive response are now known to operate by at least three different mechanisms of DNA repair.

Decades ago, it was discovered that the relationship between mutation frequency in chemically-treated *E. coli* and the length of exposure to the alkylating agent was not linear, but rather increased initially and then reached a plateau (*104-106*). Interestingly, the height of that plateau was related to the dose of alkylating agent to which the bacteria were exposed (*107*). These observations, combined with the subsequent discovery that exposure of *E. coli* to sub-lethal doses of methylating agents renders them resistant to higher concentrations of drug, led to the discovery of the so-called "adaptive response" (*108,109*). The inducible protective effect required active protein synthesis (*109*), suggesting that it was due to the production of gene products that could defend against the toxic and mutagenic effects of simple alkylating agents. Later studies demonstrated that the phenomenon involved increased expression of four genes, *ada*, *alkA*, *alkB*, and *aidB* (*110-114*).

Ada. The first protein product of the adaptive response to be identified was Ada, which was isolated as the final location of radioactively labeled methyl groups in *in vitro* repair assays (*115*). Further characterization of Ada revealed its bifunctional nature.

Ada as a Multi-Substrate Repair Protein. The bifunctional methyltransferase Ada is composed of two subunits linked by a hinge region and is capable of carrying out two types of repair reactions, one mediated by each end of the protein (Figure 1.3). The first involves the direct reversal of the stray methyl group from the highly mutagenic O^6-methylguanine lesion through transfer to a C-terminal active site cysteine residue, which as been identified as Cys321 (*62*). This mechanism can also be employed to remove the methyl group from the less frequent O^4-methylthymine lesion. A stretch of active site residues is highly conserved among DNA-methyltransferase proteins of various species, and these proteins contain a helix-turn-helix motif and an arginine finger region that are implicated in DNA binding and substrate base recognition through base-flipping, respectively (*116-118*).

The second repair reaction carried out by Ada is a similar activated-cysteine-mediated transfer of a stray methyl group, but in this case from the *S*-diastereomer of the apparently innocuous methylphosphotriester lesion. This transfer reaction takes place in the N-terminal active site of the protein, specifically with Cys38 (*119,120*), which is coordinated to a divalent zinc cation, along with three other cysteine residues, Cys42, Cys69, and Cys72 (Figure 1.3) (*6*). The specificity for the *S*-isomer is explained by the protein structure, which showed Cys38 to be solvent-exposed and the *S*-isomer to be positioned towards the free solution, while the *R* isomer projects towards the major groove of the helix, making it inaccessible to the active site of Ada (*121*). This methylation reaction alleviates the normal repulsion that exists between the negatively-charged cysteine center and the DNA backbone (*122*), increasing the affinity of Ada for DNA and converting the protein into a transcription factor under regulation by post-translational modification.

Both of these repair actions are suicidal, resulting in loss of the repair function of the Ada protein in a stoichiometric manner, and requiring the activity to be replenished

through new protein synthesis. It has been shown to be possible, however, to mutate Ada such that it becomes a catalytic enzyme. Substitution of Cys38 with a glycine residue and addition of an exogenous thiol source conferred weak enzymatic activity to the mutant protein (*120*).

Interestingly, when repair occurs via the Ada pathway, the restoration of a single alkylated base consumes an entire protein molecule. Although seemingly energetically expensive, this repair method allows reversal of damage in a single step, without compromising the integrity of the DNA backbone or the coding content of the DNA strand. For these reasons, direct reversal of base damage, such as that exhibited by the suicidal Ada protein, can be advantageous.

Ada as a Transcriptional Regulator. The adaptive response is regulated by the bifunctional methyltransferase Ada, which serves to not only repair mutagenic lesions in DNA but also remove alkyl groups from the DNA backbone. The transfer of a methyl group from one of the stereoisomers (the *S*-, but not *R*-diastereomer) of the methylphosphotriester lesion to the N-terminal active site converts the protein into a transcription factor. Ada protein methylated by this reaction has been shown to exhibit greatly increased DNA binding ability as compared to unmethylated Ada, which binds DNA only weakly (*123,124*). N-Methylated Ada has been shown to upregulate transcription of four genes at three different promoters; ada, which controls expression of both the Ada and AlkB proteins, alkA, and aidB (*123,125,126*). While this upregulation is thought to be due to facilitation of RNA polymerase binding by the bound Ada (*6*), the exact mechanisms of gene induction at these three promoters differ.

Although N-methylated Ada binds to the promoters of the ada and alkA genes at a sequence shared between the two genes (AAAGCGCA) (*123*), the consequent activation occurs differently at the two sites. At both the ada and aidB promoters, methylated Ada interacts through its C-terminal domain with a negatively-charged region of the σ^{70} subunit of RNA polymerase, functioning either to enhance the binding of RNA polymerase to the core promoter region or to facilitate the formation of the open complex, thus activating transcription (*125*). In contrast, a positively-charged region of

σ^{70} is involved in interaction with Ada at the alkA promoter, and only the N-terminal portion of Ada is required for expression of the alkA gene (*127*). This difference is attributed to the different relative locations of Ada binding at each class of promoter, as the binding site at ada and aidB is one helical turn upstream from that at alkA. The positioning of Ada in relation to the RNA polymerase is therefore different in the two cases, requiring that the proteins contact each other in different conformations. Indeed, while methylation of Ada is required to obtain transcription from the ada and aidB genes, it is not necessary for transcription of alkA, indicating that the area of Ada that facilitates interaction with RNA polymerase in the latter case is exposed regardless of methylation state (*128*).

Ethylating agents, despite producing more triesters in the DNA backbone than their corresponding methylating counterparts (*129*), are weaker at inducing the adaptive response. The lack of induction is not due to an inability of the protein to remove ethyl groups, however, as the ethylphosphotriesters are repaired by Ada (*130*) indicating that ethylation of the Cys38 residue is less efficient at activating transcriptional activity than is methylation. In addition to activation by the acquisition of methyl modifications from DNA, the adaptive response can be induced by compounds that modify proteins directly, such as methyl iodide (MeI), presumably through a direct transfer of the methyl group to the Ada active site (*131*).

The adaptive response is thought to be turned off in response to the disappearance of its substrate lesions. Increased production of Ada will enable the repair of damage until levels are reduced to the point where Ada is no longer being methylated at an appreciable rate. The accumulation of unmethylated Ada then serves to reduce gene expression by competing with the methylated protein for binding at the relevant promoter sites (*132*). This negative feedback loop to some extent mitigates the energetic burden of a suicidal protein repair system.

AlkA. Mutants in the alkA gene were first identified for their sensitivity to the lethal, but not the mutagenic, effects of alkylating agents (*112*). AlkA was later identified as a member of the 3-methyladenine glycosylase family, which is responsible for excising a

broad spectrum of lethal adducts from the genome, initiating the process of base excision repair (Figure 1.5). These glycosylases protect against base-induced toxicity by removing the abundant 3-methyladenine lesion, which is a potent block to bypass due to the positioning of the methyl adduct in the minor groove of the DNA helix (*133*). Excision of the damaged base by AlkA is followed by resection of the newly-created abasic site by AP endo- and exonucleases, and the resulting gap is filled by DNA polymerase and sealed by DNA ligase.

In addition to m3A, AlkA can excise 3-methylguanine, O^2-methylthymine, O^2-methylcytosine, 7-methylguanine, and 8-methylguanine (m8G) (*94,134-138*). Beyond these methylated bases, AlkA can also recognize damage products generated by oxidative processes, such as 1,N^6-ethenoadenine, hypoxanthine, 5-formyluracil, and a fragmented oxidized thymine product (*139*). The level of AlkA is increased 20-fold upon induction of the adaptive response, and its activity is not inhibited by high concentrations of product, in contrast to the constitutively-expressed m3A glycosylase, Tag (*94,134-137*). Despite the much greater substrate specificity of AlkA, overexpression of Tag can complement the alkylation-sensitive phenotype of an *alkA* mutant, suggesting that the actions of AlkA beyond the repair of m3A could be merely residual. The importance of m3A removal from DNA is further indicated by the fact that *tag alkA* double mutants are much more sensitive to alkylating agents than either single mutant (*133,140*). Interestingly, however, although both AlkA and Tag can recognize and excise m3A, they share no homology, suggesting that they evolved from different origins (*141,142*), which speaks to the critical nature of m3A removal for survival.

In order to remove m3A from DNA, the AlkA protein must increase the lability of this already-unstable lesion by only a few orders of magnitude (as opposed to the several orders of magnitude required by stable lesions such as uracil) (*6*). The protein contains a helix-hairpin-helix domain that aids in DNA binding and positions the catalytic Asp residue properly to initiate the hydrolytic reaction on the flipped-out base (*143,144*). The active site of AlkA is more accommodating than the tightly fitting sites of more selective glycosylases, explaining the ability of AlkA to recognize many different damaged bases.

The BER process is efficient in the recognition and removal of lethal adducts such as m3A from double-stranded DNA (dsDNA), but its utility is limited in two major ways: it is error-prone, as it depends on *de novo* DNA synthesis, and its action on lesions arising in single-stranded DNA (ssDNA) is extremely inefficient (*94*); further, BER in that context would result in a highly toxic strand break. Additionally, this pathway can remove normal bases from DNA, creating toxic and mutagenic AP sites (*145,146*).

AlkB. The alkB gene of *E. coli* was discovered in 1983, when an ethyl methane sulfonate-treated AB1157 mutant was isolated that exhibited sensitivity to methyl methane sulfonate, an S_N2-alkylating agent, but not to S_N1-alkylating agents nor to UV light, and was less able to reactivate methylated λ phage when compared to its wild-type counterpart (*113*). This strain was shown to possess both a normal growth rate and ability to induce the adaptive response, and the mutant gene was shown to lie adjacent to ada on the *E. coli* chromosome (*147*). Further studies of this mutant determined the nucleic acid sequence of the alkB gene, showed that the ada and alkB genes overlap, and found that the 216-amino acid AlkB protein has a molecular weight of ~27 kDa (*147,148*). Transfection of the alkB gene into human cells conferred the MMS-resistant phenotype (as well as resistance to DMS, but not to MNNG, MNU, or BCNU) and demonstrated that AlkB repairs, rather than prevents the formation of, DNA damage (*149*).

Insight into the possible role of AlkB in DNA repair came in 2000, when it was observed that AlkB-deficient *E. coli* were particularly defective in their ability to reactivate single-stranded phage, suggesting that the AlkB protein was required for the removal of toxic lesions formed preferentially in ssDNA (*87*). Soon after, computational analysis suggested that AlkB was a member of the 2-oxoglutarate (α-ketoglutarate, αKG) and iron-dependent dioxygenase family (*88*). This prediction was rapidly confirmed by two independent groups (*80,81*), who demonstrated that purified AlkB protein could remove 1-methyladenine (m1A) and 3-methylcytosine (m3C) lesions from single- and double-stranded DNA in a reaction dependent on the addition of molecular oxygen, α-ketoglutarate, and non-heme Fe(II) (Figure 1.4). The lesion reversal mechanism was

shown to proceed enzymatically and to utilize oxygen and iron as cofactors in an oxidative decarboxylation reaction employing αKG as a cosubstrate in order to hydroxylate the stray methyl group of m1A and m3C, with concomitant release of succinate. The hydroxylated methyl group is then released as formaldehyde, returning the base to its unadducted state. Although AlkB has been shown to bind preferentially to alkylated ssDNA over alkylated dsDNA (*150*), the protein can remove substrate lesions in both contexts (*80,81,150-152*), although the differences in relative activity vary among reports.

Soon after the mechanism of AlkB in the direct reversal of base damage was elucidated, the protein was shown to be capable of carrying out the same reaction on substrate bases in RNA (*151*). AlkB not only removed [^{3}H]-methyl groups from polyA and polyC *in vitro*, but also allowed reactivation of methylated RNA phage, demonstrating that this repair reaction can take place *in vivo*. Chemically-inactivated mRNA and tRNA can also be functionally restored by treatment with the AlkB protein, and m1A lesions are actively removed from tRNA *in vivo* (*153*). Further, single-stranded RNA (ssRNA) viruses have been shown to encode AlkB-like domains, which have been hypothesized to protect the viral genome from host-mediated methylation gene silencing or methylating pesticides (*91,151*). Whether RNA repair by AlkB represents a critical method employed by *E. coli* in maintaining the integrity of macromolecules remains to be determined.

In addition to the two originally-discovered AlkB substrates, m1A and m3C, the protein's repair capability has been extended to include dealkylation of 3-ethylcytosine (*8*), 1-methylguanine (*8,154*), 3-methylthymine (*8,154,155*), 1-ethyladenine (*84*), and 1,N^6-ethanoadenine (*156*) lesions in DNA, demonstrating the enzyme's biochemical versatility. Additionally, the protein reduces the toxic effects of agents that produce hydroxyethyl, propyl, and hydroxypropyl adducts in DNA (*157*), suggesting that these lesions may also be substrates for repair.

Chemical cross-linking studies (*158,159*) have shown interaction of AlkB with ssDNA probes containing a modified cytosine as well as with mismatched double-stranded probes, but not with correctly-paired dsDNA substrates. This observation

suggests that the protein is not capable of actively flipping bases out of a helical structure for interrogation. Instead, it relies on the conformational flexibility offered by single-stranded substrates or the lack of hydrogen-bonding ability and therefore increased extra-helical presence of methylated substrate bases in double-stranded nucleic acids.

In 2005, it was discovered that AlkB, in addition to mediating repair through oxidative demethylation, could employ a second, distinct chemical mechanism to reverse the unsaturated exocyclic lesions formed in DNA through reaction with lipid peroxidation breakdown products or exposure to vinyl chloride or its metabolites, chloroethylene oxide or chloroacetaldehyde (*82,83*). The toxicity and mutagenicity of unrepaired 1,N^6-ethenoadenine were completely abrogated in AlkB-proficient *E. coli*, although εA was demonstrated to be a less efficient substrate than m1A. Additionally, the deleterious effects of 3,N^4-ethenocytosine were shown to be somewhat abrogated by the protein. In these reactions, however, the protein inserts an oxygen atom into the double bond of the etheno lesion, creating an epoxide intermediate. The addition of water opens the epoxide into a glycol moiety, and the two-carbon bridge is then released as glyoxal, restoring the undamaged base (Figure 1.4).

A set of X-ray crystal structures of the AlkB protein in various substrate or product complexes (*160*) revealed that, in addition to the domains found in other αKG/Fe-dependent dioxygenases, AlkB contains a flexible region designated the "nucleotide recognition lid" that contacts the nucleic acid backbone. Recognition of substrates through features invariant in polynucleotides allows the protein to act on both DNA and RNA in a sequence-independent manner. The requirement for contacts with the nucleic acid backbone suggests an explanation for the previous observation that a phosphate group 5' to the methylated base was necessary for effective repair of nucleotides (*157*). In addition, the flexibility of the nucleic acid-binding region probably accounts for the enzyme's ability to recognize and correctly position substrate lesions of varying size and shape in the active site, accounting for the accommodating nature of the protein.

In addition to *E. coli*, many other species possess homologs of AlkB, including many bacteria, *Saccharomyces pombe, Caenorhabditis elegans, Drosophila*

melanogaster, mice, humans, and plant RNA viruses (*85,86,88,89,161*). Conservation of this gene is not universal, however, as other species, both bacterial and eukaryotic, lack proteins similar in sequence (*89*). Expression of yeast genes encoding proteins lacking amino acid homology to AlkB have been shown to complement the MMS-sensitive phenotype of AlkB-deficient *E. coli*, however, suggesting that perhaps an unrelated family of proteins is able to carry out this repair reaction in species that lack obvious AlkB homologs.

Two of eight identified human homologs (*90*), hABH2 and hABH3, have been shown to possess the same αKG/Fe-dependent dioxygenase ability of AlkB, although efficiencies and substrate preferences vary among studies (*84,151,152,155*). Generally, hABH2 is observed to repair more efficiently double-stranded substrates, whether they be DNA, RNA, or hybrid (*152*) and to prefer DNA to RNA (*162*). hABH3, on the other hand, is more active on single-stranded substrates, regardless of whether they are DNA or RNA (*152*). The murine homologs act similarly to their human counterparts (*162*). The two functional human homologs exhibit differential cellular localization and tissue-specific expression (*84,151*), which, paired with the different substrate preferences of the two enzymes, indicates that they may play complementary roles in protecting human cells from methylation damage.

The toxicity and mutagenicity of unrepaired AlkB substrate lesions (*8,82,87,156,163-167*) and the wide-spread existence of homologs in other species, both prokaryotic and eukaryotic (*85,86,88,89,161*), suggest the importance of the recently-established mechanism of oxidative demethylation by AlkB proteins as a nucleic acid repair process.

AidB. Despite years of effort, the function of the fourth adaptive response protein, AidB, is still unclear. The Ada-dependent upregulation of AidB *in vivo* in response to the presence of methylating agents is one of two ways in which expression of the protein can be triggered. Anaerobiosis or the presence of acetate at a slightly acidic pH induces expression of AidB in an Ada-independent fashion (*168*). This mode of induction requires a functional rpoS gene (*169*), which encodes an alternative RNA polymerase

sigma factor, σ^s, normally expressed at low levels in exponential cultures and at higher levels as *E. coli* enter the late-logarithmic and stationary growth phases. Sigma factors increase the affinity of the RNA polymerase holoenzyme for particular promoters, with each sigma factor possessing different relative affinities toward various promoters. Expression of the alternative σ^s factor in later growth phases initiates its replacement of the traditional σ^{70} factor, inducing expression of stationary phase genes and halting expression of genes involved in rapid growth. The alternative factor is additionally involved in cellular responses to carbon starvation, oxidative damage, and osmotic shock (*125*).

Some σ^s–dependent genes are induced simply by the reduction in growth rate triggered by entry into stationary phase; the slowed growth that occurs under oxygen limitation causes induction as well (*169*). AidB differs from this subset of genes, however, in that it requires anaerobic conditions in order to exhibit σ^s–dependent up-regulation. This line of evidence indicates that AidB expression is controlled specifically by oxygen availability and not by growth rate. Ada-independent AidB induction therefore has two requirements: the growth-phase dependent expression of rpoS, and some as yet undefined feature of oxygen deprivation.

Cellular growth in anaerobic conditions results in the accumulation of organic acids and a consequential decrease in internal pH. Simulation of these conditions by the addition of acetate to a slightly acidic medium (pH 6.8) caused production of AidB at approximately the same levels observed in cells grown without aeration (*170*). Furthermore, it has been shown that this acetate-pH dependent induction is independent of ada status but requires a functional rpoS gene.

Studies of an rpoS knockout have revealed low levels of AidB induction in anaerobic conditions, indicating that RNA polymerase containing the more common σ^{70} factor is able to transcribe the gene, albeit much less efficiently than that carrying the σ^s factor. A similar scenario is seen *in vitro*, as in the absence of transcription factors, aidB is transcribed efficiently by the σ^s RNA polymerase holoenzyme but at only a very low level by the σ^{70} form (*171*). The presence of methylated Ada protein, however, significantly increases transcription levels by holoenzymes containing either sigma

factor, consistent with the finding that Ada interacts with a negatively charged patch of σ^{70} that is conserved in the alternative σ^{s} (*125*). RpoS-dependent expression of the adaptive response is not limited to AidB, but seems to differ by promoter, as AlkA expression is increased as cells enter stationary phase in the absence of a methylating agent in an *rpoS* strain (*172*).

AidB knockouts have shown inconsistent phenotypes, with *aidB:lacZ* insertion mutants falling into two categories: those with alkylation sensitivity identical to that of wild-type cells, and those that demonstrate resistance to both the lethal and mutagenic effects of the S_N1 alkylating agent MNNG but show no alteration in sensitivity to other alkylating agents (*173*). This observation led to the hypothesis that perhaps AidB is involved in the activation of MNNG (*174*), which must be metabolized in order to damage DNA (*7*), and that loss of AidB function, which occurred in only a subset of insertion mutants, renders the drug innocuous. Paradoxically, however, over-expression of the AidB protein in *E. coli* also produced cells that were resistant to mutagenesis by MNNG, prompting the opposing hypothesis that the protein facilitates the metabolic inactivation of the methylating agent. Landini et al. (1994) examined the data in an attempt to resolve the apparent contradiction, and found that the *aidB:lacZ* insertion mutants studied all affected only the terminal third of the gene. The observed phenotypes would be explained if expression of the first two-thirds of the gene produced a functional protein that is either stabilized or activated in the resistant strains and similar to wild-type in the non-resistant strains. However, a complete knockout of the aidB gene has also been shown to provide MNNG resistance (Volkert, unpublished observations), a paradox that has yet to be explained.

The nucleotide sequence of the aidB open reading frame was used to predict the amino acid sequence of the AidB protein, which has an expected molecular mass of 60.5 kDa. In an attempt to elucidate the possible function of this protein, this derived amino acid sequence was entered into a homology search of the Swiss Protein and GenBank-EMBL databases. While no homology was found with any known DNA repair protein, significant homology with several mammalian acyl coenzyme A (CoA) and isovaleryl CoA dehydrogenases was uncovered. Because the protein sequence indicates 24.6%

homology with human isovaleryl-CoA dehydrogenase (IVD) precursor (*173*), an enzyme involved in leucine metabolism in mammalian cells, crude extracts from *E. coli* over-expressing AidB from an inducible plasmid were tested for IVD activity *in vitro* in an isovaleryl-CoA reduction assay utilizing the artificial electron acceptor 2,6-dichlorophenolindophenol (DCPIP) to monitor activity spectrophotometrically. The deep blue to colorless transition upon electron acceptance can be measured as a reduction in the compound's A_{600}. IVD function was detected in extracts of cells over-expressing AidB (induced MV3590), but not in extracts of cells that were not over-expressing the protein (uninduced MV3590, induced or uninduced MV3594).

Discovery of the IVD activity of AidB led to the hypothesis that the protein, like other acyl-CoA dehydrogenases, plays a role in metabolism and energy generation during times of limited oxygen availability. This hypothesized involvement in metabolism would explain the observed induction of the protein in an Ada-independent manner in response to anaerobiosis, but does not provide a simple explanation for the Ada-dependent induction of the protein in response to methyl damage.

The homology that AidB shares with human isovaleryl-CoA dehydrogenase has led to investigations into the possible role of AidB in leucine metabolism. The ability of leucine to stimulate AidB production *in vivo* has been established (*171*) and shown to be specific to the AidB promoter, and not the alternative sigma factor (σ^S). Leucine responsive protein (Lrp), which functions as a transcriptional regulator of multiple genes and most operons involving leucine in *E. coli*, binds to its target promoters and activates or inhibits transcription of the target gene. The consequence of Lrp binding depends on other factors, and leucine itself often modulates Lrp activity. Binding of pure Lrp protein to the aidB promoter was detected in a DNAseI-protection assay, but vanished with the addition of leucine or methylated Ada to the reaction, even at concentrations only half that of Lrp, indicating that competitive binding at the promoter region, rather than protein-protein interaction, is responsible for the dissociation (*171*). *In vitro* transcription experiments revealed that Lrp binding functions to prevent AidB transcription and that the presence of either leucine or methylated Ada allows transcription to take place.

Additionally, *E. coli* lacking a functional lrp gene were found to express AidB during aerobic growth in a phase-dependent manner, an expression pattern characteristic of other rpoS-dependent genes. This evidence implicates Lrp in a regulatory role, keeping AidB expression low during times of rapid growth or leucine starvation, and allowing levels to increase when cells are not dividing quickly and excess leucine is present. Such regulation is characteristic of proteins involved in the breakdown of certain amino acids, and the fact that AidB is subject to this type of negative regulation lends support to its hypothesized role in leucine metabolism.

Other mechanisms of AidB regulation are clearly at work, however, as the lack of a functional lrp gene does not affect induction by methylating agents, and growth of lrp-deficient cells in anaerobic conditions resulted in significantly higher levels of AidB expression than growth in oxygenated conditions, indicating that some method of AidB induction beyond simple de-repression is involved. Much evidence indicates that AidB regulation is complex, and without an understanding of the protein's function, the reasons for such intricate control are unclear. Further study will be required to shed light on the role of AidB in DNA repair or reactive species detoxification in *E. coli*, and to provide an explanation for these multiple layers of regulation as well as an understanding of the function of the protein in defending the human genome against DNA damage and mutation.

More recent studies of the AidB protein have demonstrated stoichiometric content of FAD, confirmed the weak IVD activity of the protein, and identified the ability of AidB to bind to dsDNA (*175*), leading to a suggested role of the protein in the direct repair of alkylated DNA.

The Adaptive Response in Other Species

Adaptive responses to various types of damaging agents beyond alkylators have been observed in prokaryotes, yeast, mammals, and plants (*176*), indicating the ubiquitous nature of this type of acquired resistance mechanism. For example, the oxidative stress corresponding to regular exercise limits the amount of lipid peroxidation produced during any one period of exercise, and occupational exposure to low levels of

ionizing radiation has been shown to increase the repair capacity of lymphocytes in those humans (*176*).

Mounting an adaptive response to alkylating agents is not uncommon, however, as *E. coli* is by no means the only species able to increase resistance to alkylating agents upon exposure to low levels. Microorganisms that show acquired enhanced resistance to MNNG toxicity and mutagenesis include *Bacillus subtilis* (*177-179*), *Bacillus thuringiensis* (*180*), *Micrococcus luteus* (*181*), *Streptomyces frodiae* (*182*), *Gloetrichia ghosei, Pseudomonas aeruginosi, Shigella sonnei,* and *Xanthomonas maltophilia* (*86*). In fact, a screen of 33 gram-negative bacterial species found the adaptive response to be a widespread phenomenon (*183,184*), although not universally conserved (*185,186*). For example, *Haemophilus influenzae, Salmonella typhimurium, Staphylococcus aureus* and *Saccharomyces cerevisiae* all seem to lack inducible repair activity (*86*). CHO fibroblasts and SV40-transformed human skin fibroblasts can mount an MNNG-induced adaptive response as well (*187*). The conservation of adaptive responses across species suggests that environmental sources of alkylation are both common and biologically relevant (*188*), and that the phenomenon of inducible repair is evolutionarily important in protecting against the naturally-occurring endogenous and exogenous alkylating agents to which all living things are continuously exposed.

Figure 1.1: Sites of DNA base methylation.

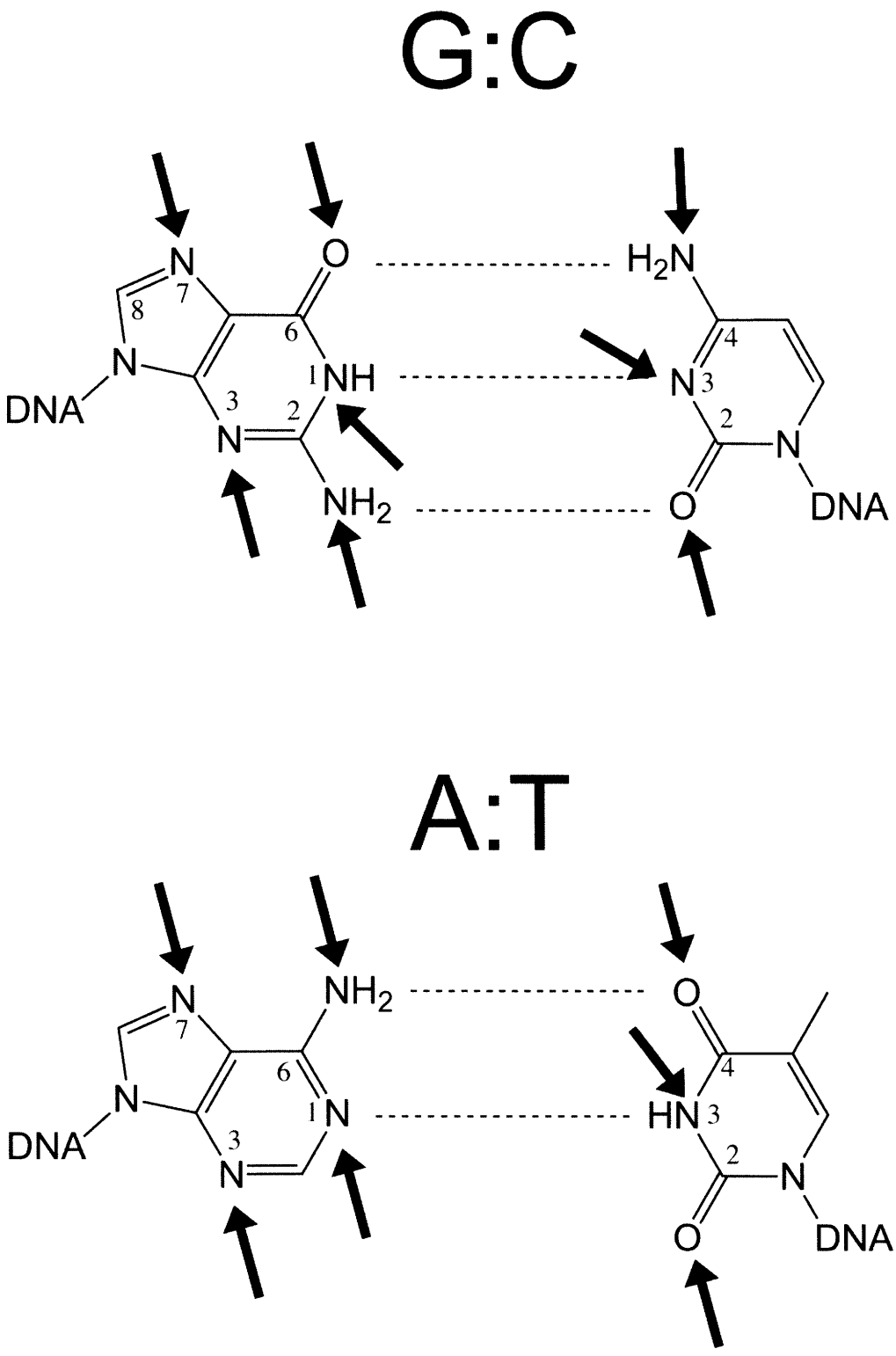
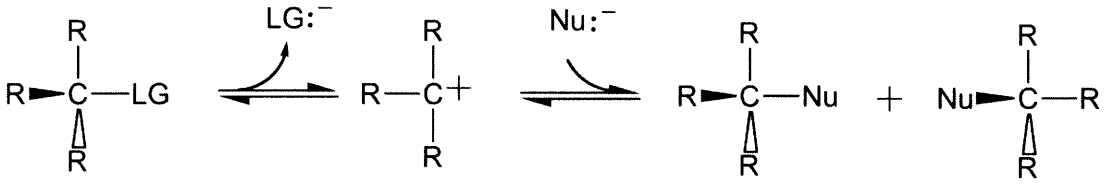


Figure 1.2: S_N1 and S_N2 reaction mechanisms through which alkylating agents can act. While the leaving group departure is unimolecular in S_N1, the reaction involves two steps. In contrast, S_N2 requires a bi-molecular interaction to cause leaving group departure, but involves a concerted reaction.

S_N1



S_N2

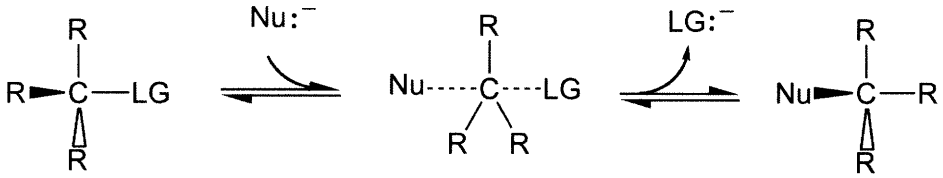


Figure 1.3: Mechanisms of methyltransferase repair by Ada.

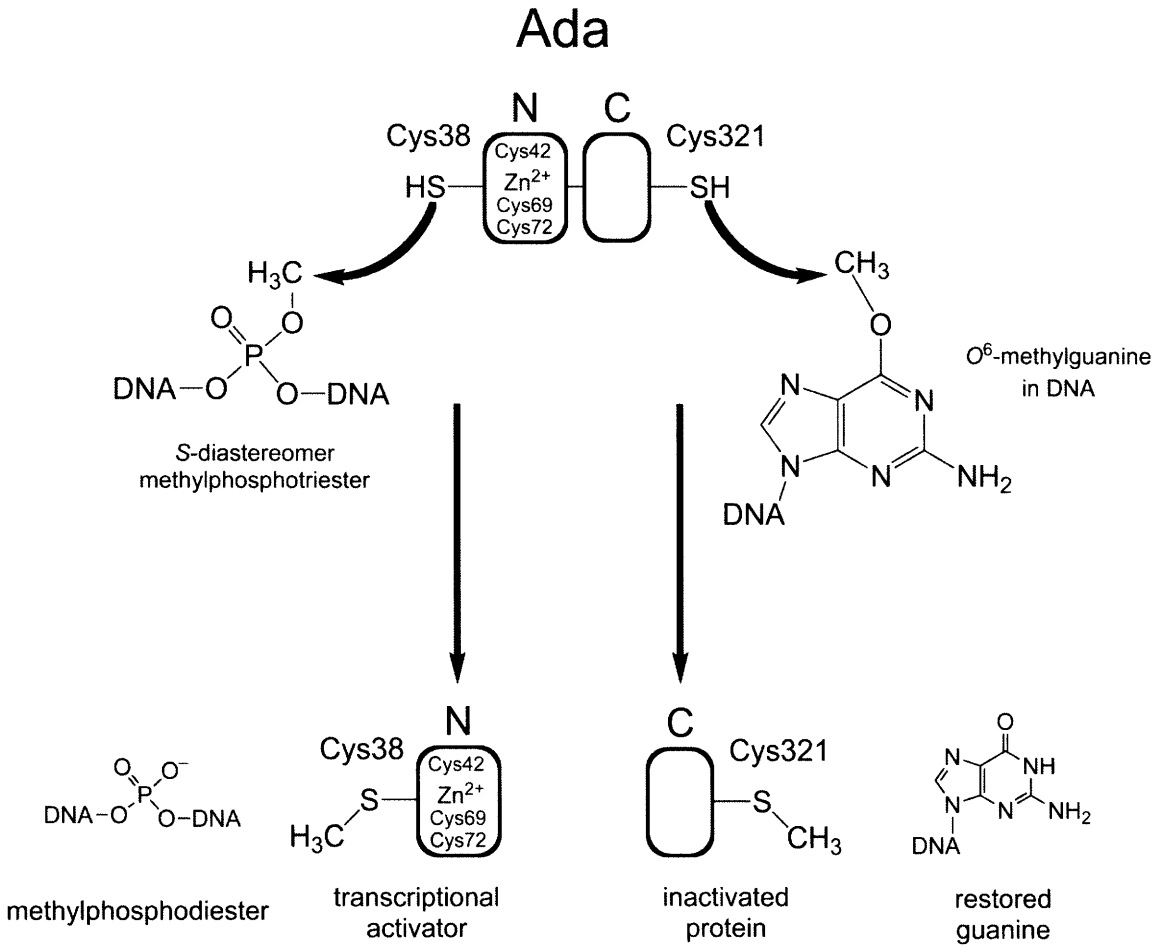


Figure 1.4: Mechanisms of repair by the *E. coli* dioxygenase AlkB.

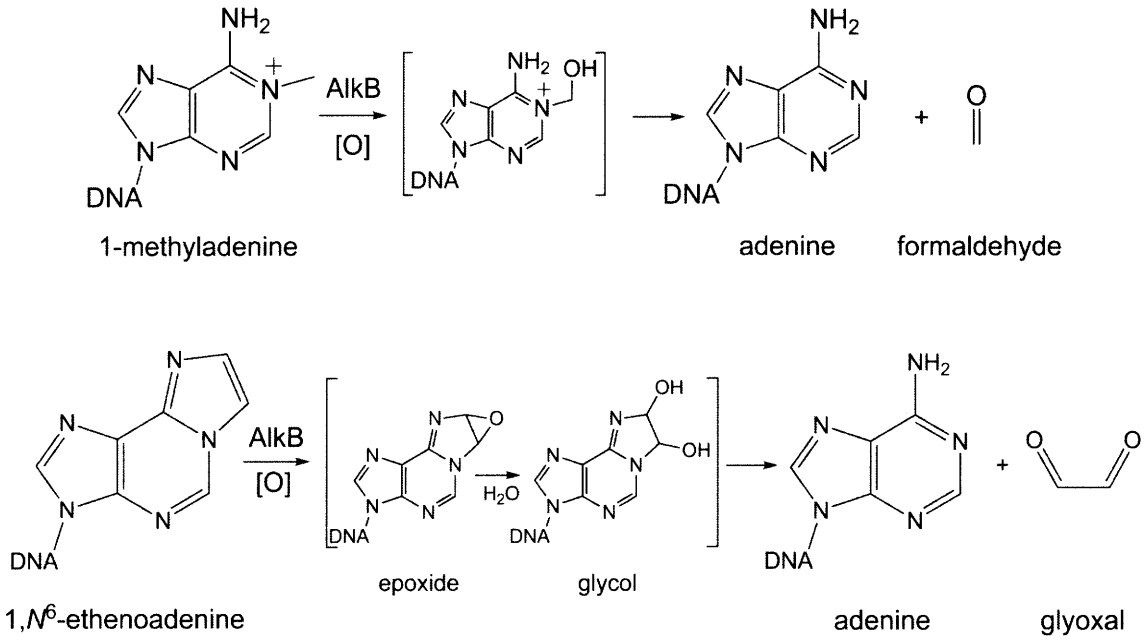


Figure 1.5: Outline of the repair of 3-methyladenine by AlkA through the base excision repair pathway.

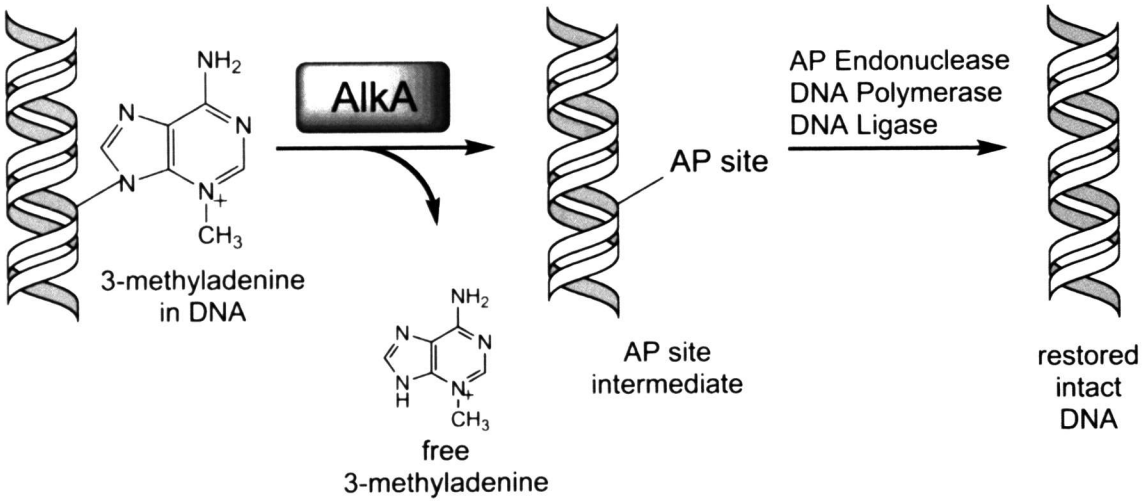


Figure 1.6: Overview of the adaptive response in *E. coli.*

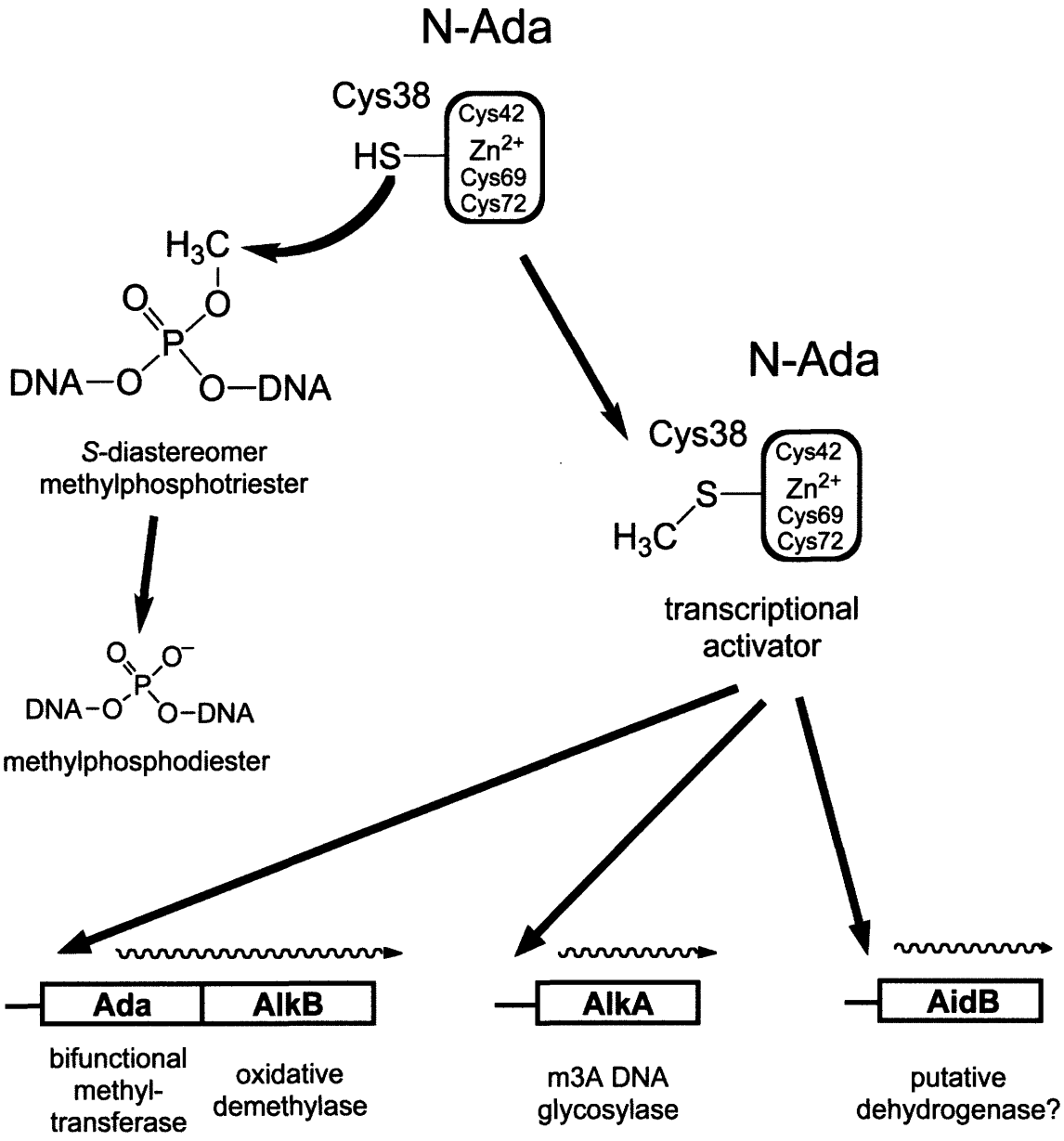


Table 1.1: Relative proportions of reaction at each base position by some common laboratory methylating agents. Adapted from (*189*).

Position of Methylation	MMS (S_N2)		MNU (S_N1)	
	ssDNA/RNA	dsDNA	ssDNA/RNA	dsDNA
Adenine				
N1	18	3.8	2.8	1.3
N3	1.4	10.4	2.6	9
N7	3.8	1.8	1.8	1.7
Guanine				
N3	~1	0.6	0.4	0.8
O^6	-	0.3	3	6.3
N7	68	83	69	67
Uracil/Thymine				
O^2	nd	nd	nd	0.11
N3	nd	nd	nd	0.3
O^4	nd	nd	nd	0.4
Cytosine				
O^2	nd	nd	nd	0.1
N3	10	<1	2.3	0.6
Diester	2	0.8	~10	17

References

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

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

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

4. Dedon, P. C. and S. R. Tannenbaum. 2004. Reactive nitrogen species in the chemical biology of inflammation. *Arch. Biochem. Biophys.* **423**:12-22.

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

6. Sedgwick, B. 2004. Repairing DNA-methylation damage. *Nat. Rev. Mol. Cell Biol.* **5**:148-157.

7. Friedberg E.C. 2006. DNA Repair and Mutagenesis. AMS Press, Washington, DC.

8. Delaney, J. C. and J. M. Essigmann. 2004. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A* **101**:14051-14056.

9. Grollman, A. P. and M. Moriya. 1993. Mutagenesis by 8-oxoguanine: an enemy within. *Trends Genet.* **9**:246-249.

10. Delaney, J. C. and J. M. Essigmann. 1999. Context-dependent mutagenesis by DNA lesions. *Chem. Biol.* **6**:743-753.

11. Lindahl, T. 2000. Cold Spring Harbor Symposia on Quantitative Biology. *Cold Spring Harbor Symp Quant Biol.*

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

13. Lutz, W. K. 1990. Endogenous genotoxic agents and processes as a basis of spontaneous carcinogenesis. *Mutat. Res.* **238**:287-295.

14. Rydberg, B. and T. Lindahl. 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.

15. Barrows, L. R. and P. N. Magee. 1982. Nonenzymatic methylation of DNA by S-adenosylmethionine in vitro. *Carcinogenesis* **3**:349-351.

16. Posnick, L. M. and L. D. Samson. 1999. Influence of S-adenosylmethionine pool size on spontaneous mutation, dam methylation, and cell growth of Escherichia coli. *J. Bacteriol.* **181**:6756-6762.

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

18. Harrison, K. L., R. Jukes, D. P. Cooper, and D. E. Shuker. 1999. Detection of concomitant formation of O6-carboxymethyl- and O6-methyl-2'-deoxyguanosine in DNA exposed to nitrosated glycine derivatives using a combined immunoaffinity/HPLC method. *Chem. Res. Toxicol.* **12**:106-111.

19. Sedgwick, B. 1997. Nitrosated peptides and polyamines as endogenous mutagens in O6-alkylguanine-DNA alkyltransferase deficient cells. *Carcinogenesis* **18**:1561-1567.

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

21. Shuker, D. E. and G. P. Margison. 1997. Nitrosated glycine derivatives as a potential source of O6-methylguanine in DNA. *Cancer Res.* **57**:366-369.

22. Marnett, L. J. and P. C. Burcham. 1993. Endogenous DNA adducts: potential and paradox. *Chem. Res. Toxicol.* **6**:771-785.

23. Lutz, W. K. and T. Fekete. 1996. Endogenous and exogenous factors in carcinogenesis: limits to cancer prevention. *Int. Arch. Occup. Environ. Health* **68**:120-125.

24. Stenvinkel, P., M. Karimi, S. Johansson, J. Axelsson, M. Suliman, B. Lindholm, O. Heimburger, P. Barany, A. Alvestrand, L. Nordfors, A. R. Qureshi, T. J. Ekstrom, and M. Schalling. 2007. Impact of inflammation on epigenetic DNA methylation - a novel risk factor for cardiovascular disease? *J. Intern. Med.* **261**:488-499.

25. Wehbe, H., R. Henson, F. Meng, J. Mize-Berge, and T. Patel. 2006. Interleukin-6 contributes to growth in cholangiocarcinoma cells by aberrant promoter methylation and gene expression. *Cancer Res.* **66**:10517-10524.

26. Shames, D. S., J. D. Minna, and A. F. Gazdar. 2007. DNA methylation in health, disease, and cancer. *Curr. Mol. Med.* **7**:85-102.

27. Crutzen, P. J. and M. O. Andreae. 1990. Biomass Burning in the Tropics: Impact on Atmospheric Chemistry and Biogeochemical Cycles. *Science* **250**:1669-1678.

28. Hamilton, J. T., W. C. McRoberts, F. Keppler, R. M. Kalin, and D. B. Harper. 2003. Chloride methylation by plant pectin: an efficient environmentally significant process. *Science* **301**:206-209.

29. Ballschmiter, K. 2003. Pattern and sources of naturally produced organohalogens in the marine environment: biogenic formation of organohalogens. *Chemosphere* **52**:313-324.

30. Mohamed, M. F., D. Kang, and V. P. Aneja. 2002. Volatile organic compounds in some urban locations in United States. *Chemosphere* **47**:863-882.

31. Goldman, R. and P. G. Shields. 2003. Food mutagens. *J. Nutr.* **133 Suppl 3**:965S-973S.

32. IARC Monographs. 1986. Tobacco smoking. IARC Monographs on the evaluation of the carcinogenic risks of chemicals to humans. IARC, Lyon.

33. Hecht, S. S. 1999. DNA adduct formation from tobacco-specific N-nitrosamines. *Mutat. Res.* **424**:127-142.

34. Olinski, R., Z. Walter, R. Wiaderkiewicz, E. Lukasova, and E. Palecek. 1980. Changes in DNA properties due to treatment with the pesticides malathion and DDVP. *Radiation and Environmental Biophysics* **18**:65-72.

35. Wooder, M. F. and A. S. Wright. 1981. Alkylation of DNA by organophosphorus pesticides. *Acta Pharmacol. Toxicol. (Copenh)* **49 Suppl 5**:51-55.

36. Humphreys, W. G., D. H. Kim, and F. P. Guengerich. 1991. Isolation and characterization of N7-guanyl adducts derived from 1,2-dibromo-3-chloropropane. *Chem. Res. Toxicol.* **4**:445-453.

37. Guengerich, F. P., W. G. Humphreys, D. H. Kim, T. Oida, and J. L. Cmarik. 1990. DNA-glutathione adducts derived from vic-dihaloalkanes: mechanisms of mutagenesis. *Princess Takamatsu Symp* **21**:101-107.

38. Tudek, B., J. Gajewska, M. Szczypka, I. Rahden-Staron, and T. Szymczyk. 1988. Screening for genotoxic activity of amitraz with short-term bacterial assays. *Mutat. Res.* **204**:585-591.

39. Hurley, L. H. 2002. DNA and its associated processes as targets for cancer therapy. *Nat. Rev. Cancer* **2**:188-200.

40. Chaney, S. G. and A. Sancar. 1996. DNA repair: enzymatic mechanisms and relevance to drug response. *J. Natl. Cancer Inst.* **88**:1346-1360.

41. Gerson, S. L. and J. K. Willson. 1995. O6-alkylguanine-DNA alkyltransferase. A target for the modulation of drug resistance. *Hematol. Oncol. Clin. North Am.* **9**:431-450.

42. Margison, G. P., M. F. Santibanez Koref, and A. C. Povey. 2002. Mechanisms of carcinogenicity/chemotherapy by O6-methylguanine. *Mutagenesis* **17**:483-487.

43. Ludlum, D. B. 1986. Nature and biological significance of DNA modification by the haloethylnitrosoureas. *IARC Sci. Publ.*71-81.

44. Ludlum, D. B. 1990. DNA alkylation by the haloethylnitrosoureas: nature of modifications produced and their enzymatic repair or removal. *Mutat. Res.* **233**:117-126.

45. Ludlum, D. B. 1986. Formation of cyclic adducts in nucleic acids by the haloethylnitrosoureas. *IARC Sci. Publ.*137-146.

46. Lober, G. and L. Kittler. 1977. Selected topics in photochemistry of nucleic acids. Recent results and perspectives. *Photochem. Photobiol.* **25**:215-233.

47. McLaren, A. D. and D. Shugar. 1964. Photochemistry of Proteins and Nucleic Acids. The Macmillan Company, New York.

48. Setlow, R. B. 1966. Cyclobutane-type pyrimidine dimers in polynucleotides. *Science* **153**:379-386.

49. Smith, K. C. H. P. C. 1969. Molecular Photobiology. Academic Press, New York and London.

50. Wang, S. Y. 1976. Photochemistry and Photobiology of Nucleic Acids. Academic Press Inc., New York.

51. Frankenberg-Schwager, M. 1990. Induction, repair and biological relevance of radiation-induced DNA lesions in eukaryotic cells. *Radiat. Environ. Biophys.* **29**:273-292.

52. Goodhead, D. T. 1989. The initial physical damage produced by ionizing radiations. *Int. J. Radiat. Biol* **56**:623-634.

53. Hutchinson, F. 1985. Chemical changes induced in DNA by ionizing radiation. *Prog. Nucleic Acid Res. Mol. Biol* **32**:115-154.

54. Teoule, R. 1987. Radiation-induced DNA damage and its repair. *Int. J. Radiat. Biol Relat Stud. Phys. Chem. Med.* **51**:573-589.

55. Ward, J. F. 1988. DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability. *Prog. Nucleic Acid Res. Mol. Biol* **35**:95-125.

56. Saffhill, R. 1986. The competitive miscoding of O6-methylguanine and O6-ethylguanine and the possible importance of cellular deoxynucleoside 5'-triphosphate pool sizes in mutagenesis and carcinogenesis. *Biochim. Biophys. Acta* **866**:53-60.

57. Basu, A. K., S. M. O'Hara, P. Valladier, K. Stone, O. Mols, and L. J. Marnett. 1988. Identification of adducts formed by reaction of guanine nucleosides with malondialdehyde and structurally related aldehydes. *Chem. Res. Toxicol.* **1**:53-59.

58. Singer, B. and J. M. Essigmann. 1991. Site-specific mutagenesis: retrospective and prospective. *Carcinogenesis* **12**:949-955.

59. Loveless, A. 1969. Possible relevance of O-6 alkylation of deoxyguanosine to the mutagenicity and carcinogenicity of nitrosamines and nitrosamides. *Nature* **223**:206-207.

60. Loechler, E. L., C. L. Green, and J. M. Essigmann. 1984. In vivo mutagenesis by O6-methylguanine built into a unique site in a viral genome. *Proc. Natl. Acad. Sci. U. S. A* **81**:6271-6275.

61. Foote, R. S., S. Mitra, and B. C. Pal. 1980. Demethylation of O6-methylguanine in a synthetic DNA polymer by an inducible activity in Escherichia coli. *Biochem. Biophys. Res. Commun.* **97**:654-659.

62. Demple, B., B. Sedgwick, P. Robins, N. Totty, M. D. Waterfield, and T. Lindahl. 1985. Active site and complete sequence of the suicidal methyltransferase that counters alkylation mutagenesis. *Proc. Natl. Acad. Sci. U. S. A* **82**:2688-2692.

63. Karran, P. 1985. Possible depletion of a DNA repair enzyme in human lymphoma cells by subversive repair. *Proc. Natl. Acad. Sci. U. S. A* **82**:5285-5289.

64. Xiao, W., B. Derfler, J. Chen, and L. Samson. 1991. Primary sequence and biological functions of a Saccharomyces cerevisiae O6-methylguanine/O4-methylthymine DNA repair methyltransferase gene. *EMBO J.* **10**:2179-2186.

65. Pegg, A. E. 2000. Repair of O(6)-alkylguanine by alkyltransferases. *Mutat. Res.* **462**:83-100.

66. Sassanfar, M., M. K. Dosanjh, J. M. Essigmann, and L. Samson. 1991. Relative efficiencies of the bacterial, yeast, and human DNA methyltransferases for the repair of O6-methylguanine and O4-methylthymine. Suggestive evidence for O4-

methylthymine repair by eukaryotic methyltransferases. *J. Biol Chem.* **266**:2767-2771.

67. Sassanfar, M. and L. Samson. 1990. Identification and preliminary characterization of an O6-methylguanine DNA repair methyltransferase in the yeast Saccharomyces cerevisiae. *J. Biol Chem.* **265**:20-25.

68. Caillet-Fauquet, P., M. Defais, and M. Radman. 1977. Molecular mechanisms of induced mutagenesis. Replication in vivo of bacteriophage phiX174 single-stranded, ultraviolet light-irradiated DNA in intact and irradiated host cells. *J. Mol. Biol* **117**:95-110.

69. Moore, P. and B. S. Strauss. 1979. Sites of inhibition of in vitro DNA synthesis in carcinogen- and UV-treated phi X174 DNA. *Nature* **278**:664-666.

70. Setlow, R. B., P. A. Swenson, and W. L. Carrier. 1963. Thymine Dimers and Inhibition of DNA Synthesis by Ultraviolet Irradiation of Cells. *Science* **142**:1464-1466.

71. Swenson, P. A. and R. B. Setlow. 1966. Effects of ultraviolet radiation on macromolecular synthesis in Escherichia coli. *J. Mol. Biol* **15**:201-219.

72. Villani, G., S. Boiteux, and M. Radman. 1978. Mechanism of ultraviolet-induced mutagenesis: extent and fidelity of in vitro DNA synthesis on irradiated templates. *Proc. Natl. Acad. Sci. U. S. A* **75**:3037-3041.

73. Sancar, A., F. W. Smith, and G. B. Sancar. 1984. Purification of Escherichia coli DNA photolyase. *J. Biol Chem.* **259**:6028-6032.

74. Sancar, G. B., F. W. Smith, M. C. Lorence, C. S. Rupert, and A. Sancar. 1984. Sequences of the Escherichia coli photolyase gene and protein. *J. Biol Chem.* **259**:6033-6038.

75. Rupert, C. S. 1975. Molecular Mechanisms for Repair of DNA. Plenum, New York.

76. Hanawalt, P. C. 1978. DNA Repair Mechanisms. Academy Press, New York.

77. Sancar, A. 1994. Structure and function of DNA photolyase. *Biochemistry* **33**:2-9.

78. Sancar, G. B. 1990. DNA photolyases: physical properties, action mechanism, and roles in dark repair. *Mutat. Res.* **236**:147-160.

79. Gibbs, P. E. and C. W. Lawrence. 1993. U-U and T-T cyclobutane dimers have different mutational properties. *Nucleic Acids Res.* **21**:4059-4065.

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

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

82. Delaney, J. C., L. Smeester, C. Wong, L. E. Frick, K. Taghizadeh, J. S. Wishnok, C. L. Drennan, L. D. Samson, and J. M. Essigmann. 2005. AlkB reverses etheno DNA lesions caused by lipid oxidation in vitro and in vivo. *Nat. Struct. Mol. Biol.* **12**:855-860.

83. Mishina, Y., C. G. Yang, and C. He. 2005. Direct repair of the exocyclic DNA adduct 1,N6-ethenoadenine by the DNA repair AlkB proteins. *J. Am. Chem. Soc.* **127**:14594-14595.

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

85. Wei, Y. F., K. C. Carter, R. P. Wang, and B. K. Shell. 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.

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

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

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

89. Falnes, P. O. and T. Rognes. 2003. DNA repair by bacterial AlkB proteins. *Res. Microbiol.* **154**:531-538.

90. Kurowski, M. A., A. S. Bhagwat, G. Papaj, and J. M. Bujnicki. 2003. Phylogenomic identification of five new human homologs of the DNA repair enzyme AlkB. *BMC. Genomics* **4**:48.

91. Bratlie, M. S. and F. Drablos. 2005. Bioinformatic mapping of AlkB homology domains in viruses. *BMC. Genomics* **6**:1.

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

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

94. McCarthy, T. V., P. Karran, and T. Lindahl. 1984. Inducible repair of O-alkylated DNA pyrimidines in Escherichia coli. *EMBO J.* **3**:545-550.

95. Pourquier, P., A. A. Pilon, G. Kohlhagen, A. Mazumder, A. Sharma, and Y. Pommier. 1997. Trapping of mammalian topoisomerase I and recombinations induced by damaged DNA containing nicks or gaps. Importance of DNA end phosphorylation and camptothecin effects. *J. Biol Chem.* **272**:26441-26447.

96. Kunkel, T. A. 1984. Mutational specificity of depurination. *Proc. Natl. Acad. Sci. U. S. A* **81**:1494-1498.

97. Loeb, L. A. and B. D. Preston. 1986. Mutagenesis by apurinic/apyrimidinic sites. *Annu. Rev. Genet.* **20**:201-230.

98. Boyce, R. P. and S. P. Howard-Flanders. 1964. Release of Ultraviolet Light-Induced Thymine Dimers from DNA in E. coli K-12. *Proc. Natl. Acad. Sci. U. S. A* **51**:293-300.

99. Setlow, R. B. and W. L. Carrier. 1964. The Dissapearance of Thymine Dimers from DNA: an Error-Correcting Mechanism. *Proc. Natl. Acad. Sci. U. S. A* **51**:226-231.

100. Howard-Flanders, P. and R. P. Boyce. 1966. DNA repair and genetic recombination: studies on mutants of Escherichia coli defective in these processes. *Radiat. Res.*Suppl.

101. Howard-Flanders, P., R. P. Boyce, and L. Theriot. 1966. Three loci in Escherichia coli K-12 that control the excision of pyrimidine dimers and certain other mutagen products from DNA. *Genetics* **53**:1119-1136.

102. Mattern, I. E., M. P. van Winden, and A. Rorsch. 1965. The range of action of genes controlling radiation sensitivity in Escherichia coli. *Mutat. Res.* **2**:111-131.

103. Murray, M. L. 1979. Substrate-specificity of uvr excision repair. *Environ. Mutagen.* **1**:347-352.

104. Cerda-Olmedo, E. and P. C. Hanawalt. 1968. Diazomethane as the active agent in nitrosoguanidine mutagenesis and lethality. *Mol. Gen. Genet.* **101**:191-202.

105. Neale, S. 1972. Effect of pH and temperature on nitrosamide-induced mutation in Escherichia coli. *Mutat. Res.* **14**:155-164.

106. Neale, S. 1976. Mutagenicity of nitrosamides and nitrosamidines in micro-organisms and plants. *Mutat. Res.* **32**:229-266.

107. Jimenez-Sanchez, A. and E. Cerda-Olmedo. 1975. Mutation and DNA replication in Escherichia coli treated with low concentrations of N-methyl-N'-nitro-N-nitrosoguanidine. *Mutat. Res.* **28**:337-345.

108. Jeggo, P., T. M. Defais, L. Samson, and P. Schendel. 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.

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

110. Schendel, P. F. and P. E. Robins. 1978. Repair of O6-methylguanine in adapted Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A* **75**:6017-6020.

111. Yamamoto, Y., M. Katsuki, M. Sekiguchi, and N. Otsuji. 1978. Escherichia coli gene that controls sensitivity to alkylating agents. *J. Bacteriol.* **135**:144-152.

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

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

114. Volkert, M. R. and D. C. Nguyen. 1984. Induction of specific Escherichia coli genes by sublethal treatments with alkylating agents. *Proc. Natl. Acad. Sci. U. S. A* **81**:4110-4114.

115. Karran, P., T. Lindahl, and B. Griffin. 1979. Adaptive response to alkylating agents involves alteration in situ of O6-methylguanine residues in DNA. *Nature* **280**:76-77.

116. Sakashita, H., T. Sakuma, Y. Akitomo, T. Ohkubo, M. Kainosho, M. Sekiguchi, and K. Morikawa. 1995. Sequence-specific DNA recognition of the Escherichia coli Ada protein associated with the methylation-dependent functional switch for transcriptional regulation. *J. Biochem. (Tokyo)* **118**:1184-1191.

117. Cheng, X. and R. J. Roberts. 2001. AdoMet-dependent methylation, DNA methyltransferases and base flipping. *Nucleic Acids Res.* **29**:3784-3795.

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

119. Sedgwick, B., P. Robins, N. Totty, and T. Lindahl. 1988. Functional domains and methyl acceptor sites of the Escherichia coli ada protein. *J. Biol. Chem.* **263**:4430-4433.

120. He, C., H. Wei, and G. L. Verdine. 2003. Converting the sacrificial DNA repair protein N-ada into a catalytic methyl phosphotriester repair enzyme. *J. Am. Chem. Soc.* **125**:1450-1451.

121. Myers, L. C., G. L. Verdine, and G. Wagner. 1993. Solution structure of the DNA methyl phosphotriester repair domain of Escherichia coli Ada. *Biochemistry* **32**:14089-14094.

122. He, C., J. C. Hus, L. J. Sun, P. Zhou, D. P. Norman, V. Dotsch, H. Wei, J. D. Gross, W. S. Lane, G. Wagner, and G. L. Verdine. 2005. A methylation-dependent electrostatic switch controls DNA repair and transcriptional activation by E. coli ada. *Mol. Cell* **20**:117-129.

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

124. Sedgwick, B. 1987. Molecular signal for induction of the adaptive response to alkylation damage in Escherichia coli. *J. Cell Sci. Suppl* **6**:215-223.

125. Landini, P. and M. R. Volkert. 2000. Regulatory responses of the adaptive response to alkylation damage: a simple regulon with complex regulatory features. *J. Bacteriol.* **182**:6543-6549.

126. Sakumi, K. and M. Sekiguchi. 1989. Regulation of expression of the ada gene controlling the adaptive response. Interactions with the ada promoter of the Ada protein and RNA polymerase. *J. Mol. Biol* **205**:373-385.

127. Akimaru, H., K. Sakumi, T. Yoshikai, M. Anai, and M. Sekiguchi. 1990. Positive and negative regulation of transcription by a cleavage product of Ada protein. *J. Mol. Biol* **216**:261-273.

128. Nakabeppu, Y. and M. Sekiguchi. 1986. Regulatory mechanisms for induction of synthesis of repair enzymes in response to alkylating agents: ada protein acts as a transcriptional regulator. *Proc. Natl. Acad. Sci. U. S. A* **83**:6297-6301.

129. Singer, B. and J. T. Kusmierek. 1982. Chemical mutagenesis. *Annu. Rev. Biochem.* **51**:655-693.

130. Margison, G. P., D. P. Cooper, and J. Brennand. 1985. Cloning of the E. coli O6-methylguanine and methylphosphotriester methyltransferase gene using a functional DNA repair assay. *Nucleic Acids Res.* **13**:1939-1952.

131. Takahashi, K. and Y. Kawazoe. 1987. Methyl iodide, a potent inducer of the adaptive response without appreciable mutagenicity in E. coli. *Biochem. Biophys. Res. Commun.* **144**:447-453.

132. Saget, B. M. and G. C. Walker. 1994. The Ada protein acts as both a positive and a negative modulator of Escherichia coli's response to methylating agents. *Proc. Natl. Acad. Sci. U. S. A* **91**:9730-9734.

133. Boiteux, S., O. Huisman, and J. Laval. 1984. 3-Methyladenine residues in DNA induce the SOS function sfiA in Escherichia coli. *EMBO J.* **3**:2569-2573.

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

135. Thomas, L., C. H. Yang, and D. A. Goldthwait. 1982. Two DNA glycosylases in Escherichia coli which release primarily 3-methyladenine. *Biochemistry* **21**:1162-1169.

136. Lindahl, T. 1976. New class of enzymes acting on damaged DNA. *Nature* **259**:64-66.

137. Laval, J. 1977. Two enzymes are required from strand incision in repair of alkylated DNA. *Nature* **269**:829-832.

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

139. Privezentzev, C. V., M. Saparbaev, A. Sambandam, M. M. Greenberg, and J. Laval. 2000. AlkA protein is the third Escherichia coli DNA repair protein excising a ring fragmentation product of thymine. *Biochemistry* **39**:14263-14268.

140. Clarke, N. D., M. Kvaal, and E. Seeberg. 1984. Cloning of Escherichia coli genes encoding 3-methyladenine DNA glycosylases I and II. *Mol. Gen. Genet.* **197**:368-372.

141. Sakumi, K., Y. Nakabeppu, Y. Yamamoto, S. Kawabata, S. Iwanaga, and M. Sekiguchi. 1986. Purification and structure of 3-methyladenine-DNA glycosylase I of Escherichia coli. *J. Biol Chem.* **261**:15761-15766.

142. Steinum, A. L. and E. Seeberg. 1986. Nucleotide sequence of the tag gene from Escherichia coli. *Nucleic Acids Res.* **14**:3763-3772.

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

144. Yamagata, Y., M. Kato, K. Odawara, Y. Tokuno, Y. Nakashima, N. Matsushima, K. Yasumura, K. Tomita, K. Ihara, Y. Fujii, Y. Nakabeppu, M. Sekiguchi, and S. Fujii. 1996. Three-dimensional structure of a DNA repair enzyme, 3-methyladenine DNA glycosylase II, from Escherichia coli. *Cell* **86**:311-319.

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

146. Memisoglu, A. and L. Samson. 2000. Base excision repair in yeast and mammals. *Mutat. Res.* **451**:39-51.

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

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

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

150. Begley, T. J. and L. D. Samson. 2003. AlkB mystery solved: oxidative demethylation of N1-methyladenine and N3-methylcytosine adducts by a direct reversal mechanism. *Trends Biochem. Sci.* **28**:2-5.

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

152. Falnes, P. O., M. Bjoras, P. A. Aas, O. Sundheim, and E. Seeberg. 2004. Substrate specificities of bacterial and human AlkB proteins. *Nucleic Acids Res.* **32**:3456-3461.

153. Ougland, R., C. M. Zhang, A. Liiv, R. F. Johansen, E. Seeberg, Y. M. Hou, J. Remme, and P. O. Falnes. 2004. AlkB restores the biological function of mRNA and tRNA inactivated by chemical methylation. *Mol. Cell* **16**:107-116.

154. Falnes, P. O. 2004. Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. *Nucleic Acids Res.* **32**:6260-6267.

155. Koivisto, P., P. Robins, T. Lindahl, and B. Sedgwick. 2004. Demethylation of 3-methylthymine in DNA by bacterial and human DNA dioxygenases. *J. Biol. Chem.* **279**:40470-40474.

156. Frick, L. E., J. C. Delaney, C. Wong, C. L. Drennan, and J. M. Essigmann. 2007. Alleviation of 1,N6-ethanoadenine genotoxicity by the Escherichia coli adaptive response protein AlkB. *Proc. Natl. Acad. Sci. U. S. A* **104**:755-760.

157. Koivisto, P., T. Duncan, T. Lindahl, and B. Sedgwick. 2003. Minimal methylated substrate and extended substrate range of Escherichia coli AlkB protein, a 1-methyladenine-DNA dioxygenase. *J. Biol. Chem.* **278**:44348-44354.

158. Mishina, Y. and C. He. 2003. Probing the structure and function of the Escherichia coli DNA alkylation repair AlkB protein through chemical cross-linking. *J. Am. Chem. Soc.* **125**:8730-8731.

159. Mishina, Y., C. H. Lee, and C. He. 2004. Interaction of human and bacterial AlkB proteins with DNA as probed through chemical cross-linking studies. *Nucleic Acids Res.* **32**:1548-1554.

160. Yu, B., W. C. Edstrom, J. Benach, Y. Hamuro, P. C. Weber, B. R. Gibney, and J. F. Hunt. 2006. Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. *Nature* **439**:879-884.

161. Drablos, F., E. Feyzi, P. A. Aas, C. B. Vaagbo, B. Kavli, M. S. Bratlie, J. Pena-Diaz, M. Otterlei, G. Slupphaug, and H. E. Krokan. 2004. Alkylation damage in DNA and RNA--repair mechanisms and medical significance. *DNA Repair (Amst)* **3**:1389-1407.

162. Lee, D. H., S. G. Jin, S. Cai, Y. Chen, G. P. Pfeifer, and T. R. O'Connor. 2005. Repair of methylation damage in DNA and RNA by mammalian AlkB homologues. *J. Biol. Chem.* **280**:39448-39459.

163. Nieminuszczy, J., A. Sikora, M. Wrzesinski, C. Janion, and E. Grzesiuk. 2006. AlkB dioxygenase in preventing MMS-induced mutagenesis in Escherichia coli: effect of Pol V and AlkA proteins. *DNA Repair (Amst)* **5**:181-188.

164. Abbott, P. J. and R. Saffhill. 1977. DNA-synthesis with methylated poly(dA-dT) templates: possible role of O4-methylthymine as a pro-mutagenic base. *Nucleic Acids Res.* **4**:761-769.

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

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

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

168. Volkert, M. R., L. I. Hajec, and D. C. Nguyen. 1989. Induction of the alkylation-inducible aidB gene of Escherichia coli by anaerobiosis. *J. Bacteriol.* **171**:1196-1198.

169. Volkert, M. R., L. I. Hajec, Z. Matijasevic, F. C. Fang, and R. Prince. 1994. Induction of the Escherichia coli aidB gene under oxygen-limiting conditions requires a functional rpoS (katF) gene. *J. Bacteriol.* **176**:7638-7645.

170. Smirnova, G. V., O. N. Oktyabrsky, E. V. Moshonkina, and N. V. Zakirova. 1994. Induction of the alkylation-inducible aidB gene of Escherichia coli by cytoplasmic acidification and N-ethylmaleimide. *Mutat. Res.* **314**:51-56.

171. Landini, P., L. I. Hajec, L. H. Nguyen, R. R. Burgess, and M. R. Volkert. 1996. The leucine-responsive regulatory protein (Lrp) acts as a specific repressor for sigma s-dependent transcription of the Escherichia coli aidB gene. *Mol. Microbiol.* **20**:947-955.

172. Landini, P. and S. J. Busby. 1999. Expression of the Escherichia coli ada regulon in stationary phase: evidence for rpoS-dependent negative regulation of alkA transcription. *J. Bacteriol.* **181**:6836-6839.

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

174. Volkert, M. R., D. C. Nguyen, and K. C. Beard. 1986. Escherichia coli gene induction by alkylation treatment. *Genetics* **112**:11-26.

175. Rohankhedkar, M. S., S. B. Mulrooney, W. J. Wedemeyer, and R. P. Hausinger. 2006. The AidB component of the Escherichia coli adaptive response to alkylating agents is a flavin-containing, DNA-binding protein. *J. Bacteriol.* **188**:223-230.

176. Crawford, D. R. and K. J. Davies. 1994. Adaptive response and oxidative stress. *Environ. Health Perspect.* **102 Suppl 10**:25-28.

177. Hadden, C. T., R. S. Foote, and S. Mitra. 1983. Adaptive response of Bacillus subtilis to N-methyl-N'-nitro-N-nitrosoguanidine. *J. Bacteriol.* **153**:756-762.

178. Morohoshi, F. and N. Munakata. 1985. Bacillus subtilis mutants deficient in the adaptive response to simple alkylating agents. *J. Bacteriol.* **161**:825-830.

179. Morohoshi, F., K. Hayashi, and N. Munakata. 1990. Bacillus subtilis ada operon encodes two DNA alkyltransferases. *Nucleic Acids Res.* **18**:5473-5480.

180. Boutibonnes, P. and Y. Auffray. 1987. Evidence that N-methyl-N'-nitro-N-nitrosoguanidine induces adaptive response in Bacillus thuringiensis. *Mutat. Res.* **190**:83-87.

181. Ather, A., Z. Ahmed, and S. Riazuddin. 1984. Adaptive response of Micrococcus luteus to alkylating chemicals. *Nucleic Acids Res.* **12**:2111-2126.

182. Baltz, R. H. and J. Stonesifer. 1985. Adaptive response and enhancement of N-methyl-N'-nitro-N-nitrosoguanidine mutagenesis by chloramphenicol in Streptomyces fradiae. *J. Bacteriol.* **164**:944-946.

183. Fernandez de Henestrosa, A. R. and J. Barbe. 1991. Induction of the alkA gene of Escherichia coli in gram-negative bacteria. *The Journal of Bacteriology* **173**:7736-7740.

184. Samson, L. 1992. The suicidal DNA repair methyltransferases of microbes. *Mol. Microbiol.* **6**:825-831.

185. Guttenplan, J. B. and S. Milstein. 1982. Resistance of Salmonella typhimurium TA 1535 to O6-guanine methylation and mutagenesis induced by low doses of N-methyl-N'-nitro-N-nitrosoguanidine: an apparent constitutive repair activity. *Carcinogenesis* **3**:327-331.

186. Kimball, R. F. 1980. Further studies on the induction of mutation in Haemophilus influenzae by N-methyl-N'-nitro-N-nitrosoguanidine: lack of an inducible error-free repair system and the effect of exposure medium. *Mutat. Res.* **72**:361-372.

187. Samson, L. and J. L. Schwartz. 1980. Evidence for an adaptive DNA repair pathway in CHO and human skin fibroblast cell lines. *Nature* **287**:861-863.

188. Vaughan, P., T. Lindahl, and B. Sedgwick. 1993. Induction of the adaptive response of Escherichia coli to alkylation damage by the environmental mutagen, methyl chloride. *Mutat. Res.* **293**:249-257.

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

CHAPTER 2:

Genotypic Characterization of the AlkB Mutant *E. coli* Strain HK82

Introduction

The *E. coli* AlkB protein is the product of one of four genes upregulated after exposure to alkylating agents in a phenomenon termed the "adaptive response" that serves to protect the bacterial genome from permanent damage. Three of the four genes are known to be involved in the repair of DNA lesions, but each employs a different mechanism (*1*). AlkB is a member of the 2-oxoglutarate (α-ketoglutarate)/iron-dependent dioxygenase family (*2*) that utilizes the oxidative mechanism characteristic of this enzyme class to hydroxylate the aberrant methyl (or larger alkyl) group of a range of DNA adducts, creating an unstable intermediate that will decompose spontaneously to release formaldehyde and the repaired base (*3,4*) (Figure 2.1). Additionally, the enzyme repairs another family of damaged bases, the exocyclic etheno- adducts produced by the reaction of DNA with certain products of lipid oxidation, through a distinct chemical mechanism that involves epoxidation of the double-bond of the two-carbon bridge connecting the exocyclic nitrogen with the ring nitrogen, creating an epoxide intermediate that can hydrolyze to form a second, glycol-containing intermediate that decomposes to the restored base and a glyoxal molecule (*5,6*).

Although the mechanism by which AlkB protects *E. coli* from the deleterious effects of exposure to alkylating agents is now well-established, the function of the protein was not always understood. Years of genetic studies demonstrated that strains lacking the protein induced the adaptive response normally, but were sensitive to the toxic effects of S_N2 alkylating agents such as methyl methanesulfonate (MMS) and that methylated λ phage were reactivated in wild-type strains to a greater extent than in mutant strains (*7*). It was later shown that the reactivation of single-stranded phage was much more drastically affected by the lack of a functional alkB gene than was reactivation of double-stranded phage, and the AlkB protein was shown to bind preferentially to single-stranded DNA (*8*). These results offered clues that AlkB was involved in DNA repair (rather than prevention of DNA damage), specifically the removal of lesions that form to a greater extent in single-stranded DNA, which suggested that 1-methyladenine (m1A) and 3-methylcytosine (m3C) would be good candidate substrates for the enzyme. Indeed, quickly after the αKG/Fe-dependent dioxygenase

homology was discovered, these two adducts were shown to be removed from DNA by the AlkB protein via its mechanism of oxidative demethylation (*4*). Since this discovery, the substrate pool of the enzyme has been greatly expanded to encompass many simple alkylated bases as well as a handful of exocyclic lesions through the two distinct chemical mechanisms outlined above (*3,5,6,9-11*).

Throughout the quest to demonstrate a mechanism responsible for the protection offered by AlkB, comparisons of phenotypes between AlkB-proficient and AlkB-deficient strains were scoured for insight. In 1983, Sekiguchi and Kataoka treated AB1157 *E. coli* with ethyl methanesulfonate (EMS) to induce mutagenesis and isolated strains that were sensitive specifically to MMS (i.e, not to S_N1 alkylators or to UV light) (*7*). A colony of one sensitive strain, HK22, also acquired a spontaneous mutation in the nearby nalA gene conferring resistance to naladixic acid; this double mutant was designated HK70. P1 transduction was then used to transfer the genetic information responsible for the S_N2-alkylator-sensitive phenotype from HK70 back into untreated AB1157 cells. A strain that acquired naladixic acid resistance (Nal^r) but remained resistant to MMS was designated HK80 (as AB1157, but *nalA*), while a strain that acquired both the Nal^r and the MMS^s phenotypes was designated HK82 (as AB1157, but *nalA alkB22)*.

These strains were utilized in 1985 by the same group to establish that the ada and alkB genes are adjacent on the *E. coli* chromosome and that they could form an operon (*12*). The next appearance of these strains in the literature occurs three years later, in a 1988 manuscript by Rebeck and coworkers (*13*). However, the "wild-type" HK80 strain had by this time been replaced by HK81, which is of the same genotype as HK80. Why the switch between strains was made has never been publicly addressed, but the replacement was permanent, as the HK80 strain makes no subsequent literary appearances.

Over the next two-and-a-half decades, many labs exploited the HK81 and HK82 strains as an isogenic pair in which to conduct genetic studies of the function of the AlkB protein, but a detailed analysis of the repair-deficient strain was never conducted. Because the mutant strain was created by EMS mutagenesis, it is reasonable to presume

that the alkB gene most likely carries a point mutation as was suggested by the original authors (*12*). Depending on its location and type, such a base change could affect the expression, folding, length, or catalytic activity of the protein, causing the strain carrying this mutation to lack the ability to remove toxic and mutagenic substrate lesions from the genome and therefore to exhibit sensitivity to agents that preferentially form these substrate lesions in DNA, such as MMS. Although this strain has been instrumental to the characterization of the AlkB protein over the past 24 years, the exact mutation that it carries had until now not been analyzed.

Materials and Methods

Bacterial Strains. The *E. coli* strains used for amplification and characterization of the alkB gene in order to compare the "wild-type" and the "mutant" sequences with the "theoretical wild-type" (NCBI gene ID 946708, sequence given in Table 2.1) were HK81 (as AB1157, but *nalA*; wild-type) and HK82 (as AB1157, but *nalA alkB22*; AlkB-deficient).

Isolation of Genomic DNA. To isolate bacterial genomic DNA, 10 mL cultures of each strain were grown with aeration overnight (to saturation) at 37 °C. Cells were isolated from 1 mL of the overnight by centrifugation for 3 min at 15,000 x g. After resuspension of the cell pellet in 570 μL TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0), 30 μL 10% SDS and 3 μL 20 mg/mL Proteinase K were added. The mixture was then incubated at 37°C for 1 h to release and digest cellular contents, following which 100 μL 5 M sodium chloride and 80 μL of a 10% cetyltrimethylammonium chloride/0.7 M sodium chloride solution were added. After a 10 min incubation at 65 °C, 750 μL phenol/chloroform/isoamyl alcohol (25:24:1, PCI) were added, and the sample was mixed well and then centrifuged for 5 min at 15,000 x g. The 750 μL aqueous layer was transferred to a new tube and the PCI extraction was repeated. The 700 μL aqueous layer of the second extraction was transferred to a fresh tube, and re-extracted using 420 μL isopropanol. Following removal of the supernatant, and a final extraction using 250 μL 70% ethanol, the supernatant was again removed. Finally, the isolated DNA was dried

under vacuum, resuspended in 100 μL H_2O, and quantified by absorbance at 260 nm. Typical yields were ~40 μg.

PCR Amplification of the AlkB Gene. Primers of sequence 5'-ATA CCT TGC GGT GAA ACC GTC AGT-3' and 5'-AGC CCG CAG TTT AAA CAT CTT CGC-3' (Integrated DNA Technologies) were used to amplify the region encoding the alkB gene of both strains (locations of primers shown in Table 2.2). The PCR Optimization Kit (Stratagene) was used to determine the specific buffering conditions that produced the highest quality product. Twelve PCR reactions were carried out on genomic DNA from each strain using Taq polymerase, and products were separated on a 1% agarose gel. Buffers 5 (pH 8.8, 1.5 mM $MgCl_2$, 25 mM KCl) and 6 (pH 8.8, 1.5 mM $MgCl_2$, 75 mM KCl) supplemented with 2% DMSO gave a single band corresponding to a 922 base-pair (bp) product (Figure 2.2), while 3 (pH 8.3, 3.5 mM $MgCl_2$, 25 mM KCl) and 11 (pH 9.2, 1.5 mM $MgCl_2$, 25 mM KCl) yielded other bands in addition to the desired product. Product DNA from the PCR reaction was isolated using the QIAquick PCR Purification Kit (Qiagen), and quantified by spectrophotometry.

Sequencing of the AlkB Gene. Approximately 12 ng PCR product was mixed with 3.2 pmol forward primer in a volume of 12 μL and submitted to the MIT Center for Cancer Research Biopolymers Laboratory for sequencing.

Results

PCR amplification of the AlkB gene from either HK81 or HK82 gave a single product of 922 bp (Figure 2.2), which was the size expected to be generated from the chosen primers (as shown in Table 2.1). Further analysis of this purified PCR product by sequencing (results in Tables 2.2 and 2.3) indicated that HK81 carries an alkB gene that is completely identical to the theoretical gene (sequence alignments given in Table 2.4). The gene carried by HK82, however, is 99.8% identical to the theoretical sequence, differing by a single base, a G→A transition at base pair 397 (Table 2.4). This substitution alters the coding information carried by the gene, as a GAT codon becomes

AAT. Accordingly, the aspartic acid residue that should reside at position 133 is replaced by an asparagine (Figure 2.3).

Discussion

The HK82 strain of *E. coli* that has been exploited for two-and-a-half decades in studies of the phenotypic and genetic consequences of the lack of a functional AlkB protein carries a single base pair substitution at bp 397 that codes for a single amino acid change at amino acid (aa) 133. This replacement of GAT with AAT, and consequently aspartic acid with asparagine, most likely inhibits catalytic activity of the protein through a reduced binding affinity for the iron cofactor necessary to the oxygenase-based mechanism through which the protein acts. The presence of a missense mutation at codon 133 indicates that it is most likely the catalytic activity of the protein that is adversely affected, as no stop codon has been introduced that would cause protein truncation and thus prohibit activity, and the coding alteration is not in a regulatory region that would inhibit normal expression of the gene. Whether the protein is simply misfolded due to the loss of negative charge or whether the properly-structured protein is unable to carry out its repair chemistry due to the substitution of a residue critical for catalysis remains to be determined.

Analysis of the recently-published crystal structure (*14*) shows that D133 is one of three residues critical for binding the catalytic iron molecule in the active site of the enzyme (Figure 2.4). Presumably, the loss of this amino acid prohibits or inhibits binding of the iron atom required to initiate the free-radical chemistry employed by AlkB in its direct reversal mechanisms. If this is the case, the HK82 strain most likely produces normal amounts of a full-length but non-functional protein. Indeed, the mutation of the equivalent residue in human and mouse ABH2 and ABH3 to alanine completely abolished the activity of all four proteins (*15*), underscoring the importance of the aspartic acid residue at this position in co-factor binding and thus enzymatic activity. The non-functional AlkB protein encoded by HK82, then, would likely possess unmodified DNA-binding and base-flipping properties, but would be unable to initiate repair chemistry once the adducted base was properly positioned in the active site. This

unique characteristic could make protein produced by the HK82 strain extremely useful in further structural studies of AlkB.

The possibility that the aspartic acid residue is critical not only for iron-binding but also for structural stability cannot be ruled out. Crystallographic studies of the mutant protein would shed light on the protein's conformation and perhaps the exact mechanism by which this single amino acid change eliminates protein activity. It is also possible that the mutant protein merely has reduced affinity for iron and actually retains a low level of activity. Comparison of methylation sensitivity of an *alkB* strain that lacks the entire gene to this mutant strain in addition to *in vitro* activity studies using AlkB protein purified from HK82 cells would be helpful in ascertaining whether this is the case.

If the protein is indeed unaffected in all aspects beyond catalytic activity, this mutant could be useful in exploration of enzymatic properties such as substrate location and identification as well as activity regulation. Provided that the mutant protein possesses normal DNA-binding and substrate-positioning activities, insight could be gained into how the protein initially locates methylated bases as well as how (or whether) the protein discriminates between bases that have been damaged by methylation and those which have been intentionally modified by the enzymatic addition of a methyl group. The question of whether the protein is capable of differentiating between the two classes of methylated bases at the level of binding could be answered, offering a piece of the puzzle that could be completed through studies of functional protein. Such studies could also reveal whether regulation could occur at the level of mechanism initiation. Similar experiments could be conducted using methylated RNA to ascertain whether the protein's characteristics differ depending on the context in which a methylated base occurs. Further investigation of AlkB activity and mechanism exploiting this mutant protein could shed light on many of the questions currently surrounding substrate location, identification, and discrimination by the AlkB protein.

Figure 2.1: Chemical mechanisms employed by the *E. coli* adaptive response protein AlkB in the direct reversal of DNA and RNA base damage. Oxidative demethylation used to remove simple alkyl lesions (methyl, ethyl, propyl, etc.) involves hydroxylation of the alkyl group followed by spontaneous release of formaldehyde and restored adenine. Epoxidation of the double-bond of the cyclic etheno-lesions allows hydrolysis to a glycol intermediate, followed by decomposition to the restored adenine and released glyoxal.

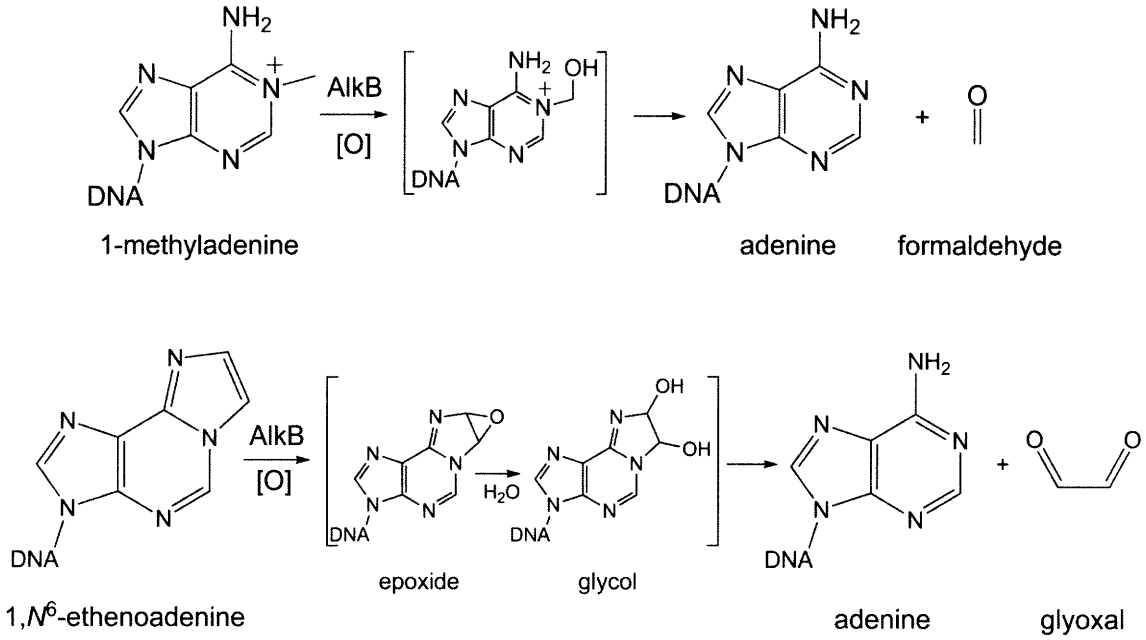


Figure 2.2: Separation of PCR products on a 1% agarose gel gives a single 922 bp product, as expected from gene sequence and chosen primers. Lanes represent different PCR buffering conditions utilized by the Stratagene PCR Optimization Kit. Buffers 5 (pH 8.8, 1.5 mM $MgCl_2$, 25 mM KCl) and 6 (pH 8.8, 1.5 mM $MgCl_2$, 75 mM KCl) supplemented with 2% DMSO gave a single band corresponding to a 922 base-pair (bp) product, while buffers 3 (pH 8.3, 3.5 mM $MgCl_2$, 25 mM KCl) and 11 (pH 9.2, 1.5 mM $MgCl_2$, 25 mM KCl) yielded additional products.

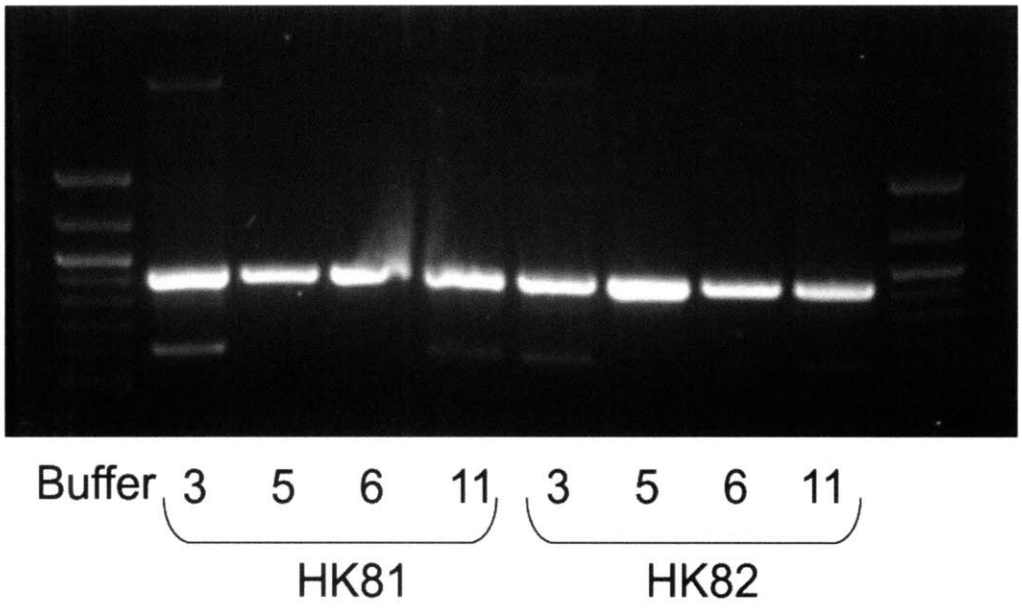


Figure 2.3: Amino acid change caused by the G to A mutation at base pair 397 that is carried by HK82.

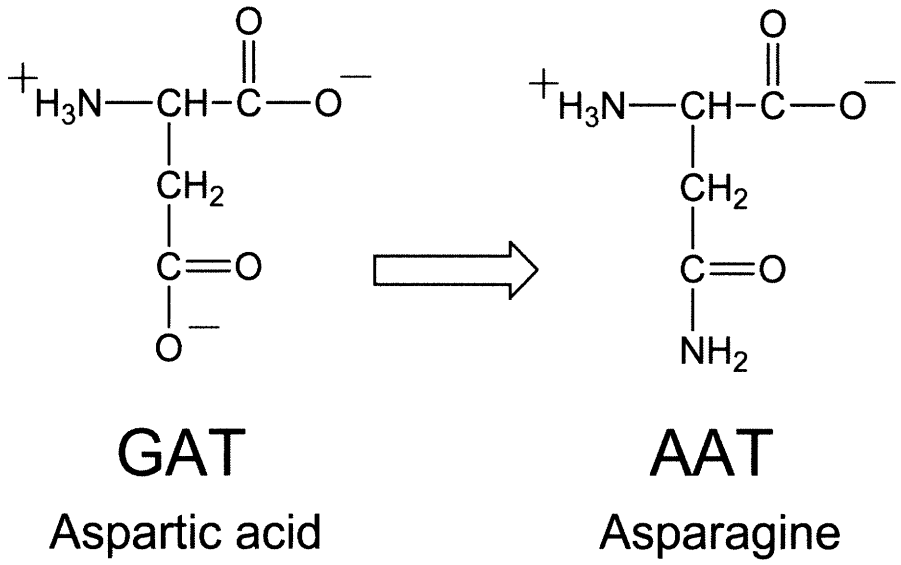


Figure 2.4: Coordination of catalytic iron by three key amino acids and the α-ketoglutarate co-substrate in the active site of AlkB. The mutated Asp133 is in the upper right hand corner.

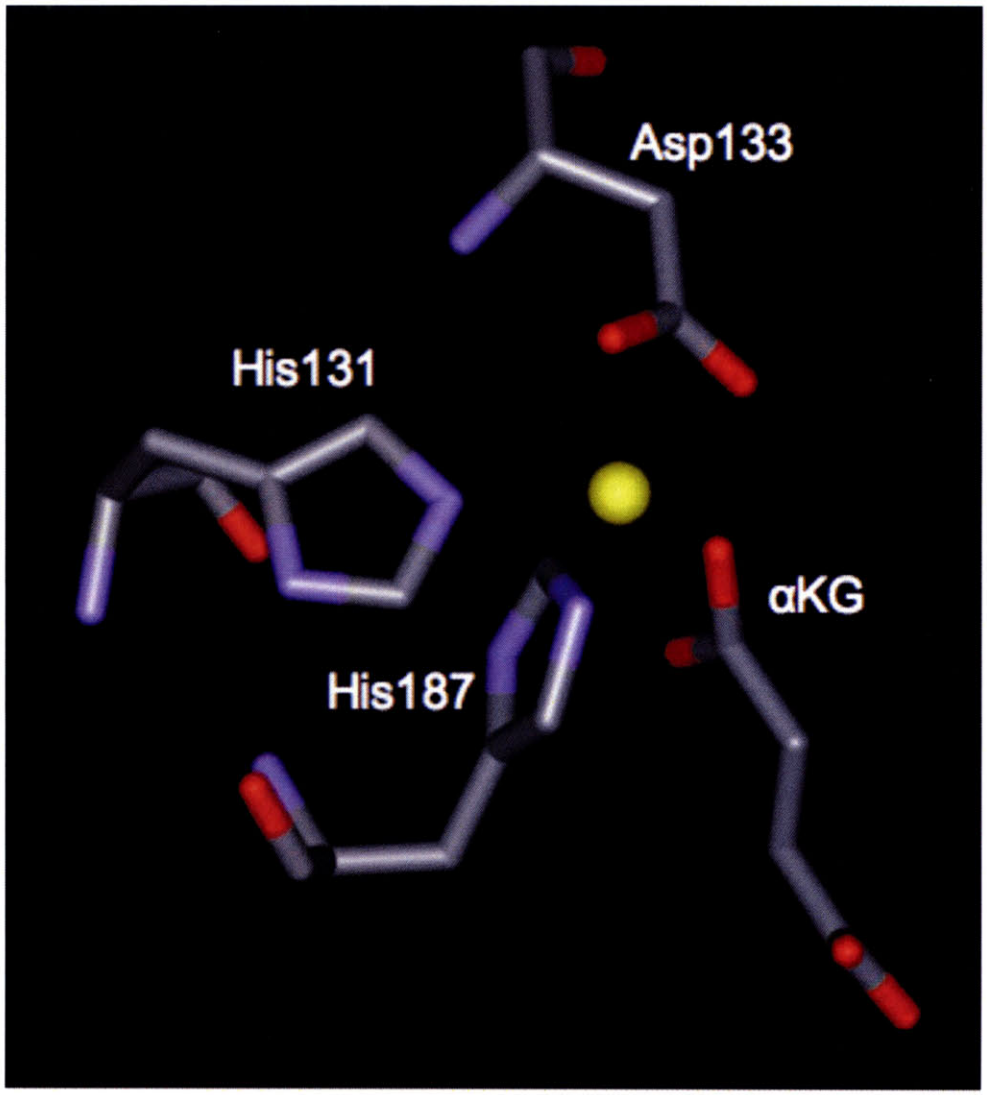


Table 2.1: Theoretical sequence of the alkB gene (NCBI Gene ID 946708) including upstream and downstream sequence. The gene is shown in black, the surrounding sequences in gray, and the primers used for PCR amplification in bold.

A TACCTTGCGG TGAAACCGTC AGT

ATGTTGGATC TGTTTGCCGA TGCTGAACCG TGGCAAGAGC CACTGGCGGC TGGTGCGGTA
ATTTTACGGC GTTTTGCTTT TAACGCTGCG GAGCAACTGA TCCGCGATAT TAATGACGTT
GCCAGCCAGT CGCCGTTTCG CCAGATGGTC ACCCCCGGGG GATATACCAT GTCGGTGGCG
ATGACCAACT GTGGGCATCT GGGCTGGACG ACCCATCGGC AAGGTTATCT CTATTCGCCC
ATTGATCCGC AAACAAATAA ACCGTGGCCC GCCATGCCAC AGAGTTTTCA TAATTTATGT
CAACGTGCGG CTACGGCGGC GGGCTATCCA GATTTCCAGC CAGATGCTTG TCTTATCAAC
CGCTACGCTC CTGGCGCGAA ACTGTCGCTG CATCAGGATA AAGACGAACC GGATCTGCGC
GCGCCAATTG TTTCTGTTTC TCTGGGCTTA CCCGCGATTT TTCAATTTGG CGGCCTGAAA
CGAAATGATC CGCTCAAACG TTTGTTGTTG GAACATGGCG ATGTGGTGGT ATGGGGCGGT
GAATCGCGGC TGTTTTATCA CGGTATTCAA CCGTTGAAAG CGGGGTTTCA TCCACTCACC
ATCGACTGCC GCTACAACCT GACATTCCGT CAGGCAGGTA AAAAAGAATA A

GCG AAGATGTTTA AACTGCGGGC T

Table 2.2: The alkB gene of HK81 (wild-type) as sequenced.

```
ATGTTGGATC TGTTTGCCGA TGCTGAACCG TGGCAAGAGC CACTGGCGGC TGGTGCGGTA
ATTTTACGGC GTTTTGCTTT TAACGCTGCG GAGCAACTGA TCCGCGATAT TAATGACGTT
GCCAGCCAGT CGCCGTTTCG CCAGATGGTC ACCCCCGGGG GATATACCAT GTCGGTGGCG
ATGACCAACT GTGGGCATCT GGGCTGGACG ACCCATCGGC AAGGTTATCT CTATTCGCCC
ATTGATCCGC AAACAAATAA ACCGTGGCCC GCCATGCCAC AGAGTTTTCA TAATTTATGT
CAACGTGCGG CTACGGCGGC GGGCTATCCA GATTTCCAGC CAGATGCTTG TCTTATCAAC
CGCTACGCTC CTGGCGCGAA ACTGTCGCTG CATCAGGATA AAGACGAACC GGATCTGCGC
GCGCCAATTG TTTCTGTTTC TCTGGGCTTA CCCGCGATTT TTCAATTTGG CGGCCTGAAA
CGAAATGATC CGCTCAAACG TTTGTTGTTG GAACATGGCG ATGTGGTGGT ATGGGGCGGT
GAATCGCGGC TGTTTTATCA CGGTATTCAA CCGTTGAAAG CGGGGTTTCA TCCACTCACC
ATCGACTGCC GCTACAACCT GACATTCCGT CAGGCAGGTA AAAAAGAATA A
```

Table 2.3: The alkB gene of HK82 (AlkB-deficient) as sequenced.

ATGTTGGATC TGTTTGCCGA TGCTGAACCG TGGCAAGAGC CACTGGCGGC TGGTGCGGTA
ATTTTACGGC GTTTTGCTTT TAACGCTGCG GAGCAACTGA TCCGCGATAT TAATGACGTT
GCCAGCCAGT CGCCGTTTCG CCAGATGGTC ACCCCCGGGG GATATACCAT GTCGGTGGCG
ATGACCAACT GTGGGCATCT GGGCTGGACG ACCCATCGGC AAGGTTATCT CTATTCGCCC
ATTGATCCGC AAACAAATAA ACCGTGGCCC GCCATGCCAC AGAGTTTTCA TAATTTATGT
CAACGTGCGG CTACGGCGGC GGGCTATCCA GATTTCCAGC CAGATGCTTG TCTTATCAAC
CGCTACGCTC CTGGCGCGAA ACTGTCGCTG CATCAGAATA AAGACGAACC GGATCTGCGC
GCGCCAATTG TTTCTGTTTC TCTGGGCTTA CCCGCGATTT TTCAATTTGG CGGCCTGAAA
CGAAATGATC CGCTCAAACG TTTGTTGTTG GAACATGGCG ATGTGGTGGT ATGGGGCGGT
GAATCGCGGC TGTTTTATCA CGGTATTCAA CCGTTGAAAG CGGGGTTTCA TCCACTCACC
ATCGACTGCC GCTACAACCT GACATTCCGT CAGGCAGGTA AAAAAGAATA A

Table 2.4: Alignment of sequencing results of the alkB genes from strains HK81 and HK82 with the theoretical wild-type sequence. All three sequences are identical except for a G to A mutation at bp 397 in HK82, shown in bold.

```
CLUSTAL FORMAT for T-COFFEE Version_1.41, CPU=1.57 sec, SCORE=74, Nseq=3, Len=651

theoretical    1 ATGTTGGATCTGTTTGCCGATGCTGAACCGTGGCAAGAGCCACTGGCG GCTGGTGCGGTA
HK81             ATGTTGGATCTGTTTGCCGATGCTGAACCGTGGCAAGAGCCACTGGCG GCTGGTGCGGTA
HK82             ATGTTGGATCTGTTTGCCGATGCTGAACCGTGGCAAGAGCCACTGGCG GCTGGTGCGGTA
                 ************************************************ ************

theoretical   61 ATTTTACGGCGTTTTGCTTTTAACGCTGCGGAGCAACTGATCCGCGAT ATTAATGACGTT
HK81             ATTTTACGGCGTTTTGCTTTTAACGCTGCGGAGCAACTGATCCGCGAT ATTAATGACGTT
HK82             ATTTTACGGCGTTTTGCTTTTAACGCTGCGGAGCAACTGATCCGCGAT ATTAATGACGTT
                 ************************************************ ************

theoretical  121 GCCAGCCAGTCGCCGTTTCGCCAGATGGTCACCCCCGGGGGATATACC ATGTCGGTGGCG
HK81             GCCAGCCAGTCGCCGTTTCGCCAGATGGTCACCCCCGGGGGATATACC ATGTCGGTGGCG
HK82             GCCAGCCAGTCGCCGTTTCGCCAGATGGTCACCCCCGGGGGATATACC ATGTCGGTGGCG
                 ************************************************ ************

theoretical  181 ATGACCAACTGTGGGCATCTGGGCTGGACGACCCATCGGCAAGGTTAT CTCTATTCGCCC
HK81             ATGACCAACTGTGGGCATCTGGGCTGGACGACCCATCGGCAAGGTTAT CTCTATTCGCCC
HK82             ATGACCAACTGTGGGCATCTGGGCTGGACGACCCATCGGCAAGGTTAT CTCTATTCGCCC
                 ************************************************ ************

theoretical  241 ATTGATCCGCAAACAAATAAACCGTGGCCCGCCATGCCACAGAGTTTT CATAATTTATGT
HK81             ATTGATCCGCAAACAAATAAACCGTGGCCCGCCATGCCACAGAGTTTT CATAATTTATGT
HK82             ATTGATCCGCAAACAAATAAACCGTGGCCCGCCATGCCACAGAGTTTT CATAATTTATGT
                 ************************************************ ************

theoretical  301 CAACGTGCGGCTACGGCGGCGGGCTATCCAGATTTCCAGCCAGATGCT TGTCTTATCAAC
HK81             CAACGTGCGGCTACGGCGGCGGGCTATCCAGATTTCCAGCCAGATGCT TGTCTTATCAAC
HK82             CAACGTGCGGCTACGGCGGCGGGCTATCCAGATTTCCAGCCAGATGCT TGTCTTATCAAC
                 ************************************************ ************

theoretical  361 CGCTACGCTCCTGGCGCGAAACTGTCGCTGCATCAG GATAAAGACGAACCGGATCTGCGC
HK81             CGCTACGCTCCTGGCGCGAAACTGTCGCTGCATCAG GATAAAGACGAACCGGATCTGCGC
HK82             CGCTACGCTCCTGGCGCGAAACTGTCGCTGCATCAG AATAAAGACGAACCGGATCTGCGC
                 ************************************ ********** ************

theoretical  421 GCGCCAATTGTTTCTGTTTCTCTGGGCTTACCCGCGATTTTTCAATTT GGCGGCCTGAAA
HK81             GCGCCAATTGTTTCTGTTTCTCTGGGCTTACCCGCGATTTTTCAATTT GGCGGCCTGAAA
HK82             GCGCCAATTGTTTCTGTTTCTCTGGGCTTACCCGCGATTTTTCAATTT GGCGGCCTGAAA
                 ************************************************ ************

theoretical  481 CGAAATGATCCGCTCAAACGTTTGTTGTTGGAACATGGCGATGTGGTG GTATGGGGCGGT
HK81             CGAAATGATCCGCTCAAACGTTTGTTGTTGGAACATGGCGATGTGGTG GTATGGGGCGGT
HK82             CGAAATGATCCGCTCAAACGTTTGTTGTTGGAACATGGCGATGTGGTG GTATGGGGCGGT
                 ************************************************ ************

theoretical  541 GAATCGCGGCTGTTTTATCACGGTATTCAACCGTTGAAAGCGGGGTTT CATCCACTCACC
HK81             GAATCGCGGCTGTTTTATCACGGTATTCAACCGTTGAAAGCGGGGTTT CATCCACTCACC
HK82             GAATCGCGGCTGTTTTATCACGGTATTCAACCGTTGAAAGCGGGGTTT CATCCACTCACC
                 ************************************************ ************

theoretical  601 ATCGACTGCCGCTACAACCTGACATTCCGTCAGGCAGGTAAAAAAGAA TAA
HK81             ATCGACTGCCGCTACAACCTGACATTCCGTCAGGCAGGTAAAAAAGAA TAA
HK82             ATCGACTGCCGCTACAACCTGACATTCCGTCAGGCAGGTAAAAAAGAA TAA
                 ************************************************ ***
```

References

1. Sedgwick, B. 2004. Repairing DNA-methylation damage. *Nat. Rev. Mol. Cell Biol.* **5**:148-157.

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

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

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

5. Delaney, J. C., L. Smeester, C. Wong, L. E. Frick, K. Taghizadeh, J. S. Wishnok, C. L. Drennan, L. D. Samson, and J. M. Essigmann. 2005. AlkB reverses etheno DNA lesions caused by lipid oxidation in vitro and in vivo. *Nat. Struct. Mol. Biol.* **12**:855-860.

6. Mishina, Y., C. G. Yang, and C. He. 2005. Direct repair of the exocyclic DNA adduct 1,N6-ethenoadenine by the DNA repair AlkB proteins. *J. Am. Chem. Soc.* **127**:14594-14595.

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

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

9. Delaney, J. C. and J. M. Essigmann. 2004. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A* **101**:14051-14056.

10. Koivisto, P., T. Duncan, T. Lindahl, and B. Sedgwick. 2003. Minimal methylated substrate and extended substrate range of Escherichia coli AlkB protein, a 1-methyladenine-DNA dioxygenase. *J. Biol. Chem.* **278**:44348-44354.

11. Falnes, P. O. 2004. Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. *Nucleic Acids Res.* **32**:6260-6267.

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

13. Rebeck, G. W., S. Coons, P. Carroll, and L. Samson. 1988. A second DNA methyltransferase repair enzyme in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A* **85**:3039-3043.

14. Yu, B., W. C. Edstrom, J. Benach, Y. Hamuro, P. C. Weber, B. R. Gibney, and J. F. Hunt. 2006. Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. *Nature* **439**:879-884.

15. Lee, D. H., S. G. Jin, S. Cai, Y. Chen, G. P. Pfeifer, and T. R. O'Connor. 2005. Repair of methylation damage in DNA and RNA by mammalian AlkB homologues. *J. Biol. Chem.* **280**:39448-39459.

CHAPTER 3:

AlkB Reverses Etheno DNA Lesions Caused by Lipid Oxidation *in Vitro* and *in Vivo*

Reproduced in part from: Delaney et al. *Nature Structural & Molecular Biology* 12, 855 - 860 (2005)

Abstract

Oxidative stress converts lipids into DNA damaging agents that form complex genomic lesions, including 1,N^6-ethenoadenine (εA) and 3,N^4-ethenocytosine (εC), in which two carbons of the lipid alkyl chain form an exocyclic adduct with a DNA base. Here we show that the newly characterized enzyme AlkB repairs εA and εC. The potent toxicity and mutagenicity of εA in *E. coli* lacking AlkB was completely reversed in $AlkB^+$ cells; AlkB also mitigated the effects of εC. *In vitro*, AlkB cleaved the lipid-derived alkyl chain from DNA, returning εA and εC to adenine and cytosine, respectively. Mass spectrometry revealed putative intermediates along the εA repair pathway consistent with epoxidation of the 1,N^6 double bond followed by hydrolysis to the glycol and release of the dialdehyde glyoxal. This new class of substrate lesions demonstrates a previously unrecognized chemical versatility of AlkB. In mammals, the corresponding AlkB homologs may be a critical defense against aging, cancer and oxidative stress.

Introduction

The DNA repair protein AlkB is an α-ketoglutarate dioxygenase that requires non-heme iron (II), oxygen, and α-ketoglutarate to repair 1-alkylpurines and 3-alkylpyrimidines in DNA (*1-8*). AlkB couples the oxidative decarboxylation of α-ketoglutarate with hydroxylation of alkylated DNA to yield lesion-free DNA and a jettisoned aldehyde. Remarkably, this enzyme and its human homolog hABH3 can also catalyze the repair of RNA (*9,10*), making alkB and its human homologs versatile components for maintaining the fidelity of replication, transcription, and translation during times of alkylation stress (*11*).

Oxidative stress is another source of cellular damage that has been linked to cancer promotion, inflammation, neurological disease, and aging (*12-16*). Aerobic respiration enhances the efficiency of ATP generation, but generates reactive oxygen species (ROS) byproducts that, along with ROS produced by macrophages, can attack DNA both directly and indirectly, generating a variety of DNA lesions; if left unrepaired, those lesions can be both cytotoxic and mutagenic (*13,15-17*). 1,N^6-Ethenoadenine (εA) and 3,N^4-ethenocytosine (εC) (Fig. 3.1) result from the reaction of adenine and cytosine with the breakdown products of oxidatively damaged unsaturated lipids (*17-21*). These complex DNA lesions arise endogenously in both rodents and humans (*22,23*); moreover, they are efficiently induced in chronically inflamed human tissues and in rodents and possibly humans exposed to vinyl chloride and its metabolites (*24*). *In vitro*, εA (*25*) and εC (*26*) in duplex DNA can be repaired by glycosylases of the base excision repair pathway. Although it is clear that mammalian cells can repair etheno adducts by this pathway *in vivo* (*27,28*), the co-involvement of other pathways has not been ruled out. Repair *in vivo* of etheno adducts in *E. coli* is less clearly understood; one study showed that neither base excision repair nor nucleotide excision repair plays a major role in etheno adduct repair (*29*). Interestingly, *E. coli* induced for the adaptive response are resistant to induction of mutation by the vinyl chloride metabolite chloroacetaldehyde (CAA) (*30,31*) suggesting that one or more of the four genes induced during adaptation may play a physiologically important protective role against its DNA adducts. Here, we

show that etheno DNA base lesions are repaired by the *E. coli* AlkB protein, which is produced as part of the adaptive response (*11*).

Materials and Methods

Oligonucleotide synthesis. Oligonucleotides of sequence 5' GAAGACCTXGGCGTCC 3', where X is εA or εC, were made using phosphoramidites (Chemgenes, Wilmington, MA) and were deprotected in concentrated NH_4OH at room temperature (2 h for *in vivo* studies, 1 h for *in vitro* studies), after which they were lyophilized, gel (*in vivo* studies) or anion exchange (*in vitro* studies) purified, and characterized by MALDI-TOF as described (7). Oligonucleotide masses for the εA- and εC-containing oligonucleotides were 4,914.20 (4,914.20 calculated) and 4,890.19 (4,890.18 calculated), respectively. DNA concentration was determined by UV absorbance using the extinction coefficients at 260 nm of 172,700 $L \cdot mol^{-1} \cdot cm^{-1}$ for the εA 16-mer, and 164,600 $L \cdot mol^{-1} \cdot cm^{-1}$ for the εC 16-mer. The oligonucleotides containing the lesions in Fig. 3.1 were inserted into single-stranded M13 bacteriophage as described, but using a 1/1.25/1.5 ratio of linearized vector/scaffolds/oligonucleotide insert (*7*).

Genotoxicity and mutagenicity analysis. The replication-inhibition data were generated using a modification of the CRAB assay (*7*) which obviates the need for counting plaques and allows both lesion bypass and mutagenesis analysis from a single transfection (Fig. 3.2).

Genotoxicity analysis. The concentrations of viable lesion-bearing and control (T at lesion site) genomes containing the sequence 5' GAAGACCTXGGCGTCC 3' were equalized (Fig. 3.3) and mixed with ~10% of a competitor M13 genome containing the sequence 5' GAAGACCTGGTAGCGCAGG 3', after which the mixture was electroporated into HK81 (AB1157, *nalA*) and HK82 (as HK81 but *alkB22*) *E. coli* in triplicate using a multi-well HT-200 electroporator (one pulse for 5 samples) set at 2500 V, 125 Ω, 50 μF (Harvard Apparatus). Cells were grown in liquid culture and worked up as described (*7*). If the lesion blocks replication, there will be proportionately more

output from the competitor. Output was measured by PCR amplification of DNA from the progeny. The primer 5'-YCAGCTATGACCATGATTCAGTGGAAGAC-3' (Y is an aminoethoxyethyl ether group) is identical in sequence to the vector and 5' end of the lesion-bearing, control, and competitor oligonucleotide inserts, allowing for specific amplification of these genomes. The primer 5'-YCAGGGTTTTCCCAGTCACGAC GTTGTAA-3' anneals only to the invariant region of the M13 vector and does not span the insert region. PCR products were worked up as described for the REAP assay (*7*), whereby the output mixture was cut with *Bbs*I, 5'-end-labeled at the newly exposed site that contained the lesion in the progenitor genome, trimmed with *Hae*III, and electrophoresed on a 20% polyacrylamide gel. Band intensity from the lesion output (18-mer) was divided by that of the competitor output (21-mer), and this ratio was divided by that obtained from a non-lesion (T at lesion site) and competitor control mixture to obtain the percentage of lesion bypass *in vivo*. Genomes containing each lesion were electroporated in triplicate for each cell type, and the average is reported with one standard deviation in Figure 3.4.

Mutagenicity analysis. The REAP assay (Fig 3.2) ultimately radiolabels the site that originally contained the lesion and partitions the radioactivity on a TLC plate, revealing the post-replication base composition and, hence, the mutation frequency and specificity of the lesion (*7*). Because the intensity of the 5'-^{32}P-(lesion site)-18-mer output signal was substantially less in situations in which the lesion was strongly blocking, the primers described above for the CRAB bypass assay could not be employed in the REAP assay. Equally robust signals were therefore achieved by specific amplification of progeny from the lesion-bearing genome using the primers 5'-YCAGCTATGACCATGATTCAGTG GAAGAC-3' and 5'-YTGTAAAACGACGGCCAGTGAATTGGACG-3', which were specific for both vector-insert junctions, but mismatched in the latter primer when annealed to the competitor. Increasing the annealing temperature from 64°C to 67°C eliminated the trace amount of competitor signal still present. After digestion with *Bbs*I to expose the site that had contained the lesion in the progenitor genome, radiolabeling, *Hae*III trimming, and PAGE purifying of the 5'-^{32}P-18-mer, the oligonucleotide was gel

purified and digested to 5'-^{32}P-dNMPs, which were separated on a TLC plate and quantified by PhosphorImagery as described (*7*). The average is reported with one standard deviation in Figure 3.5.

AlkB cloning and purification. The *E. coli* AlkB open reading frame was cloned into a 6× N-terminal histidine-tagged pET28(a) expression vector (Novagen, Madison, WI) by PCR using Platinum Taq (Invitrogen, Carlsbad, CA) and *E. coli* genomic DNA as a template. Two preparations of AlkB were used for generating the *in vitro* mass spectral data (Figs. 3.6 - 3.9). Preparation A was used to generate the data in Figs. 3.7a, 3.8 and 3.9. Preparation B was used in Figs. 3.6 and 3.7b.

***Preparation A*:** The His-tagged *E. coli* AlkB expression construct was transformed into BL21(DE3) *E. coli* (Novagen), and expression was induced by adding 1 mM IPTG to a culture at A_{600} 0.4 and continuing incubation for 4 h at 37°C. Cells were harvested by centrifugation and lysed in 50 mM Tris pH 8.0, 200 mM NaCl, 1 mM EDTA, and 0.1 mg ml^{-1} phenylmethanesulfonyl fluoride by sonication on ice. Lysates were centrifuged at 20,000 rpm for 15 min at 4°C and the soluble fractions were collected. Tagged proteins were purified using 'His-Spin Protein Miniprep' columns (Zymo Research, Orange, CA) according to manufacturer's instructions and were eluted in 50 mM sodium phosphate pH 7.7, 300 mM sodium chloride, and 50 mM imidazole. The concentration of AlkB (50 pmol μl^{-1}) was determined on a microfluidic 2100 Bioanalyzer workstation using a Protein 200 Plus LabChip Kit (Agilent, Palo Alto, CA).

***Preparation B*:** AlkB in Rosetta 2(DE3) cells were grown in LB media at 37°C, induced with 1 mM IPTG at 0.6 OD_{600} and grown at 30°C for 4 h. Cells were harvested by centrifugation, homogenized in lysis buffer (10 mM Tris pH 7.34, 300 mM NaCl, 5% glycerol, 2 mM $CaCl_2$, 10 mM $MgCl_2$, 1 mM 2-mercaptoethanol) by sonication, and centrifuged at 35,000*g* for 20 min. The lysate was immediately loaded onto a Ni-NTA column, washed with 10-column volumes (CV) of 10 mM imidazole in lysis buffer, followed by 5 CV of 20 mM imidazole in lysis buffer, and eluted in 5 CV of 70 mM and

5 CV of 250 mM imidazole in lysis buffer. The 70 mM and 250 mM imidazole fractions containing AlkB were dialyzed against 4 l of buffer A (50 mM 2-[(tris(hydroxymethyl)methyl)amino]-1-ethanesulfonic acid (TES), pH 7.1) overnight. The supernatant was loaded onto a SP Sepharose cation exchange column (Amersham Pharmacia, 25 ml) that had been equilibrated with buffer A. AlkB was eluted with a linear gradient of NaCl (0.0–1.0 M in 75 ml). Fractions containing AlkB were pooled and purity was established by SDS-PAGE. The concentration of AlkB (57 pmol μl^{-1}) was determined by ultraviolet absorbance at 280 nm.

***In vitro* reactions with AlkB.** The kinetics experiment in Fig. 3.6 was performed at 22°C, while all other AlkB reactions were performed at 37°C as described (*2*). All used 45 mM Tris (pH 8.0), 67 μM $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$, 0.9 mM α-ketoglutarate, and 1.8 mM ascorbate. Repair reactions with DNA were with the 16mer 5' GAAGACCTXGGCGTCC 3', (X is the lesion). Reactions using AlkB from preparation A to be analyzed by ESI-TOF mass spectrometry were performed for the indicated times with 20 μM DNA and 10 μM AlkB in 10 μL (Figs. 3.7a, 3.9); preparation A was also used to generate Fig. 3.8. Reactions were kept on dry ice until ESI-TOF analysis. Reactions using AlkB from preparation B to be analyzed by GC-MS were performed with 40 μM DNA and 20 μM AlkB in 25 μL for 30 min (Figs. 3.7b, 3.7c)); this preparation was also used in the kinetics experiment in Fig. 3.4, which employed 2.5 μM DNA and 1.75 μM AlkB, with 25 pmol DNA analyzed by MALDI-TOF for conversion of m1A or εA to A (the reactions were quenched by cofactor removal with a C18-ZipTip (Millipore) and were worked up as described (*7*)). Reconstruction experiments showed expected ratios of m1A:A and εA:A for the tested ratios of 1:3, 1:1, and 3:1. Standards of formaldehyde (made fresh from paraformaldehyde (Sigma)) and glyoxal (40% wt/vol (Aldrich); concentration verified by ^1H NMR integration after derivatization with hydroxylamine (Aldrich) and comparison with an internal standard of acetonitrile in D_2O) were also treated with AlkB for standard curve generation (r^2 = 0.99, peer reviewed data not shown). After 30 min, an internal standard consisting of 3 nmol glutaraldehyde (Fluka) were mixed with the reaction, and 60 μL of a 0.2 M solution of

PFBHA was added. After a 20 min room temperature incubation, the samples were extracted with 100 μL ethyl acetate, and the organic layer to be used for GC-MS analysis was then extracted with 100 μL of 50 mM Tris (pH 8.0).

E. coli survival assay. To ascertain the relative contributions of AlkA and AlkB to the protection of cells from CAA (Fig. 3.10), a pBAD24 vector conferring carbenicillin resistance was transformed into the isogenic pairs MV1161 *(AB1157* derivative, AlkA-proficient) and MV1174 (as MV1161 *alkA1,* AlkA-deficient) (*2*), and HK81 (AB1157 derivative, AlkB-proficient) and HK82 (as HK81 *alkB22*, AlkB-deficient) (*7*). After growth overnight at 37°C in LB containing carbenicillin (50 μg ml^{-1}), cells were plated on LB containing carbenicillin (50 μg ml^{-1}) and arabinose (0.02%) at target dilutions that, when plated, would yield either 250 or 2500 cells per 150 mm diameter control plate (that did not contain chloroacetaldehyde). Plates containing CAA (Aldrich) were made in large batches and were used within 12 h. After incubation at 37°C, surviving colonies were counted, and the average of two or more replicates with between 10 and 300 colonies per plate was used to generate Figure 3.10, with % survival expressed as the number of colonies formed on a CAA plate divided by that of the CAA-free control plate times 100%. Additional survival assays are described in Supplemental Information.

Instrumentation.

MALDI-TOF MS. Experiments were performed using a PerSeptive Biosystems Voyager-DE STR BioSpectrometry Workstation in the negative ion linear mode as described (*7*).

LC-ESI-TOF MS. Experiments were done on an Agilent 1100 MSD ESI-TOF, an electrospray time-of-flight mass spectrometer that allows simultaneous introduction of sample and a reference solution via separate spray assemblies. The instrument was operated in the negative-ion mode and calibrated with a commercial tuning mixture just prior to analyses; resolution and mass accuracy were typically higher than 10,000 and less than 1 ppm, respectively. Capillary voltage was –3500 V, nebulizing gas pressure

was 35 psi, drying gas flow was 12 l min^{-1} at 325°C. The entire reaction volume was introduced with an Agilent 1100 Series thermostatted autosampler via a C-18 column (Agilent Poroshell 300SB-C18 5 μm; 2.1 mm × 75 mm) at 200 μl min^{-1} with a gradient from 3%–15% B over 5 min followed by 15%–3% B over 5 min with a 10 min equilibration time (A: 10 mM ammonium acetate; B: 100% acetonitrile). Robust signals were obtained for all of the compounds of interest, and deconvolutions for multiple-charge-states were done with in-house software, reporting values for the monoisotopic (100% ^{12}C) peaks.

GC-MS. Formaldehyde and glyoxal release were quantified as their PFBHA derivatives with an Agilent 5973N GC-MS system consisting of an Agilent 5890N Network gas chromatograph (GC) and an Agilent 5973N Mass Selective Detector (MSD) using methane as the moderating gas. The GC column consisted of a 30 meter HP-5MS, 5% cross-linked phenyl methyl siloxane stationary phase with 0.25 μm film thinness and 0.25 mm i.d. All the injections were done with an autosampler (1.5 μl for NCI mode or 2.0 μl for PCI mode) in pulsed splitless mode with an injection pressure of 50 psi for 1 min and temperature of 250°C, with helium as a carrier gas at a flow rate of 1 ml min^{-1}. The GC temperature program was 50°C for 1.5 min, ramped to 310°C at 10°C min^{-1}, and finally held at 310°C for 5 min. Transfer line temperature was set at 280°C and MSD source at 150°C for NCI mode (used for quantification) or 250°C for PCI mode (used for confirming the parent ion for each standard). MSD was scanned from 50–500 Da (NCI mode) or 50–550 Da (PCI mode) with a scan rate of 3.25 scans s^{-1}. The GC-MS system control, data collection and data processing were done using Agilent MSD Chemstation software. The characteristic daughter ions obtained from fragmentation of the PFBHA-aldehyde complex in the NCI mode was used for quantification. The extracted ion traces shown in Fig. 3.7b correspond to formaldehyde (205.1, black), glyoxal (267.1, green), and glutaraldehyde (450.1, blue).

Results

To measure the influence of AlkB on the biological response toward etheno lesions, we placed εA, εC, and control lesions into single-stranded M13 viral genomes, which were passaged through *E. coli* of three types: AlkB proficient, AlkB deficient, or AlkB deficient but induced for the SOS bypass polymerases by prior treatment of the cells with UV light as described (*7*) (Fig. 3.1). Using the competitive replication of adduct bypass (CRAB) assay and the restriction endonuclease and postlabeling analysis of mutation frequency (REAP) assay (*7*) (Figs. 3.2 and 3.3), we measured the genotoxicity and mutagenicity of each lesion in these three types of *E. coli.*

Over a decade ago, we evaluated the mutagenicity of εA and εC lesions in a single-stranded context in wild-type (DL7) *E. coli.* The mutation frequencies were only 0.1% and 2%, respectively (*32*). Here we show that the mutagenicity of εA in an M13 single-stranded viral genome processed in *E. coli* lacking AlkB is increased by over 2 orders of magnitude with respect to that found in previous studies of AlkB positive cells. As shown in Fig. 3.2, εA does not appear to be very toxic towards M13 replication in wild-type cells, as εA-containing genomes were replicated 85% as efficiently as a non-lesion control; however, the 5% replication efficiency in AlkB-deficient cells demonstrates that this lesion is extremely toxic to M13 replication when unrepaired. Although this toxicity can be partially overcome by SOS bypass polymerases, AlkB is clearly the main mechanism for avoiding εA-induced genotoxicity within single-stranded DNA. We found εA to be negligibly mutagenic (<0.5%) in wild-type cells, agreeing well with all previous studies performed in AlkB-proficient *E. coli* (<0.3%) (*29,32-34*); however, we show εA to be 35% mutagenic in AlkB-deficient cells, yielding 25% A to T, 5% A to G and 5% A to C mutations (Fig. 3.5).

We also found AlkB to play a major role in reducing mutations from the pyrimidine etheno adduct εC (Fig. 3.5), and to have a modest effect on the bypass of εC (Fig. 3.4). We found the mutation frequency of 82% for εC in AlkB-deficient cells to be higher than in previous studies performed in cells containing AlkB (*32,35,36*); this value dropped to 37% in unirradiated AlkB-proficient cells with a proportional reduction in C to A and C to T mutations (Fig. 3.5). SOS-induction led to an increase in dAMP incorporation

opposite both εA and εC adducts, which obeys the "A-rule" for lesions lacking instructional hydrogen bonds (Fig. 3.5).

Based on the lesion bypass (Fig. 3.4) and mutagenesis (Fig. 3.5) data, the substrate preference of AlkB for the single-stranded lesions studied is εA > m1G > εC (m1G is a known substrate for AlkB (*7*)). This preference was confirmed *in vitro*, as a 16-mer containing εA was repaired much better by AlkB than one containing εC (Fig. 3.8). Concurrent analysis of trans-lesion synthesis past a chemically stable tetrahydrofuran (THF) abasic site revealed that, in the absence of AlkB, εA was as powerful a block to replication as was THF. By contrast, εC was more easily bypassed, while m1G was the strongest block to DNA polymerases.

Our *in vivo* experiments with site-specifically modified viral genomes containing m1A (*7*) and εA (this work) show that both lesions are repaired by AlkB before they are encountered by the DNA polymerase. To establish definitively that AlkB is able to repair εA at a relevant rate and to a significant extent, we reconstituted the repair reaction *in vitro* using 16-mer single-stranded oligonucleotides containing either a single εA or m1A lesion. This system facilitated comparison between this new substrate lesion (εA) and a substrate previously established (m1A). Monitoring the fraction of substrate repaired over two hours by MALDI-TOF mass spectrometry of the oligonucleotides revealed rapid and efficient repair of m1A (Fig. 3.6), as expected. εA was also repaired, albeit somewhat more slowly, and the extent of repair plateaued at >60% after one hour of reaction (Fig. 3.6). Longer incubation with AlkB resulted in only a minor increase in repair. While m1A is the overall better substrate, as it is completely repaired after only five minutes of incubation, the two lesions were repaired at comparable rates at early timepoints, indicating that AlkB is able to provide protection against εA lesions on par with that provided against m1A.

We also used the *in vitro* system to establish that the decrease in both the lethality and mutagenicity of εA was through a direct reversal mechanism. We monitored the loss of mass in the oligonucleotide by both MALDI-TOF (Figs. 3.6 and 3.8) and ESI-TOF (Fig. 3.7a and 3.9) mass spectrometry. As shown in Fig. 3.7a, ESI-TOF revealed that

incubation of the εA oligonucleotide with AlkB caused a loss of 24 amu, consistent with removal of the etheno bridge linking N6 and N1 of adenine (i.e., conversion to adenine). Over 30 min, εA disappears and is converted largely to A, as well as two less abundant species, which appear as the reaction progresses. One species is 16 amu heavier than the starting material, consistent with the εA-epoxide, while the other is 34 amu heavier, consistent with the εA-glycol.

These intermediates are consistent with the reaction scheme shown in Fig. 3.11, in which an εA epoxide and glycol are intermediates in the pathway. For reasons under investigation, the putative epoxide is stable, and not fully converted to product. Control reactions of 16mer oligonucleotides containing N^6-methyladenine (m6A), m1A, or A incubated with AlkB showed no repair of m6A (nor a change in spectra for A), and complete repair of m1A, with no epoxide or glycol signatures. These signature masses were present even after a 28 h incubation of the εA 16mer with AlkB, or after the initial reaction was supplemented with a fresh dose of cofactors and AlkB for a 1 h incubation; the epoxide signal was greatly diminshed upon heating the intermediate with concentrated ammonium hydroxide (peer reviewed data not shown). Although the pathway shown is chemically reasonable, it is also possible that the epoxide and glycol are co-products and not formal intermediates.

Figure 3.7b characterizes the final step in the release of the etheno exocyclic bridge from DNA. εA, m1A, and A 16-mers were incubated with AlkB. When the starting material was consumed, an internal standard (glutaraldehyde) and a trapping reagent (pentafluorobenzyl hydroxylamine - PFBHA) were added to the reaction mixture. The trapping reagent volatilized the reaction products released from DNA, as well as glutaraldehyde. These were analyzed by GC-MS. The results show that a product identical to the PFBHA derivative of glyoxal is trapped following release of the etheno group from etheno adenine. Hence, glyoxal is the final reaction product. Stoichiometrically, 0.77 nmol of PFBHA-glyoxal is recovered from one nmol of εA 16-mer, while 0.98 nmol of PFBHA-formaldehyde is recovered from one nmol of the control m1A 16-mer (Fig. 3.7c). The non-stoichiometric recovery of glyoxal is attributed to the stability of the εA-epoxide intermediate.

Having established that AlkB from *E. coli* can repair εA by direct reversal, and that AlkB is the predominant repair system for εA in single-stranded DNA within cells, we next sought to ascertain the relative contributions of AlkA and AlkB to *E. coli* survival upon CAA challenge, in which the majority of etheno adducts formed would presumably be in double-stranded DNA. As shown in Fig. 3.10, a deficiency in either repair protein renders mutant cells more sensitive to CAA, with loss of AlkA conferring a more severe phenotype than loss of AlkB. Much work went into establishing this result, however, the details of which are included as Supplementary Information.

Discussion

Life in an oxygen rich environment has its liabilities. The incomplete reduction of oxygen during metabolism generates reactive oxygen species. Moreover, a plethora of reactive oxygen and nitrogen species is generated during the process of inflammation (*15*). Much emphasis has recently been placed on the adverse genetic consequences that arise from the interaction of reactive oxygen and nitrogen species with DNA (*13-16,19*). Less studied has been damage to lipids, although it is known that etheno-DNA adducts are spontaneous lesions derived from lipid breakdown products. Lipid oxidation (for example, in an inflammatory reaction) gives rise to organic species that react with DNA to form toxic and mutagenic DNA adducts (*17,18,20,21*). Recently, the pathways by which etheno and related 9-carbon cyclic DNA adducts form have been elucidated (*21*). These analytical studies show that the genomes of organisms carry, in addition to formally oxidized DNA bases (e.g., 8-oxoguanine), a formidable burden of cyclic etheno, ethano and possibly 4-hydroxynonenal derived exocyclic DNA adducts. The etheno adducts are also formed by potent industrial carcinogens such as vinyl chloride and CAA (*24,37*). In the present work, we show that εA and εC, the prototypic exocyclic DNA adducts, are potently mutagenic and toxic to *E. coli* in the absence of AlkB. AlkB positive cells reverse the damage and display virtually no mutagenicity and toxicity for ethenoadenine; most of the mutagenicity of ethenocytosine is also eliminated, and there is some reduction in toxicity.

AlkB is shown here to be an amazingly versatile DNA repair enzyme. Not only can it suppress the toxicity and mutagenicity of simple alkyl residues, it can reverse the biological effects of cyclic DNA adducts previously thought to be repaired solely by base excision repair. Both AlkA and AlkB can repair εA, and under most circumstances AlkA seems to be the more important enzyme (Fig. 3.10). AlkB, which repairs damaged single-stranded DNA at a physiologically relevant rate (Figs. 3.4 and 3.5), is likely to play a specialized but important role in the cellular defense arsenal. It is not unique to have multiple enzymes repair the same lesion, but, whenever such a case arises, it is useful to ask if there is biological rationale for this biochemical redundancy. Most of the genome is duplex and not in the process of being transcribed or replicated. In such a scenario, AlkA probably plays the predominant role in repairing εA lesions. Indeed, rampant repair by AlkB could pose a hazard through the generation of glyoxal, which is itself known to be toxic. However, in situations where repair by AlkA is not feasible or would result in a toxic single-strand break (single-stranded regions in promoters, transcription bubbles, replication forks, and loop structures), the lethality and mutagenicity of such lesions can be prevented very capably by AlkB. Thus, AlkB and AlkA may have specialized niches in which they biochemically complement each other.

The only substrates reported previously to be repaired by AlkB all have the commonality that the alkyl chain or methyl group is docked at a single position, to the N1 purine or N3 pyrimidine base (*2-8*). AlkB most likely ablates these appendages via hydroxylation of the carbon vicinal to the base (the hydroxylated intermediate has yet to be detected), ultimately generating a terminal aldehyde or formaldehyde from the jettisoned alkyl group, and a lesion-free base (Fig. 3.11). Unlike the aforementioned lesions, ethenoadenine and ethenocytosine have an ethylene bridge attached to the base at two positions (Figs. 3.1 and 3.11), and we show that AlkB can repair these lesions (which are chemically dissimilar to simpler adducts such as m1A) via a different and novel mechanism. Some α-ketoglutarate dioxygenases, such as thymine hydroxylase from *Rhodotorula glutinis* (*38*), have been shown not only to hydroxylate unactivated C-H bonds, but also to epoxidize C=C double bonds. An attractive mechanistic hypothesis for repair of εA or εC by AlkB, therefore, would be epoxidation of the exocyclic double

bond followed by epoxide ring opening and ultimate liberation of the dialdehyde glyoxal. The mass spectral data depicted in Fig. 3.7 support this scenario.

It is possible that AlkB and its human homologs could act through an epoxidation mechanism on other lesions containing exocyclic double bonds anchored to the nitrogens of DNA bases. Similarly, lesions anchored at the N1-purine or N3-pyrimidine position by straight chain or cyclic alkane moieties may be hydroxylated by AlkB and its homologs, restoring DNA or generating secondary lesions.

The function and mechanism of AlkB elucidated here may contribute to our understanding of the etiology of several human diseases. Wilson's disease and primary hemochromatosis are caused by hepatic accumulation of copper and iron, respectively (*37*). Such metal overloading triggers increased lipid peroxidation which, in turn, elevates the spontaneous rate of εA and εC formation in the genome; for example, the levels of etheno adducts correlate well with the copper content of the liver in Wilson's disease (*39*). Similarly, inflammatory diseases implicated in carcinogenesis, such as Crohn's disease and ulcerative colitis, show an increase in the etheno load (*40*). As demonstrated here, the etheno DNA lesions represent a new class of DNA damage that AlkB can repair. By using the chemistry described here, it is likely that one or more of the human homologs of AlkB will also be found to defend the genome against these and similar lesions formed by the constant barrage of endogenous lipid peroxidation metabolites and by environmental agents.

Figure 3.1: Structures of alkylated bases and site-specific *in vivo* approach used. Oligonucleotides containing each lesion were ligated into the single-stranded genome of M13mp7(L2). The adduct-containing genome was replicated within *E. coli* cells that were wild-type, *alkB*, or SOS induced *alkB*. Lesion bypass and mutation frequencies were evaluated as described in Figs. 3.4 and 3.5, respectively.

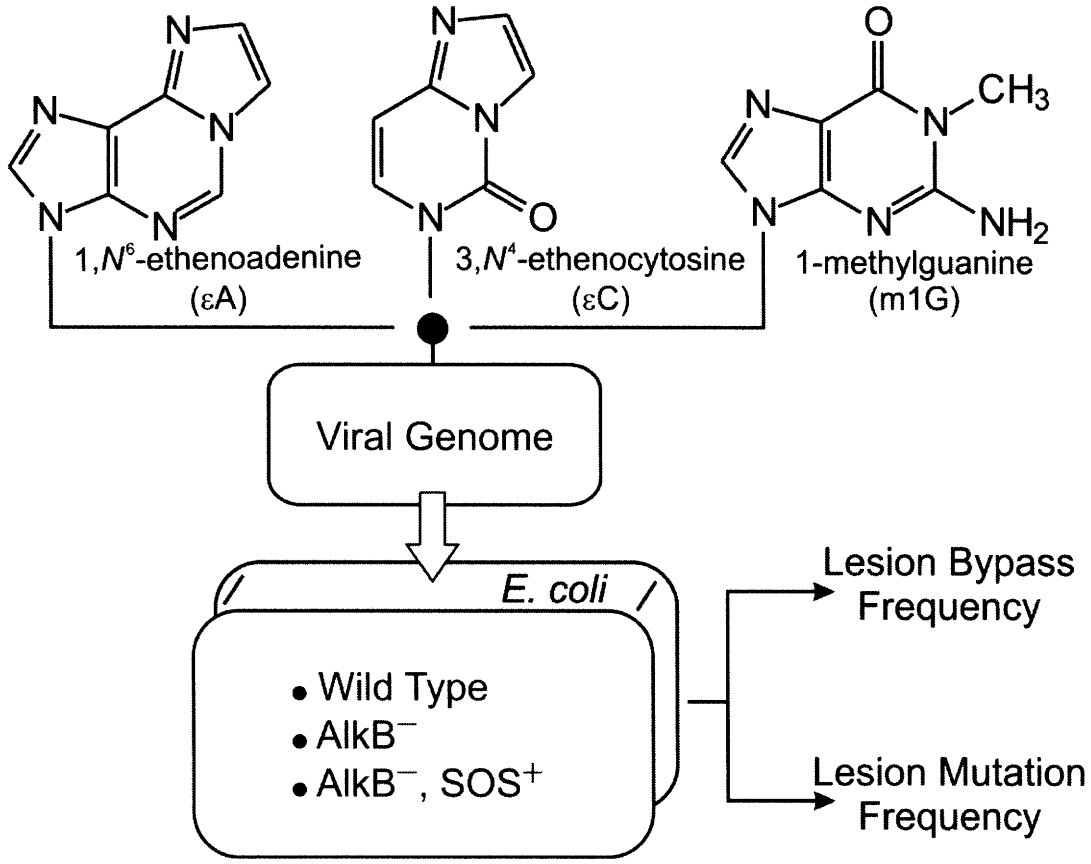


Figure 3.2: Schematic for lesion bypass and mutagenesis studies *in vivo*.

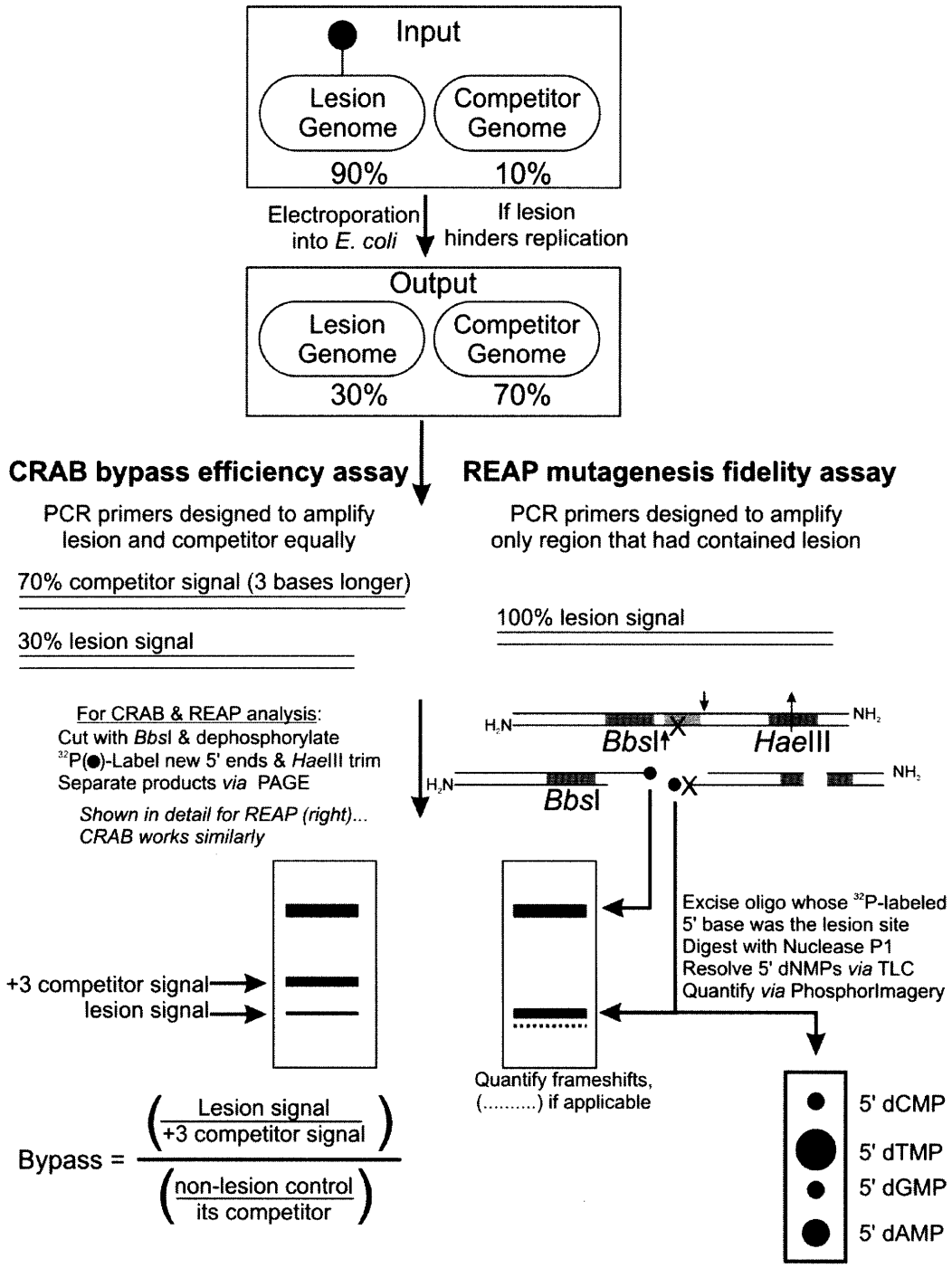


Figure 3.3: Schematic for quantifying fully ligated genomes bearing lesion used in the CRAB lesion bypass assay.

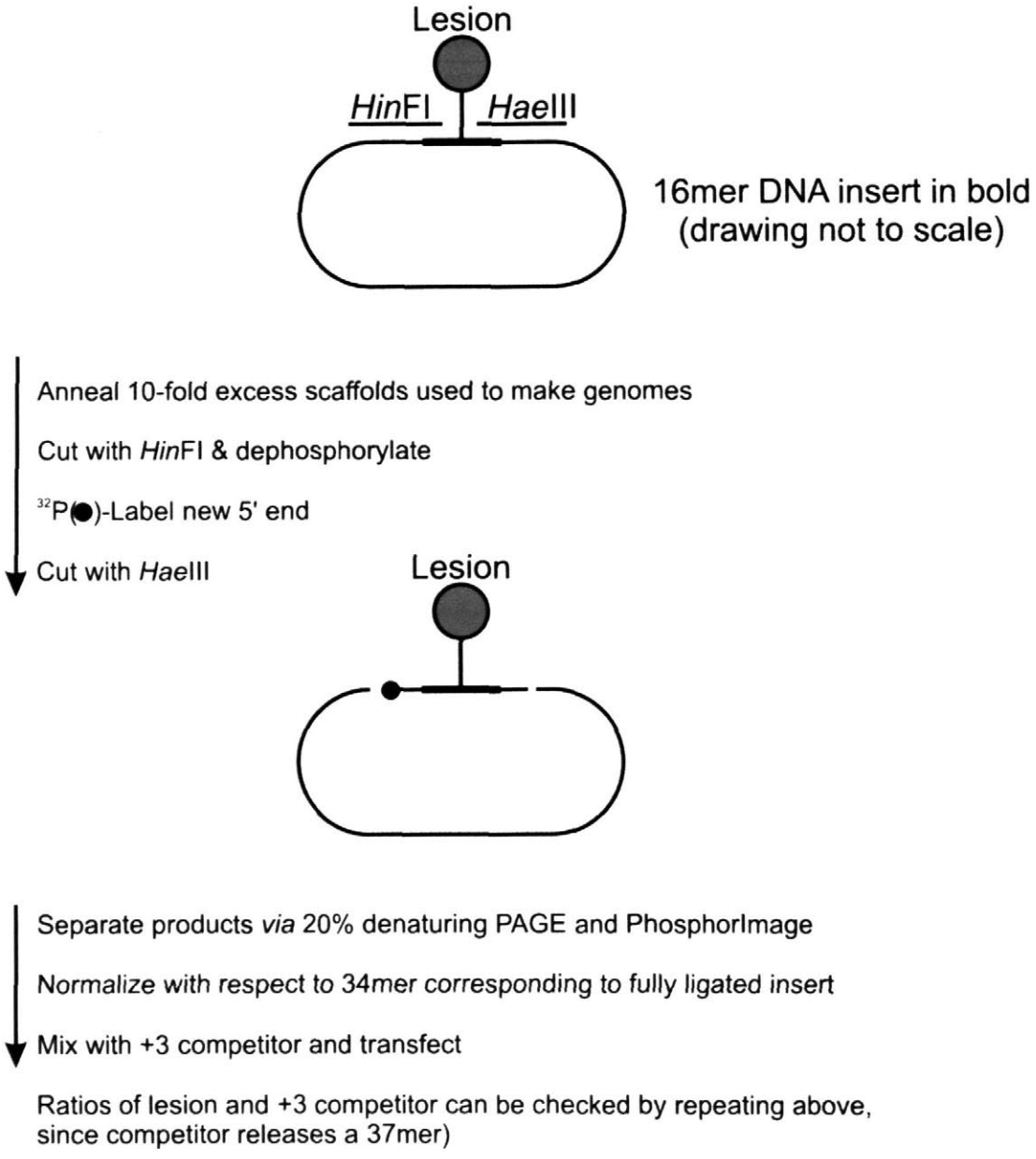


Figure 3.4: Lethality of εA, εC, m1G, and THF. The competitive replication of adduct bypass (CRAB) assay was used to determine a change in the output ratio from an electroporated mixture of lesion-bearing and non-lesion competitor genomes (Methods). If a lesion is a strong block to replication ($SOS^-/AlkB^-$ *E. coli*), a large increase in the proportion of competitor output occurs, resulting in a very low percentage of lesion bypass. In contrast, partial blockade relief through lesion repair by AlkB or induction of SOS bypass DNA polymerases results in a less drastic increase in competitor output, yielding a higher percentage of lesion bypass. Modified and control genomes (T at lesion site) were constructed and normalized to one another before being mixed with competitor genome and electroporated in triplicate. Each mixture was transformed into wild-type ($SOS^-/AlkB^+$, black), AlkB-deficient ($SOS^-/AlkB^-$, cross-hatched) and UV-irradiated AlkB-deficient ($SOS^+/AlkB^-$, white) *E. coli* and bypass was scored as described (Methods), and reported with one standard deviation of error. The large drop in the percentage of lesion bypass for εA in the absence of AlkB suggests that in AlkB-proficient cells, the enzyme is able to repair the lesion prior to replication by DNA polymerase.

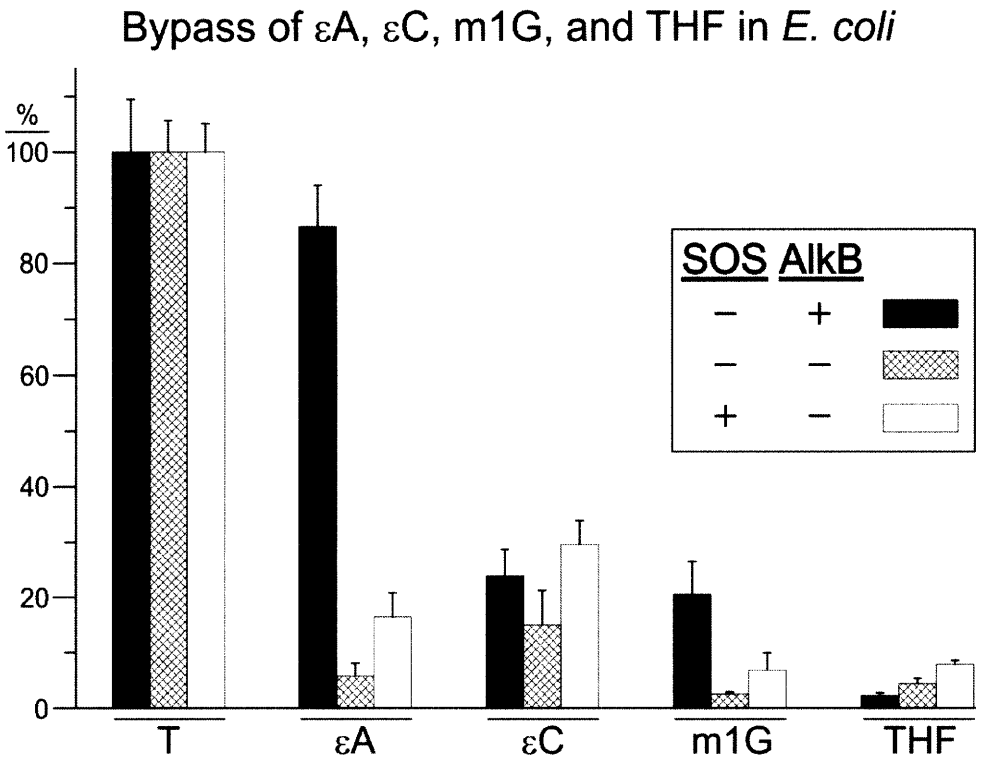


Figure 3.5: Mutagenesis of εA, εC, and m1G in *E. coli*. The restriction endonuclease and postlabeling analysis of mutation frequency (REAP) assay was used to determine the base composition (mutation frequency and specificity) at the lesion site by radiolabeling the site after cellular processing and partitioning the radioactive bases on a TLC plate (Methods). The output from the REAP assay is reported with one standard deviation of error. The percentage of G (black), A (cross-hatched), T (white), and C (hatched) at the lesion site reveals that the absence of AlkB causes a substantial increase in εA and m1G mutagenicity, as well as a quantitative increase in εC mutagenicity.

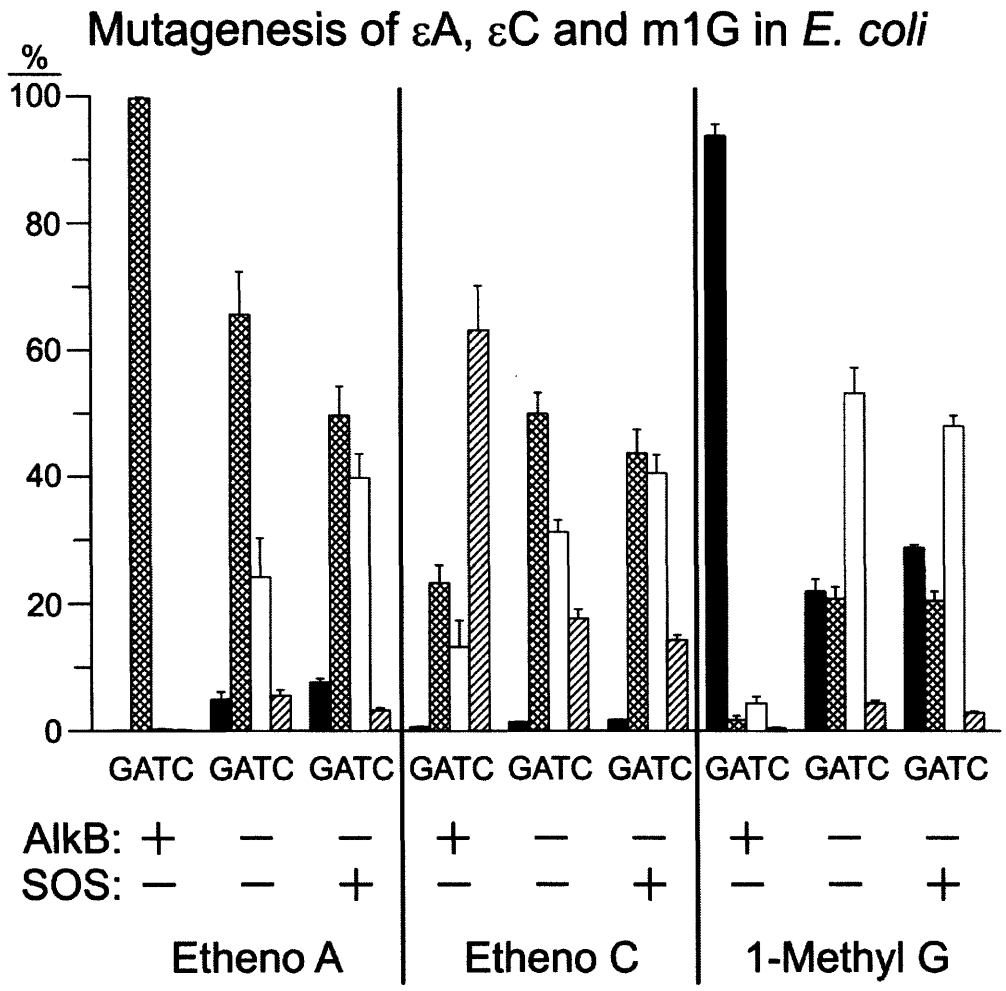


Figure 3.6: AlkB-mediated direct reversal of εA and m1A *in vitro*. The kinetics of εA vs. m1A repair by AlkB was followed by MALDI-TOF. Reactions were performed in triplicate (the 60 min point was done in duplicate), with one standard deviation of error shown.

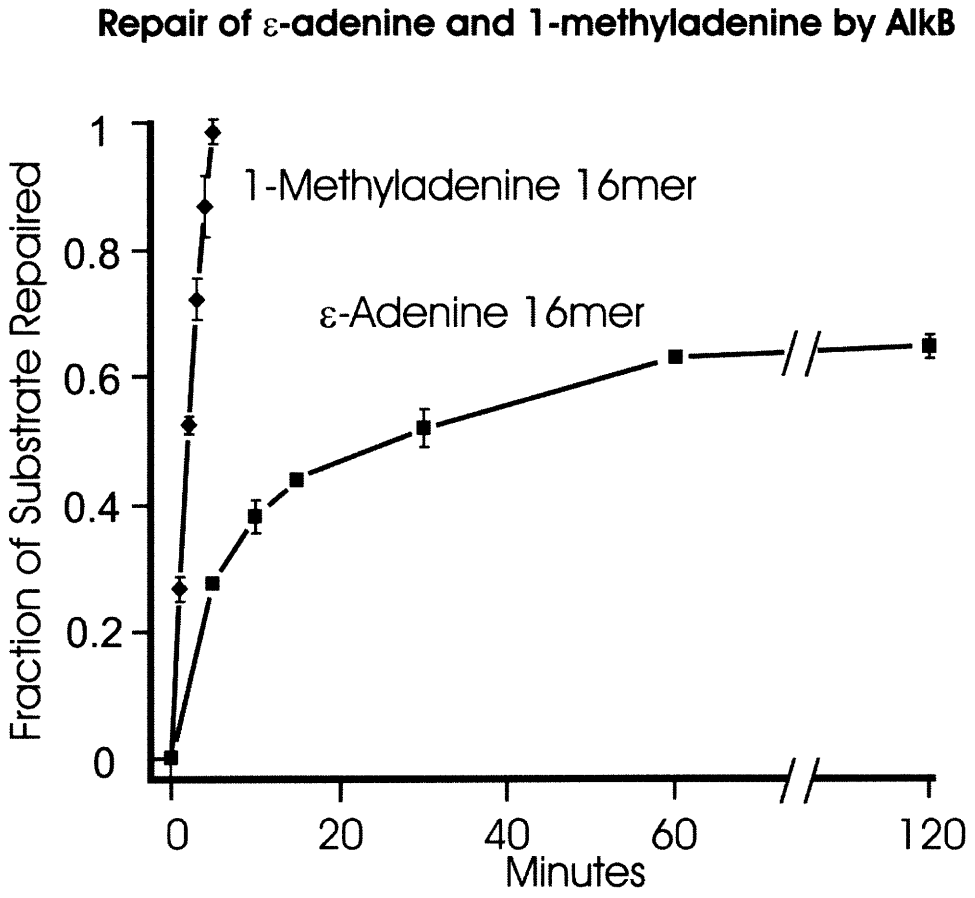
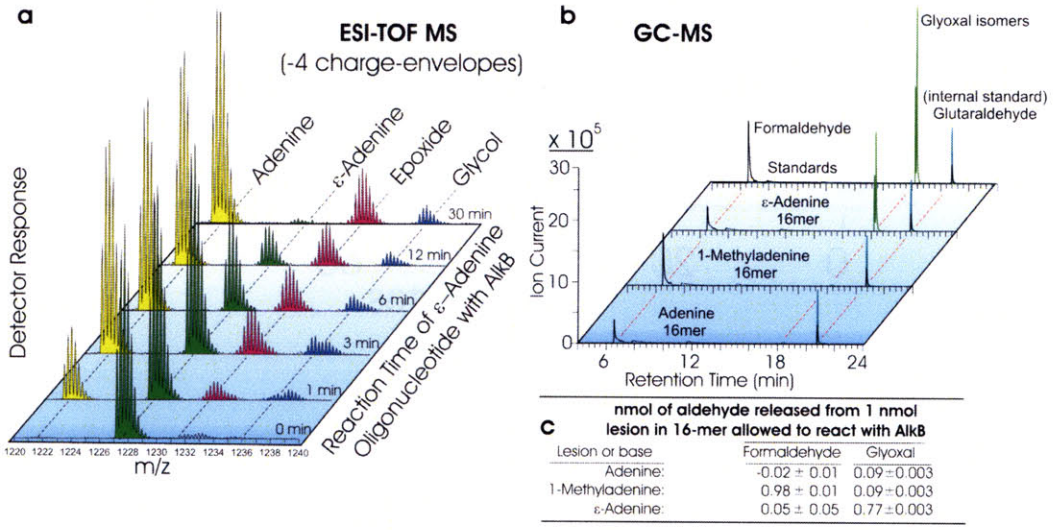


Figure 3.7: AlkB-mediated direct reversal of εA *in vitro*. **a,** Electrospray ionization time-of-flight mass spectral analysis of 16-mers during a 30 minute reaction with AlkB. Products and εA-epoxide and εA-glycol intermediates were evident as early as 1 min. Data represent the -4 charge-envelopes; multiple ion mass peaks associated with each envelope reflect the number of ^{13}C atoms in each ionic species. **b,** Liberation of glyoxal from εA by AlkB. Oligonucleotides containing A, m1A, or εA were allowed to react with AlkB until no m1A or εA starting material was left, after which time an internal standard of glutaraldehyde was added, followed by the carbonyl derivatization reagent pentafluorobenzyl hydroxylamine. PFBHA-derivatized products were separated by gas chromatography and detected in a mass spectrometer connected in tandem. Repair of εA yielded glyoxal (which forms a doublet), while repair of m1A formed formaldehyde. **c,** The amount of formaldehyde and glyoxal released during the AlkB reaction was determined from linear regression analysis equations based on standard curves for formaldehyde and glyoxal (peer reviewed data not shown).



c nmol of aldehyde released from 1 nmol lesion in 16-mer allowed to react with AlkB

Lesion or base	Formaldehyde	Glyoxal
Adenine:	-0.02 ± 0.01	0.09±0.003
1-Methyladenine:	0.98 ± 0.01	0.09±0.003
ε-Adenine:	0.05 ± 0.05	0.77±0.003

Figure 3.8: MALDI-TOF of εA and εC 16mers allowed to react with purified AlkB protein for 90 min. The -1 charge species are shown.

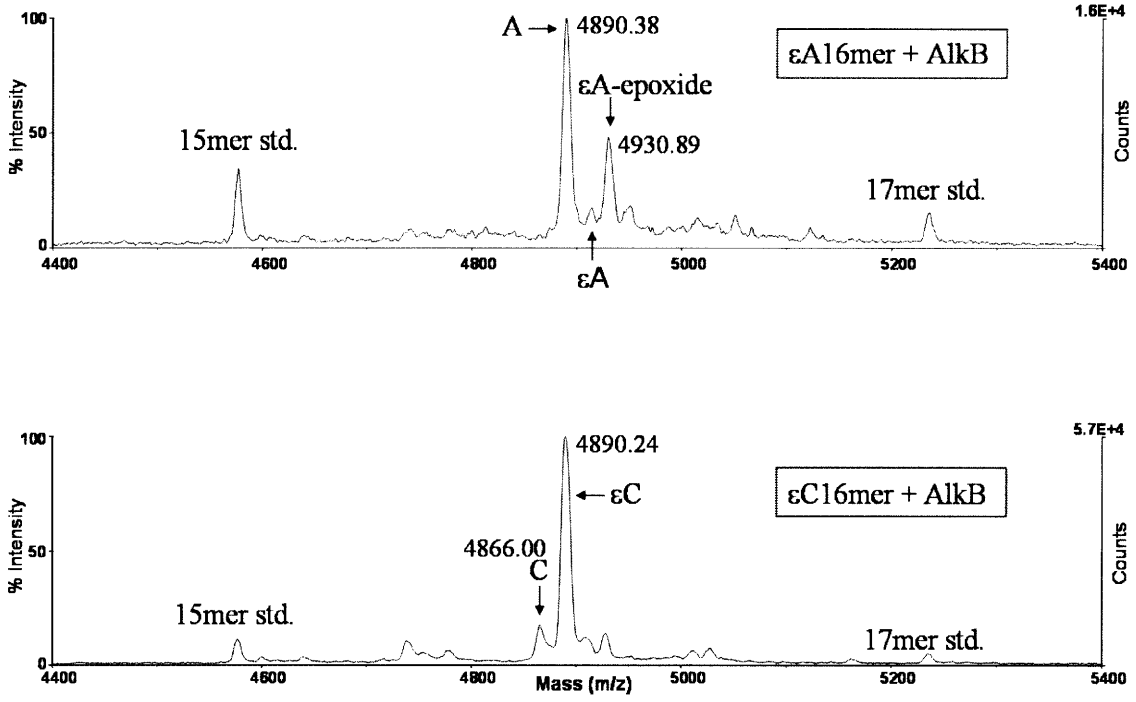


Figure 3.9: ESI-TOF of εA 16mer allowed to react with AlkB (12 min time point in Figure 3.7a).

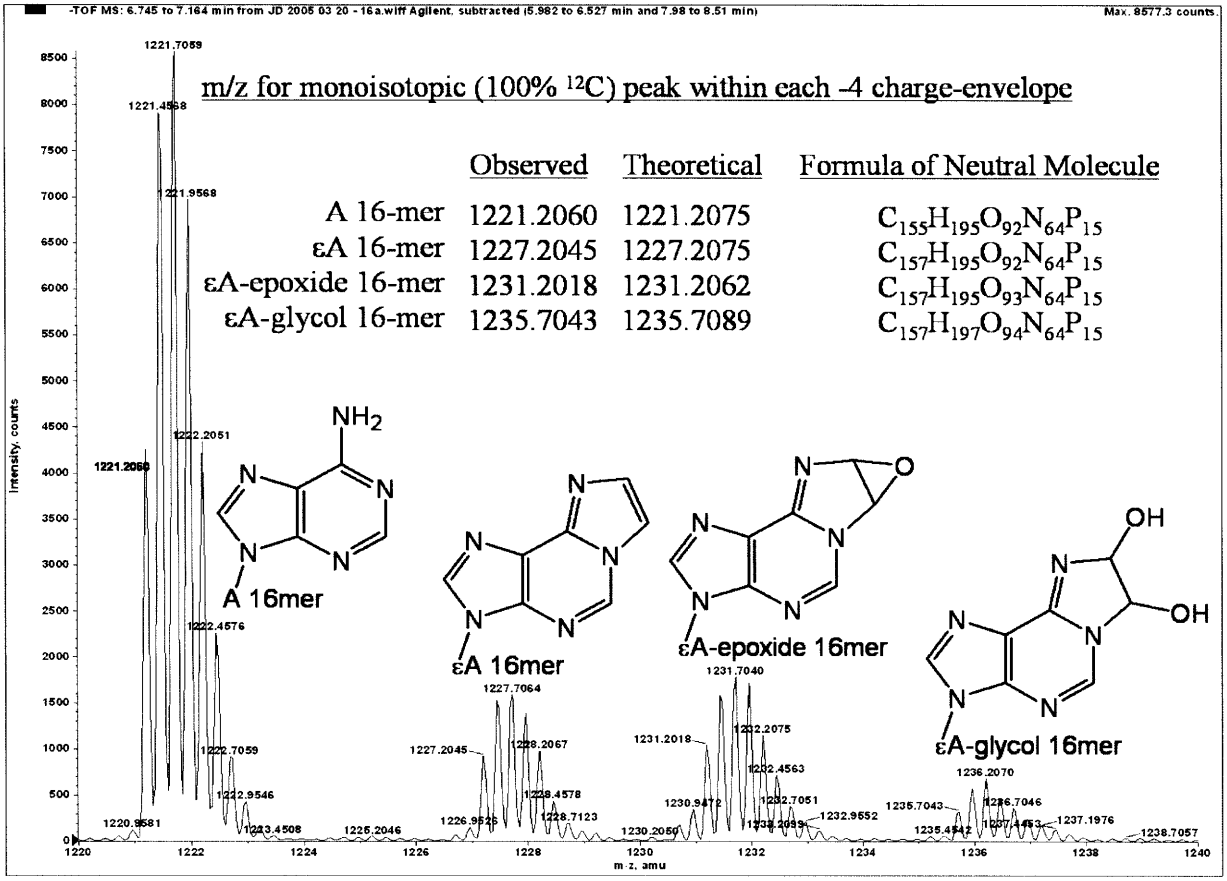


Figure 3.10: Sensitivity of *E. coli* deficient in AlkA or AlkB to CAA. Isogenic pairs of cells proficient or deficient in AlkA or AlkB were plated on LB agar containing increasing concentrations of CAA. Although loss of AlkA confers a more severe phenotype than loss of AlkB, deficiency in either repair protein renders mutant cells more sensitive to CAA.

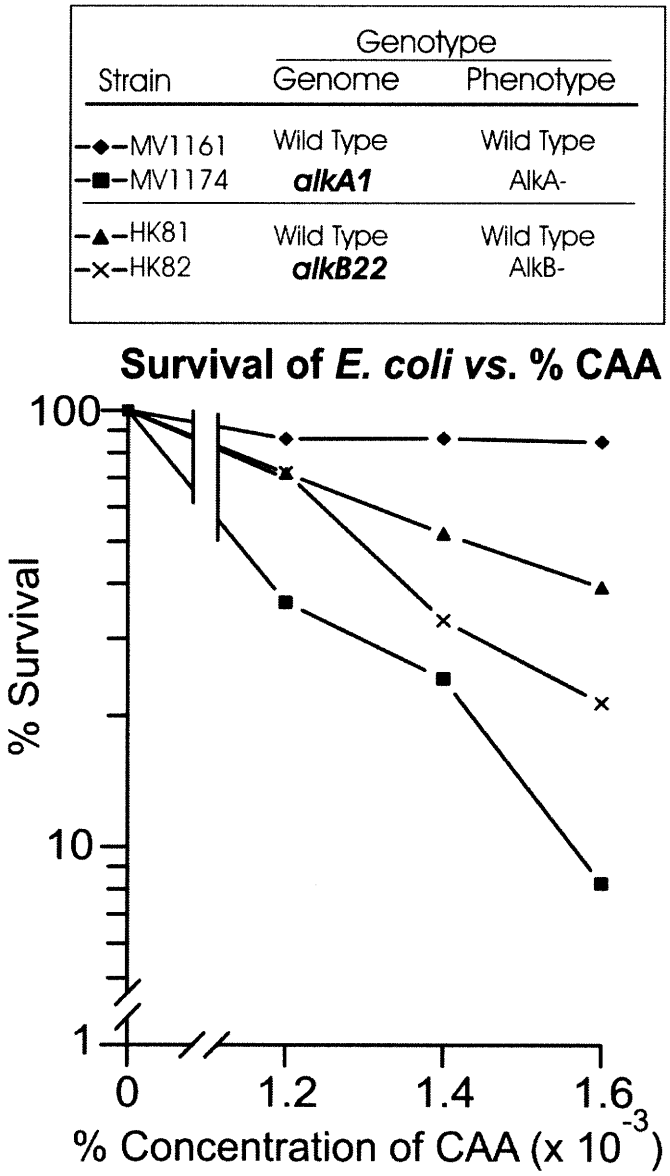
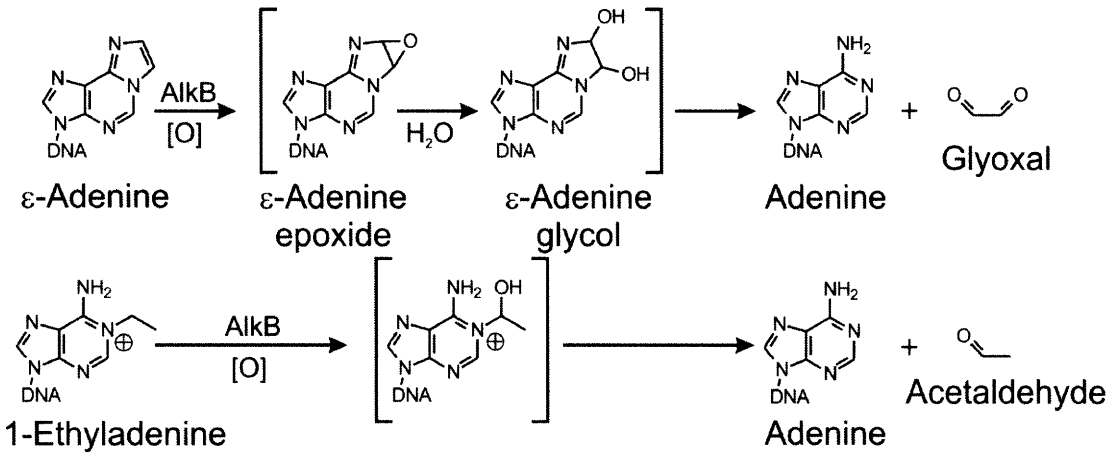


Figure 3.11: Likely AlkB reaction mechanism for repair of etheno vs. ethyl adducts.



References

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

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

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

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

5. Koivisto, P., T. Duncan, T. Lindahl, and B. Sedgwick. 2003. Minimal methylated substrate and extended substrate range of Escherichia coli AlkB protein, a 1-methyladenine-DNA dioxygenase. *J. Biol. Chem.* **278**:44348-44354.

6. Koivisto, P., P. Robins, T. Lindahl, and B. Sedgwick. 2004. Demethylation of 3-methylthymine in DNA by bacterial and human DNA dioxygenases. *J. Biol. Chem.* **279**:40470-40474.

7. Delaney, J. C. and J. M. Essigmann. 2004. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A* **101**:14051-14056.

8. Falnes, P. O. 2004. Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. *Nucleic Acids Res.* **32**:6260-6267.

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

10. Ougland, R., C. M. Zhang, A. Liiv, R. F. Johansen, E. Seeberg, Y. M. Hou, J. Remme, and P. O. Falnes. 2004. AlkB restores the biological function of mRNA and tRNA inactivated by chemical methylation. *Mol. Cell* **16**:107-116.

11. Sedgwick, B. 2004. Repairing DNA-methylation damage. *Nat. Rev. Mol. Cell Biol.* **5**:148-157.

12. Ames, B. N., M. K. Shigenaga, and T. M. Hagen. 1993. Oxidants, antioxidants, and the degenerative diseases of aging. *Proc. Natl. Acad. Sci. U. S. A* **90**:7915-7922.

13. Bjelland, S. and E. Seeberg. 2003. Mutagenicity, toxicity and repair of DNA base damage induced by oxidation. *Mutat. Res.* **531**:37-80.

14. Klaunig, J. E. and L. M. Kamendulis. 2004. The role of oxidative stress in carcinogenesis. *Annu. Rev. Pharmacol. Toxicol.* **44**:239-267.

15. Dedon, P. C. and S. R. Tannenbaum. 2004. Reactive nitrogen species in the chemical biology of inflammation. *Arch. Biochem. Biophys.* **423**:12-22.

16. Evans, M. D., M. Dizdaroglu, and M. S. Cooke. 2004. Oxidative DNA damage and disease: induction, repair and significance. *Mutat. Res.* **567**:1-61.

17. el Ghissassi, F., A. Barbin, J. Nair, and H. Bartsch. 1995. Formation of 1,N6-ethenoadenine and 3,N4-ethenocytosine by lipid peroxidation products and nucleic acid bases. *Chem. Res. Toxicol.* **8**:278-283.

18. Chung, F. L., H. J. Chen, and R. G. Nath. 1996. Lipid peroxidation as a potential endogenous source for the formation of exocyclic DNA adducts. *Carcinogenesis* **17**:2105-2111.

19. Marnett, L. J. 2000. Oxyradicals and DNA damage. *Carcinogenesis* **21**:361-370.

20. Blair, I. A. 2001. Lipid hydroperoxide-mediated DNA damage. *Exp. Gerontol.* **36**:1473-1481.

21. Lee, S. H., J. A. Arora, T. Oe, and I. A. Blair. 2005. 4-Hydroperoxy-2-nonenal-induced formation of 1,N2-etheno-2'-deoxyguanosine adducts. *Chem. Res. Toxicol.* **18**:780-786.

22. Nair, J., A. Barbin, Y. Guichard, and H. Bartsch. 1995. 1,N6-ethenodeoxyadenosine and 3,N4-ethenodeoxycytine in liver DNA from humans and untreated rodents detected by immunoaffinity/32P-postlabeling. *Carcinogenesis* **16**:613-617.

23. Barbin, A., H. Ohgaki, J. Nakamura, M. Kurrer, P. Kleihues, and J. A. Swenberg. 2003. Endogenous deoxyribonucleic Acid (DNA) damage in human tissues: a comparison of ethenobases with aldehydic DNA lesions. *Cancer Epidemiol. Biomarkers Prev.* **12**:1241-1247.

24. Barbin, A. 2000. Etheno-adduct-forming chemicals: from mutagenicity testing to tumor mutation spectra. *Mutat. Res.* **462**:55-69.

25. Saparbaev, M., K. Kleibl, and J. Laval. 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.

26. Saparbaev, M. and J. Laval. 1998. 3,N4-ethenocytosine, a highly mutagenic adduct, is a primary substrate for Escherichia coli double-stranded uracil-DNA glycosylase

and human mismatch-specific thymine-DNA glycosylase. *Proc. Natl. Acad. Sci. U. S. A* **95**:8508-8513.

27. Engelward, B. P., G. Weeda, M. D. Wyatt, J. L. Broekhof, W. J. de, I. Donker, J. M. Allan, B. Gold, J. H. Hoeijmakers, and L. D. Samson. 1997. Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase. *Proc. Natl. Acad. Sci. U. S. A* **94**:13087-13092.

28. Ham, A. J., B. P. Engelward, H. Koc, R. Sangaiah, L. B. Meira, L. D. Samson, and J. A. Swenberg. 2004. New immunoaffinity-LC-MS/MS methodology reveals that Aag null mice are deficient in their ability to clear 1,N6-etheno-deoxyadenosine DNA lesions from lung and liver in vivo. *DNA Repair (Amst)* **3**:257-265.

29. Pandya, G. A., I. Y. Yang, A. P. Grollman, and M. Moriya. 2000. Escherichia coli responses to a single DNA adduct. *J. Bacteriol.* **182**:6598-6604.

30. Mroczkowska, M. M., I. K. Kolasa, and J. T. Kusmierek. 1993. Chloroacetaldehyde-induced mutagenesis in Escherichia coli: specificity of mutations and modulation by induction of the adaptive response to alkylating agents. *Mutagenesis* **8**:341-348.

31. Borys, E., M. M. Mroczkowska-Slupska, and J. T. Kusmierek. 1994. The induction of adaptive response to alkylating agents in Escherichia coli reduces the frequency of specific C-->T mutations in chloroacetaldehyde-treated M13 glyU phage. *Mutagenesis* **9**:407-410.

32. Basu, A. K., M. L. Wood, L. J. Niedernhofer, L. A. Ramos, and J. M. Essigmann. 1993. Mutagenic and genotoxic effects of three vinyl chloride-induced DNA lesions: 1,N6-ethenoadenine, 3,N4-ethenocytosine, and 4-amino-5-(imidazol-2-yl)imidazole. *Biochemistry* **32**:12793-12801.

33. Pandya, G. A. and M. Moriya. 1996. 1,N6-ethenodeoxyadenosine, a DNA adduct highly mutagenic in mammalian cells. *Biochemistry* **35**:11487-11492.

34. Levine, R. L., I. Y. Yang, M. Hossain, G. A. Pandya, A. P. Grollman, and M. Moriya. 2000. Mutagenesis induced by a single 1,N6-ethenodeoxyadenosine adduct in human cells. *Cancer Res.* **60**:4098-4104.

35. Moriya, M., W. Zhang, F. Johnson, and A. P. Grollman. 1994. Mutagenic potency of exocyclic DNA adducts: marked differences between Escherichia coli and simian kidney cells. *Proc. Natl. Acad. Sci. U. S. A* **91**:11899-11903.

36. Palejwala, V. A., D. Simha, and M. Z. Humayun. 1991. Mechanisms of mutagenesis by exocyclic DNA adducts. Transfection of M13 viral DNA bearing a site-specific adduct shows that ethenocytosine is a highly efficient RecA-independent mutagenic noninstructional lesion. *Biochemistry* **30**:8736-8743.

37. Singer, B. and H. Bartsch. 1999. Exocyclic DNA Adducts in Mutagenesis and Carcinogenesis. International Agency for Research on Cancer, Lyon, France.

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

39. Nair, J., P. L. Carmichael, R. C. Fernando, D. H. Phillips, A. J. Strain, and H. Bartsch. 1998. Lipid peroxidation-induced etheno-DNA adducts in the liver of patients with the genetic metal storage disorders Wilson's disease and primary hemochromatosis. *Cancer Epidemiol. Biomarkers Prev.* **7**:435-440.

40. Bartsch, H. and J. Nair. 2004. Oxidative stress and lipid peroxidation-derived DNA-lesions in inflammation driven carcinogenesis. *Cancer Detect. Prev.* **28**:385-391.

Supplemental Information - Additional *E. coli* Survival Assays

Motivation

The discovery of a new chemical mechanism possessed by the *E. coli* DNA repair protein AlkB was made through the observation of mitigation of toxicity and mutagenicity of a single 1,N^6-ethenoadenine lesion in a viral genome processed through bacterial cells either proficient or deficient in the AlkB protein. These results established that AlkB could repair a site-specific adduct within cells, but also raised multiple important questions, including: 1. Does AlkB protect against agents that form these lesions in DNA (i.e, does loss of AlkB-mediated repair render cells sensitive to such agents?) 2. How does the magnitude of the protection conferred by AlkB compare with that conferred by the previously-discovered repair protein, AlkA? and 3. Can any of the human homologs of AlkB provide similar protection?

We first set out to address whether this method of etheno base repair is a significant contributor to the removal of these lesions from cellular DNA. It is already established that the glycosylases of the base excision repair pathway recognize and excise εA and εC *in vitro* (*1,2*) and *in vivo* (*3,4*). Thus, the ability of AlkB to reverse these lesions directly may be merely coincidental and not critical to their repair *in vivo*. To establish that the ability of AlkB to repair the etheno-adducts is not merely accidental but is in fact critical to the maintenance of a lesion-free genome (perhaps specifically the single-stranded regions), it would be necessary to demonstrate that the lack of this protein confers an increased sensitivity to agents that form these lesions. Consequently, a comparison of the survival of an *alkB* strain with its wild-type counterpart upon treatment with various concentrations of chloroacetaldehyde (CAA; a metabolic product of the occupational carcinogen vinyl chloride that reacts with DNA bases to form the exocyclic etheno-adducts (*5*)) was conducted to detect any sensitive phenotype conferred by the lack of AlkB-mediated repair. Additionally, sensitivity analyses of strains lacking either AlkA or AlkB in reference to the corresponding isogenic wild-type for each mutant were conducted to ascertain the magnitude of the phenotype conferred by loss of each repair pathway, and thereby the relative contribution each makes to removing toxic etheno-bases from the genome.

Because etheno adducts are generated endogenously in humans under normal circumstances (*6,7*) and at especially high concentrations at sites of inflammation (*8-11*), and because some humans are exposed to vinyl chloride and its metabolites industrially, this new mechanism of repair by AlkB would be particularly relevant if it were conserved in one or more of the human homologs of the protein. To investigate this possibility, a set of bacterial strains were created from an AlkB-deficient origin that contained an inducible plasmid carrying a wild-type copy of either the *E. coli* AlkB gene, or one of the human homologs (hABH1-3) to assay for the ability of these proteins to complement any sensitive phenotype observed.

Materials and Methods

Bacterial Strains and Plasmids. *E. coli* strains used in the survival experiments included HK81 (as AB1157, but *nalA*; wild-type), HK82 (as HK81, but *alkB22*; AlkB-deficient), MV1161 (*thr-1 araC14 leuB6(Am) Δ(gpt-proA)62 lacY1 tsx-33 supE44(AS) galK2(Oc) hisG4(Oc) rfbD1 mgl-51 rpoS396(Am) rpsL31(StrR) kdgK51 xylA5 mtl-1 argE3(Oc) thi-1 rfa-550*; wild-type), and MV1174 (as MV1161, but *alkA1*; AlkA-deficient). The open reading frames of the AlkB, hABH1, hABH2, and hABH3 genes were cloned into the pBAD24 plasmid (Figure 3S.1), which confers carbenicillin resistance and allows gene expression to be induced by the presence of arabinose. The pBAD vector and its derivatives were transfected into the isogenic pairs HK81 and HK82 and MV1161 and MV1174 to create HK81pSham (containing the empty pBAD24 vector), HK82pSham, HK82pAlkB, HK82phABH1, HK82phABH2, HK82phABH3, MV1161pSham, and MV1174pSham.

Standard Survival Curves. Cultures were grown overnight in LB medium containing 0.02% arabinose and 50 μg/mL carbenicillin at 37 °C on a roller drum to saturation, and then diluted 1:100 into fresh LB supplemented with 10 uL 20% arabinose to induce expression of the plasmid gene. Dilutions were grown to mid-log phase (OD_{600} ~0.5, about 2.5 h), at which point cells were pelleted by centrifugation. Following resuspension in 6 mL M9 minimal salts, 1 mL aliquots of cells were distributed into 1.5

mL Eppendorf tubes and washed with M9 to remove any traces of LB. Final resuspensions were in 1mL M9 supplemented with the appropriate dose of CAA (Aldrich; 0.005, 0.01, 0.015, 0.02, 0.03%) and incubated at 37 °C for 2 h. Drug was removed by dilution of the cells 100-fold into fresh M9. Further dilutions of 10^{-2} and 10^{-3} were made as appropriate for the plating of each strain, and a portion of each dilution (aiming for ~500 colonies per plate) was plated in triplicate on LB agar and allowed to grow overnight at 37 °C. Untreated controls for each strain were plated in quadruplicate. After about 16 h, colonies were counted and survival calculated as a fraction of the corresponding untreated control.

Fixed Concentration Plate Survival Curves. Saturated overnight cultures containing 0.02% arabinose and 50 μg/mL carbenicillin were diluted 1:100 into fresh LB supplemted with arabinose and grown with aeration at 37 °C to mid-log phase (OD_{600} ~0.5, about 2.5 h). Serial 1:10 dilutions (10^{-1} to 10^{-5}) were made of both the stationary overnight and the exponentially-growing cultures by dilution into M9 medium. Dilutions were then plated in triplicate (or quadruplicate for the untreated control) on LB plates already containing 0.02% arabinose and the appropriate concentration of CAA. Following incubation at 37 °C for 15-36 h (growth rate varied widely among drug concentrations), colonies were counted and survival calculated relative to the untreated control for each strain (i.e. cells plated on LB containing arabinose but lacking CAA).

Modified Growth and Survival Curves. Pilot growth curves were constructed for the three strains most likely to reveal any growth advantage offered by the AlkB proteins (HK82pSham, HK82pAlkB, and HK82phABH3) in order to establish a dose range at which survival experiments with all strains would be conducted. Saturated overnight cultures containing 0.02% arabinose and 50 μg/mL carbenicillin were diluted 1:10 into M9 medium containing CAA (0.0001 – 0.0010%) and incubated at 37 °C for 2 h to allow for adduct formation. Treated cells were diluted 1:100 into fresh LB medium containing 0.02% arabinose and allowed to grow on a wheel at 37 °C for 4-7 h. OD_{600} measurements were taken at 4, 5, 6, and 7 h to monitor differences in growth rate among

strains. Because the largest difference in growth was found at 5 h, this time-point was chosen for the survival assay. This experiment was carried out exactly as the pilot growth curve assay, except that CAA doses were 0, 0.0001, 0.0003, 0.0005, and 0.0007% and after 5 h of growth, cultures were diluted serially to 10^{-5} and plated in triplicate (or quadruplicate for the untreated control) on LB agar. After ~16 h growth at 37 °C, surviving colonies were counted and relative survival was calculated as a fraction of the untreated control for each strain.

Results

Standard Survival Curves. Although HK81 and MV1161 are both derivatives of AB1157 that are considered "wild-type", MV1161 consistently survives CAA treatment better than HK81 (Figure 3S.2). Even the *alkA* derivative of MV1161, MV1174, seems to be more resistant to the drug than HK81. Additionally, no substantial or reproducible difference could be detected in the sensitivities of the AlkA-deficient strain and its wild-type parent in these experiments, perhaps indicating that the MV1161 strains are more resistant to CAA in general and that these assays were conducted at doses of CAA too low to cause a differential in survival caused by the lack of AlkA in MV1174.

HK81pSham (AlkB-proficient) consistently survived better than HK82pSham (AlkB-deficient), although the magnitude of the difference varied among replicates (Figure 3S.2). While expression of *E. coli* AlkB from the plasmid in the AlkB-deficient strain complemented the sensitive phenotype displayed by the mutant, none of the human homologs were capable of increasing survival to any significant extent (compare HK82pSham, HK82phABH1, HK82phABH2, and HK82phABH3 in Figure 3S.2). Expression of AlkB (either genomically-encoded or expressed exogenously) seems to provide about one log of protection to the toxic effects of CAA.

Fixed Concentration Plate Survival Curves. Plating exponentially-growing cultures on LB agar containing various doses of CAA also showed a substantial difference in growth between the AlkB-proficient (HK81) and the AlkB-deficient (HK82) strains (Figure 3S.3). The magnitude of this difference varied greatly among experiments, however, as

did the amount of killing caused by a particular dose on different days. The sensitivity of these strains to CAA also becomes apparent, as the difference in killing of all strains between 1.15 and 1.2 μL CAA per 25 mL LB (where 1.15 μL = 0.0023% and 1.2 μL = 0.0024%) is quite substantial. None of the plasmid-encoded AlkB family proteins conferred a large protective effect on the mutant strain, as only small increases in survival were observed, and these increases were not consistently reproducible among replicates. Only AlkB caused a weakly-protective effect with any constancy.

Stationary cultures analyzed in the same manner showed similar variable phenotypes among replicates (Figure 3S.4). HK81 was consistently resistant to the range of CAA doses used, while HK82 exhibited sensitivity. Expression of AlkB in this case was able to partially complement the sensitive phenotype of the HK82 mutant, although the effect was much smaller than that seen in the traditional killing curve experiments. HK82phABH3 also survived better than HK82pSham in most replicates, suggesting that hABH3 might also confer a mild protective effect against CAA. The strains carrying hABH1 and hABH2 did not generate consistent results, as they seemed to offer protection in some experiments, but not in others.

Modified Growth and Survival Curves. Stationary phase cultures treated with CAA in M9 minimal medium to allow formation of adducts were grown in LB with aeration to determine whether a change in growth rate, rather than survival, is the major effect of CAA exposure. Indeed, at three of the four time-points monitored, the strains expressing an AlkB protein (either *E. coli* AlkB or human ABH3) exhibited slightly faster growth (Figure 3S.5). By 7 h, all three strains began to approach the optical density values of the untreated controls, and only the AlkB-containing strain showed any semblance of a growth advantage. Even this small difference was limited to high doses of CAA.

When cells were treated, allowed to grow for 5 h, and then plated for survival, no large differences were seen among any of the strains (Figure 3S.6), including the "wild-type" HK81, which was just as sensitive as the strains lacking the AlkB protein. HK82pAlkB survived slightly better than the other strains at higher doses of CAA, but the effect was minor.

Discussion

The differing AlkA status of the otherwise-isogenic pair MV1161 and MV1174 did not produce a detectable alteration in survival when mid-log cultures were treated with CAA and then plated on LB agar. Both strains survived well at the doses employed here, most likely indicating that higher concentrations of CAA would be necessary to distinguish any variance in survival ability conferred by the lack of AlkA in the experiments as they were conducted here. Because no phenotype was observed, these strains were not included in further studies.

In all three types of assays used to examine the correlation of CAA-sensitivity and AlkB status, only the *E. coli* AlkB protein conferred a substantial and reproducible protective effect, regardless of whether it was encoded in the *E. coli* genome (as in HK81pSham) or carried on an external plasmid (as in HK82pAlkB). The human homolog ABH3 seemed to increase survival in some cases, but the effect was not as great as that generated by AlkB, nor was it observed by all the methods used. Additionally, the hABH1 and hABH2-containing strains did not exhibit enhanced survival, indicating that these two homologs do not repair the type of damage created by CAA, that the human proteins when synthesized in bacteria are perhaps mis-folded or lack post-translational modifications essential for activity, or that another method of detection or a lower dose range is necessary to pick up any mild protection they might confer.

The threshold effect exhibited by these strains upon exposure to CAA caused widely-varying phenotypes to be produced by very slightly differing doses in all strains. Additionally, although triplicates of the same experiment gave similar results, the same dose in two different experimental replicates sometimes gave different orders of magnitude of killing. This variance led to quantitatively different results among experiments, complicating the statistical analysis and limiting interpretation to qualitative trends. The reasons behind the inconsistencies observed were not determined, as a similar experiment conducted in another lab succeeded, and this work was consequently abandoned.

Figure 3S.1: Structure of the pBAD24 vector used to express AlkB and hABH1-3 in AlkB-deficient cells.

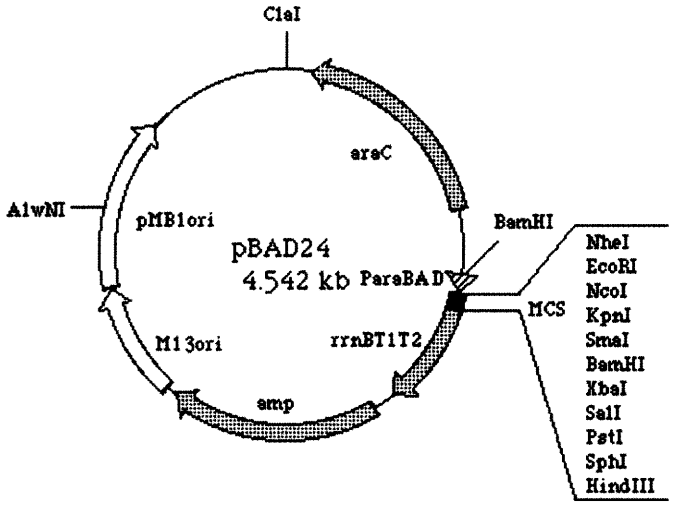


Figure 3S.2: Standard CAA survival curves. Cells were grown to mid-log phase, treated in liquid culture, diluted, and plated on LB agar. Individual replicates are shown only to highlight the variance between experiments.

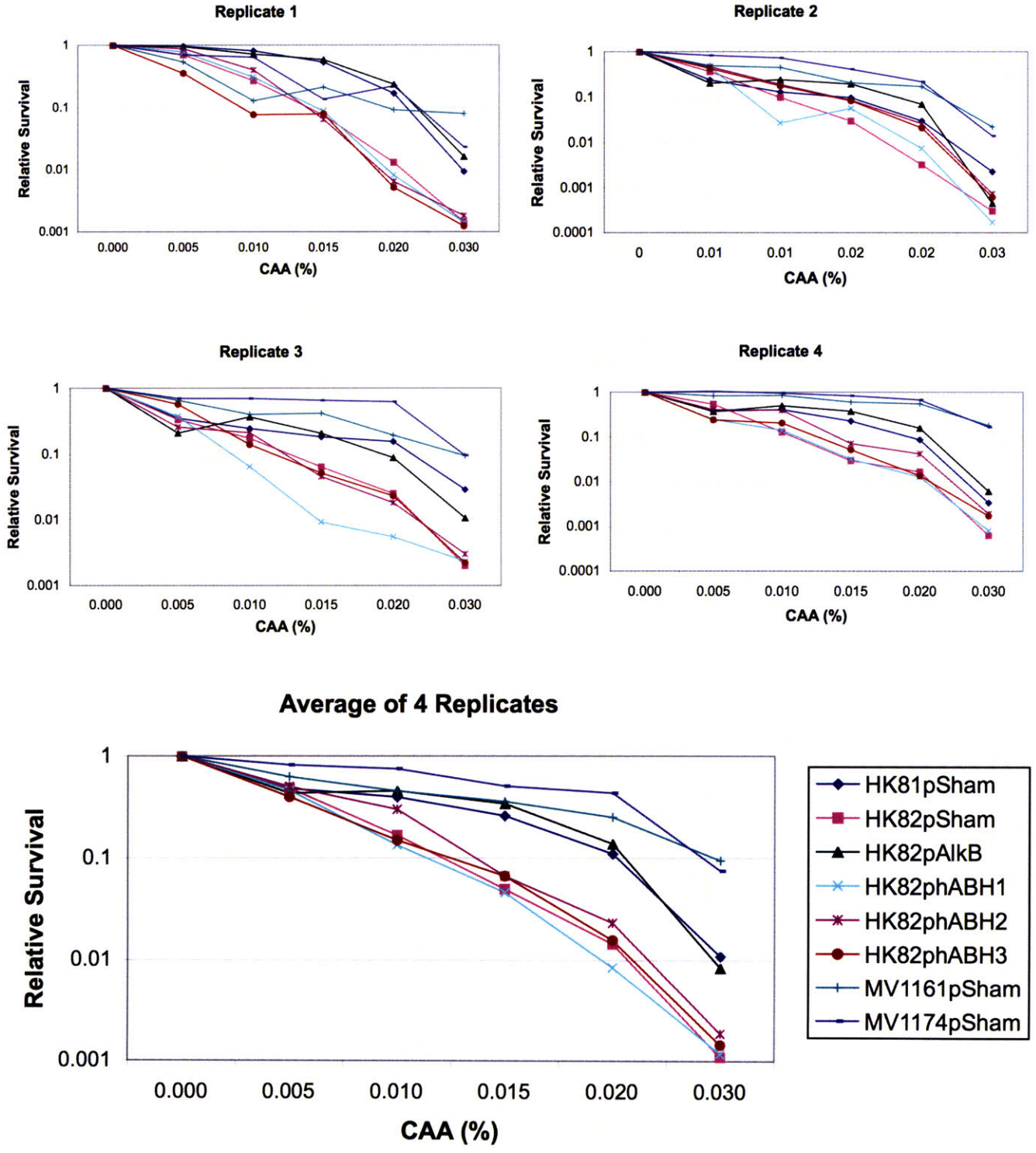
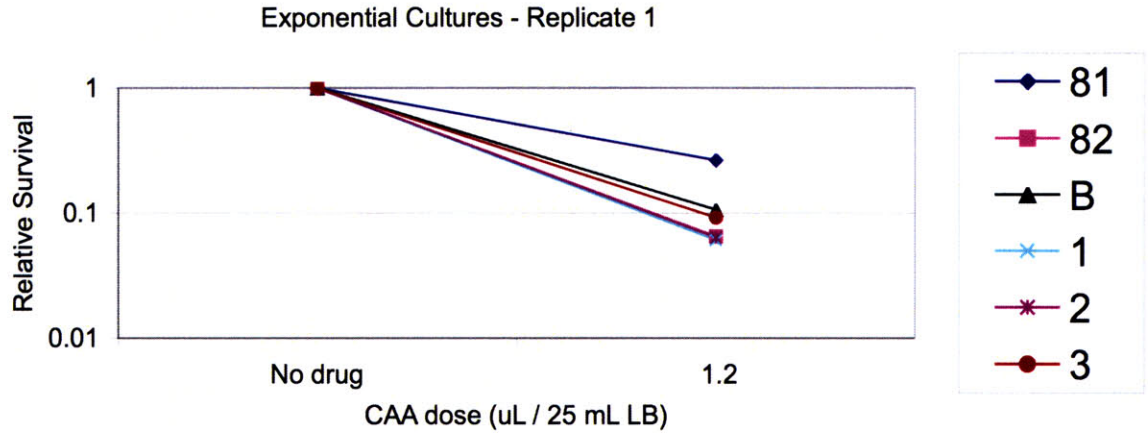
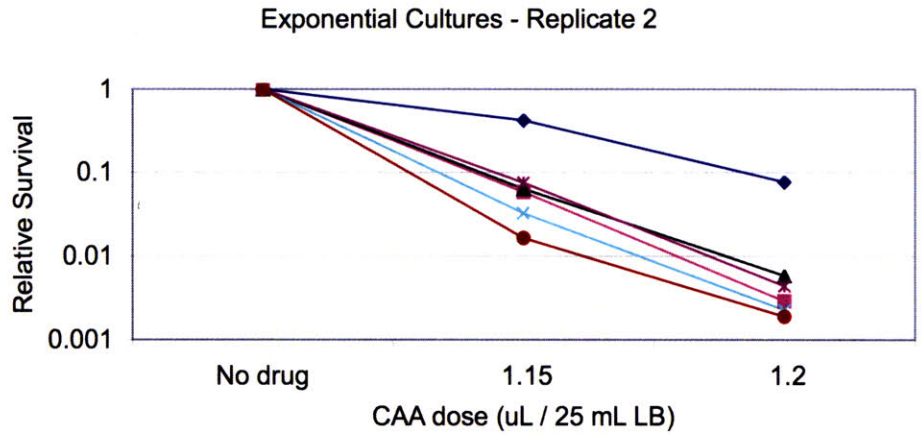


Figure 3S.3: Fixed concentration plate CAA survival curves. Cells were grown to mid-log phase, diluted, and plated on LB agar containing CAA (and arabinose). 81 = HK81pSham, 82 = HK82pSham, B = HK82pAlkB, 1 = HK82phABH1, 2 = HK82phABH2, 3 = HK82phABH3.





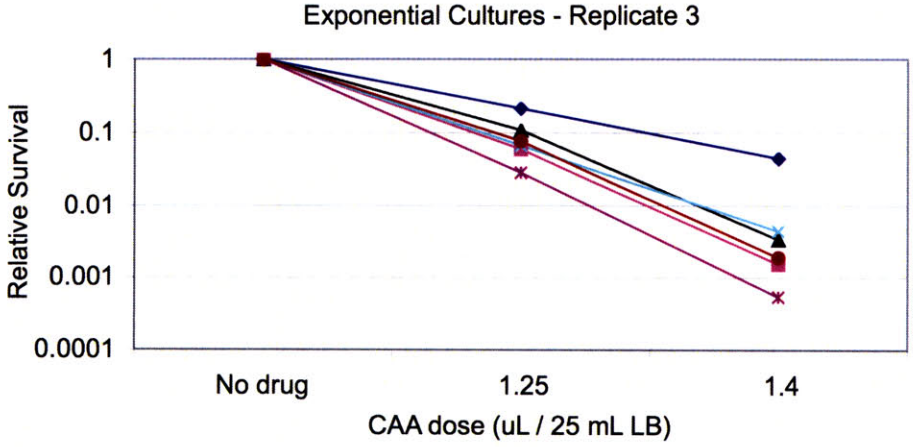


Figure 3S.4: Fixed concentration plate survival curves. Saturated overnight cultures in the stationary phase were diluted and plated on LB agar containing CAA. 81 = HK81pSham, 82 = HK82pSham, B = HK82pAlkB, 1 = HK82phABH1, 2 = HK82phABH2, 3 = HK82phABH3.

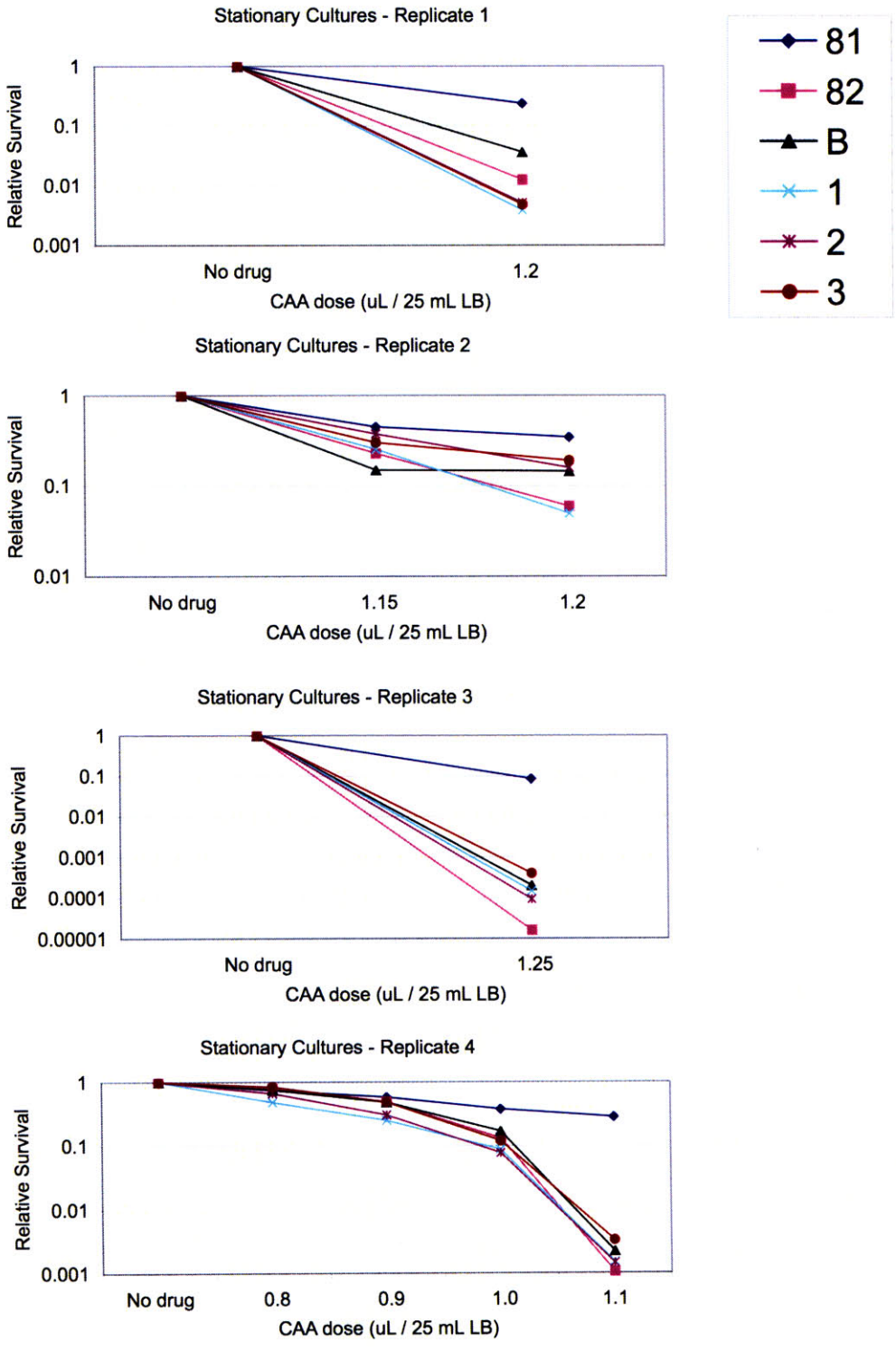


Figure 3S.5: Modified growth curves. Cells were treated with CAA in liquid culture in M9, diluted, and allowed to grow in LB for several hours. Growth was monitored to assess differences in growth ability among the strains. 82 = HK82pSham, B = HK82pAlkB, 3 = HK82phABH3.

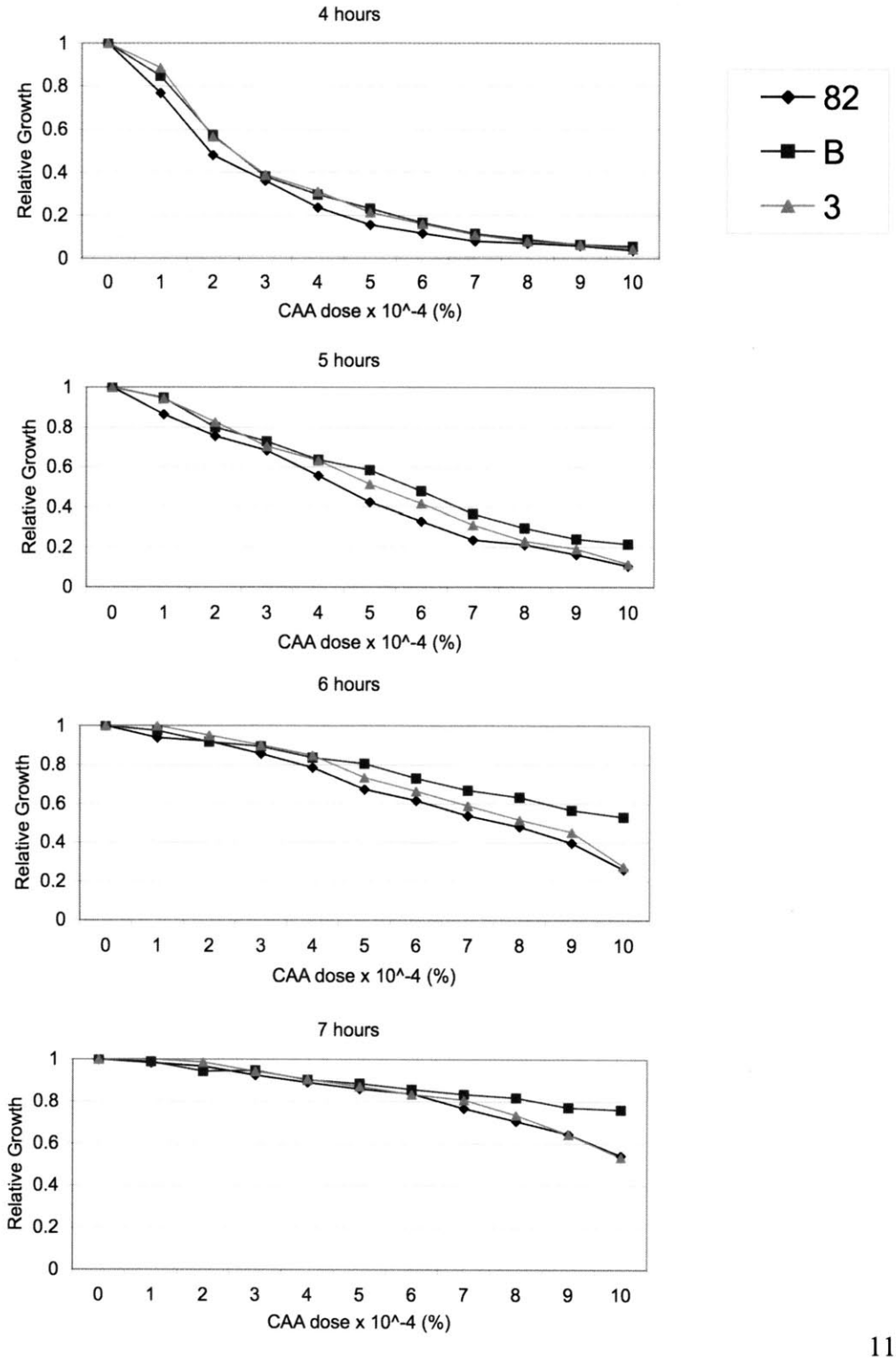
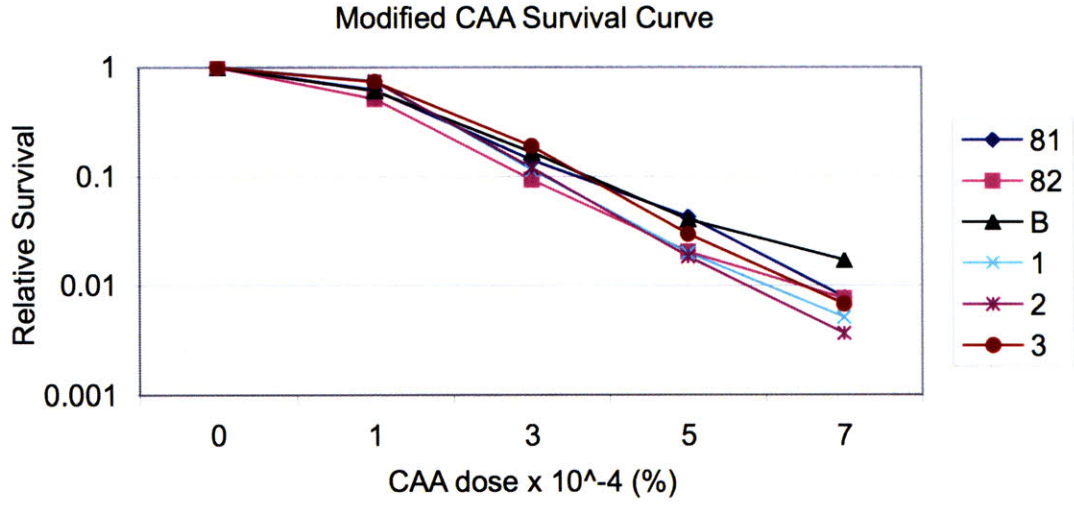


Figure 3S.6: Modified survival curve. Stationary-phase cells treated with CAA in M9 were diluted, allowed to recover for 5 h, and plated on LB agar. 81 = HK81pSham, 82 = HK82pSham, B = HK82pAlkB, 1 = HK82phABH1, 2 = HK82phABH2, 3 = HK82phABH3.



Supplemental References

1. Saparbaev, M., K. Kleibl, and J. Laval. 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.

2. Saparbaev, M. and J. Laval. 1998. 3,N4-ethenocytosine, a highly mutagenic adduct, is a primary substrate for Escherichia coli double-stranded uracil-DNA glycosylase and human mismatch-specific thymine-DNA glycosylase. *Proc. Natl. Acad. Sci. U. S. A* **95**:8508-8513.

3. Engelward, B. P., G. Weeda, M. D. Wyatt, J. L. Broekhof, W. J. de, I. Donker, J. M. Allan, B. Gold, J. H. Hoeijmakers, and L. D. Samson. 1997. Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase. *Proc. Natl. Acad. Sci. U. S. A* **94**:13087-13092.

4. Ham, A. J., B. P. Engelward, H. Koc, R. Sangaiah, L. B. Meira, L. D. Samson, and J. A. Swenberg. 2004. New immunoaffinity-LC-MS/MS methodology reveals that Aag null mice are deficient in their ability to clear 1,N6-etheno-deoxyadenosine DNA lesions from lung and liver in vivo. *DNA Repair (Amst)* **3**:257-265.

5. Barbin, A. 2000. Etheno-adduct-forming chemicals: from mutagenicity testing to tumor mutation spectra. *Mutat. Res.* **462**:55-69.

6. Nair, J., A. Barbin, Y. Guichard, and H. Bartsch. 1995. 1,N6-ethenodeoxyadenosine and 3,N4-ethenodeoxycytine in liver DNA from humans and untreated rodents detected by immunoaffinity/32P-postlabeling. *Carcinogenesis* **16**:613-617.

7. Barbin, A., H. Ohgaki, J. Nakamura, M. Kurrer, P. Kleihues, and J. A. Swenberg. 2003. Endogenous deoxyribonucleic Acid (DNA) damage in human tissues: a comparison of ethenobases with aldehydic DNA lesions. *Cancer Epidemiol. Biomarkers Prev.* **12**:1241-1247.

8. el Ghissassi, F., A. Barbin, J. Nair, and H. Bartsch. 1995. Formation of 1,N6-ethenoadenine and 3,N4-ethenocytosine by lipid peroxidation products and nucleic acid bases. *Chem. Res. Toxicol.* **8**:278-283.

9. Chung, F. L., H. J. Chen, and R. G. Nath. 1996. Lipid peroxidation as a potential endogenous source for the formation of exocyclic DNA adducts. *Carcinogenesis* **17**:2105-2111.

10. Marnett, L. J. 2000. Oxyradicals and DNA damage. *Carcinogenesis* **21**:361-370.

11. Blair, I. A. 2001. Lipid hydroperoxide-mediated DNA damage. *Exp. Gerontol.* **36**:1473-1481.

CHAPTER 4:

Alleviation of 1,N^6-ethanoadenine Genotoxicity by the *Escherichia coli* Adaptive Response Protein AlkB

Reproduced from: Frick et al. *PNAS* 104, 755 – 760 (2007)

Abstract

1,N^6-Ethanoadenine (EA) forms through the reaction of adenine in DNA with the antitumor agent 1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU), a chemotherapeutic used to combat various brain, head, and neck tumors. Previous studies of the toxic and mutagenic properties of the DNA adduct EA have been limited to *in vitro* experiments using mammalian polymerases and have revealed the lesion to be both miscoding and genotoxic. This work explores lesion bypass and mutagenicity of EA replicated *in vivo* and demonstrates that EA is neither toxic nor mutagenic in wild-type *E. coli*. Although the base excision repair glycosylase enzymes of both humans and *E. coli* possess a weak ability to act on the lesion *in vitro*, an *in vivo* repair pathway has not yet been demonstrated. Here we show that an enzyme mechanistically unrelated to DNA glycosylases, the adaptive response protein AlkB, is capable of acting on EA *via* its canonical mechanism of oxidative dealkylation. The reaction alleviates the unrepaired adduct's potent toxicity through metabolism at the C8 position (attached to N1 of adenine), producing a non-toxic and weakly-mutagenic N^6 adduct. This is the first example of a repair protein reducing the toxicity of a DNA adduct by conversion to a secondary lesion.

Introduction

1,N^6-ethanoadenine (EA) is a DNA base adduct in which the exocyclic nitrogen of adenine is connected to the N1 ring nitrogen by a saturated two-carbon bridge, creating a five-membered ring (Fig. 1A). This adduct is formed by the reaction of DNA with 1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU) (*1-3*) and is possibly responsible for the mutagenic property of this anticancer agent. EA is repaired to some extent by the human alkyladenine DNA glycosylase (MPG, also called AAG, APNG, ANPG) (*4*), although its excision is 65-fold less efficient than that of the structurally related 1,N^6-ethenoadenine (εA) (Fig. 4.1A). Similarly, the *E. coli* repair protein 3-methyladenine DNA glycosylase (AlkA) recognizes and excises the EA adduct, but with 20-fold lower activity than towards εA (*5*). Because EA lacks the structural features needed to form instructional hydrogen bonds and diverges from adenine in both size and shape, it is likely to be both a toxic and a mutagenic lesion.

In the absence of repair, EA both blocks polymerase bypass and miscodes during attempted replication by mammalian DNA polymerases *in vitro* (*6*). The analogous εA adduct, which is missing the identical hydrogen bonding capabilities, is both toxic and mutagenic in *E. coli* when unrepaired (*7*). Because the EA adduct is likely to be mutagenic and genotoxic when unrepaired and is excised only inefficiently by base excision repair glycosylases, another method of repair beyond the action of AlkA would greatly aid the cell in preserving genomic stability and viability. Given that the structurally similar εA lesion is repaired by AlkB, a protein upregulated in response to alkylating agents, it was hypothesized that EA may also be a substrate.

The *E. coli* AlkB protein is the product of one of four genes upregulated in response to methylating agents in a biochemical network termed the adaptive response (*8*). Ada is a bifunctional methyltransferase that serves not only to reverse DNA damage directly by irreversibly accepting methyl groups from O^6-methylguanine and one isomer of backbone methylphosphotriesters, but also as a transcriptional regulator, inducing expression at three promoters after methylation of Cys38. AlkA, as mentioned above, is a 3-methyladenine DNA glycosylase responsible for excision of damaged bases as the first step in the base excision repair pathway. AidB is homolgous to human isovaleryl

coenzyme A dehydrogenase (IVD), a central player in the leucine metabolism pathway, but its functional role in the adaptive response has not yet been elucidated beyond its ability to bind double-stranded DNA (*9*).

AlkB is an α-ketoglutarate dioxygenase that uses non-heme iron (II) as a cofactor and molecular oxygen and α-ketoglutarate as co-substrates to reverse directly, through oxidative demethylation, 1-methyladenine (m1A) and 3-methylcytosine (m3C), lesions formed predominantly in single-stranded DNA under conditions of alkylative stress (*10,11*). Hydroxylation of the DNA-bound methyl group is coupled with decarboxylation of α-ketoglutarate, releasing succinate and carbon dioxide while producing an unstable hydroxymethyl intermediate that decomposes, releasing formaldehyde and the restored adenine or cytosine. While homologs of AlkB have been found in many other bacterial species, as well as in *S. pombe*, *C. elegans, Drosophila,* mice, humans, and plant RNA viruses (*12-14*), proteins similar in sequence are apparently lacking in many other species, both bacterial and eukaryotic (*13*). Interestingly, however, expression of several *S. cerevisiae* genes that possess no amino acid homology to AlkB in AlkB-deficient *E. coli* has been shown to complement the MMS-sensitive phenotype imparted by the mutation (*15*). The mouse AlkB homologs, mABH2 and mABH3 (*16*), and two of the eight identified human homologs, hABH2 and hABH3 (*17-20*), have been shown to possess the same ability to reverse methylated DNA base damage, although substrate preferences vary among enzymes (*21,22*). Recently, mABH2 has been demonstrated to remove endogenously-formed m1A adducts in mice (*23*).

Since the initial discovery of AlkB's mechanism of action in the repair of m1A and m3C, the substrate range of the protein has expanded to include 3-ethylcytosine (*24*), 1-methylguanine (*17,24*), 3-methylthymine (*17,18,24*), and 1-ethyladenine (*20*) lesions in DNA, demonstrating the enzyme's biochemical versatility. The recently solved X-ray crystal structure (*25*) shows a particularly malleable region within the nucleotide-recognition lid that may account for the accommodating nature of the protein.

In addition to these methylated and ethylated substrates, AlkB has been recently shown to directly repair εA and εC. Here, the protein employs a second, distinct chemical mechanism of direct reversal: epoxidation of the double bond of etheno lesions

resulting in the release of glyoxal and the undamaged base (*7*). The human homolog hABH3 has also been shown to reverse εA directly, albeit inefficiently, *in vitro* (*26*). In addition to the broad pool of DNA substrates that AlkB repairs, the protein and its human homolog hABH3 are also able to remove methyl groups from RNA, allowing reactivation of methylated RNA phage (*27*) and restoring the function of methylated mRNA and tRNA molecules *in vitro* (*28*).

Given the structural similarity of EA to known substrates for AlkB (εA and 1-alkyladenines), we hypothesized that it could itself be a substrate for this versatile enzyme. To investigate this possibility, a 16-mer oligonucleotide containing a site-specific EA was inserted into a single-stranded M13 phage genome, and the ability of DNA polymerase to replicate past the lesion was measured in wild-type and *alkB E. coli* (Fig. 4.1B). While replication of the adduct-containing genome was at least as efficient as that of an unmodified control genome in AlkB-containing cells, replication was severely impaired in an AlkB-deficient strain, indicating that the protein does indeed play a role in alleviating the toxicity of the damaged base. *In vitro* reconstitution of the repair reaction followed by mass spectral analysis revealed that AlkB is able to hydroxylate EA oxidatively, resulting in a product that rapidly converts to a reactive ring-opened aldehydic isomer that can form secondary adducts with primary amines. Despite the demonstrated ability of AlkB to metabolize EA, the primary role of AlkB *in vivo* was to counter lesion toxicity, rather than lesion mutagenesis, as EA mutagenicity was quantifiable but low in both AlkB-proficient and -deficient cells.

Materials and methods

Bacterial strains. The *E. coli* strains used for transfection and replication of lesion-containing viral genomes were HK81 (as AB1157, but *nalA*; wild-type) and HK82 (as AB1157, but *nalA alkB22*; AlkB deficient). AlkB status was confirmed by PCR amplification and sequencing analysis of the *alkB* gene of both strains. NR9050 *E. coli* were used for quantification of transfection efficiency, while the bacterial strain used for regrowth of progeny phage was SCS110 (Stratagene).

Oligonucleotides. Oligonucleotides were synthesized using phenoxyacetal-protected phosphoramidites and capping reagents from Glen Research (*24*). Approximately 50 mg of EA phosphoramidite (Chemgenes), which has previously been site-specifically incorporated into an oligonucleotide (*29*), were used during a 30 min coupling with the solid support. Following synthesis, the resin was incubated with concentrated ammonium hydroxide for one hour prior to lyophilization. The 16-mer insert oligonucleotide 5'-GAAGACCTXGGCGTCC-3' (X=EA) was purified by anion exchange and reversed phase HPLC, and the molecular weight was verified by MALDI-TOF mass spectrometry (4916.21 observed, 4916.22 calculated). Sixteen-mer oligonucleotides of the same sequence, but where X = εA or T, were synthesized and purified as described (*7*). Scaffold oligonucleotides (of sequence 5'-GGTCTTCCACTGAATCATGGTCATAGC-3' and 5'-AAAACGACGGCCAGTGAA TTGGACGC-3') used in genome construction were obtained from Integrated DNA Technologies.

Genome construction. Single-stranded M13 phage genomes containing a single site-specific lesion were constructed as described (*24,30*). Briefly, 20 pmol of M13mp7L2 single-stranded bacteriophage genomic DNA were linearized by digestion with *EcoRI*, which incises the hairpin region at a unique restriction site. Two oligonucleotide scaffolds, one complementary to the 5' end of the linearized phage genome and the 3' end of the 16-mer insert, and the other complementary to the 3' end of the genome and the 5' end of the insert, were annealed to the M13 DNA. Sixteen-mer oligonucleotide inserts containing EA, εA, or T at the lesion site, were phosphorylated by T4 polynucleotide kinase. Phosphorylated inserts were added to the genome:scaffold mixture and genomes were recircularized by ligation with T4 DNA ligase. After removal of the scaffolds by T4 DNA polymerase, genomes were extracted with phenol:chloroform:isoamyl alcohol (25:24:1), desalted in four washes on Centricon-100 spin-dialysis columns (Millipore), and stored at -20 °C.

Genome concentration normalization. Genome concentrations were normalized by annealing the scaffolds used for construction to the recircularized genomes, digesting the partially double-stranded product with *HinFI*, and dephosphorylating the resulting fragment with shrimp alkaline phosphatase (Roche). Genomes were then 5'-radiolabeled using [γ-^{32}P]-ATP and T4 polynucleotide kinase (USB), digested with *HaeIII*, and products were separated by denaturing polyacrylamide gel electrophoresis (PAGE). Band intensities were quantified using phosphorimagery (Molecular Dynamics) and the concentration of each genome was normalized to that of the most dilute by the addition of water.

Preparation of electrocompetent cells. Three baffled flasks each containing 150 mL of LB media were inoculated with 1.5 mL saturated overnight culture of the strain to be transformed (HK81, wild-type, or HK82, AlkB deficient) and grown with shaking at 37 °C to an OD_{600} of about 0.5 (~2.5 h). The three cultures were then pelleted and the cells pooled, followed by 3 resuspension/repelleting cycles at 4 °C using 175 mL water. The final resuspension of cells was in 6 mL 10% glycerol, after which the cells were transformed within 2 h.

Examination of lesion bypass efficiency. Lesion-containing genomes were mixed to a ratio of ~10:1 with an internal standard competitor genome containing a T instead of a lesion in an insert that is three bases longer than that bearing the lesion. One hundred μL of electrocompetent cells were mixed with 4.8 μL of genome mixture, transferred to a pre-chilled 2 mm gap electroporation cuvette, and exposed to ~2.5 kV and ~125 Ω. After electroporation, cells were immediately transferred to 10 mL fresh LB and grown for 6 h at 37 °C. Cells were then pelleted, and the supernatant containing progeny phage was removed and stored at 4 °C. The number of successful transformation events was monitored by immediately plating 10 μL of the 10 mL culture onto a lawn of NR9050 indicator bacteria. Each independent event will result in a plaque, enabling quantification of initial events and ensuring a robust progeny population for analysis. All

electroporation procedures reported here produced $>10^4$ events, with the vast majority resulting in $>10^5$.

Because not all copies of the lesion-bearing genome were incorporated into cells during the electroporation process, the isolated supernatant contained some unreplicated genomes in addition to the post-replication progeny phage. In order to eliminate analysis of these residual genomes, an amplification of the viable progeny phage was conducted. One hundred μL progeny phage ($\sim 10^9$) were re-grown for 6 h in 10 mL LB inoculated with 10 μL of an overnight culture of SCS110 *E. coli* (Stratagene). Only viable progeny phage, i.e. those previously replicated within a cell, would be amplified, while the unincorporated genomes would not be affected, effectively increasing the signal to noise ratio by several orders of magnitude. Cells were again pelleted and progeny-containing supernatant was isolated and stored at 4 °C.

Prior to PCR amplification, single-stranded DNA was isolated from the progeny phage using a QIAprep Spin M13 kit (Qiagen). Amplification of the region of interest was performed using PCR primers (5'-YCAGCTATGACCATGATTCAGTGGAAGAC-3' and 5'-YCAGGGTTTTCCCAGTCACGACGTTGTAA-3' where Y signifies amino modification of the 5' nucleotide) as described (*7,24,28,30*). After phenol:chloroform:isoamyl alcohol extraction and removal of salt and organic residue with Sephadex G50 fine resin (Amersham), the PCR product was digested with *BbsI* and dephosphorylated with shrimp alkaline phosphatase. T4 polynucleotide kinase was used to 5'-radiolabel the resulting fragment, which was further digested by *HaeIII*. Radiolabeled products were then separated by PAGE, and the intensity of each band quantified by phosphorimagery. Comparison of the intensity of the 18-mer band (corresponding to the lesion-containing genome) to that of the 21-mer band (representing the internal standard) gave a measure of lesion bypass.

Construction and Purification of AlkBΔN11. The AlkB open reading frame, excluding the first 11 amino acids, was PCR-amplified using the oligonucleotides 5'-GGAATTC**CATATG**CAAGAGCCACTGGCGG (*Nde*I restriction site), and 5'-CCG**CTCGAG**GCCTTGAAAATATAGGTTTTCTTTTTTACCTGCCTG (*Xho*I site).

The PCR product was purified, digested with *Nde*I and *Xho*I (NEB), and ligated into a C-terminal-histidine-tag-containing pET24a vector (Novagen) that had been digested by *Nde*I and *Xho*I. The ligation mixture was transformed into DH5α cells for verification by sequencing. pET24a-AlkBΔN11 was transformed into BL21(DE3) (Novagen) cells for expression.

pET24a-AlkBΔN11 cells were grown at 37 °C to an A_{600} of 0.4, when isopropyl-β-D-thiogalactoside (IPTG) was added at 1 mM and the temperature was lowered to 30 °C. Cells were harvested after 4 h by centrifugation, resuspended in 10 mM Tris pH 7.3, 300 mM NaCl, 2 mM $CaCl_2$, 10 mM $MgCl_2$, 5% glycerol and 1 mM BME (lysis buffer), and homogenized by sonication. Lysate was recovered by centrifugation. The supernatant was loaded onto an Ni-NTA column (Qiagen), washed (20 mM) and eluted (70 and 250 mM) in imidazole in lysis buffer. The eluent was dialyzed against 50 mM TES pH 7.1 and loaded onto an SP Sepharose cation exchange column (Amersham Pharmacia). AlkBΔN11 was eluted by a linear gradient of 0-1 M NaCl over 75 mL. The fractions containing AlkBΔN11 were pooled and purity was established by SDS-PAGE. As previously reported (25), AlkBΔN11 has similar activity to the full-length protein in standard AlkB assays (data not shown).

Examination of reaction products by MALDI-TOF mass spectrometry. Purified AlkBΔN11 protein (10.5 pmol) was incubated with 25 pmol of EA 16-mer in 10 μL buffer containing 70 μM $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$, 0.9 mM α-ketoglutarate, 1.8 mM L-ascorbate, and 46.5 mM HEPES (pH 8.0) and molecular weight standards (15-mer and 17-mer) for 30 min at 37 °C. Reactions were quenched by application of the sample to a $ZipTip_{C18}$ pipette tip (Millipore) followed by 0.1 M triethylamine washes to remove cofactors. In one reaction, approximately 2 nmol of the trapping reagent PFBHA was added prior to cofactor removal. Quenched samples were then analyzed by matrix assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry.

Analysis of EA mutational spectrum. The mutational specificity and frequency of the EA adduct was determined exactly as in the lesion bypass assay described above up to

the PCR step. A different set of primers (5'-YCAGCTATGACCATGAT TCAGTGGAAGAC-3' and 5'YTGTAAAACGACGGCCAGTGAATTGGACG-3') are used to obtain specific amplification of progeny from genomes that originally contained the lesion. The second primer is mismatched when annealed to the internal standard, which is thus not amplified. Samples are processed as in the lesion bypass assay through the PAGE step, at which point the 18-mer bands are excised from the gel, and the DNA fragments extracted by crushing and soaking. Importantly, the restriction enzyme used in this assay incises the DNA to expose the base that corresponds to the original site of the lesion at the 5' end, which is then radiolabeled. This site-specific labeling allows for precise inquiry into the base present at the lesion site after replication. The extracted DNA is desalted on Sephadex G50 fine resin (GE Healthcare), lyophilized, and digested to free nucleotide monophosphates by nuclease P1 (United States Biological). The digested nucleotides are separated on a polyethyleneimine thin layer chromatography (TLC) plate, and the resulting spots, each representing a normal base, are quantified by phosphorimagery. Because only the 5'- base was radiolabeled, the TLC represents the base composition at the lesion site after replication by *E. coli*, enabling calculation of both mutation frequency and specificity.

Results

Bypass of EA *in vivo*. EA is a non-toxic adduct when replicated in wild-type (AlkB+) cells, as it shows a level of bypass comparable to that seen for the control T (Fig. 4.2). In striking contrast, when a viral genome containing a single EA lesion is replicated in cells deficient in AlkB-mediated repair, the adduct reveals itself to be extremely toxic, as it is bypassed only 14% as well as the control. This vast difference in lesion tolerance between wild-type and *alkB E. coli* is suggestive of a role for AlkB in mitigating lesion toxicity through repair of this adduct. The bypass characteristics of εA were investigated again in this work as an established benchmark against which to compare EA. The data obtained here are in agreement with those previously observed for εA (*7*), as bypass was comparable to controls in wild-type cells but dropped to only 9% in AlkB-deficient cells.

***In vitro* repair/metabolism of EA in DNA.** In order to determine the cause of this large change in bypass, the repair reaction was reconstituted *in vitro*. Purified AlkBΔN11 protein was incubated with the EA-containing 16-mer oligonucleotide used for genome construction in HEPES buffer containing the requisite co-factors and co-substrates to achieve enzymatic activity. The products of the reaction were monitored using MALDI-TOF mass spectrometry. Incubation of the oligonucleotide and two molecular weight markers in buffer lacking protein produced three peaks, the flanking two corresponding to the 15-mer and 17-mer controls, and the central peak representing the 16-mer substrate containing the unaltered lesion (Fig. 4.3A). After incubation with protein for 30 min at 37 °C, the substrate DNA was converted entirely to two products, one with a mass of +16 relative to EA and the other +135 (Fig. 4.3B). The +16 peak is consistent with hydroxylation of EA (Fig. 4.4A). The regiochemistry shown (hydroxylation at the C8 position) mimics the product of the oxidative hydroxylation reaction mechanism that AlkB has been shown to employ in the repair of simple methylated lesions such as 1-methyladenine (*10,11,31*). Additionally, this C8-hydroxylated product could rearrange to form the ring-opened aldehydic isomer (Fig. 4.4A). The +135 peak is consistent with the Schiff base product of this ring-opened aldehyde with the anthranilic acid (molecular weight 137.1) component of the matrix used for ionization (Fig. 4.4B). Indeed, the addition of a 100-fold molar excess of the trapping reagent *O*-(2,3,4,5,6-pentafluorobenzyl)hyroxylamine (PFBHA, molecular weight 213.1) prior to matrix exposure produced quantitative conversion to a +211 product, consistent with the oxime product of the ring-opened EA aldehyde and PFBHA (Fig. 4.4C).

Mutational frequency and specificity of EA. Analysis of the base composition of progeny phage at the site that originally contained the lesion reveals that EA is only weakly mutagenic in wild-type cells. DNA polymerase is able to replicate a transfected genome containing the EA adduct faithfully more than 98% of the time (Fig. 4.5). Similarly, εA is also replicated correctly, as >97% A is seen at the site of interest in this work and >99% A has been seen previously (*7*). The 2% mutation frequency of EA in wild-type cells is composed of ~1% EA to C and ~0.5% each EA to G and EA to T.

Perhaps surprisingly, however, given the dramatic alteration in toxicity, the mutational signature of EA does not change substantially when the adduct is replicated in *alkB* cells, in which case the observed base composition at the lesion site is ~96% A, 2% C, and 1% each G and T, yielding only ~4% total mutagenesis. In contrast, εA is seen to be ~37% mutagenic in a repair-deficient strain, producing 22% εA→T, 8% εA→C, and 7% εA→G. The results generated by the εA control are in agreement with previous studies of εA mutagenicity (*7*).

Discussion

The anticancer chemotherapeutic agent BCNU is a bifunctional electrophile that reacts with DNA to form a range of adducts, including chloroethyl- and hydroxyethyl-monosubstituted bases, saturated ethano adducts, and inter- and intrastrand crosslinks (*32-35*). The compound has been shown to exhibit both genotoxic and mutagenic properties and, while the potent toxicity of the drug is attributed to quantitatively minor 1-(3-cytosinyl)-2-(1-guanyl)ethane interstrand crosslinks (*34*), the mutagenic characteristics are thought to be due to other damaged bases (*35*). It has been hypothesized that EA is to some extent responsible for the mutagenicity of BCNU, and this investigation into the mutagenic signature of EA when replicated within cells was intended to shed light on that question.

The EA adduct is a new addition to the continuously-expanding list of substrates of the promiscuous direct reversal repair enzyme AlkB. While the unrepaired adduct is exceedingly toxic, AlkB is able to mitigate that toxicity completely, rendering the lesion as easily bypassed as a normal base in *E. coli* that possess functional protein. Such lesion bypass evidence has been exploited in the past to establish an adduct as a substrate for AlkB. For example, the structurally related adduct 1,N^6-ethenoadenine was seen to exhibit the same characteristic non-toxic properties in wild-type cells (85% bypass) while allowing replication only 5% of the time in *alkB* cells (*7*). Through further studies, this alleviation in toxicity was demonstrated to be attributable to direct reversal of the εA adduct by AlkB.

Indeed, mass spectral analysis of an *in vitro* reconstitution of the repair reaction revealed that AlkB is capable of metabolizing the EA adduct, most likely through the protein's well-established mechanism of oxidative hydroxylation at the carbon connected to the N1 of adenine. While the data do not rule out hydroxylation of EA at the C7 position (that attached to the exocyclic nitrogen), this scenario is unlikely as AlkB did not demonstrate repair activity of 6-methyladenine (data not shown), and an adduct attached at the N1 position of adenine would likely block replication (as m1A (*24*)), making its presence inconsistent with the lack of toxicity observed here.

Application of the canonical repair mechanism of AlkB in this situation diverges from the usual pathway in that the oxidized carbon that is normally released as formaldehyde is in this case covalently tethered to the base as an acetaldehyde residue (Fig. 4.6), providing opportunity for further reaction. Indeed, the hydroxylated adduct seems to isomerize to yield a ring-opened aldehydic form, which, at least *in vitro*, can then readily react with primary amines in the system, as demonstrated by the observed MALDI peaks possessing molecular weights consistent with the Schiff base product of the aldehydic ring-opened EA metabolite and either anthranilic acid from the matrix or PFBHA added exogenously as a trapping reagent. Reaction with anthranilic acid speaks to the high reactivity of the EA product, as the DNA is mixed with the matrix immediately prior to analysis, and yet the metabolite reacts extensively with an aromatically substituted amine in that brief amount of time. If this rapid reactivity holds *in vivo*, the action of AlkB on EA could lead to the formation of a range of secondary adducts, possibly including various small molecule lesions, intra- or interstrand DNA cross-links, and protein-DNA cross-links.

The mutational spectrum produced by replication of EA in wild-type *E. coli* is similar to that observed for the structurally-related εA adduct. Neither is substantially miscoding, as >98% and >97% A is present at the lesion site after replication of EA and εA, respectively. Given the analogous structures of the two adducts, however, it was expected that their mutational signatures would resemble each other, as they are impaired in hydrogen bonding capability in the same manner at identical sites. As has been demonstrated to be the case for εA, it was suspected that the low mutation frequency of

EA could be due to repair of the adduct before it is encountered by a replicative polymerase. Consequently, the coding characteristics of EA were also examined in an AlkB-deficient strain of bacteria. Despite displaying potent toxicity when unrepaired, EA is poorly mutagenic, even in this cellular context. The coding specificity of EA in this repair-deficient background was observed to be 96% A, 2% C, and 1% each G and T, in contrast to approximately 64% A, 22% T, 8% C, and 7% G for εA. The large differences in mutagenicity are most likely attributable to differential recognition by replicative polymerase of the EA lesion as compared to εA.

Although neither εA nor EA can participate in hydrogen bonding at the N^6 position, there are other characteristics of the two adducts that could account for their differential mutagenic signatures. While εA is planar, the saturated portion of the imidazole ring of EA may cause the adduct to take on an altered conformation within the active site of DNA polymerase. Additionally, the aromatic nature of εA could stabilize stacking interactions with amino acid side-chains of the polymerase, possibly resulting in differential positioning of the lesion within the active site relative to EA and leading to variation in what is incorporated opposite the adduct. Indeed, observed differences in the abilities of the base excision repair glycosylases of both humans and *E. coli* to recognize εA and EA have been attributed to these characteristics (*4,5*).

Although AlkB is able to act on EA, this is the first known instance of incomplete direct reversal of a damaged base (Fig. 4.6), resulting in the formation of a secondary adduct that may react further to create a variety of DNA lesions. Interestingly, this conversion to an N^6-adduct completely alleviates the toxicity of EA, indicating that neither it nor any downstream product blocks the progress of polymerase during replication. Additionally, replication past an EA adduct in the presence of AlkB results in incorporation of T more than 98% of the time, suggesting that the N^6-adduct and any derivatives it may form are able to code mainly as A, perhaps in part due to the restoration of hydrogen bonding ability (Fig. 4.6). Also surprising is the finding that in the absence of AlkB, unaltered EA, which lacks an instructional hydrogen bond, is still able to code as A 96% of the time, indicating that EA possesses unique characteristics that enable it to be read correctly, either by replicative or bypass polymerases, even

without employing hydrogen bonds to stabilize pairing with an incoming T nucleotide. While a similar phenomenon has been seen in other contexts (*36,37*), the relative contributions of size, shape and hydrogen bonding ability to faithful replication is not yet clear. Although the AlkA pathway is active in these cells, it is not likely responsible for the observed weakly-mutagenic phenotype given its poor affinity for EA and the single-stranded context in which the adduct exists within cells in this system (*5,34,35*). While the link between BCNU and EA has been established firmly only *in vitro*, the failure to detect EA in DNA isolated from treated cells or tissues is not entirely surprising. The rapidity of metabolism by AlkB (and perhaps the protein's human homologs) seen here would hinder a search for intact adduct by altering the structure of before analysis could be performed. Even a small quantity of the unmodified adduct, however, would carry major biological significance due to its extreme toxicity.

The poor mutagenesis induced by replication of EA seen here is in stark contrast to the miscoding nature of the adduct when copied by mammalian polymerases *in vitro* (*6*), suggesting either that further studies into alternative repair pathways that could be active on EA in *E. coli* are warranted or that fundamental dissimilarities in the replicative polymerases of mammals and bacteria cause the adduct to be read differently in the two systems. Although this study does not speak to the mutagenicity of EA, an investigation into the ability of hABH3 to metabolize EA would be useful to determine whether the adduct is in fact responsible for the mutagenic property of BCNU in humans. It may also be possible that AlkB and its mammalian homologs employ this novel pathway of metabolism to a secondary lesion to alleviate the toxicity and/or mutagenicity of additional cyclic lesions formed by this or other compounds.

Figure 4.1: Experimental outline. A. Structures of adenine (A), 1,N^6-ethanoadenine (EA) and 1,N^6-ethenoadenine (εA). B. Overview of experimental procedure for mutation detection.

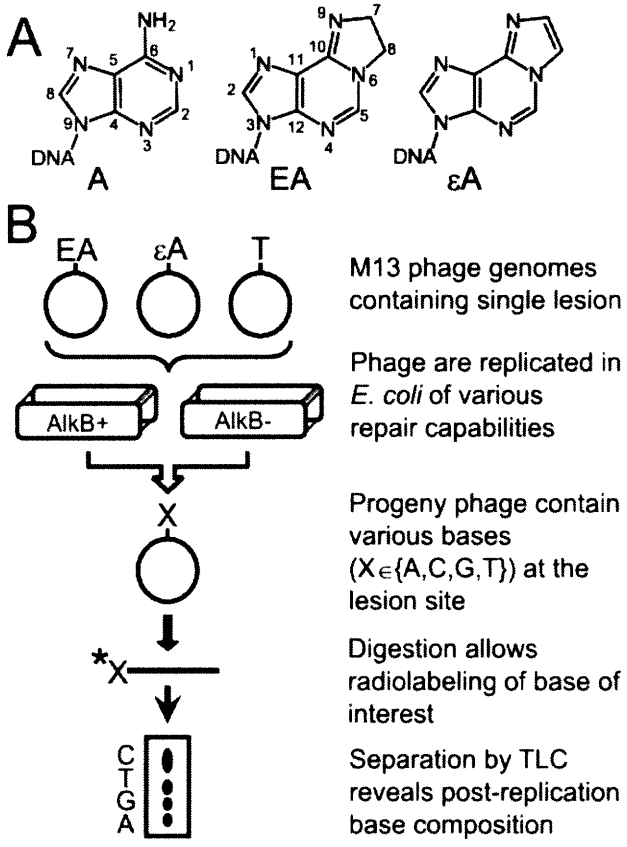


Figure 4.2: Lesion bypass of EA and εA in wild-type and *alkB E. coli* as compared to a lesion-free control containing T at the relevant site.

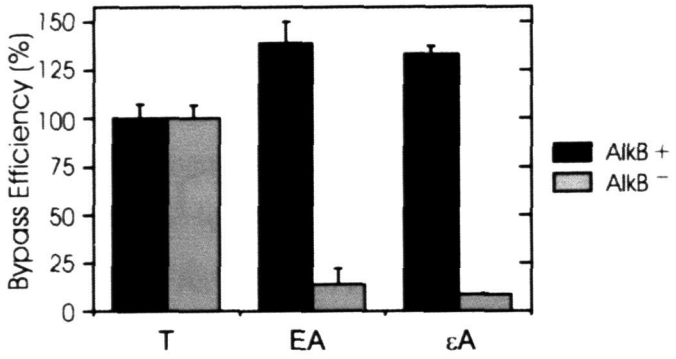


Figure 4.3: MALDI-TOF mass spectra of *in vitro* repair reaction of AlkB with a 16-mer oligonucleotide containing a single EA residue. A. Control incubation of oligonucleotide and molecular weight standards without AlkB. B. Complete conversion of EA to +16 and +135 products after a 30 min incubation at 37°C. C. Complete conversion to a single +211 product by the post-incubation addition of a 100-fold excess of PFBHA.

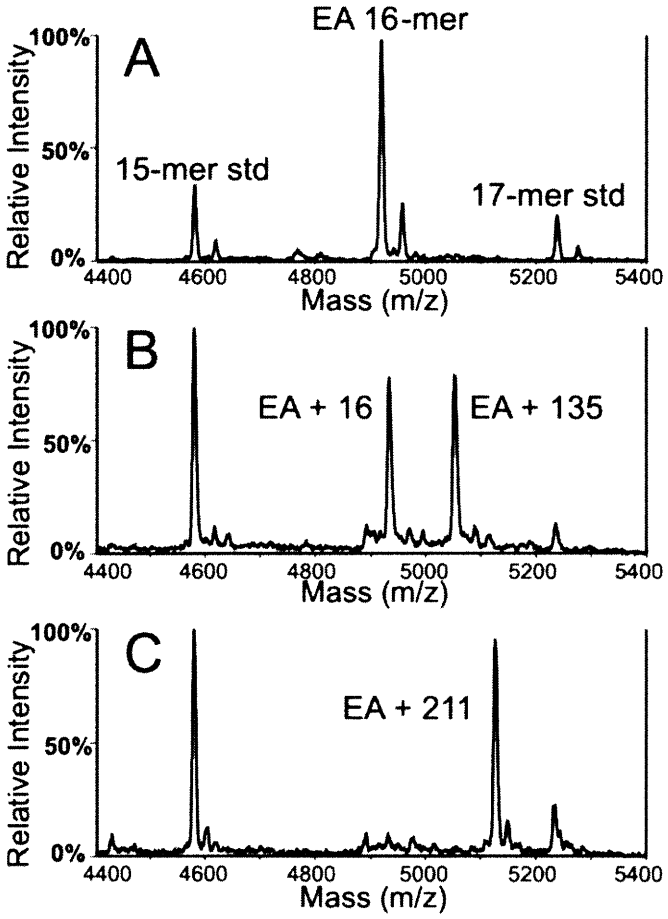


Figure 4.4: Putative structures of repair reaction products as seen by mass spectrometry. A. Hydroxylated product (EA+16) of AlkB metabolism of EA and its ring-opened aldehydic isomer. B. Product (EA+135) of ring-opened isomer reaction with anthranilic acid. C. Product (EA+211) of ring-opened isomer reaction with PFBHA.

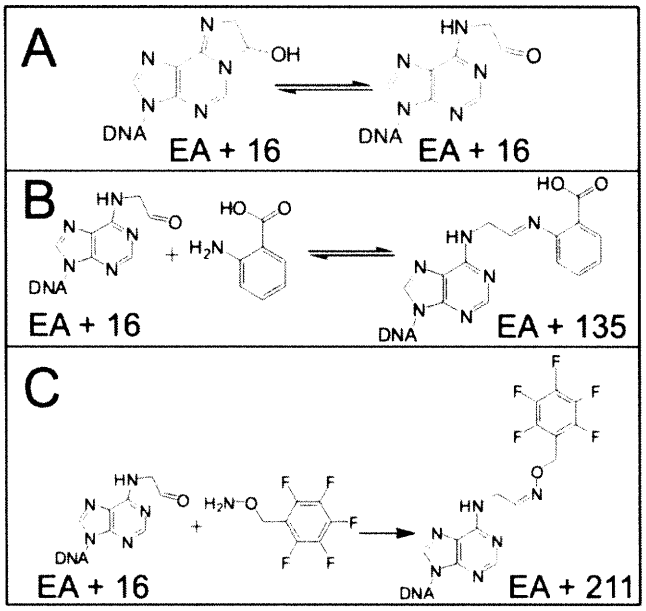


Figure 4.5: Mutational frequency and specificity of EA and εA replicated in wild-type and *alkB E. coli*.

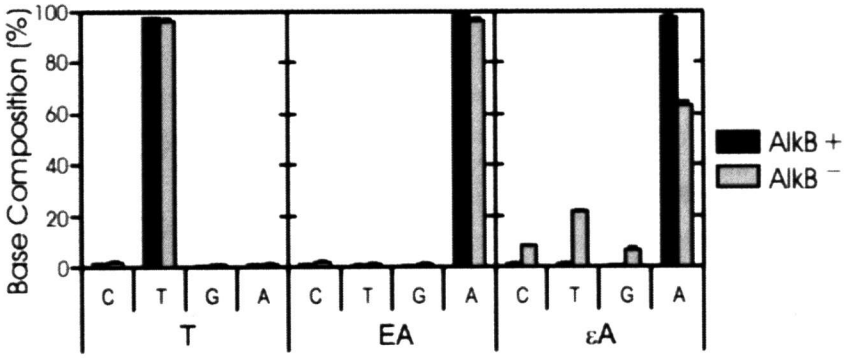
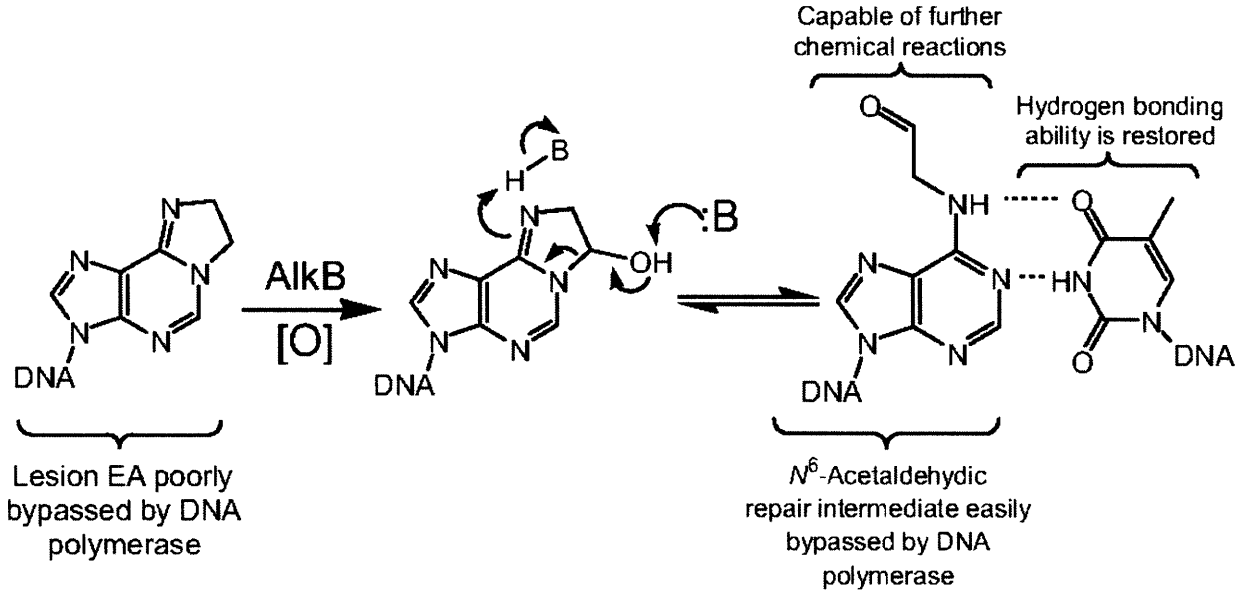


Figure 4.6: Proposed reaction pathway through which AlkB metabolizes EA. Restoration of hydrogen bonding ability might favor formation of the reactive ring-opened aldehydic species in double-stranded DNA.



References

1. Ludlum, D. B. 1986. Nature and biological significance of DNA modification by the haloethylnitrosoureas. *IARC Sci. Publ.*71-81.

2. Ludlum, D. B. 1986. Formation of cyclic adducts in nucleic acids by the haloethylnitrosoureas. *IARC Sci. Publ.*137-146.

3. Ludlum, D. B. 1990. DNA alkylation by the haloethylnitrosoureas: nature of modifications produced and their enzymatic repair or removal. *Mutat. Res.* **233**:117-126.

4. Guliaev, A. B., B. Hang, and B. Singer. 2002. Structural insights by molecular dynamics simulations into differential repair efficiency for ethano-A versus etheno-A adducts by the human alkylpurine-DNA N-glycosylase. *Nucleic Acids Res.* **30**:3778-3787.

5. Guliaev, A. B., B. Singer, and B. Hang. 2004. Chloroethylnitrosourea-derived ethano cytosine and adenine adducts are substrates for Escherichia coli glycosylases excising analogous etheno adducts. *DNA Repair (Amst)* **3**:1311-1321.

6. Hang, B., A. Chenna, A. B. Guliaev, and B. Singer. 2003. Miscoding properties of 1,N6-ethanoadenine, a DNA adduct derived from reaction with the antitumor agent 1,3-bis(2-chloroethyl)-1-nitrosourea. *Mutat. Res.* **531**:191-203.

7. Delaney, J. C., L. Smeester, C. Wong, L. E. Frick, K. Taghizadeh, J. S. Wishnok, C. L. Drennan, L. D. Samson, and J. M. Essigmann. 2005. AlkB reverses etheno DNA lesions caused by lipid oxidation in vitro and in vivo. *Nat. Struct. Mol. Biol.* **12**:855-860.

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

9. Rohankhedkar, M. S., S. B. Mulrooney, W. J. Wedemeyer, and R. P. Hausinger. 2006. The AidB component of the Escherichia coli adaptive response to alkylating agents is a flavin-containing, DNA-binding protein. *J. Bacteriol.* **188**:223-230.

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

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

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

13. Falnes, P. O. and T. Rognes. 2003. DNA repair by bacterial AlkB proteins. *Res. Microbiol.* **154**:531-538.

14. Drablos, F., E. Feyzi, P. A. Aas, C. B. Vaagbo, B. Kavli, M. S. Bratlie, J. Pena-Diaz, M. Otterlei, G. Slupphaug, and H. E. Krokan. 2004. Alkylation damage in DNA and RNA--repair mechanisms and medical significance. *DNA Repair (Amst)* **3**:1389-1407.

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

16. Lee, D. H., S. G. Jin, S. Cai, Y. Chen, G. P. Pfeifer, and T. R. O'Connor. 2005. Repair of methylation damage in DNA and RNA by mammalian AlkB homologues. *J. Biol. Chem.* **280**:39448-39459.

17. Falnes, P. O. 2004. Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. *Nucleic Acids Res.* **32**:6260-6267.

18. Koivisto, P., P. Robins, T. Lindahl, and B. Sedgwick. 2004. Demethylation of 3-methylthymine in DNA by bacterial and human DNA dioxygenases. *J. Biol. Chem.* **279**:40470-40474.

19. Sedgwick, B. 2004. Repairing DNA-methylation damage. *Nat. Rev. Mol. Cell Biol.* **5**:148-157.

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

21. Falnes, P. O., M. Bjoras, P. A. Aas, O. Sundheim, and E. Seeberg. 2004. Substrate specificities of bacterial and human AlkB proteins. *Nucleic Acids Res.* **32**:3456-3461.

22. Koivisto, P., T. Duncan, T. Lindahl, and B. Sedgwick. 2003. Minimal methylated substrate and extended substrate range of Escherichia coli AlkB protein, a 1-methyladenine-DNA dioxygenase. *J. Biol. Chem.* **278**:44348-44354.

23. Ringvoll, J., L. M. Nordstrand, C. B. Vagbo, V. Talstad, K. Reite, P. A. Aas, K. H. Lauritzen, N. B. Liabakk, A. Bjork, R. W. Doughty, P. O. Falnes, H. E. Krokan, and A. Klungland. 2006. Repair deficient mice reveal mABH2 as the primary oxidative demethylase for repairing 1meA and 3meC lesions in DNA. *EMBO J.* **25**:2189-2198.

24. Delaney, J. C. and J. M. Essigmann. 2004. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A* **101**:14051-14056.

25. Yu, B., W. C. Edstrom, J. Benach, Y. Hamuro, P. C. Weber, B. R. Gibney, and J. F. Hunt. 2006. Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. *Nature* **439**:879-884.

26. Mishina, Y., C. G. Yang, and C. He. 2005. Direct repair of the exocyclic DNA adduct 1,N6-ethenoadenine by the DNA repair AlkB proteins. *J. Am. Chem. Soc.* **127**:14594-14595.

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

28. Ougland, R., C. M. Zhang, A. Liiv, R. F. Johansen, E. Seeberg, Y. M. Hou, J. Remme, and P. O. Falnes. 2004. AlkB restores the biological function of mRNA and tRNA inactivated by chemical methylation. *Mol. Cell* **16**:107-116.

29. Chenna, A., H. Maruenda, and B. Singer. 1999. Synthesis of para-benzoquinone and 1,3-bis(2-chloroethyl)nitrosourea adducts and their incorporation into oligonucleotides. *IARC Sci. Publ.*89-101.

30. Delaney, J. C. and J. M. Essigmann. 2006. Assays for Determining Lesion Bypass Efficiency and Mutagenicity of Site-Specific DNA Lesions *In Vivo*. *Methods in Enzymology* **408A**:1-15.

31. Begley, T. J. and L. D. Samson. 2003. AlkB mystery solved: oxidative demethylation of N1-methyladenine and N3-methylcytosine adducts by a direct reversal mechanism. *Trends Biochem. Sci.* **28**:2-5.

32. Lemoine, A., C. Lucas, and R. M. Ings. 1991. Metabolism of the chloroethylnitrosoureas. *Xenobiotica* **21**:775-791.

33. Eisenbrand, G., N. Muller, E. Denkel, and W. Sterzel. 1986. DNA adducts and DNA damage by antineoplastic and carcinogenic N-nitrosocompounds. *J. Cancer Res. Clin. Oncol.* **112**:196-204.

34. Tong, W. P., M. C. Kirk, and D. B. Ludlum. 1982. Formation of the cross-link 1-[N3-deoxycytidyl),2-[N1-deoxyguanosinyl]ethane in DNA treated with N,N'-bis(2-chloroethyl)-N-nitrosourea. *Cancer Res.* **42**:3102-3105.

35. Wiencke, J. K. and J. Wiemels. 1995. Genotoxicity of 1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU). *Mutat. Res.* **339**:91-119.

36. Potapova, O., C. Chan, A. M. DeLucia, S. A. Helquist, E. T. Kool, N. D. Grindley, and C. M. Joyce. 2006. DNA polymerase catalysis in the absence of Watson-Crick hydrogen bonds: analysis by single-turnover kinetics. *Biochemistry* **45**:890-898.

37. Kim, T. W., J. C. Delaney, J. M. Essigmann, and E. T. Kool. 2005. Probing the active site tightness of DNA polymerase in subangstrom increments. *Proc. Natl. Acad. Sci. U. S. A* **102**:15803-15808.

CHAPTER 5:

Cross-linking in Duplex DNA by the Metabolite of 1,N^6-ethanoadenine Created by the *Escherichia coli* Adaptive Response Protein AlkB

Abstract

The 1,N^6-ethanoadenine (EA) adduct is formed in DNA through the reaction of adenine with 1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU), a chemotherapeutic agent used to combat various brain, head, and neck tumors. The *Escherichia coli* adaptive response protein AlkB is capable of acting on EA via its canonical mechanism of oxidative dealkylation. The reaction alleviates the unrepaired adduct's potent toxicity through metabolism at the C8 position (attached to N1 of adenine), producing a non-toxic and weakly-mutagenic N^6 adduct. The secondary lesion formed by the action of this repair protein contains a reactive aldehyde moiety, which has been shown in other contexts to be capable of forming cross-links. Here, we explore the ability of the EA metabolite generated by AlkB to cross-link *in vitro*, and demonstrate that the "repair" product can mediate the formation of interstrand cross-links in certain sequence contexts of duplex DNA.

Introduction

The *E. coli* AlkB protein is the product of one of four genes upregulated in response to cellular exposure to methylating agents in a biochemical network called the adaptive response (*1*). The rest of the network comprises Ada, a bifunctional methyltransferase that not only reverses DNA damage directly by accepting methyl groups from O^6-methylguanine and the *S*-diastereomer of backbone methylphosphotriesters through irreversible transfer, but also regulates transcription, inducing expression at three promoters after methylation of Cys38, AlkA, a 3-methyladenine DNA glycosylase responsible for the recognition and hydrolysis of damaged bases as the first step in the base excision repair pathway, and AidB, which is homologous to human isovaleryl coenzyme A dehydrogenase (IVD, a central player in the metabolism of leucine), but the function of which has not yet been determined beyond its ability to bind double-stranded DNA (*2*).

The AlkB protein is a member of the α-ketoglutarate dioxygenase family of enzymes that uses non-heme iron (II), oxygen, and α-ketoglutarate as cofactors and cosubstrates to remove the aberrant methyl group from 1-methyladenine (m1A) and 3-methylcytosine (m3C), lesions formed predominantly in single-stranded DNA under conditions of alkylative stress (*3,4*), through a mechanism termed oxidative demethylation. Hydroxylation of the DNA-bound methyl group is coupled with decarboxylation of α-ketoglutarate, which releases succinate and carbon dioxide and produces an unstable hydroxymethyl intermediate that then decomposes, releasing formaldehyde and the repaired adenine or cytosine (Figure 5.1). While proteins homologous to AlkB exist in many other bacterial species, as well as in *S. pombe*, *C. elegans, Drosophila,* mice, humans, and plant RNA viruses (*5-7*), proteins of similar sequence are apparently lacking in many other species, both bacterial and eukaryotic (*6*). Interestingly, however, the expression of several *S. cerevisiae* genes that possess no amino acid homology to AlkB in *E. coli* lacking functional AlkB complements the MMS-sensitive phenotype imparted by the deficiency (*8*). The mouse AlkB homologs, mABH2 and mABH3 (*9*), and two of the eight identified human homologs, hABH2 and hABH3 (*10-13*), are also capable of mediating the direct reversal of methylated DNA base

damage, although substrate preferences vary among enzymes (*14,15*). Recently, mABH2 was shown to be responsible for the removal of endogenously-formed m1A adducts in mice (*16*).

In addition to the originally-discovered substrates, m1A and m3C, AlkB has been shown to remove the alkyl groups from 3-ethylcytosine (*17*), 1-methylguanine (*10,17*), 3-methylthymine (*10,11,17*), 1-ethyladenine (*13*), EA (*18*), εA (*19,20*), and εC (*19*) lesions in DNA, demonstrating the enzyme's biochemical versatility. The etheno bases are reversed by a second chemical mechanism in which the double bond is epoxidized, hydrolyzed, and released as glyoxal, restoring the undamaged base (*19*). Additionally, AlkB and its human homolog hABH3 are able to remove methyl groups from RNA, reactivating methylated RNA phage (*21*) and restoring the function of methylated mRNA and tRNA molecules *in vitro* (*22*). The recently solved X-ray crystal structure (*23*) shows a particularly malleable region within the nucleotide-recognition lid that may explain the ability of the protein to recognize many lesions differing in both size and shape.

The base adduct 1,N^6-ethanoadenine (EA) forms when, upon reaction of DNA with 1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU) (*24-26*), the exocyclic nitrogen of adenine becomes connected to the N1 ring nitrogen by a saturated two-carbon bridge, creating a five-membered ring (Figure 5.1). The lesion is a weak substrate for the human alkyladenine DNA glycosylase (MPG, also AAG, APNG, or ANPG) (27), while the structurally related 1,N^6-ethenoadenine (εA) (Figure 5.1) is recognized and repaired with 65-fold greater efficiency. Similarly, the *E. coli* repair protein 3-methyladenine DNA glycosylase (AlkA) is to some extent capable of recognizing and excising EA, but with 20-fold lower activity than εA (*28*). Recently, the lesion has been shown to be a substrate for the AlkB protein both *in vitro* and *in vivo* (*18*).

While the adduct is potently toxic but only weakly mutagenic when replicated *in vivo* in AlkB-deficient *E. coli*, replication of a lesion-containing viral genome in an AlkB-proficient strain proceeds as efficiently as that of a thymine-containing control genome (*18*). AlkB alleviates the toxicity of EA completely through hydroxylation at the C8 position (that attached to N1 of adenine, Figure 5.1), creating a metabolite that

equilibrates with a ring-opened, aldehydic isomer. This isomer was shown to react with various primary amines to form an array of adducts *in vitro*, suggesting that perhaps the EA-metabolite could also cross-link *in vivo* with a more biologically-relevant species - i.e. protein or DNA.

Aldehydes are reactive chemical species that have been established as cross-linking agents in many contexts (*29-36*) and are even used in the generation of synthetic cross-links (*37*). These compounds are capable of forming cross-links of several types, including intrastrand DNA, interstrand DNA, and DNA-protein conjugations. For example, acetaldehyde, a natural product of cellular glycolysis that is usually detoxified through reduction to ethanol by alcohol dehydrogenases, covalently binds DNA to form an interstrand linkage (*38,39*). Additionally, aldehydic species that are produced at high concentrations at sites of inflammation, where levels of lipid peroxidation are elevated, are thought to form interstrand cross-links in DNA (*38*). DNA-protein cross-links are formed easily by aldehydes at sites such as histones where the two reside in close proximity. Reactive aldehydic species found in car exhaust and tobacco smoke and generated as a breakdown product of oxidized lipids, such as acrolein and crotonaldehyde, are also known to generate both interstrand DNA and DNA-protein cross-links (*35,38*).

Interstrand cross-links are a particularly grave form of DNA damage, as they severely impede cellular processes by preventing the separation of the complementary strands necessary for replication and transcription. It is just this toxic property of interstrand cross-links that enables chemicals that form them efficiently to be used as chemotherapeutic agents. Nitrous acid, mitomycin, nitrogen mustards, cisplatin, and some psoralens are examples of cross-linking drugs that have been used to combat cancer clinically (*38,40-43*). Beyond the physical hazard of binding the strands of the double helix together, interstrand cross-links present an informational danger because they affect the coding ability of both strands of DNA. Nucleotide excision repair paired with homologous recombination is one mechanism by which cells attempt to repair interstrand cross-links (*38*). While this method can be effective, it proceeds through an intermediate involving a single-strand break, which can be quite deleterious if it is encountered during cellular replication and converted into a double-stranded break. The previously observed

interaction of the aldehydic metabolite created by the action of AlkB on the EA adduct with primary amines *in vitro*, paired with the established ability of aldehydes to form interstrand DNA cross-links and the potential toxicity of those cross-links *in vivo*, motivated an exploration of the reactivity of the EA metabolite to determine whether it was capable of forming cross-links in double-stranded DNA.

Materials and Methods

Oligonucleotides. Oligonucleotides were synthesized using phenoxyacetal-protected phosphoramidites and capping reagents from Glen Research. Approximately 50 mg of EA phosphoramidite (Chemgenes), which has previously been site-specifically incorporated into an oligonucleotide (*44*), were used during a 30 min coupling with the solid support. Following synthesis, the resin was incubated with concentrated ammonium hydroxide for one hour prior to lyophilization. The 16-mer insert oligonucleotide 5'-GAAGACCTXGGCGTCC-3' (X=EA) was purified by anion exchange and reversed phase HPLC, and the molecular weight was verified by MALDI-TOF mass spectrometry (4916.21 observed, 4916.22 calculated). Sixteen-mer oligonucleotides of the same sequence, but where X = εA, or A) were synthesized and purified as described (*19*). Purified complementary sixteen-mer oligonucleotides that were either entirely correctly paired or mismatched at one site with the sequence above (either one base 5' or 3' to the lesion site) were obtained from Integrated DNA Technologies (Coralville, IA). Complement sequences were 5'-GGACGCCTAGGTCTTC-3', denoted C1; 5'-GGACGCCT-2AP-GGTCTTC-3', where 2AP = 2-aminopurine, denoted C2; and 5'-GGACGCTTAGGTCTTC-3', denoted C3 (Table 5.1).

Construction and Purification of AlkBΔN11. The AlkB open reading frame, excluding the first 11 amino acids, was PCR-amplified using the oligonucleotides 5'-GGAATTC**CATATG**CAAGAGCCACTGGCGG (*Nde*I restriction site), and 5'-CCG**CTCGAG**GCCTTGAAAATATAGGTTTTCTTTTTTACCTGCCTG (*Xho*I site). The PCR product was purified, digested with *Nde*I and *Xho*I (NEB), and ligated into a C-

terminal-histidine-tag-containing pET24a vector (Novagen) that had been digested by *Nde*I and *Xho*I. The ligation mixture was transformed into DH5α cells for verification by sequencing. pET24a-AlkBΔN11 was transformed into BL21(DE3) (Novagen) cells for expression.

pET24a-AlkBΔN11 cells were grown at 37 °C to an A_{600} of 0.4, when isopropyl-β-D-thiogalactoside (IPTG) was added at 1 mM and the temperature was lowered to 30 °C. Cells were harvested after 4 h by centrifugation, resuspended in 10 mM Tris pH 7.3, 300 mM NaCl, 2 mM $CaCl_2$, 10 mM $MgCl_2$, 5% glycerol and 1 mM BME (lysis buffer), and homogenized by sonication. Lysate was recovered by centrifugation. The supernatant was loaded onto an Ni-NTA column (Qiagen), washed (20 mM) and eluted (70 and 250 mM) in imidazole in lysis buffer. The eluent was dialyzed against 50 mM TES pH 7.1 and loaded onto an SP Sepharose cation exchange column (Amersham Pharmacia). AlkBΔN11 was eluted by a linear gradient of 0-1 M NaCl over 75 mL. The fractions containing AlkBΔN11 were pooled and purity was established by SDS-PAGE. As previously reported (*23*), AlkBΔN11 has similar activity to the full-length protein in standard AlkB assays (data not shown).

Standard *in vitro* cross-linking. In the basic protocol (Figure 5.2), a 16-mer oligodeoxynucleotide of sequence 5'-GAAGACCTXGGCGTCC-3' (where X = EA, εA, or A) was radiolabeled using T4 polynucleotide kinase, purified using Micro Bio-Spin chromatography columns containing Bio-Gel P-6 (Bio-Rad Laboratories) to remove unincorporated ATP, and annealed to its complement strand containing T opposite the lesion (C1; 5'-GGACGCCTAGGTCTTC-3'). Double-stranded substrates (1 pmol) were then incubated with 2 pmol purified AlkBΔN11 in 10 μL reaction buffer (70 μM $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$, 0.9 mM α-ketoglutarate, 1.8 mM L-ascorbate, and 46.5 mM HEPES (pH 8.0)) for 30 min at 37 °C. Reactions were halted by the addition of 10 μL denaturing loading buffer and products were separated by 20% denaturing PAGE. Bands were visualized using phosphorimagery and quantified using ImageQuant 5.2 software. Variations on the standard protocol included the use of an equal amount of unannealed single-stranded lesion-bearing substrate (rather than the duplex), a post-AlkB-reaction

incubation lasting 1 h – 7 d at 30 - 37 °C with 50 - 500 mM $NaCNBH_3$, competition with a 1000-fold molar excess of the trapping reagent *O*-(2,3,4,5,6-pentafluorobenzyl)hydroxylamine (PFBHA), digestion with 0.3 U Proteinase K at 37 °C for 2 h prior to addition of loading buffer, and radiolabeling of the complement, rather than the lesion-containing, DNA strand.

***In vitro* cross-linking with mismatched substrates.** Three complement oligonucleotides (of sequence 5'-GGACGCCTAGGTCTTC-3', denoted C1; 5'-GGACGCCT-2AP-GGTCTTC-3', denoted C2; 5'-GGACGCTTAGGTCTTC-3', denoted C3) were radiolabeled using T4 polynucleotide kinase, and purified as above. Mixtures containing various ratios of complement strands (100% C1, 75% C1 + 25% C2, 50% C1 + 50% C2, 25% C1 + 75% C2, 100% C2 and analogous combinations of C1 and C3) were annealed to the EA-containing 16-mer substrate strand. These partially mismatched double-stranded substrates were incubated with AlkBΔN11, separated by PAGE, and quantified using phosphorimagery.

Results

***In vitro* cross-linking by the EA-repair product.** Incubation of EA*:C1 (where * indicates the radiolabeled strand) with AlkBΔN11 caused ~30% of the radioactive label to run more slowly than the free 16-mer (Figure 5.3). Two types of retarded signal were observed, a distinct band and a diffuse smear. Because this interaction is most likely mediated by the formation of a Schiff base between the aldehydic portion of the EA "repair product" and an amine in the system, which is a reversible interaction, it is hypothesized that the band and smear most likely correspond to the fully cross-linked product and product that was merely temporarily cross-linked, respectively. The A*:C1 and εA*:C1 substrates did not cross-link efficiently (<4%), nor did EA*:C1 duplex when incubated without AlkB or with heat-inactivated protein.

Reducing agent does not increase observed cross-linking. The addition of $NaCNBH_3$ to the reaction at various times, concentrations, temperatures, and pHs and for various

incubation periods did not increase the amount of observed cross-linking, nor did it increase the amount of signal in the distinct band, which would correspond to reduced covalently cross-linked product (Figure 5.4 and data not shown).

Cross-links are not DNA-protein. Digestion of the AlkB protein post-incubation by Proteinase K did not diminish the amount of shifting observed (Figure 5.5), indicating that the slower-running complex is not due to an interaction of the EA-repair product with the protein itself.

Cross-links are not DNA-small molecule. Adding a small molecule trapping reagent *O*-(2,3,4,5,6-pentafluorobenzyl)hydroxylamine (PFBHA) during the AlkB reaction completely abrogates the shift in radioactivity, while addition post-AlkB-reaction causes a ~65% decrease in shifting (Figure 5.6).

Cross-links require both strands of DNA. Incubation of unannealed EA-containing single-stranded oligonucleotide with protein produced no observed shift of radioactive signal (Figure 5.7), indicating that single-stranded substrates are not capable of forming the cross-link observed when both strands are present, and suggesting that the interaction is not intramolecular or between two strands of the same sequence. Additionally, the amount of cross-linking observed for EA*:C1 is identical to that generated by EA:C1* (Figure 5.7), suggesting that both strands of the duplex are involved in the interaction.

Cross-links occur with C in opposite strand. In the titration experiment using mismatched complements (Figure 5.8), all ratios of EA:C1* to EA:C2* (the 2-aminopurine-containing complement) produced the same amount of shifted radioactivity (~25%; Figure 5.8). In contrast, the amount of shifted radioactivity was observed to correlate directly with the amount of C1 substrate in the C1/C3 (the T-containing complement) experiment (Figure 5.9), providing evidence that the cytosine one position 5' to the base opposite the lesion in the complement strand is essential for cross-linking.

Discussion

The ability of an EA-containing oligonucleotide to cause the appearance of slower-running radioactive signal upon metabolism by the AlkB protein indicates that the ring-opened aldehydic product is in fact capable of forming secondary adducts in DNA, most likely through a reversible Schiff base interaction with an amine group. The lack of shifting observed for control substrates (containing A or εA) and in control reactions (with absent or heat-inactivated protein) suggests that the responsible interaction is specific to the AlkB-mediated product of EA. Such shifting could in principle be the product of a DNA-protein adduct, a DNA-small molecule complex, an intrastrand DNA-DNA interaction, or an interstrand DNA-DNA cross-link, either between two copies of the same sequence or between the two strands of the double helix. Evidence is provided here that the radiolabeled substrate of lower mobility corresponds to an interstrand DNA-DNA cross-link that tethers together the two complementary strands.

Despite the established use of $NaCNBH_3$ to reduce imine linkages such as that produced here to allow for characterization of Schiff base mediated cross-links (*32-35,37*), attempts to reduce the Schiff base product created here to a covalent adduct identifiable by mass spectrometry were unsuccessful. Many sets of reaction conditions and incubation times were explored, but none produced a covalently-linked product as monitored by the intensity of the distinct band running approximately as a 32-mer on a denaturing gel. Relatively extreme conditions or incubation times have been required in the past for the success of this reaction (*33,35*), and perhaps reduction in this case would be observed with greatly extended incubation times or altered pH or temperature.

One potential source of a reduced-mobility species from the reaction of AlkB with an EA-containing duplex is a cross-link between the aldehyde of EA that is created by the AlkB reaction with a lysine residue in the active site of the protein itself. Such reaction of a DNA repair protein with a partially-repaired substrate has precedent, at least in the case of human O^6-alkylguanine DNA alkyltransferase (*45*). Furthermore, the cross-linked DNA-protein product has been detected by this gel shift method in the past (*30,32,45*). If this were the reaction responsible for the shifting of signal seen here, digestion of the protein by a protease would reduce or eliminate the shifting seen, as the

migration of the full protein adduct would be expected to be different from the migration of a peptide adduct. In this case, however, digestion of the AlkB protein by a protease after the repair reaction did not affect the amount or pattern of observed shifting, suggesting that the shift is not a DNA-protein cross-link.

Although it has been suggested previously (*34*) that the interaction of an aldehyde-containing oligonucleotide with Tris base from gel running buffer could slow the mobility of that oligonucleotide and cause a shifted band to be visualized on a gel, we find evidence to the contrary here. In fact, the addition of a large molecular excess of PFBHA, which forms a covalent linkage with the aldehyde formed by the action of AlkB on EA, is capable of completely abolishing the shift in radioactive signal. The observed shift, then, must be due to a reversible interaction with something much larger (or otherwise mobility-inhibiting) than PFBHA, as the complex of the oligonucleotide with PFBHA is not different enough in mass or mobility to be observed on the gel but is capable of titrating the reactive oligodeoxynucleotide away from the interaction that produces the shifted radioactivity. The cross-links observed are then most likely not the result of reaction with a small molecule.

The cross-link could also derive from an intramolecular interaction that constrains the motion of and thereby reduces the mobility of the oligonucleotide or from an interaction between two copies of the adducted DNA strand. However, the lack of shifting observed for the unannealed single-stranded EA substrate (upon which AlkB should be even more active than on the double-stranded equivalent (*21,46*)), provides evidence against the interaction being intramolecular or between multiple copies of the substrate strand. This lack of reactivity in the absence of the complementary strand also indicates that both strands of the duplex are required for cross-linking to occur.

If the shifting is in fact due to an interstrand cross-link involving both the lesion-bearing strand and its complement, then which of the two strands bears the radioactive label should not in any way impact the slower-running bands. Indeed, radiolabeling of the complement strand rather than the lesion-containing substrate strand produces the same amount and pattern of shifting, indicating that both strands of DNA are

incorporated in an interstrand cross-link that appears as a more slowly-running band on the denaturing gel.

To determine which base in the complement strand was responsible for interacting with the aldehyde group of the EA-metabolite, duplexes were constructed that were lacking certain proximal candidate amino groups, and the amount of cross-linking produced by these modified substrates upon incubation with AlkB was observed. Elimination of the exocyclic amino group of the adenine 3' to the lesion site in the complementary strand by substitution with 2-aminopurine had no effect on the ability of the duplex to cross-link upon incubation with functional AlkB protein. Removal of the cytosine residue one base 5' to the lesion (again in the complementary strand) by replacement with a thymine, however, caused a decrease in observed cross-linking proportional to the amount of thymine-containing substrate in the reaction. The observed interaction between the aldehyde and the amino group of cytosine could perhaps have been predicted by the fact that the aldehydic group of the EA metabolite resides in the major groove of the helix and is oriented in the 3' direction when the adduct is modeled into duplex DNA (Figure 5.9), which would position it well to interact with an exocyclic amine of the base one position 5' in the complement strand. The predicted placement of the secondary adduct created by the action of AlkB on EA in the major groove would not be surprising, as many adducts involving the N^6 position of adenine have been shown to take on such a conformation in duplex DNA (*47-51*).

Schiff base-mediated interstrand cross-links have been observed in other contexts in DNA (*31-33,35,37*), indicating that this type of interaction may have relevance that extends beyond the AlkB-mediated repair product of EA. DNA cross-links are exceedingly toxic to cells, as they prevent strand separation and are therefore potent blocks to replication, and a single unrepaired adduct is reported to be capable of killing a eukaryotic cell (*38,52*). In fact, it is just this toxic characteristic that is exploited in the use of efficient cross-linking agents such as cisplatin, nitrogen mustards, and mitomycin as chemotherapeutics (*38,40-43*). Even BCNU, the compound responsible for the formation of the EA adduct in DNA, is capable of inducing cell death through the

formation of toxic interstrand cross-links between N3 of cytosine and N1 of the opposing guanine (*24-26,53,54*).

These therapeutic agents form irreversible cross-links, however, while the imine species that form from aldehydes in DNA are readily hydrolyzable and their formation is perhaps therefore of less drastic consequence to the cell. Although up to 25% of the labeled oligonucleotide containing the reactive aldehydic EA metabolite can be shifted through formation of a Schiff base *in vitro*, the potent toxicity of the intact EA adduct was abrogated by metabolism by AlkB *in vivo*, indicating either that the cross-link observed *in vitro* does not form to a significant extent *in vivo* or that it is reversed readily enough during replication of the phage genome that its presence is not detected as a drop in replication efficiency. Further work to elucidate the kinetics exhibited by these cross-links *in vivo* will be necessary to fully understand their impact on cellular replication, transcription, and survival.

Figure 5.1: Top panel: Structures of A and the two exocyclic adducts EA and εA. The differential numbering of A and EA is shown for clarity. Bottom panel: Mechanism of metabolism of EA by the adaptive response protein AlkB. The hydroxylated product can ring-open to form a reactive secondary adduct.

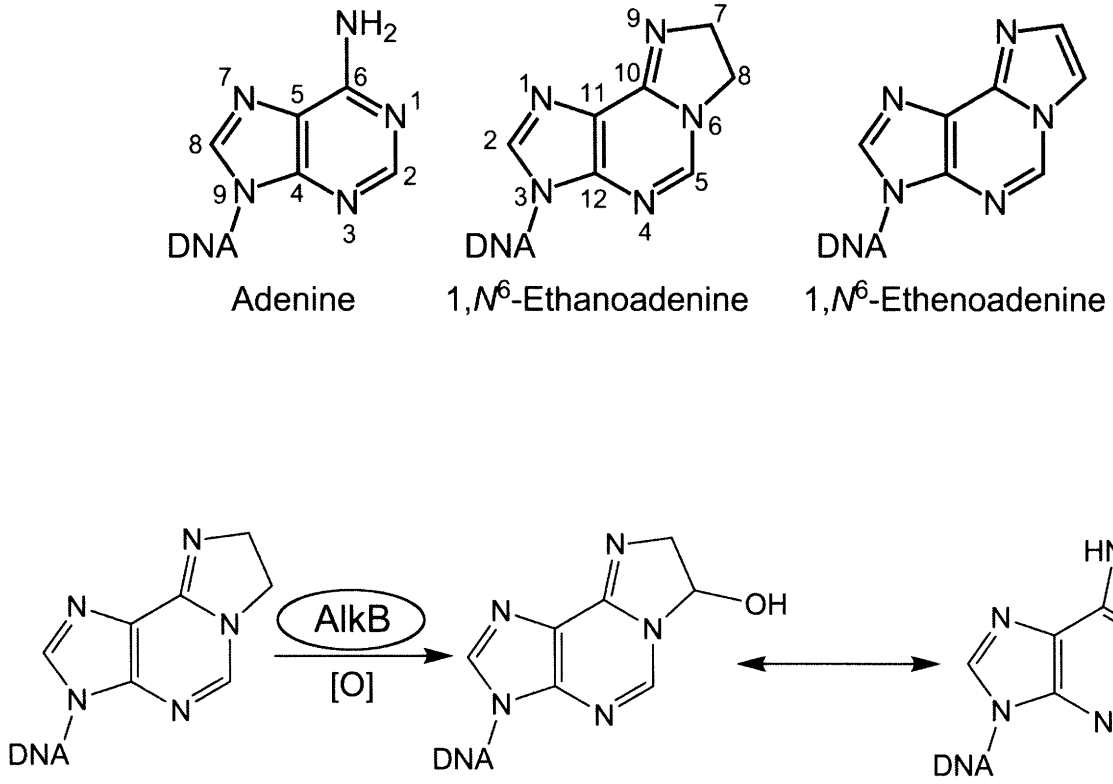


Figure 5.2: Experimental outline of *in vitro* cross-linking studies

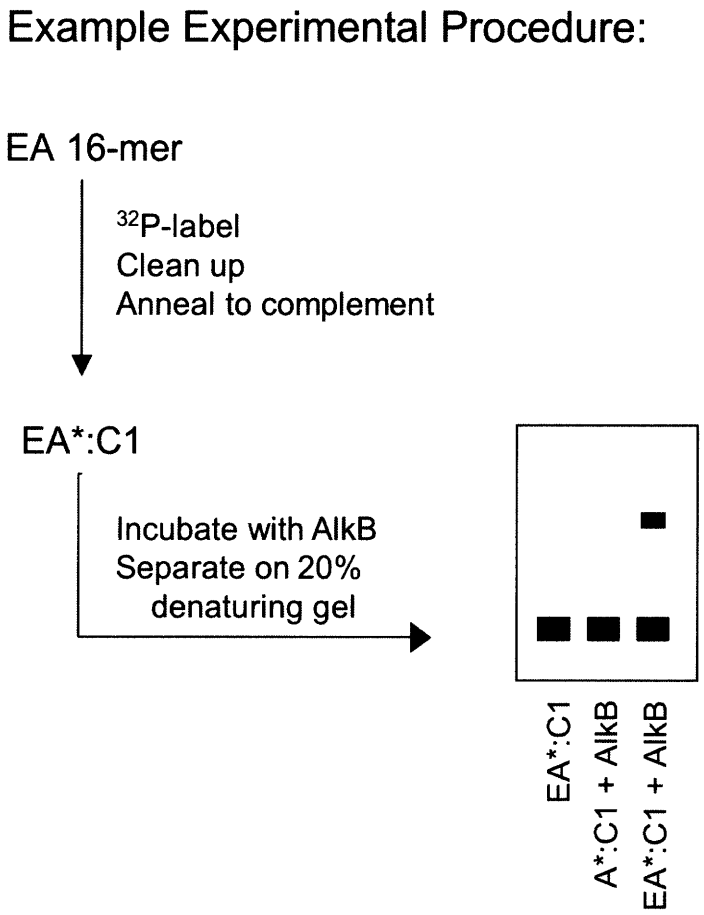
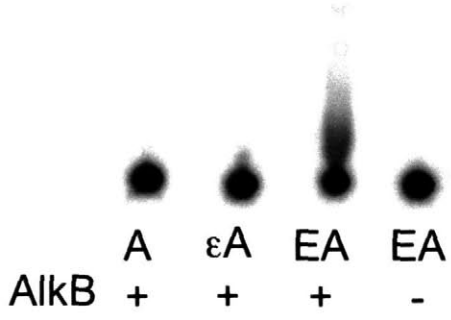


Figure 5.3: Cross-linking experiments with lesion-containing oligonucleotides annealed to the correctly base paired complement C1. Cross-linking requires functional AlkB and is specific to oligonucleotide substrates containing EA.



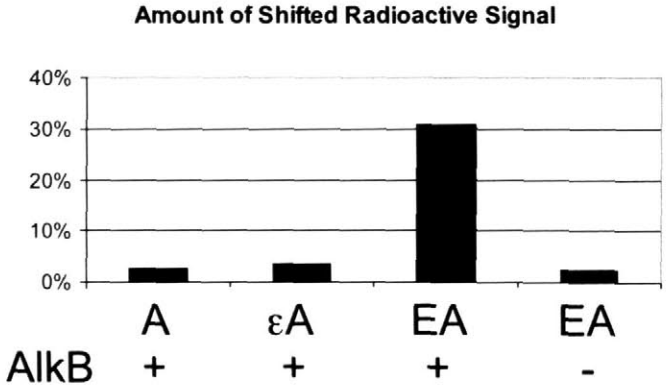
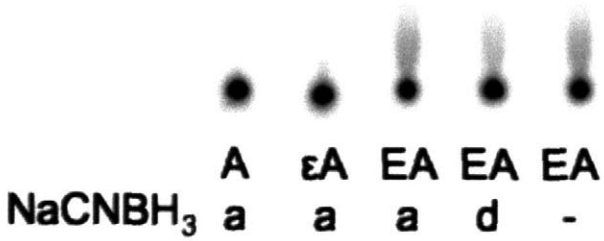


Figure 5.4: Cross-linking experiments with lesion-containing oligonucleotides annealed to the correctly base paired complement C1. Cross-linking is not enhanced by reduction with 50 mM $NaCNBH_3$ added during (d) or after (a) the AlkB repair reaction.



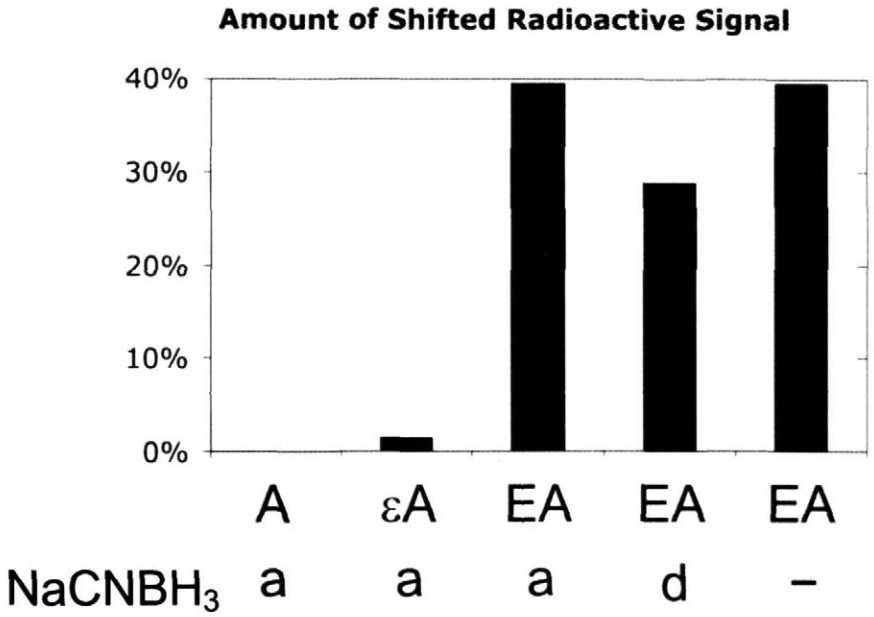
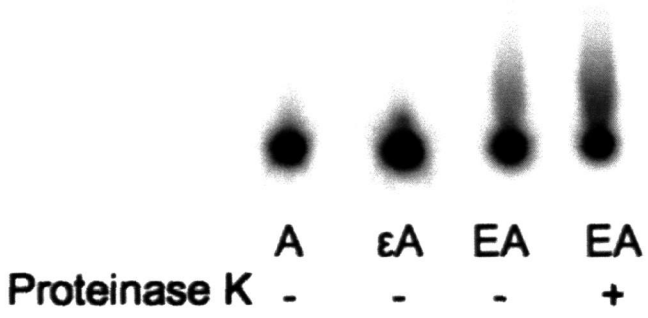


Figure 5.5: Cross-linking experiments with lesion-containing oligonucleotides annealed to the correctly base paired complement C1. Cross-linking is not diminished by post-repair protein digestion with 0.6 U Proteinase K.



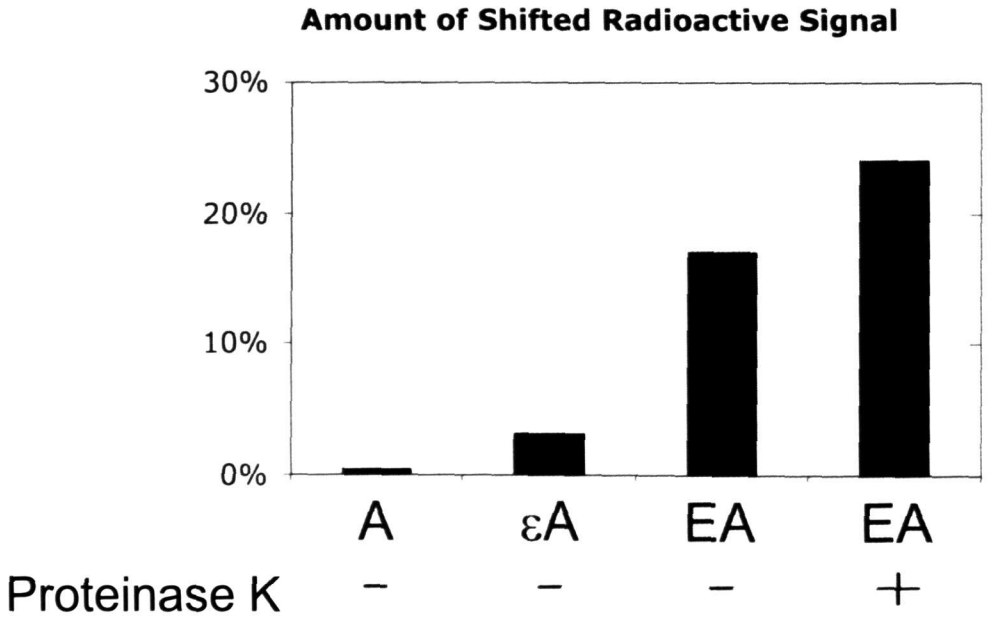
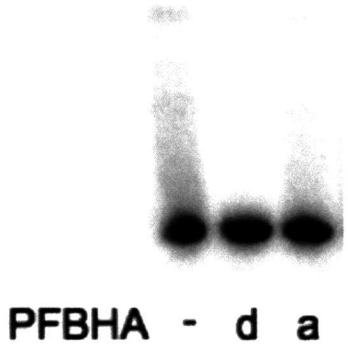


Figure 5.6: Cross-linking experiments with lesion-containing oligonucleotides annealed to the correctly base paired complement C1. Addition of a 1000-fold molar excess PFBHA during (d) the AlkB reaction abolishes cross-linking, while addition after (a) reduces cross-link formation.



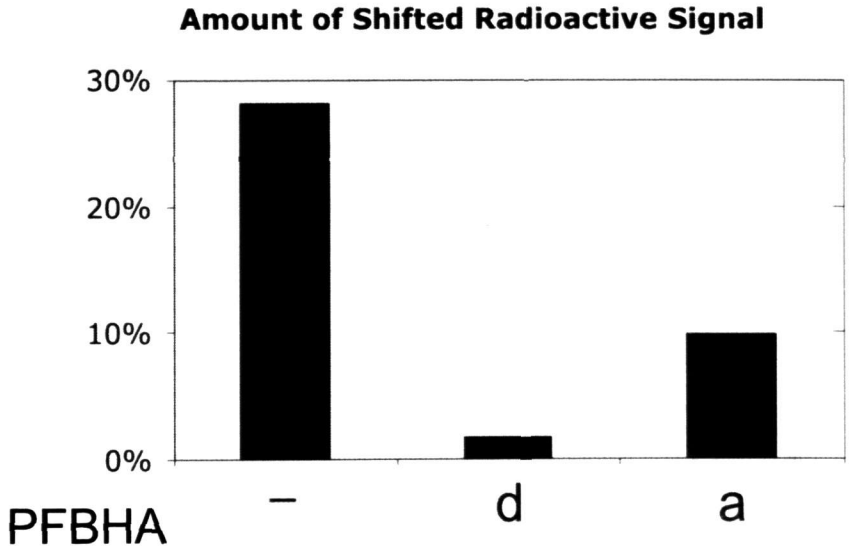
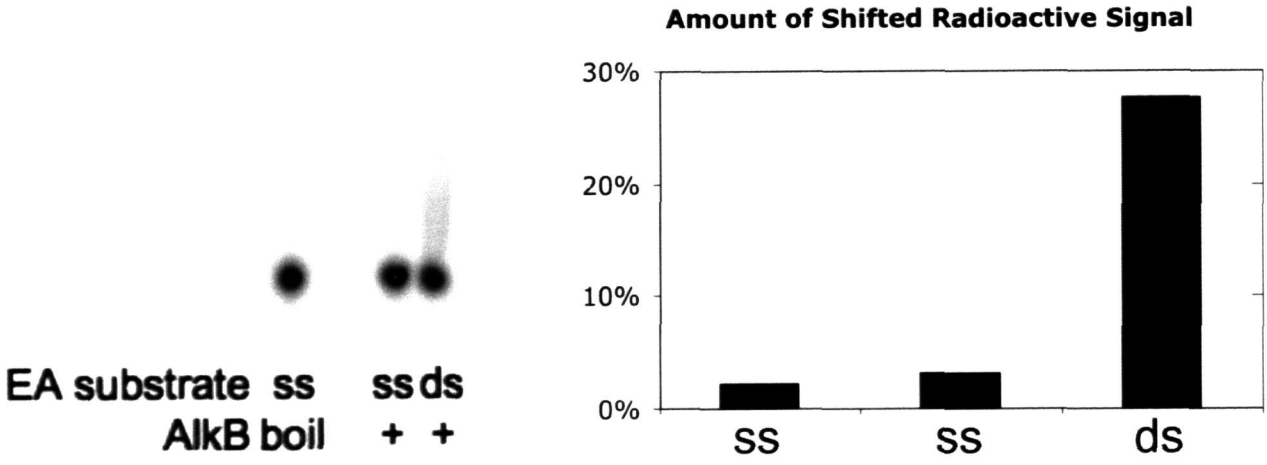


Figure 5.7: Cross-linking experiments with lesion-containing oligonucleotides annealed to the correctly base paired complement C1. Cross-linking requires a double-stranded substrate. Cross-linking is independent of which strand carries the radiolabel.



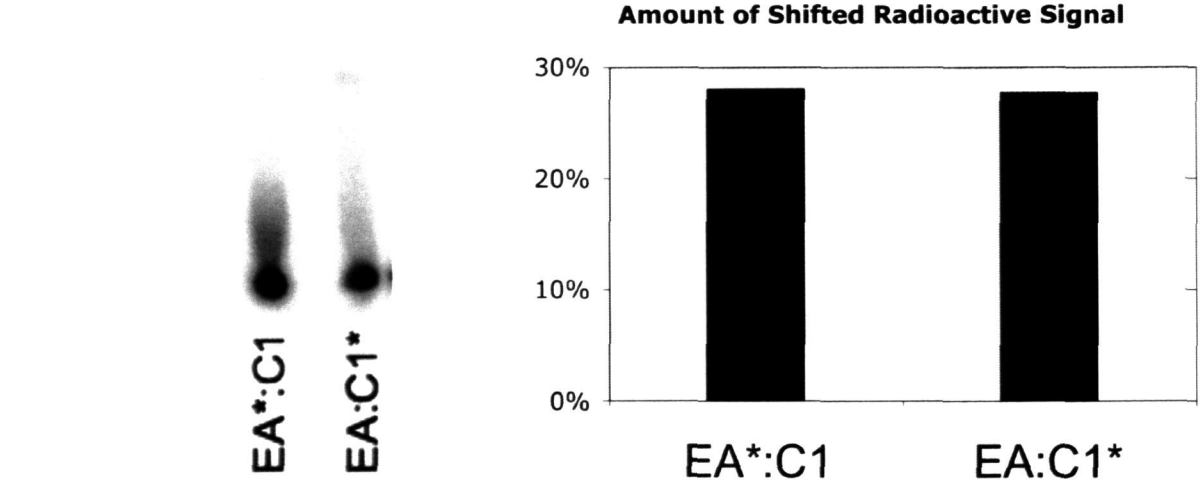


Figure 5.8: Cross-linking experiments with lesion-containing oligonucleotides annealed to mismatched radiolabeled complements C2 and C3 to allow for identification of specific bases critical to cross-link formation. Structures of the relevant portion of the EA lesion-containing strand and of the correctly-matched and singly mismatched complements to which it is annealed.

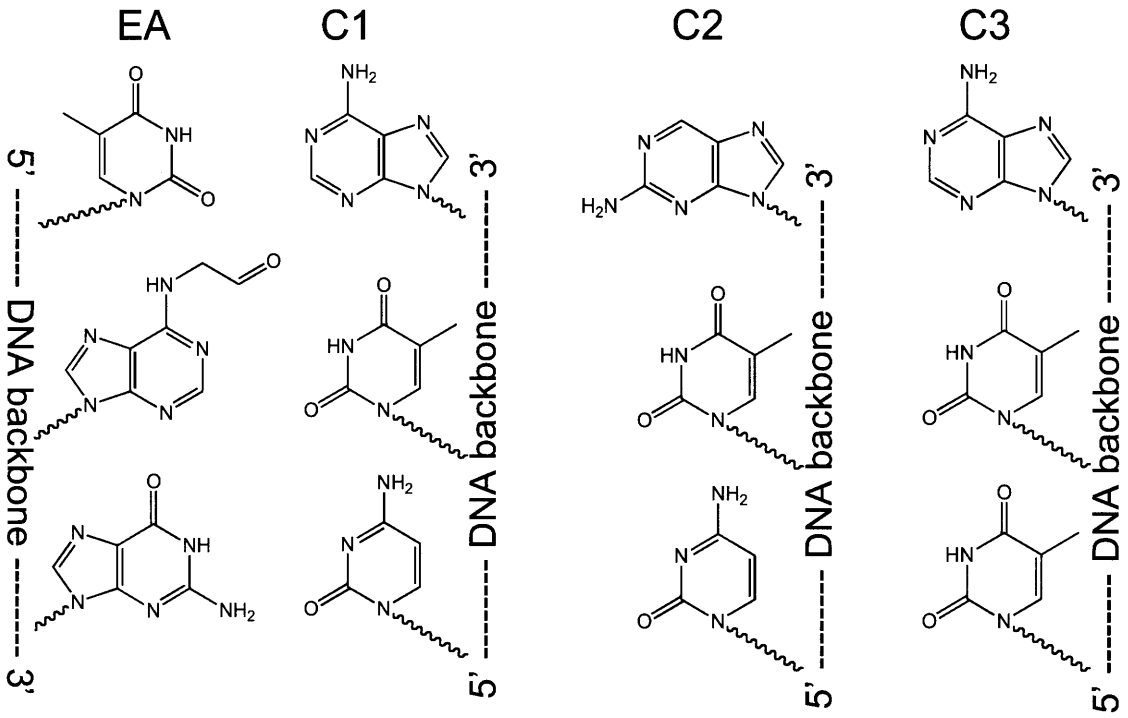
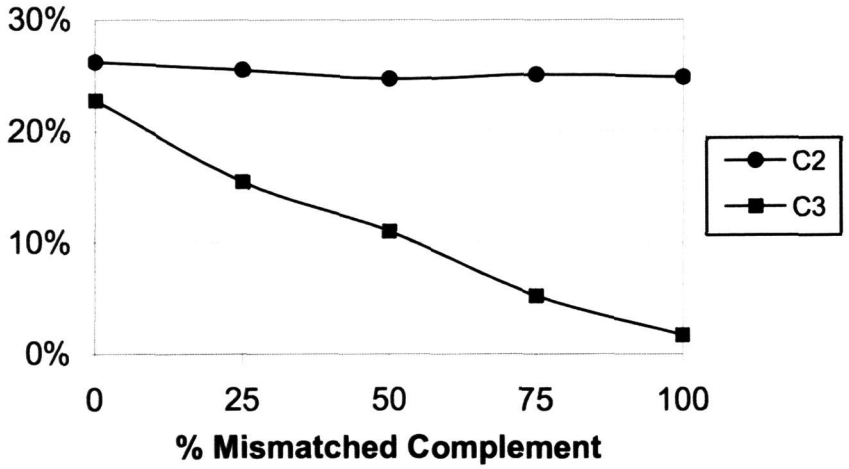


Figure 5.9: Replacing C1 with C2 does not affect the amount of cross-link formed, while replacing C1 with C3 reduces the observed cross-linking proportionally, indicating that the C one base 5' to the lesion site is involved in cross-linking. The putative structure of the cross-linked product is shown.



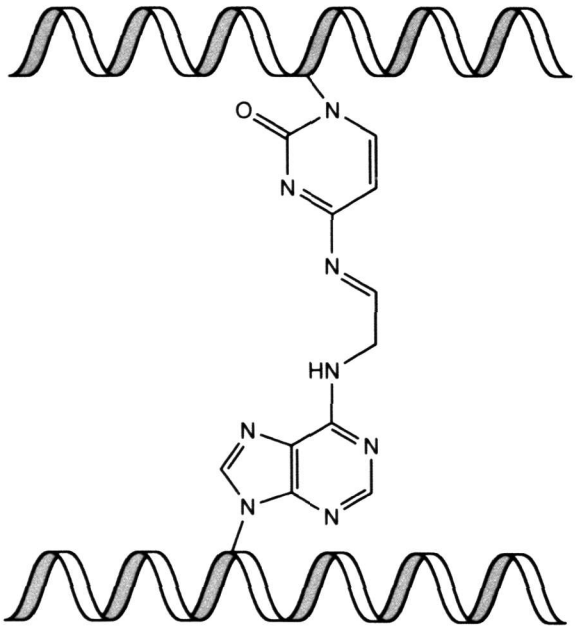


Figure 5.9: Molecular model showing the predicted positioning and orientation of the ring-opened aldehydic EA metabolite in duplex DNA. The slight 3' orientation positions the aldehyde well to react with the exocyclic amine group of the cytosine one base 5' in the opposite strand.

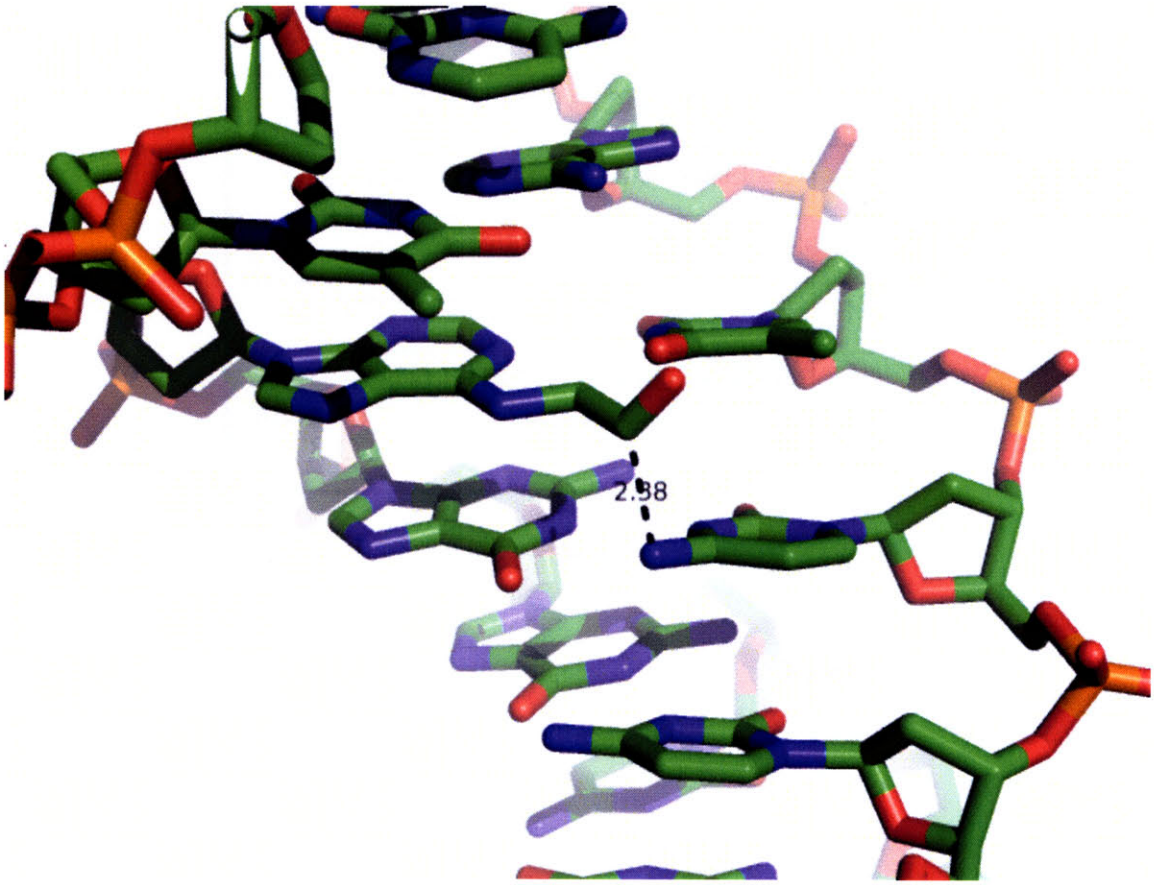


Table 5.1: Oligodeoxynucleotide substrates used for cross-linking experiments.

Lesion-Bearing Oligonucleotides	
A	5'- GAA GAC CT **A** GGC GTC C-3'
EA	5'- GAA GAC CT **EA** GGC GTC C-3'
εA	5'- GAA GAC CT **εA** GGC GTC C-3'
Complement Oligonucleotides	
C1	5'-G GAC GCC T AG GTC TTC-3'
C2	5'-G GAC GCC T **2AP**G GTC TTC-3'
C3	5'-G GAC GC**T** T AG GTC TTC-3'

References

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

2. Rohankhedkar, M. S., S. B. Mulrooney, W. J. Wedemeyer, and R. P. Hausinger. 2006. The AidB component of the Escherichia coli adaptive response to alkylating agents is a flavin-containing, DNA-binding protein. *J. Bacteriol.* **188**:223-230.

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

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

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

6. Falnes, P. O. and T. Rognes. 2003. DNA repair by bacterial AlkB proteins. *Res. Microbiol.* **154**:531-538.

7. Drablos, F., E. Feyzi, P. A. Aas, C. B. Vaagbo, B. Kavli, M. S. Bratlie, J. Pena-Diaz, M. Otterlei, G. Slupphaug, and H. E. Krokan. 2004. Alkylation damage in DNA and RNA--repair mechanisms and medical significance. *DNA Repair (Amst)* **3**:1389-1407.

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

9. Lee, D. H., S. G. Jin, S. Cai, Y. Chen, G. P. Pfeifer, and T. R. O'Connor. 2005. Repair of methylation damage in DNA and RNA by mammalian AlkB homologues. *J. Biol. Chem.* **280**:39448-39459.

10. Falnes, P. O. 2004. Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. *Nucleic Acids Res.* **32**:6260-6267.

11. Koivisto, P., P. Robins, T. Lindahl, and B. Sedgwick. 2004. Demethylation of 3-methylthymine in DNA by bacterial and human DNA dioxygenases. *J. Biol. Chem.* **279**:40470-40474.

12. Sedgwick, B. 2004. Repairing DNA-methylation damage. *Nat. Rev. Mol. Cell Biol.* **5**:148-157.

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

14. Falnes, P. O., M. Bjoras, P. A. Aas, O. Sundheim, and E. Seeberg. 2004. Substrate specificities of bacterial and human AlkB proteins. *Nucleic Acids Res.* **32**:3456-3461.

15. Koivisto, P., T. Duncan, T. Lindahl, and B. Sedgwick. 2003. Minimal methylated substrate and extended substrate range of Escherichia coli AlkB protein, a 1-methyladenine-DNA dioxygenase. *J. Biol. Chem.* **278**:44348-44354.

16. Ringvoll, J., L. M. Nordstrand, C. B. Vagbo, V. Talstad, K. Reite, P. A. Aas, K. H. Lauritzen, N. B. Liabakk, A. Bjork, R. W. Doughty, P. O. Falnes, H. E. Krokan, and A. Klungland. 2006. Repair deficient mice reveal mABH2 as the primary oxidative demethylase for repairing 1meA and 3meC lesions in DNA. *EMBO J.* **25**:2189-2198.

17. Delaney, J. C. and J. M. Essigmann. 2004. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A* **101**:14051-14056.

18. Frick, L. E., J. C. Delaney, C. Wong, C. L. Drennan, and J. M. Essigmann. 2007. Alleviation of 1,N6-ethanoadenine genotoxicity by the Escherichia coli adaptive response protein AlkB. *Proc. Natl. Acad. Sci. U. S. A* **104**:755-760.

19. Delaney, J. C., L. Smeester, C. Wong, L. E. Frick, K. Taghizadeh, J. S. Wishnok, C. L. Drennan, L. D. Samson, and J. M. Essigmann. 2005. AlkB reverses etheno DNA lesions caused by lipid oxidation in vitro and in vivo. *Nat. Struct. Mol. Biol.* **12**:855-860.

20. Mishina, Y., C. G. Yang, and C. He. 2005. Direct repair of the exocyclic DNA adduct 1,N6-ethenoadenine by the DNA repair AlkB proteins. *J. Am. Chem. Soc.* **127**:14594-14595.

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

22. Ougland, R., C. M. Zhang, A. Liiv, R. F. Johansen, E. Seeberg, Y. M. Hou, J. Remme, and P. O. Falnes. 2004. AlkB restores the biological function of mRNA and tRNA inactivated by chemical methylation. *Mol. Cell* **16**:107-116.

23. Yu, B., W. C. Edstrom, J. Benach, Y. Hamuro, P. C. Weber, B. R. Gibney, and J. F. Hunt. 2006. Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. *Nature* **439**:879-884.

24. Ludlum, D. B. 1986. Nature and biological significance of DNA modification by the haloethylnitrosoureas. *IARC Sci. Publ.*71-81.

25. Ludlum, D. B. 1986. Formation of cyclic adducts in nucleic acids by the haloethylnitrosoureas. *IARC Sci. Publ.*137-146.

26. Ludlum, D. B. 1990. DNA alkylation by the haloethylnitrosoureas: nature of modifications produced and their enzymatic repair or removal. *Mutat. Res.* **233**:117-126.

27. Guliaev, A. B., B. Hang, and B. Singer. 2002. Structural insights by molecular dynamics simulations into differential repair efficiency for ethano-A versus etheno-A adducts by the human alkylpurine-DNA N-glycosylase. *Nucleic Acids Res.* **30**:3778-3787.

28. Guliaev, A. B., B. Singer, and B. Hang. 2004. Chloroethylnitrosourea-derived ethano cytosine and adenine adducts are substrates for Escherichia coli glycosylases excising analogous etheno adducts. *DNA Repair (Amst)* **3**:1311-1321.

29. Toth, J. and M. D. Biggin. 2000. The specificity of protein-DNA crosslinking by formaldehyde: in vitro and in drosophila embryos. *Nucleic Acids Res.* **28**:e4.

30. Murata-Kamiya, N. and H. Kamiya. 2001. Methylglyoxal, an endogenous aldehyde, crosslinks DNA polymerase and the substrate DNA. *Nucleic Acids Res.* **29**:3433-3438.

31. Cho, Y. J., I. D. Kozekov, T. M. Harris, C. J. Rizzo, and M. P. Stone. 2007. Stereochemistry modulates the stability of reduced interstrand cross-links arising from R- and S-alpha-CH3-gamma-OH-1,N2-propano-2'-deoxyguanosine in the 5'-CpG-3' DNA sequence. *Biochemistry* **46**:2608-2621.

32. Cho, Y. J., H. Wang, I. D. Kozekov, A. J. Kurtz, J. Jacob, M. Voehler, J. Smith, T. M. Harris, R. S. Lloyd, C. J. Rizzo, and M. P. Stone. 2006. Stereospecific formation of interstrand carbinolamine DNA cross-links by crotonaldehyde- and acetaldehyde-derived alpha-CH3-gamma-OH-1,N2-propano-2'-deoxyguanosine adducts in the 5'-CpG-3' sequence. *Chem. Res. Toxicol.* **19**:195-208.

33. Dutta, S., G. Chowdhury, and K. S. Gates. 2007. Interstrand cross-links generated by abasic sites in duplex DNA. *J. Am. Chem. Soc.* **129**:1852-1853.

34. Kurtz, A. J. and R. S. Lloyd. 2003. 1,N2-deoxyguanosine adducts of acrolein, crotonaldehyde, and trans-4-hydroxynonenal cross-link to peptides via Schiff base linkage. *J. Biol Chem.* **278**:5970-5976.

35. Kozekov, I. D., L. V. Nechev, M. S. Moseley, C. M. Harris, C. J. Rizzo, M. P. Stone, and T. M. Harris. 2003. DNA interchain cross-links formed by acrolein and crotonaldehyde. *J. Am. Chem. Soc.* **125**:50-61.

36. Abe, M., M. Takahashi, K. Horiuchi, and A. Nagano. 2003. The changes in crosslink contents in tissues after formalin fixation. *Anal. Biochem.* **318**:118-123.

37. Dohno, C., A. Okamoto, and I. Saito. 2005. Stable, specific, and reversible base pairing via Schiff base. *J. Am. Chem. Soc.* **127**:16681-16684.

38. Friedberg E.C., W. G. C. S. W. W. R. D. S. R. A. E. T. 2006. DNA Repair and Mutagenesis. AMS Press, Washington, DC.

39. Ristow, H. and G. Obe. 1978. Acetaldehyde induces cross-links in DNA and causes sister-chromatid exchanges in human cells. *Mutat. Res.* **58**:115-119.

40. Borowy-Borowski, H., R. Lipman, and M. Tomasz. 1990. Recognition between mitomycin C and specific DNA sequences for cross-link formation. *Biochemistry* **29**:2999-3006.

41. Chu, G. 1994. Cellular responses to cisplatin. The roles of DNA-binding proteins and DNA repair. *J. Biol Chem.* **269**:787-790.

42. Eastman, A. 1987. The formation, isolation and characterization of DNA adducts produced by anticancer platinum complexes. *Pharmacol. Ther.* **34**:155-166.

43. Kohn, K. W., C. L. Spears, and P. Doty. 1966. Inter-strand crosslinking of DNA by nitrogen mustard. *J. Mol. Biol* **19**:266-288.

44. Chenna, A., H. Maruenda, and B. Singer. 1999. Synthesis of para-benzoquinone and 1,3-bis(2-chloroethyl)nitrosourea adducts and their incorporation into oligonucleotides. *IARC Sci. Publ.*89-101.

45. Loeber, R., M. Rajesh, Q. Fang, A. E. Pegg, and N. Tretyakova. 2006. Cross-linking of the human DNA repair protein O6-alkylguanine DNA alkyltransferase to DNA in the presence of 1,2,3,4-diepoxybutane. *Chem. Res. Toxicol.* **19**:645-654.

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

47. Li, Z., H. Mao, H. Y. Kim, P. J. Tamura, C. M. Harris, T. M. Harris, and M. P. Stone. 1999. Intercalation of the (-)-(1R,2S,3R, 4S)-N6-[1-benz[a]anthracenyl]-2'-

deoxyadenosyl adduct in an oligodeoxynucleotide containing the human N-ras codon 61 sequence. *Biochemistry* **38**:2969-2981.

48. Li, Z., H. Y. Kim, P. J. Tamura, C. M. Harris, T. M. Harris, and M. P. Stone. 1999. Intercalation of the (1S,2R,3S,4R)-N6-[1-(1,2,3,4-tetrahydro-2,3, 4-trihydroxybenz[a]anthracenyl)]-2'-deoxyadenosyl adduct in an oligodeoxynucleotide containing the human N-ras codon 61 sequence. *Biochemistry* **38**:16045-16057.

49. Scholdberg, T. A., L. V. Nechev, W. K. Merritt, T. M. Harris, C. M. Harris, R. S. Lloyd, and M. P. Stone. 2005. Mispairing of a site specific major groove (2S,3S)-N6-(2,3,4-trihydroxybutyl)-2'-deoxyadenosyl DNA Adduct of butadiene diol epoxide with deoxyguanosine: formation of a dA(anti).dG(anti) pairing interaction. *Chem. Res. Toxicol.* **18**:145-153.

50. Wang, L., B. E. Hingerty, R. Shapiro, and S. Broyde. 2004. Structural and stereoisomer effects of model estrogen quinone-derived DNA adducts: N6-(2-hydroxyestron-6(alpha,beta)-yl)-2'-deoxyadenosine and N2-(2-hydroxyestron-6(alpha,beta)-yl)-2'-deoxyguanosine. *Chem. Res. Toxicol.* **17**:311-324.

51. Wang, L., M. Wu, S. F. Yan, D. J. Patel, N. E. Geacintov, and S. Broyde. 2005. Accommodation of a 1S-(-)-benzo[c]phenanthrenyl-N6-dA adduct in the Y-family Dpo4 DNA polymerase active site: structural insights through molecular dynamics simulations. *Chem. Res. Toxicol.* **18**:441-456.

52. Reddy, M. C. and K. M. Vasquez. 2005. Repair of genome destabilizing lesions. *Radiat. Res.* **164**:345-356.

53. Eisenbrand, G., N. Muller, E. Denkel, and W. Sterzel. 1986. DNA adducts and DNA damage by antineoplastic and carcinogenic N-nitrosocompounds. *J. Cancer Res. Clin. Oncol.* **112**:196-204.

54. Wiencke, J. K. and J. Wiemels. 1995. Genotoxicity of 1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU). *Mutat. Res.* **339**:91-119.

CHAPTER 6:

AlkB Repairs 3-alkylated Pyrimidines in DNA *in Vivo*

Introduction

Nucleic acids within cells are constantly exposed to a range of compounds with which they can react, creating covalent chemical nucleic acid adducts. RNA and DNA show similar patterns of reactivity towards most alkylative damaging agents, although the spectrum formed by a given agent in single-stranded nucleic acids differs from that formed in a double-stranded context (*1*).

When it evades repair, 3-methylcytosine in DNA is a toxic and mutagenic adduct (*2-4*). Because the 3-position of cytosine is protected from reaction in duplexes due to its involvement in Watson-Crick hydrogen bonding, this lesion arises more readily in single-stranded than in double-stranded nucleic acids. The DNA repair protein AlkB allows cells to avoid the deleterious consequences of this adduct by reversing it directly, through its mechanism of oxidative demethylation (*5,6*). This direct reversal of base damage releases formaldehyde and the restored cytosine, allowing replication of the genome to proceed unhindered and unaltered.

Maintenance of the structural and informational integrity of the genome is a major commitment for all life forms. In addition to withstanding exposure to exogenous damaging agents, spontaneous reactions take place constantly that delete or alter the coding information of the genome (*7*). Hydrolytic depurination or depyrimidination of a nucleotide, in which the information-carrying base is lost and an AP (apurinic or apyrimidinic) site is created, results in complete loss of coding information. The deamination of normal amine-containing bases (cytosine → uracil or to a lesser extent adenine → hypoxanthine) occurs at a measurable rate in double-stranded DNA (*7*), and these events alter coding information, leading to mutations if the deaminated products are copied during replication. To combat the loss of information from deaminated cytosines, cells encode a uracil-DNA glycosylase that is responsible for excising uracil from DNA, initiating the base excision repair process, and allowing the intact, properly-coding sequence to be restored.

Although both of these processes (depurination and deamination) take place at significant rates under physiological conditions with the four common coding bases, modification of the bases, for example by alkylation, can cause increased susceptibility to

these hydrolytic reactions (*1*). For example, the half-life of deoxyguanosine in DNA decreases from 1.5×10^7 h to 155 h when it is methylated at the 7 position, a drop of 5 orders of magnitude. Similarly, cytosine methylated at the 3 position deaminates at a rate much faster than unmodified cytosine (*1,8-10*), and the increase in the deamination rate caused by the presence of this single methyl group has been estimated to be several thousand-fold under physiological conditions (*10*). Given that the rate of unmodified cytosine deamination to uracil is sufficient to warrant the existence of a dedicated repair enzyme (uracil DNA glycosylase, UNG, also called UDG or DGU, herein referred to as UDG for consistency with the nomenclature given to the purified enzyme by its manufacturer), it is reasonable to look for a repair mechanism for 3-methyluracil (m3U). Although it is produced from a rare species (m3C), the rapidity with which it is formed suggests that it would likely exist at a non-zero frequency in DNA under conditions of alkylation stress, and its removal from the genome would be advantageous, as it is likely a highly toxic and mutagenic lesion. While the presence of m3U in DNA has not been detected, this is not entirely surprising considering that it is expected to be a rare lesion and that repair could be taking place. Whether m3U is a substrate for the UDG enzyme that guards against the presence of uracil in DNA has not been explored, nor has there been any reported investigation into other methods of repair that could be active on this lesion.

While 3-methyluracil has not been isolated from DNA, its natural presence in RNA has been documented in mammalian, yeast, (*11*) and plant cells (*12*). This modified base could be either an intentional modification (*13*), playing a structural or regulatory role, or the result of alkylation damage, and probably occurs in both contexts *in vivo*. The reaction of RNA bases with various methylating agents creates patterns bearing strong resemblance to those seen in DNA, where 3-substituted cytosine and thymine occur at a low but biologically significant (at least in the case of 3-methylcytosine) level. Because RNA exists more often in a single-stranded form, the 3-positions of uracil and cytosine are unprotected by hydrogen bonding and are more solvent-exposed, and therefore susceptible to attack by a methylating agent, than they are in double-stranded nucleic acid, suggesting that perhaps 3-alkyl-damage is more

prevalent in cellular RNA than DNA. Reaction of various alkylating agents with poly(U) *in vitro* has shown that the 3-position is a favored site of reaction, accounting for up to 85% of observed modifications (*1*), and suggesting that m3U could form at consequential levels *in vivo*. Indeed, 3-methyluracil has been isolated from mRNA and rRNA from many sources (*11,12*).

RNA contains many natural modifications, including about 20 types of methylation at sites within specific sequences, some of which are 1-methyladenine and 3-methylcytosine (*13*). These altered bases are introduced enzymatically and are thought to play roles in controlling expression of genes at both the transcriptional and translational levels as well as in regulating interactions between RNA and other RNAs or proteins and in ensuring proper structure and folding of an RNA. The existence of a pool of m3C unrelated to alkylation damage within cells paired with the rapid deamination that the base undergoes suggests that m3U may be formed at some level constantly in cellular RNA even in the absence of alkylative stressors.

Although RNA methylation is exploited for various purposes, aberrant methylation can confer deleterious effects. In the case of 3-methylcytosine and 1-methyladenine, which are modified on their hydrogen-bonding faces, inhibited translation or mis-translation can result, creating truncated or non-functional proteins. Additionally, ribosomes can stall upon encountering such a lesion, causing sequestration of functional ribosomes and global inhibition of translation and protein production. The RNAs within the ribosomal complex are also subject to methylation damage, and a non-functional ribosome can carry with it the dominant negative consequences associated with ribosome sequestration as described above. Traditionally, RNA has been considered an expendable molecule, in which mistakes and damage were not fixed, but rather the molecule was destroyed and a replacement synthesized. It is possible, however, that the energetic cost of disassembling and degrading large RNA-containing molecules simply due to a minor diversion from normalcy is too high. Perhaps cells avoid the negative outcomes following RNA methylation through repair of the affected molecule.

Although the biological relevance of RNA repair by the AlkB family of proteins is unclear, that the process can take place has been demonstrated both *in vitro* and *in vivo*

(*14-23*). 3-Methylcytosine, 1-methyladenine, and 1-methylguanine have been isolated and identified by HPLC analysis of [^{3}H]-methylated RNA both before and after incubation with AlkB or hABH3 protein, and hABH3 has been shown to carry out the oxidative demethylation repair reaction in RNA just as efficiently as in DNA (*16*). Additionally, mRNAs functionally inactivated by global methylation have been restored by incubation with either the AlkB or hABH3 protein (*17*), and tRNA repair by AlkB was observed within bacterial cells. Further, [^{3}H]-methylated homopolymers (dA, dC, A, and C) were incubated with AlkB, hABH2, or hABH3 and repair, which was characterized as HPLC-detected radioactivity released as formaldehyde, was observed with both AlkB and hABH3 (*14*). Similarly, *E. coli* expressing either the AlkB or the hABH3 protein have been shown to reactivate methylated single-stranded RNA phage (*14*), indicating that repair of the damaged RNA is taking place within cells.

The case for RNA repair in general has been summarized recently (*16,19-21,23,24*). Because RNA is essential to the success of basic biological processes, and because a single damaged base can destroy an entire RNA molecule or even a large conglomerate of which RNA is just one part, such as a ribosome, the efficient repair of RNA would be of great energetic benefit to the cell. The contribution of RNA damage to cytotoxicity induced by chemotherapeutic agents has been established (*20*), and repair of such damage would be critical to avoiding cell death upon exposure to such agents. Other mechanisms of RNA repair beyond the action of AlkB have also been observed, including the removal of methyl groups from RNA bases by the DNA methyltransferases (*25*), the reversal of *cis-syn* U-U cyclobutane dimers in RNA by the *E. coli* DNA photolyase (*26*), and the restoration of truncated tRNA molecules (*27*). Consequently, although it has in the past largely been assumed that damaged RNA is merely degraded and a replacement molecule synthesized, this view is widening to consider the possibility that faulty RNA molecules are repaired much as their DNA counterparts, ensuring the fidelity and functionality of these molecules.

Although the majority of cellular m3U adducts probably reside in RNA, the presence of even just a few lesions in DNA could be biologically relevant. The genetic consequences of the existence of 3-methyluracil in DNA have never been examined, and

it would be worthwhile to determine whether the lesion is as toxic and mutagenic as the analogously modified 3-methylcytosine or 3-methylthymine. Additionally, the possible repair of this lesion when it does arise in DNA has not yet been investigated, and here we explore whether the adduct is a substrate for UDG *in vitro* or for AlkB *in vivo* and *in vitro*. Because AlkB-mediated repair of m3U would produce uracil in DNA, which would be removed by the UDG enzyme, possibly complicating analysis, *ung* derivatives of the AlkB+ and AlkB- *E. coli* strains were made by PCR-mediated gene replacement for use in the lesion bypass and mutagenesis assays. The replication efficiencies and coding specificities of seven modified or unmodified bases (Figure 6.1; U, m3U, e3U, m3C, e3C, m3T and T) were examined in these AlkB-proficient or –deficient *ung* mutants. 3-Methyluracil repair is further studied in RNA *in vitro* to determine whether AlkB is capable of removing this modified base in its presumably more common conformation *in vivo*.

Materials and Methods

Oligonucleotides. Oligodeoxynucleotides of sequence 5'-GAAGACCTXGGCGTCC-3' (where X = U, m3C, or e3C) were obtained from Integrated DNA Technologies (uracil-containing-16-mer; Coralville, IA) or Chemgenes (m3C and e3C 16-mers; Wilmington, MA). Fractions of the m3C and e3C oligodeoxynucleotides (30 nmol in 200 μL of a 10 mM Na_3PO_4 solution (pH ~12)) were converted to the deaminated equivalents (m3U and e3U) by incubation at 80 °C for 5 h, after which the DNA was purified by anion exchange HPLC and desalted, followed by desalting on a SepPak cartridge. Under the conditions used, the retention times of the methylated and ethylated species were nearly identical. Starting material and conversion product were well-resolved, with a retention time of 19.5 min for the 16-mer containing 3-(m)ethyldeoxycytidine, while that of the conversion product was 22.0 min. Approximately 90% conversion was obtained for both species as judged by the UV absorbance from the HPLC peak integration values with conversion of the methyl-containing 16-mer taking place more rapidly than the ethyl oligodeoxynucleotide. The RNA oligonucleotide of sequence 5'-AAAAAAAAXAA AAAAA-3' (where X = m3U) was obtained from Midland Certified Reagent Company

(Midland, TX) and synthesized using phosphoramidite from Chemgenes (Wilmington, MA).

Bacterial Strains. The *E. coli* strains used for construction of ung-deficient derivatives were HK81 (as AB1157, but *nalA*; wild-type) and HK82 (as AB1157, but *nalA alkB22*; AlkB-deficient). AlkB status of these strains was confirmed by PCR amplification and sequencing of the region encoding the *alkB* gene. The lack of the UDG coding sequence in these derived strains (HK81 ung$^+$ and HK82 ung$^-$) was also verified by PCR amplification and sequencing of the corresponding region. These ung-deficient strains were then used for transfection and replication of lesion-containing viral genomes for bypass and mutagenesis analysis. Additionally, NR9050 *E. coli* were used to measure transfection efficiency and SCS110 cells (Stratagene, La Jolla, CA) were used for regrowth of progeny phage.

UDG Assay. Uracil-DNA Glycosylase (UDG) was obtained from New England Biolabs (Ipswich, MA). Purified oligonucleotides containing a single uracil or 3-methyluracil residue were radiolabeled at the 5' end using [γ-^{32}P]ATP and T4 polynucleotide kinase (USB, Cleveland, OH). After removal of unincorporated nucleotides by passage of the labeled oligonucleotides through Micro Bio-Spin™ Chromatography Columns (BioRad, Hercules, CA) packed with Bio-Gel P-6, the eluted DNA was either used without further processing or annealed to the 16-mer complement containing an A opposite the lesion site by incubation with a 1.25-fold excess of complement at 90 °C for 3 min followed by bench-top cooling to room temperature. DNA was divided into 1 pmol aliquots, which were then digested with UDG (0.5 U in 10 uL 1X NEBuffer 3) at 37 °C for 30 min followed by cleavage of the resulting AP site by the addition of 1 U EndoIV (New England Biolabs, Ipswich, MA) and incubation at 37 °C for 1 h. Reaction products were separated on a 20% denaturing PAGE gel and visualized using phosphorimagery.

Creation of *ung* Strains by PCR-mediated Gene Replacement.

Isolation of E. coli Genomic DNA. Genomic DNA was isolated from saturated overnight bacterial cultures basically as described in Current Protocols (*28*). Ten mL cultures of each strain were grown with aeration overnight (to saturation) at 37 °C. Cells were isolated from 1 mL of the overnight culture by centrifugation for 3 min at 15,000 x g. After resuspension of the cell pellet in 570 μL TE buffer (10 mM Tris, 1mM EDTA, pH 8.0), 30 μL 10% SDS and 3 μL 20 mg/mL Proteinase K were added. The mixture was then incubated at 37 °C for 1 hour to release and digest cellular contents, following which 100 μL 5 M sodium chloride and 80 μL of a 10% cetyltrimethylammonium chloride/0.7 M sodium chloride solution were added. After a 10 minute incubation at 65°C, 750 μL phenol/chloroform/isoamyl alcohol (25:24:1, PCI) were added, and the sample was mixed well and centrifuged for 5 min at 15,000 x g. The 750 μL aqueous layer was transferred to a new tube and the PCI extraction was repeated. The 700 μL aqueous layer of the second extraction was transferred to a fresh tube, and a third extraction using 420 μL isopropanol was conducted. Following removal of the supernatant, a final extraction using 250 μL 70% ethanol was carried out, and the supernatant was again removed. Finally, the isolated DNA was dried under vacuum, resuspended in 100 μL H_2O, and quantified by absorbance at 260 nm. Typical yields were ~40 μg.

Creation of Recombination Substrates. The region containing the *Tn10* marker, which encodes the TetR_N and TetR_C genes necessary to achieve tetracycline resistence, was amplified from genomic DNA isolated from GM5555 using PCR primers that contained both *Tn10* and ung sequences to create a product containing the Tn10 gene flanked on both sides by portions of the ung gene (Figure 6.4). Amplification was carried out in 50 μL containing 10 μM each primer, 10 mM dNTPs, 200 ng genomic DNA, 2% DMSO and 2.5 U *Taq* DNA polymerase (NEB, Ipswich, MA) in 10 mM Tris-HCl pH 9.2, 5 nM EDTA, 3.5 mM $MgCl_2$, and 25 mM KCl. PCR cycling parameters were: 94 °C for 3 min followed by thirty sets of 94 °C for 1 min, 55 °C for 1 min, 72 °C for 1.5 min, ending

with 72 °C for 8 min. Separation on a 1% agarose gel revealed a single 2 kb product that was purified with a QIAquick PCR Purification Kit (Qiagen) and verified by sequencing.

Transfection of Hyper-recombinant Strain. Saturated overnight cultures of KM22 cells (a gift of Peter T. Rye) were diluted 1:100 in 10 mL LB medium and supplemented with 1 mM IPTG. Diluted cultures were grown with aeration at 37 °C to mid-log phase (A_{600} = 0.6), at which point cells were pelletted by centrifugation. After two washes with 15 mL cold H_2O, cells were resuspended in 75 μL 1 mM MOPS/20% glycerol. One hundred microliters of cells were mixed with 500 ng purified recombination substrate in a pre-chilled electroporation cuvette and exposed to a 2.5 kV 129 Ohm pulse. Treated cells were immediately transferred to culture tubes containing 3 mL LB supplemented with 1 mM IPTG and allowed to recover on a roller drum at 37 °C for 4 hours. Surviving cells were pelleted by centrifugation and resuspended in 100 μL LB before plating on LB plates containing 4 μg/mL tetracycline. After overnight incubation at 37 °C, isolated colonies representing successful transductants were picked and grown overnight in LB containing 12 μg/mL tetracycline. Genomic DNA was then isolated as above and the region of interest amplified by PCR using two sets of ung-specific primers (Table 6.1) and sequenced to verify correct genotype.

Transfer of Genotype via P1 Transduction. KM22 cells of the proper genotype were grown overnight in LB containing 12 μg/mL tetracycline, diluted 1:100 into 5 mL LB containing 5 mM $CaCl_2$, and grown until A_{600} = 0.5. One hundred microliters of cells were then mixed with 10 μL of each of a series of dilutions of P1 virus (10^0 to 10^{-5} of the 10^9 pfu/uL stock, obtained from Peter T. Rye) or 10 μL LB as a control and incubated at 37 °C for 20 min. Infected cells were then mixed with 3.5 mL 52 °C R-top agar (10 g Bacto tryptone, 1 g Bacto yeast extract, 8 g NaCl, 8 g Difco agar per L, supplemented with 2 mL 1 M $CaCl_2$ and 5 mL 20 % glucose after autoclaving) and plated on R plates (as R-top but 12 g Difco agar) before overnight incubation at 37 °C. Virus was isolated from the top layer of the plate bearing the largest number of still-separate plaques by scraping off the soft agar layer with a cell lifter and transferring it into a 15 mL tube. The

plate was washed with two 1 mL volumes of LB to gather any residual virus, which was added to the agar-containing tube. Three hundred μL chloroform were added and the tube vortexed vigorously for 1 min before centrifugation at 6.5 krpm for 10 min to isolate the virus-containing supernatant, which was removed, supplemented with 50 μL chloroform, and stored at 4 °C in a sterile 2 mL glass vial.

Recipient cells (HK81 and HK82) were diluted 1:100 from a saturated overnight culture into 10 mL LB and grown to mid-log phase (A_{600} ~ 0.5), when cultures were supplemented with 500 μL of a 1 M $MgCl_2$/50 mM $CaCl_2$ solution, vortexed, and incubated at room temperature for 20 min. The isolated virus was diluted serially (10^0 to 10^{-5}) in LB medium and 100 μL cells were mixed with 10 μL each dilution or LB (non-viral control). Following incubation at 37 °C for 20 min, 100 μL 250 mM sodium citrate were added to stop the infection and cells were allowed to recover at 37 °C for 3 h. All ~210 μL were then plated on LB plates containing 4 μg/mL tetracycline and incubated at 37 °C overnight. Isolated colonies of each strain were grown overnight in 12 μg/mL tetracycline before extraction of genomic DNA, PCR amplification of the ung gene using two different primer sets, and sequencing of the PCR products to verify the presence of the interrupted gene. As all sequenced successful transductants possessed the correct genotype, one was arbitrarily chosen as the strain used in all subsequent experiments.

Phage Genome Construction. Single-stranded M13 phage genomes containing an oligonucleotide insert bearing a single, site-specific lesion (Figure 6.1) were constructed as previously described (*29*). For each lesion-containing genome to be constructed, 20 pmol of wild-type M13mp7L2 genome were digested with EcoRI, which cleaves at a unique restriction site in the hairpin structure, resulting in linearized genome. Two oligonucleotides are then annealed as scaffolds, one that hybridizes to the 5' end of the digested phage DNA and the 3' end of the 16-mer lesion-containing insert, and one that bridges the 3' end of the genome and the 5' end of the insert. Simultaneously, the 16-mer oligonucleotide inserts containing U, m3U, e3U, m3C, e3C, m3T, or T at the lesion site were phosphorylated using T4 polynucleotide kinase. Following phosphorylation, the kinased inserts were added to the pre-annealed scaffold-genome mixture and ligated

in place by T4 DNA ligase, thus restoring viable circular phage genome. The scaffolds were then digested with T4 DNA polymerase, and genomes were extracted with phenol:chloroform:isoamyl alcohol (25:24:1), desalted on Sephadex G50 fine resin, and stored at -20 °C.

Genome Normalization. The concentrations of the individually constructed genomes were normalized to the most dilute by first acquiring approximate concentrations based on A_{260} as measured by UV-visible spectroscopy. Roughly 0.35 pmol of each genome were annealed to the same scaffold oligonucleotides used in genome construction to create a double-stranded region that was digested with HinFI. The resulting fragment was dephosphorylated with shrimp alkaline phosphatase (Roche, Basel, Switzerland) and 5'-^{32}P-radiolabeled using T4 polynucleotide kinase (USB, Cleveland, Ohio) and [γ-^{32}P]ATP. Following a final digestion with HaeIII (New England Biolabs, Ipswich, MA), products were separated by denaturing PAGE and the appropriate band intensities were quantified using phosphorimagery (GE Healthcare, Piscataway, NJ). Water was then added to the more concentrated genomes such that the final concentration of all genomes was the same.

Electrocompetent Cell Preparation. 150 mL LB medium were added to each of three baffled flasks and inoculated with 1.5 mL of a saturated overnight culture of the strain to be transformed (ung-deficient HK81 (AlkB-proficient) or ung-deficient HK82 (AlkB-deficient)). Cultures were grown with shaking at 37°C for 2.5 h, (to an OD_{600} of ~0.5). The three cell preparations were then pelleted by centrifugation, pooled, resuspended in 1 mL and washed three times with 175 mL cold sterile H_2O. The final resuspension was in 4 mL 10% glycerol, following which the cells were stored at 4 °C and used for transfection within 16 h.

Assay of Lesion Bypass Efficiency. Prior to electroporation into cells, normalized genomes were mixed with a known amount of a competitor genome that serves as an internal standard for replication. The competitor contains an insert that is three

nucleotides longer than the lesion-containing insert (19 nucleotides instead of 16) and that contains a thymine nucleobase at the lesion site. After mixing genome with competitor in a 9:1 ratio, ~0.3 pmol genome in 4.18 μL were combined with 100 μL electrocompetent cells in a 2 mm-gap electroporation cuvette on ice. Immediately after application of approximately 2.5 kV and 125 Ω, the cells were transferred to a culture tube containing 10 mL LB medium, 10 μL of which were plated on a lawn of NR9050 indicator bacteria to quantify the number of independent initial events, each of which is visualized by the formation of a plaque. In all cases reported here, the electroporation procedure resulted in $>10^6$ initial events, indicating statistical robustness. The 10 mL cultures were grown on a roller drum at 37 °C for 6 h. Cultures were then pelleted, and the progeny phage-containing supernatant was removed and stored at 4 °C.

During the electroporation procedure, a small number of copies of genomes are not incorporated into cells, but remain in the external medium. To prevent analysis of these un-replicated genomes, 100 μL of the phage-containing supernatant were mixed with 10 μL of a saturated overnight culture of SCS110 cells in 10 mL LB medium and grown for 7 h on a roller drum at 37 °C. In this step, only progeny phage capable of infecting SCS110 cells will be amplified, thus reducing by several orders of magnitude the aberrant signal that might come from residual unreplicated genome. Following regrowth, cultures were pelleted and the supernatant containing progeny phage was isolated, titered, and stored at 4 °C. Regrowths typically contained 10^{11} pfu/mL for analysis.

Single-stranded DNA was isolated from 0.7 mL progeny phage using a QIAprep Spin M13 kit (Qiagen, Valencia, CA) and the region of interest was amplified by PCR using primers (5'-YCAGCTATGACCATGATTCAGTGGAAGAC-3' and 5'-YCAGGGTTTTCCCAGTCACGAVGTTGTAA-3', where Y indicates amino modification of the 5' nucleotide) as described (*30*). Purified DNA was isolated by extraction with phenol/chloroform/isoamyl alcohol (25:24:1) followed by desalting with Sephadex G50 fine resin. Isolated PCR product was digested with BbsI and dephosphorylated with shrimp alkaline phosphatase, and then radiolabeled by T4 polynucleotide kinase with [γ-^{32}P] ATP. Following digestion with HaeIII, the products

were separated by denaturing PAGE and visualized by phosphorimagery. Band intensities were quantified using ImageQuant software, and lesion bypass was measured by comparison of the 18-mer band intensity (lesion signal) to the 21-mer intensity (competitor signal).

Mutational Spectrum Analysis. PCR amplification is carried out as in the lesion bypass assay described above, but with a different primer set (5'-YCAGCTATGA CCATGATTCAGTGGAAGAC-3' and 5'-YTGTAAAACGACGGCCAGTGAATTGG ACG-3'), which allows for specific amplification of only the phage DNA derived from lesion-containing genomes (i.e. the second primer is mismatched when annealed to the competitor, thus preventing amplification of that DNA). PCR products are phenol/chloroform/isoamyl alcohol extracted, desalted, BbsI digested, shrimp alkaline-dephosphorylated, 5'-radiolabeled, and PAGE separated all as in the lesion bypass assay. Following visualization of the gels by phosphorimagery, the 18-mer bands were excised from the gel, and the DNA fragments contained therein were extracted by crushing and soaking overnight in H_2O. Recovered DNA was desalted with Sephadex G50 fine resin and digested to single nucleotide monophosphates by nuclease P1 (USB). The system was designed such that cleavage of the PCR product with BbsI would expose at the 5' end the site that originally contained the lesion. The radiolabeled nucleotide monophosphates, then, correspond to the base composition generated by replication of the lesion. The digested nucleotides were separated on a polyethyleneimine thin layer chromatograpy plate (J. T. Baker) and the resulting four spots, each of which represents a normal base, were visualized by phosphorimagery. Quantification of the spots allowed measurement of the mutational frequency and specificity of the lesion studied.

Construction and Purification of AlkBΔN11. The AlkB open reading frame, excluding the first 11 amino acids, was PCR-amplified using the oligonucleotides 5'-GGAATTCCATATGCAAGAGCCACTGGCGG (*Nde*I restriction site), and 5'-CCGCTCGAGGCCTTGAAAATATAGGTTTTCTTTTTTACCTGCCTG (*Xho*I site). The PCR product was purified, digested with *Nde*I and *Xho*I (NEB), and ligated into a C-

terminal-histidine-tag-containing pET24a vector (Novagen) that had been digested by *Nde*I and *Xho*I. The ligation mixture was transformed into DH5α cells for verification by sequencing. pET24a-AlkBΔN11 was transformed into BL21(DE3) (Novagen) cells for expression.

pET24a-AlkBΔN11 cells were grown at 37 °C to an A_{600} of 0.4, when isopropyl-β-D-thiogalactoside (IPTG) was added at 1 mM and the temperature was lowered to 30 °C. Cells were harvested after 4 h by centrifugation, resuspended in 10 mM Tris pH 7.3, 300 mM NaCl, 2 mM $CaCl_2$, 10 mM $MgCl_2$, 5% glycerol and 1 mM BME (lysis buffer), and homogenized by sonication. Lysate was recovered by centrifugation. The supernatant was loaded onto an Ni-NTA column (Qiagen), washed (20 mM) and eluted (70 and 250 mM) in imidazole in lysis buffer. The eluent was dialyzed against 50 mM TES pH 7.1 and loaded onto an SP Sepharose cation exchange column (Amersham Pharmacia). AlkBΔN11 was eluted by a linear gradient of 0-1 M NaCl over 75 mL. The fractions containing AlkBΔN11 were pooled and purity was established by SDS-PAGE.

***In vitro* AlkB Repair Reactions.** Two hundred pmol of RNA oligonucleotide of sequence 5'-AAAAAAAAXAAAAAAA-3' (where X = m3U) or 25 pmol of DNA oligodeoxynucleotide of sequence 5'-GAAGACCTXGGCGTCC-3' (where X = m3C or m3U) were incubated with ~25 pmol purified AlkBΔN11 in 50 μL buffer containing 70 μM $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$, 0.9 mM α-ketoglutarate, 1.8 mM L-ascorbate, and 46.5 mM HEPES (pH 8.0) at 37 °C for 2 h. Following reaction, oligonucleotides were analyzed by electrospray ionization time of flight (ESI-TOF) mass spectrometry in negative ion mode (capillary voltage was –3000 V, nebulizing gas pressure was 35 psi, drying gas flow rate was 12 L min^{-1} at 325°C). Samples were introduced through an Agilent Zorbax C-18 column at a rate of 300 μL min^{-1} using a gradient of 0-50% B over 20 min followed by 50-80% B over 5 min, followed by 80-0% B over 5 min, where A = 10 mM ammonium acetate and B = 100% acetonitrile.

Results

UDG Assay. Analysis of cleavage products created by the sequential action of UDG and EndoIV on uracil- or 3-methyluracil-containing oligonucleotides demonstrates that m3U is not an efficient substrate for the UDG enzyme. Lanes 1-4 of Figure 6.2 correspond to single-stranded uracil-containing substrates and demonstrate that the AP site created by the removal of uracil by UDG is subject to a low level of decomposition, regardless of the presence of EndoIV (as expected, as EndoIV does not cleave single-stranded substrates), as demonstrated by the appearance of a faint band that runs as an 8-mer (lanes 2 and 3). The presence of EndoIV alone causes the appearance of a second band that runs just slightly faster than the 16-mer parent band (lanes 4, 8, and 12), the cause of which is unclear. Lanes 5-8, corresponding to the single-stranded m3U substrate, show no 8-mer products, indicating that m3U is not being removed in this context and the cleavable AP site is therefore not being formed. Lanes 9-12 depict double-stranded U substrates. Lane 10, which contains only UDG, shows the same spontaneous AP site decomposition observed in the equivalent single-stranded experiments. Lane 11, which represents the sequential reaction with UDG and EndoIV, shows almost complete cleavage of substrate by EndoIV, indicating that uracil was removed with ~100% efficiency. Lanes 13-16 correspond to the double-stranded m3U substrates. The lack of background band in lane 14 and the minimal cleavage product observed in lane 15 indicate that m3U is not an efficient substrate for UDG, as a spontaneously or enzymatically cleavable AP site is not being formed in this substrate.

Quantification of cleavage products was performed using ImageQuant software, and results are shown in Figure 6.3. The top panel represents quantification of only the intact and centrally-cleaved substrates (i.e, the 16- and 8-mer bands, ignoring the unidentified nicked substrate). Detectable cleavage was observed in lane 2 (ssU UDG+ EndoIV-; 9.7%), lane 3 (ssU UDG+ EndoIV+; 8.2%), lane 10 (dsU UDG+ EndoIV-; 13.2%), lane 11 (dsU UDG+ EndoIV+; 99.7%) and lane 15 (dsm3U UDG+ EndoIV+; 1.8%). Including the nicked band in the quantification generated the bottom panel in Figure 6.3 and affected the results only of lanes in which incubation was with EndoIV, but not UDG (lanes 4, 8, 12, and 16). This unidentified nicked product accounted for

8.2% 19%, 9.2% and 4.5% of the total radioactivity of lanes 4, 8, 12, and 16 respectively. In none of these lanes was cleaved product (8-mer) observed.

***In Vivo* Lesion Bypass Efficiencies of 3-alkylated Pyrimidines.** The seven bases studied here show different patterns of polymerase bypass in AlkB-proficient cells as compared to AlkB-deficient cells, both of which were rendered UDG-deficient by PCR-mediated gene replacement (Figure 6.4). The presence of uracil in DNA within cells appears to be well-tolerated (Figure 6.5), as the uracil-containing genome is replicated just as well as a control T in AlkB+ cells, and is fairly non-toxic in AlkB- cells as well, with an observed bypass frequency of ~90% that of T. In contrast, m3U and e3U are both severely toxic in AlkB-deficient cells, allowing bypass less than 10% of the time. A mild alleviation of that toxicity is seen when the m3U lesion is replicated in cells possessing functional AlkB, however, indicating that the protein is capable of repairing this methylated base, at least to some extent. The same alleviation of toxicity is not seen with the ethylated uracil. Similarly to m3U and e3U, the m3C and e3C lesions, which are modified in the same relative position, show potent toxicity in AlkB-deficient cells. Both of these lesions are good substrates for the AlkB protein, however, as demonstrated by the large increases in bypass seen in wild-type cells. The ~6% bypass efficiency of m3C increases to approximately 90% when protein function is intact, while alleviation of the extreme toxicity of e3C in mutant cells (~4% bypass) is slightly smaller, increasing to ~80% in repair-proficient cells. The m3T lesion shows similar toxicity in repair-deficient cells to the methylated and ethylated U and C adducts, in the range of 7%, with a slight increase observed in AlkB+ cells (~11%), demonstrating that AlkB can repair m3T, but that this lesion is a relatively poor substrate for the protein. These data agree with previous studies (*2*) done in AlkB+ and AlkB- strains that contained functional UDG.

Mutational Frequency and Specificity of 3-alkylated Pyrimidines. Analysis of the base composition present at the lesion site after genome replication in AlkB+ or AlkB- cells allows calculation of a mutation frequency and generation of a mutational spectrum for each lesion. In addition to its lack of toxicity in either context, U demonstrates a

similar lack of miscoding potential regardless of the AlkB status of the cells in which it is replicated, as 100% T is observed at the lesion site in both strains (Figure 6.6A). In contrast, the alkylated derivatives prove to be potently mutagenic (Figures 6.6B and 6.6C). m3U generates a spectrum of 6% C, 36% T, 2% G, and 56% A in repair-deficient cells, which becomes 2% C, 69% T, 2% G, and 28% A in cells which possess functional AlkB, supporting the finding that AlkB is able to repair the lesion *in vivo,* thus mitigating, although not eliminating, its toxic and mutagenic characteristics. e3U shows a similar mutational spectrum to m3U in *alkB* cells (13% C, 37% T, 2% G, 48% A), but the change induced by the presence of the AlkB protein is less pronounced in this case (9% C, 44% T, 2% G, 44% A), further demonstrating that e3U is a poorer substrate for AlkB than is m3U. The mutational signatures of m3C and e3C also resemble one another (Figures 6.6D and 6.6E), but in this case, both change dramatically with AlkB status, underscoring the ease with which AlkB can repair these adducts. The potently mutagenic m3C adduct (18% C, 40% T, 8% G, 35% A in AlkB- cells) is rendered virtually non-mutagenic when replicated in AlkB+ cells (99% C, <1% each T, G, and A). A similar scenario is seen with e3C, as the 27% C, 46% T, 7% G, 20% A signature in AlkB- cells becomes >98% C, <1% each T and G, 1% A when the protein is present during replication. The m3T lesion generates 7% C, 46% T, 1% G, 46% A when replicated in repair-deficient cells (Figure 6.6F), and the slight increase in T seen in repair-proficient cells (4% C, 51% T, 2% G, 43% A) demonstrates that the m3T adduct is a weak substrate for the AlkB protein. As expected, the control genome containing T at the lesion site produces 100% T at that site after replication in either strain (Figure 6.6G).

***In vitro* Repair Assay.** Although m3C in DNA was completely repaired by AlkB in the 2 h reaction time allowed, repair was not detected for m3U in DNA or in RNA (Figure 6.7).

Discussion

The alkylation damage product 3-methylcytosine occurs largely in regions of single-stranded DNA at times of alkylation stress and is markedly more susceptible to

deamination than is normal cytosine. The result of deamination of m3C would be the generation of m3U in DNA. Cells encode a repair activity for the occurrence of uracil in DNA, however none has been established for m3U. The possibility that the protein responsible for the removal of U, UDG, could also be capable of removing m3U was examined and found to not be the case. Incubation of 16-mer oligodeoxynucleotides containing either uracil or m3U in either a single- or double-stranded form with UDG and EndoIV generated three distinct bands when products were separated by denaturing PAGE, a 16-mer band corresponding to the intact substrate, an 8-mer band representing substrate that had been cleaved at the centrally-located lesion site, and a band running slightly faster than the intact substrate, designated the nicked band, that has not been identified. The appearance of this latter product appeared only upon incubation with EndoIV alone. Although EndoIV has a known 3'-diesterase activity that can cause the release of phosphoglycoaldehyde, deoxyribose 5-phosphate and phosphate from the 3' terminus of DNA (NEB technical literature), none of these reactions should take place with the 3'-hydroxylated oligonucleotide employed in this study. Importantly, while UDG removes uracil from double-stranded DNA almost 100% of the time (99.7% cleavage by EndoIV was observed), the protein generates a cleavable substrate from a m3U-containing substrate only 1.8% of the time. Indeed, it seems that the base excision repair process initiated by UDG is not a viable pathway through which cells can protect themselves from the deleterious consequences of m3U, should it arise in DNA.

The ability of a cell to replicate past each of the seven bases shown in Figure 6.1 was examined in UDG-deficient *E. coli* either containing or lacking functional AlkB. The efficiencies of replication of the non-canonical bases (U, m3U, e3U, m3C, e3C, and m3T) were compared with that past T, and results expressed as a relative percentage. Uracil was replicated as well as T in AlkB+ cells and with an efficiency just slightly lower than that of T in cells lacking AlkB (Figure 6.5). These results indicate that *E. coli* tolerate the presence of U in DNA quite well, as these cells are lacking the repair system to remove this abnormal base, and yet no great toxicity is observed. Additionally, the replication of uracil in DNA is not mutagenic, as 100% thymine was found at the lesion site after processing within cells regardless of the presence of AlkB (Figure 6.6A). In

contrast, the alkylated derivatives of uracil, m3U and e3U, are both potent blocks to replication when encountered by polymerase. In cells lacking AlkB, neither lesion allows replication to proceed more than 10% of the time, and when the lesion is successfully passed, miscoding is common, as both m3U and e3U are >60% mutagenic in this context (Figures 6.6B and C, right sides). When AlkB is present, blockage by m3U is alleviated somewhat, while that by e3U is almost unaffected (Figure 6.5). Both lesions are rendered slightly less mutagenic by functional AlkB protein, as m3U is only ~30% miscoding and e3U ~55% miscoding in cells proficient in AlkB-mediated repair (Figures 6.6B and C, left sides). Taken together, the lesion bypass and post-replication base composition data indicate that m3U is a substrate for repair by AlkB, which presumably mitigates the toxicity and mutagenicity of the unrepaired adduct through the originally-discovered mechanism of oxidative demethylation, converting m3U to U, which is treated in these cells as T. In contrast, only weak evidence for repair of e3U by AlkB is seen here. A slight change in mutational spectrum and a very minor increase in bypass ability induced by AlkB indicate that the protein might be capable of dealkylating this base to some extent, but that the reaction is extremely inefficient. It has been suggested previously that m3U could be a substrate for the AlkB protein (*24*), and the evidence presented here indicates that this is indeed the case.

Although the genetic consequences of m3C, e3C, and m3T have been examined in the past (*2*), they were reinvestigated here as a known baseline against which to compare the results of the novel lesions as well as to determine whether the lack of UDG in the cells used here caused any intrinsic phenotypic differences in the processing of lesions. Additionally, examining the complete set of 3-alkylated pyrimidines simultaneously was intended to minimize experimental variability to allow the strongest possible conclusions to be drawn. The alkylated cytosine adducts, m3C and e3C, were both potently toxic in AlkB-deficient cells, and that toxicity could be largely mitigated by the presence of AlkB (Figure 6.5). This pattern was also seen in the previous study, although all replication efficiencies were slightly higher in that work (*2*). Further evidence for fast and fairly complete repair of the 3-alkylated cytosines is provided by the mutagenic spectra, in which lesions that are strongly promutagenic in repair deficient

cells (~82 and 73% miscoding for m3C and e3C respectively) code only as C when AlkB is functional (Figure 6.6D and E). Again these results are in good agreement with those found previously (*2*). While the replication efficiencies observed for m3T are extremely close to those seen in the past study, with AlkB enabling a small increase from ~7% to ~11%, the decrease in observed mutagenicity from ~54 to ~49% was somewhat smaller than that seen before (approximately 60 to 40%), although the trends agreed.

3-Methyluracil can now definitively be added to the list of alkylated bases that AlkB can repair *in vivo*, although it falls into the category of weaker substrates along with m1G and m3T (*2,8*) and was not detectably repaired *in vitro* in either DNA or RNA (Figure 6.7). Perhaps alternate incubation conditions, longer reaction times, or larger amounts of substrate oligonucleotide would allow mass spectral detection of m3U repair by AlkB, and these possibilities will be explored in the future. 3-Ethyluracil, however, does not seem to be repaired effectively by AlkB *in vivo*, unlike 3-ethylcytosine, which is a good substrate for the enzyme. This finding is not entirely surprising given the substantial difference in the relative efficiencies of repair of the methylated analogs (C>>U). Although AlkB is not as active upon 3-methyluracil as on other substrate bases, perhaps this low level of *in vivo* activity is enough to protect cells against the deleterious consequences of this presumably rare lesion in genomic DNA.

Figure 6.1: Structures of the bases used in the UDG assay and the lesion bypass and mutagenesis experiments.

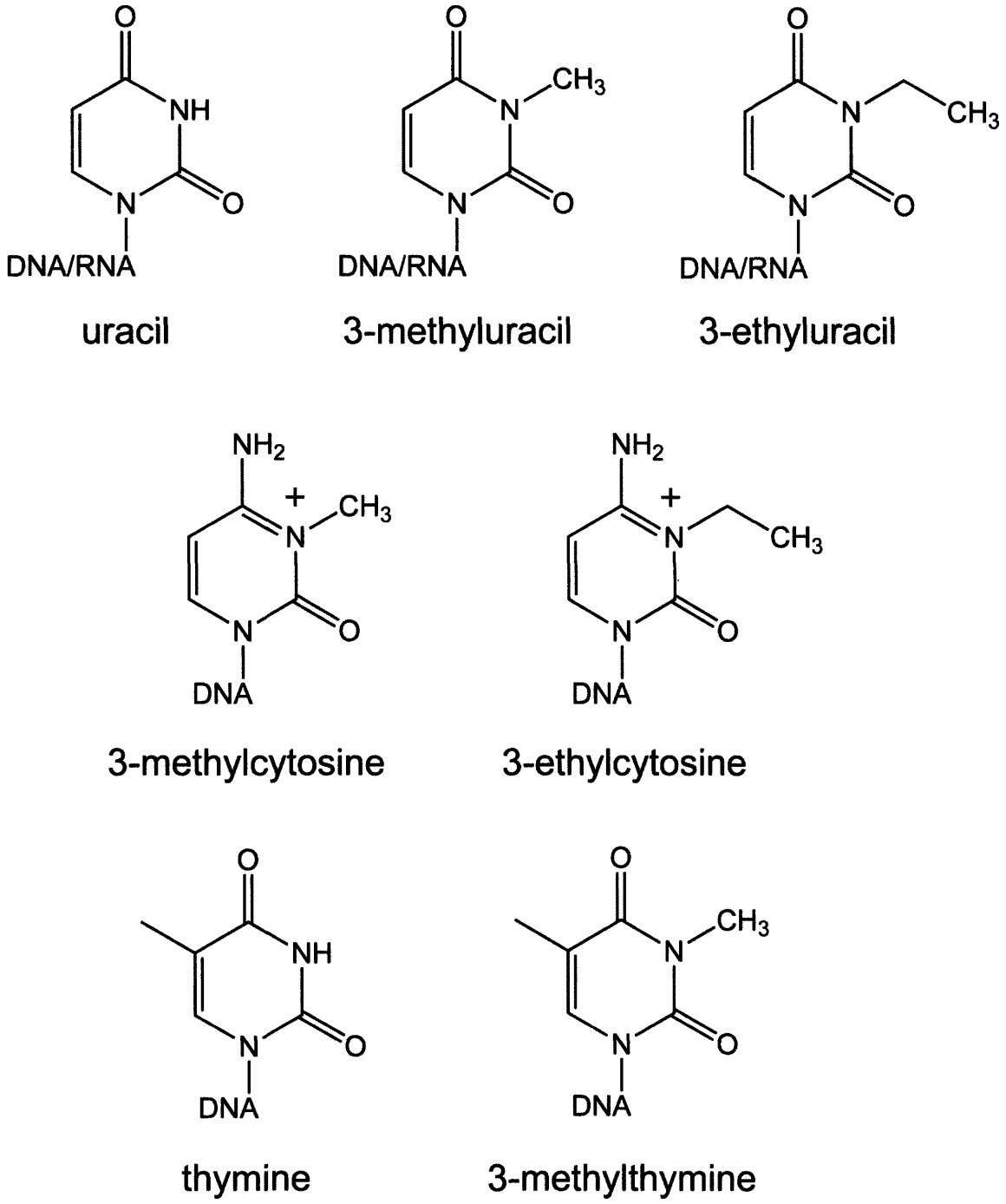


Figure 6.2: 3-methyluracil is not an efficient substrate for *E. coli* Uracil-DNA Glycosylase in either a single- or double-stranded oligodeoxynucleotide. 5'-radiolabeled oligonucleotides containing a single m3U or U residue were either used immediately or annealed to a complementary strand containing A and treated as indicated before separation of products on a 20% denaturing gel. UDG incubations were at 37 °C for 30 min, and EndoIV incubations were at 37 °C for 1 h.

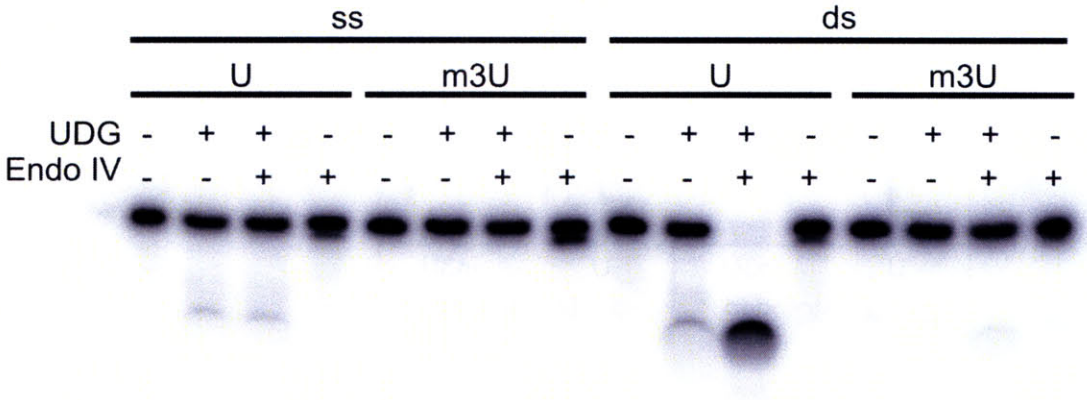


Figure 6.3: Quantification of strand cleavage by UDG and EndoIV. **A.** Excluding the unidentified nicked band. **B.** Quantification of cleavage products including the unidentified nicked band.

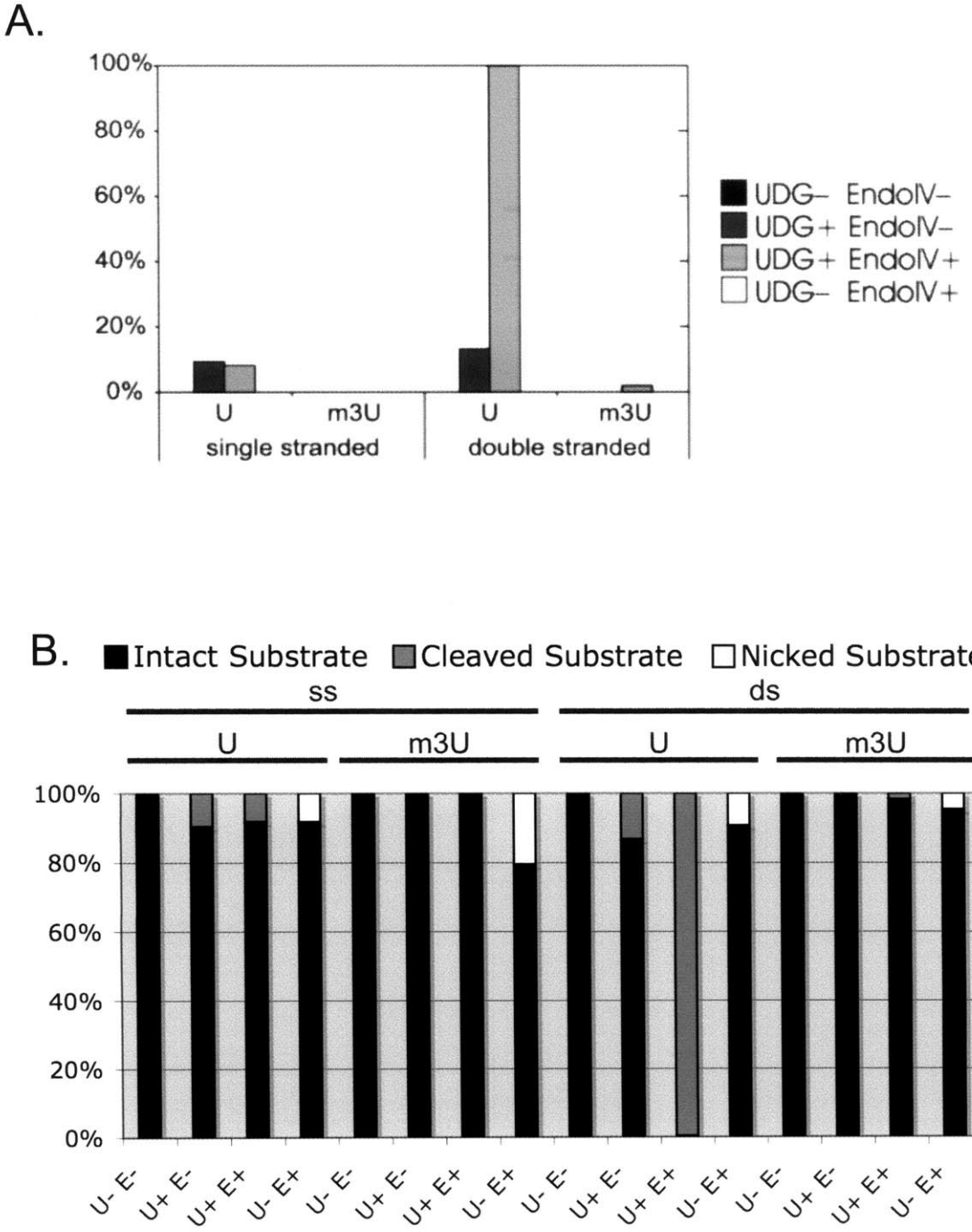


Figure 6.4: PCR-mediated gene replacement used to create ung-deficient derivatives of HK81 and HK82.

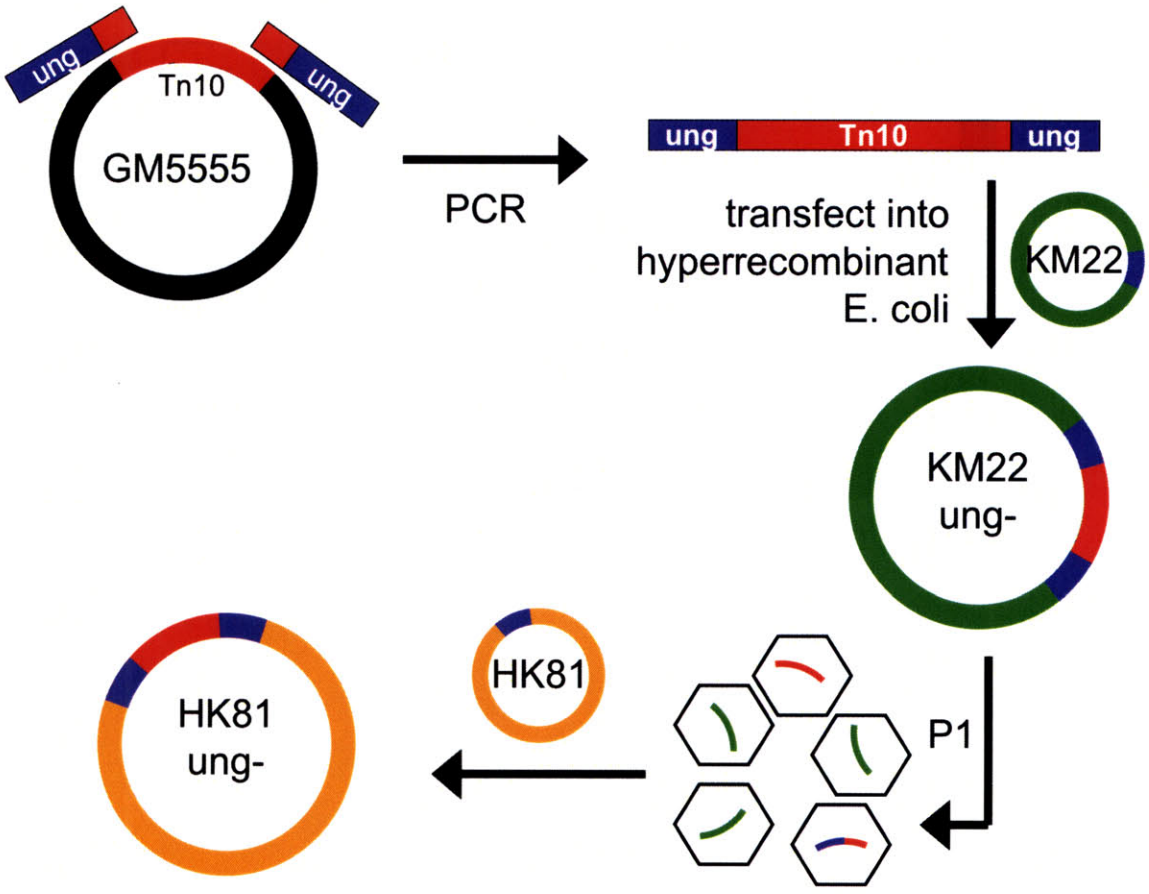


Figure 6.5: Efficiency of replication in *E. coli* of single-stranded M13 genomes containing one site-specific lesion of interest either producing AlkB (AlkB+, HK81, black bars) or lacking AlkB (AlkB-, HK82, gray bars).

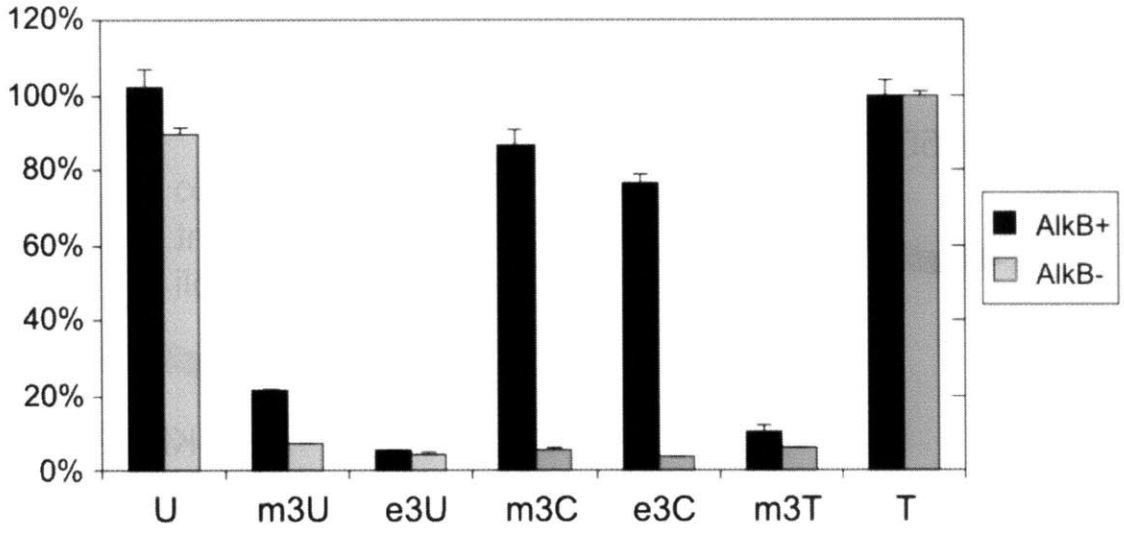


Figure 6.6: Mutational frequencies and specificities for the seven lesions when replicated in cells either proficient (HK81) or deficient (HK82) in AlkB.

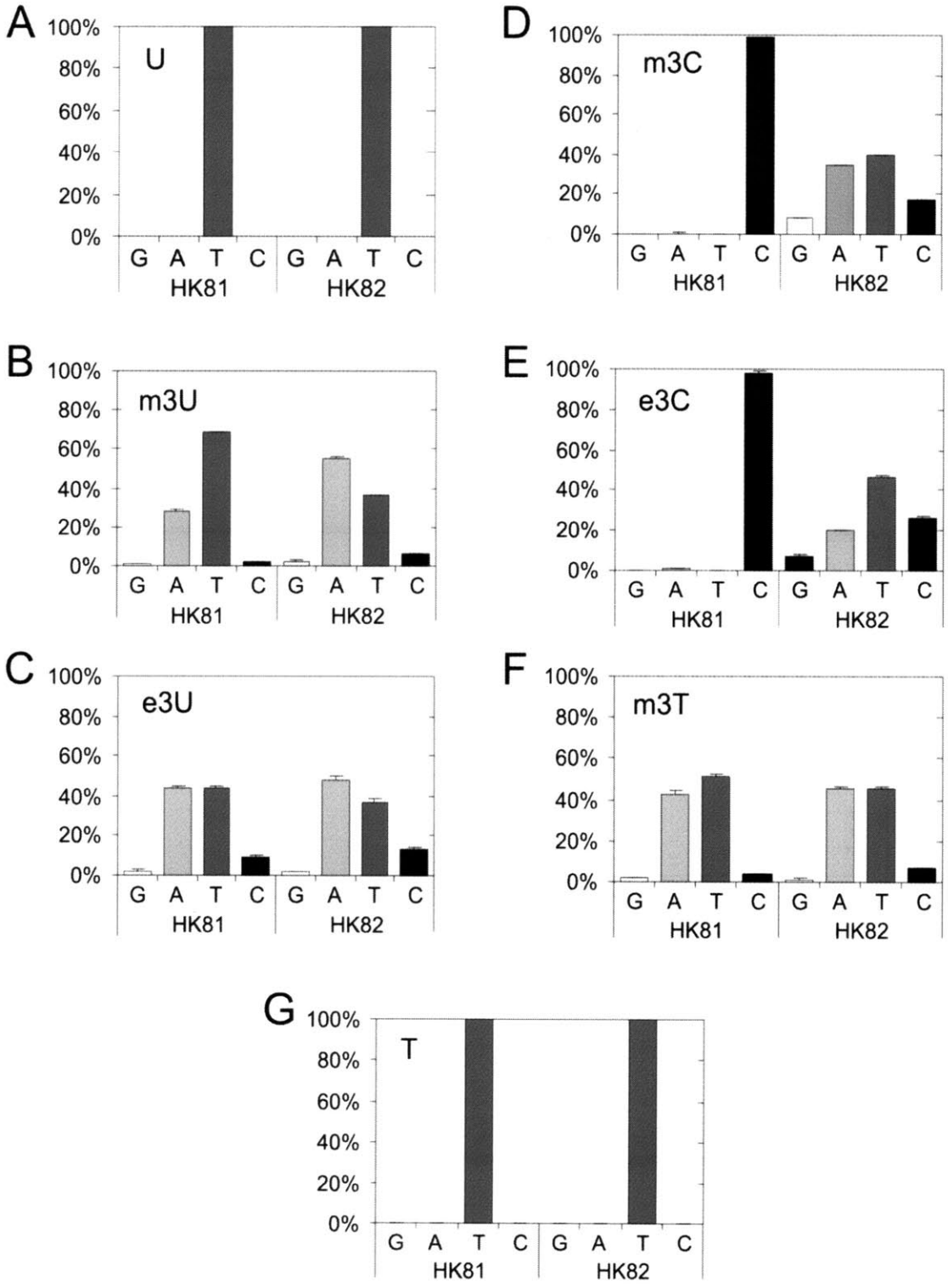


Figure 6.7: Although m3C in DNA was completely repaired by AlkB in 2 h, no reversal of m3U was detected in DNA or in RNA.

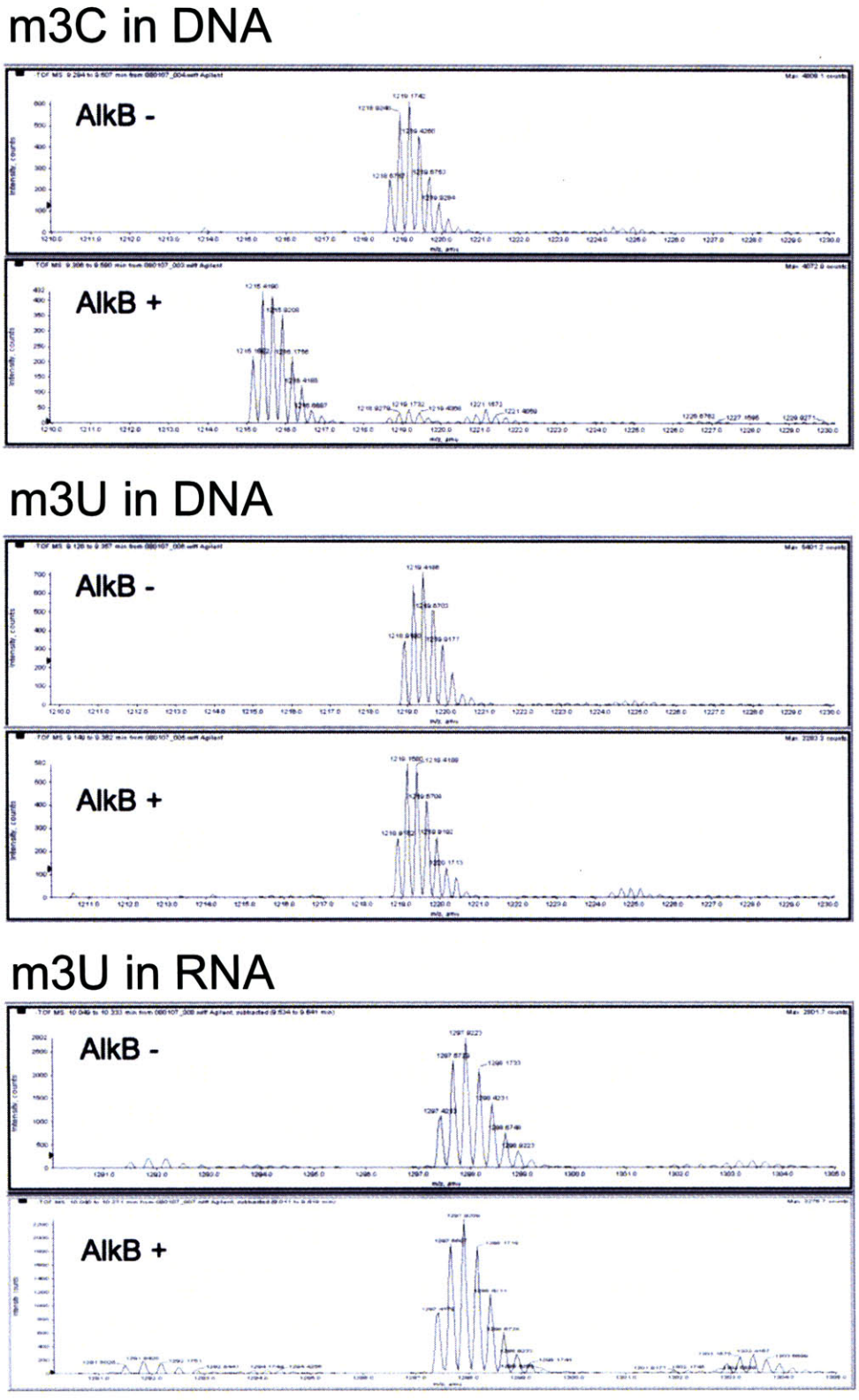


Table 6.1: Sequence of the ung gene (black) including upstream and downstream regions (light gray) showing the locations of the primers used for amplification for genotyping (forward primers highlighted in black, reverse in gray) and those used in the construction of the recombination substrate (forward and reverse in bold).

GCAAAATTAT TTGATTTGTT CAGCCTGTCG CGGCCAATTG GTAAAACCAT
TGTTGCTTGA GTGTATATAT ACTCCTCAAA CACCCTTGAA TCTTTGATTT
AAATCAATAA AAACCACACA TCAAGTATGG TCGCAAATGG ATTTTATTGT
TTTACATCAA CTTATGCGGG TGTGAAATTT TACCAATTTA CATTTTTTTG
CACTCGTTTA AGTCTAAAAA TAAAAAATGA GCATGATTTT GTTCTGTAGA
AAGAAGCAGT TAAGCTAGGC GGATTGAAGA TTCGCAGGAG AGCGAGatgG
CTAACGAATT AACCTGGCAT GACGTGCTGG CTGAAGAGAA GCAGCAACCC
TATTTTCTTA ATACCCTTCA GACCGTCGCC AGCGAGCGGC AGTCCGGCGT
CACTATCTAC CCACCACAAA AAGATGTCTT TAACGCGTTC CGCTTTACAG
AGTTGGGTGA CGTTAAAGTG GTGATTCTCG GCCAGGATCC TTATCACGGA
CCGGGACAGG CGCATGGTCT GGCATTTTCC GTTCGTCCCG GCATTGCCAT
TCCTCCGTCA TTATTGAATA TGTATAAAGA GCTGGAAAAT ACTATTCCGG
GCTTCACCCG CCCTAATCAT GGTTATCTTG AAAGCTGGGC GCGTCAGGGC
GTTCTGCTAC TCAATACTGT GTTGACGGTA CGCGCAGGTC AGGCGCATTC
CCACGCCAGC CTCGGCTGGG AAACCTTCAC CGATAAAGTG ATCAGCCTGA
TTAACCAGCA TCGCGAAGGC GTGGTGTTTT TGTTGTGGGG ATCGCATGCG
CAAAAGAAAG GGGCGATTAT AGATAAGCAA CGCCATCATG TACTGAAAGC
ACCGCATCCG TCGCCGCTTT CGGCGCATCG TGGATTCTTT GGCTGCAACC
ATTTTGTGCT GGCAAATCAG TGGCTGGAAC AAGGTGGCGA GACGCCGATT
GACTGGATGC CAGTATTACC GGCAGAGAGT GAGTAAATTT GCGGGGAAAT
GCCGGATGGC AGAGTTGCCA

References

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

2. Delaney, J. C. and J. M. Essigmann. 2004. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A* **101**:14051-14056.

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

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

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

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

7. Friedberg E.C., 2006. DNA Repair and Mutagenesis. AMS Press, Washington, DC.

8. Sedgwick, B., P. A. Bates, J. Paik, S. C. Jacobs, and T. Lindahl. 2007. Repair of alkylated DNA: recent advances. *DNA Repair (Amst)* **6**:429-442.

9. Leutzinger, E. E., P. S. Miller, and L. S. Kan. 1982. Studies on the hydrolysis of 3-methyl-2'-deoxycytidine in aqueous solution A synthesis of 3-methyl-2'-deoxyuridine. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **697**:243-251.

10. Sowers, L. C., W. D. Sedwick, and B. R. Shaw. 1989. Hydrolysis of N3-methyl-2'-deoxycytidine: model compound for reactivity of protonated cytosine residues in DNA. *Mutat. Res.* **215**:131-138.

11. Cong, H. N., O. Bertaux, R. Valencia, T. Becue, T. Fournier, D. Biou, and D. Porquet. 1994. Separation and characterization of the main methylated nucleobases from nuclear, cytoplasmic and poly (A)+ RNA by high-performance liquid chromatography and mass spectrometry. *J. Chromatogr. B Biomed. Appl.* **661**:193-204.

12. Cecchini, J. P. and R. Miassod. 1979. An improved method for the purification of higher plant ribosomal RNA labeled on the methyl groups. *Anal. Biochem.* **100**:70-76.

13. Rozenski, J., P. F. Crain, and J. A. McCloskey. 1999. The RNA Modification Database: 1999 update. *Nucleic Acids Res.* **27**:196-197.

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

15. Falnes, P. O. 2004. Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. *Nucleic Acids Res.* **32**:6260-6267.

16. Falnes, P. O., M. Bjoras, P. A. Aas, O. Sundheim, and E. Seeberg. 2004. Substrate specificities of bacterial and human AlkB proteins. *Nucleic Acids Res.* **32**:3456-3461.

17. Ougland, R., C. M. Zhang, A. Liiv, R. F. Johansen, E. Seeberg, Y. M. Hou, J. Remme, and P. O. Falnes. 2004. AlkB restores the biological function of mRNA and tRNA inactivated by chemical methylation. *Mol. Cell* **16**:107-116.

18. Lee, D. H., S. G. Jin, S. Cai, Y. Chen, G. P. Pfeifer, and T. R. O'Connor. 2005. Repair of methylation damage in DNA and RNA by mammalian AlkB homologues. *J. Biol. Chem.* **280**:39448-39459.

19. Begley, T. J. and L. D. Samson. 2003. AlkB mystery solved: oxidative demethylation of N1-methyladenine and N3-methylcytosine adducts by a direct reversal mechanism. *Trends Biochem. Sci.* **28**:2-5.

20. Bellacosa, A. and E. G. Moss. 2003. RNA repair: damage control. *Curr. Biol* **13**:R482-R484.

21. Drablos, F., E. Feyzi, P. A. Aas, C. B. Vaagbo, B. Kavli, M. S. Bratlie, J. Pena-Diaz, M. Otterlei, G. Slupphaug, and H. E. Krokan. 2004. Alkylation damage in DNA and RNA--repair mechanisms and medical significance. *DNA Repair (Amst)* **3**:1389-1407.

22. Bratlie, M. S. and F. Drablos. 2005. Bioinformatic mapping of AlkB homology domains in viruses. *BMC. Genomics* **6**:1.

23. Falnes, P. O. 2005. RNA repair--the latest addition to the toolbox for macromolecular maintenance. *RNA. Biol* **2**:14-16.

24. Falnes, P. O., A. Klungland, and I. Alseth. 2007. Repair of methyl lesions in DNA and RNA by oxidative demethylation. *Neuroscience* **145**:1222-1232.

25. Karran, P. 1985. Possible depletion of a DNA repair enzyme in human lymphoma cells by subversive repair. *Proc. Natl. Acad. Sci. U. S. A* **82**:5285-5289.

26. Gibbs, P. E. and C. W. Lawrence. 1993. U-U and T-T cyclobutane dimers have different mutational properties. *Nucleic Acids Res.* **21**:4059-4065.

27. Reichert, A. S. and M. Morl. 2000. Repair of tRNAs in metazoan mitochondria. *Nucleic Acids Res.* **28**:2043-2048.

28. 2001. Current Protocols in Molecular Biology. John Wiley & Sons, Inc., Hoboken, NJ.

29. Delaney, J. C. and J. M. Essigmann. 2006. Assays for Determining Lesion Bypass Efficiency and Mutagenicity of Site-Specific DNA Lesions *In Vivo*. *Methods in Enzymology* **408A**:1-15.

30. Frick, L. E., J. C. Delaney, C. Wong, C. L. Drennan, and J. M. Essigmann. 2007. Alleviation of 1,N6-ethanoadenine genotoxicity by the Escherichia coli adaptive response protein AlkB. *Proc. Natl. Acad. Sci. U. S. A* **104**:755-760.

APPENDIX 1:

Oxidative Stress Does Not Induce the *Escherichia coli* Adaptive Response Regulated by Ada

Introduction

The adaptive response to alkylating agents was discovered in 1977 when it was observed that bacteria exposed to low levels of *N*-methyl-*N*'-nitro-nitrosoguanidine (MNNG) displayed resistance to both killing and mutation caused by subsequent higher doses (*1*). These experiments also showed that this protective effect required protein synthesis and was probably due to a novel DNA repair activity rather than prevention of DNA damage, as previously-damaged phage survived better in induced than in non-induced cells.

Since that time, the adaptive response to alkylation damage has been shown to comprise four genes (*2-6*). Ada is a bifunctional methyltransferase that serves to reverse DNA damage as well as to regulate gene expression (*7*). The protein is composed of two methyltransferase regions with separate activity. The carboxy-terminal portion is responsible for the DNA repair activity, removing DNA damage by accepting methyl groups from O^6-methylguanine and O^4-methylthymine through irreversible transfer of that methyl group to Cys321 in the C-terminal active site. The amino-terminal domain confers regulatory functionality to the protein through its ability to accept a methyl group from the *Sp* isomer of backbone methylphosphotriesters via methyl transfer to the Cys38 residue. Methylation at this site confers to Ada the ability to bind to and induce expression at three promoters, thus up-regulating the four genes of the adaptive response (Ada and AlkB share a promoter, while AlkA and AidB each have their own).

AlkB is an α-ketoglutarate dioxygenase that uses non-heme iron (II) as a cofactor and molecular oxygen and α-ketoglutarate as cosubstrates to reverse directly a range of damaged DNA bases (*8,9*). The enzyme operates through two distinct chemical mechanisms. The first, oxidative demethylation, removes methyl groups from 1-methyladenine and 3-methylcytosine, lesions formed predominantly in single-stranded DNA under conditions of alkylative stress, by hydroxylating the extraneous methyl group, creating an unstable hydroxymethyl intermediate, which decomposes to release formaldehyde and the repaired base (*8,9*). AlkB has since been shown to also repair 3-ethylcytosine (*10*), 1-methylguanine (*10,11*), 3-methylthymine (*10-12*), and 1-ethyladenine (*13*) as well as hydroxyethyl, propyl, and hydroxypropyl (*14*) lesions in

DNA through this mechanism. The second chemical pathway through which AlkB can act involves the epoxidation of double-bond-containing exocyclic lesions, such as 1,N^6-ethenoadenine and 3,N^4-ethenocytosine, resulting in the release of glyoxal and the undamaged base (*15,16*).

AlkA is the inducible 3-methyladenine glycosylase responsible for excision of damaged bases as the first step in the base excision repair pathway. Substrate bases include 3-methyladenine, 7-methylguanine, 3-methylguanine, O^2-methylcytosine, O^2-methylthymine, 8-methylguanine, hypoxanthine, 5-formyluracil, N^2,3-ethenoguanine and 1,N^6-ethenoadenine (*17-24*).

AidB is homologous to human isovaleryl CoA dehyrdogenase (*25*), a central player in the leucine metabolism pathway, but its functional role in the adaptive response has not yet been elucidated beyond its ability to bind double-stranded DNA (*26*).

In addition to the adaptive response to alkylating agents, *E. coli* possess other inducible sets of proteins, one of which allows the bacteria to survive by metabolizing lactose. Bacteria grown in the absence of lactose contain very few molecules of the two proteins necessary to carry out this process (*27*), β-galactosidase, which breaks the disaccharide lactose into its monosaccharide components galactose and glucose (Figure A1.1A), and galactoside permease, which transports lactose from the environment into the cell. Exposure to lactose causes an ~1000-fold increase in the expression levels of these proteins, enabling the efficient uptake and breakdown of lactose in *E. coli* (*27*).

The β-galactosidase enzyme is capable of carrying out its chemistry on other substrates as well as on lactose. For example, the colorless compound 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) is broken down by β-galactosidase into β-D-galactoside and 5-bromo-4-chloro-3-hydroxyindole (Figure A1.1B), which upon oxidation combines with another molecule, forming a conjugated blue substance that allows for easy identification of enzymatic activity. Similarly, the colorless o-nitrophenyl-β-D-galactoside (ONPG) is split by β-galactosidase into β-D-galactoside and o-nitrophenol (Figure A1.1C), which is yellow and absorbs 420 nm light. The absorbance of released o-nitrophenol can be measured to indicate the amount of β-galactosidase activity present (*28*).

These simple methods of enzymatic activity measurement have been combined with gene fusions and exploited to allow quantification of the expression levels of various proteins and investigation into the mechanisms of regulation of those proteins within a cell (*29-35*). Fusion of the β-galactosidase gene (called lacZ) to a particular promoter of interest will result in the regulated production of a measurable activity under the control of that promoter's specific conditions of regulation.

The induction of the *E. coli* adaptive response to alkylation has been probed by this method using an ada-LacZ operon fusion (*29,30*). Plasmids containing either the entire ada gene or an ada-lacZ fusion in which the lacZ gene is regulated by the ada promoter were introduced into bacterial strains of varying Ada status. Induction of β-galactosidase by exposure to MNNG was observed only in ada+ strains, indicating that functional Ada is necessary for upregulation of genes controlled by the ada promoter and that Ada is therefore autoregulatory. Additionally, it was discovered that the mechanism by which Ada regulates its own expression is different from that which it employs in the regulation of its other inducible genes (specifically alkA). Truncated forms of Ada were constitutive activators of ada-lacZ but inducible activators of alkA-lacZ. Additionally, the N-terminal half of the Ada protein induced alkA transcription in cells exposed to MNNG, but this protein portion proved insufficient to induce ada transcription, suggesting the presence of differences in Ada binding and function between the two promoters.

The *E. coli* adaptive response is induced by the presence of the very damage that its constituent proteins repair. However, in addition to the methylated bases subject to three different known forms of adaptive response-mediated repair, other lesions have been shown to be substrates for the repair activities of this inducible regulon. Because the proteins encoded by the genes under control of the Ada protein are capable of removing many types of damage from cellular DNA, such as that created under conditions of oxidative stress, it was hypothesized that perhaps these other stresses (beyond exposure to simple S_N1 alkylating agents) could analogously induce expression of the proteins necessary to mitigate the consequences of the DNA lesions that they create, thus facilitating survival. The studies described here were intended to utilize this

ada:LacZ fusion system to further investigate the possibility that other types of cellular stress, in addition to alkylative stress, could induce this protective response.

Materials and Methods

Bacterial Strains and Plasmids. The *E. coli* strains used were BW140 (as AB1157, but *lacU169 rpsL*; plasmid source), HK81 (as AB1157, but *nalA*; wild-type), and HK82 (as HK81, but *alkB22*; AlkB-deficient). The plasmids used were pGW2620 (an ada+ derivative of pSE101) and pGW2622 (a derivative of pGW2620 that contains an ada'-lacZ+ fusion). These plasmids were gifts of Graham C. Walker.

Plasmids were isolated from pelleted 10 mL cultures of host strains (grown in LB supplemented with 50 μg/mL kanamycin) using a QIAprep Spin Miniprep Kit (Qiagen) and following the protocol for a low copy number plasmid. Isolated plasmids were then transfected into strains HK81 and HK82 using a $CaCl_2$ solution. Briefly, 10 mL saturated overnight cultures of HK81 and HK82 were diluted 1:100 in 10 mL fresh LB and grown to mid-log phase (A_{600} = 0.5). Cells were pelleted by centrifugation, washed twice with $CaCl_2$ solution (15% glycerol, 60 mM $CaCl_2$, 10 mM PIPES pH 9) and finally resuspended in 1 mL $CaCl_2$ solution on ice. 100 μL cells (HK81 or HK82) were mixed with 5 μL isolated plasmid (pGW2620 or pGW2622) or H_2O as a control and allowed to sit on ice for 10 min. Following a 1 min incubation at 42 °C, the cells were returned to ice for an additional 10 min, and then each mixture was transferred to 10 mL LB. After a recovery period of growth for 1 h at 37 °C with aeration, kanamycin was added to a final concentration of 50 μg/mL, and cultures were grown overnight. All cultures containing plasmid grew with selection, and the following day were streaked on 50 μg/mL kanamycin/LB plates and allowed to grow overnight at 37 °C. Single well-isolated colonies were then picked, grown in liquid culture with selection overnight, and used for all subsequent experiments.

Growth Inhibition by Various Alkylating or Oxidative Agents. Growth curves were created in order to determine optimal dosing conditions for a range of alkylating and oxidative agents. Saturated overnight cultures of plasmid-containing HK81 and HK82

were diluted 1:100 in 10 mL LB media supplemented with 50 μg/mL kanamycin. In some experiments, cultures were treated immediately with various concentrations of the drug of interest (0-0.01% chloroacetaldehyde (CAA), 0-0.15% methyl methanesulfonate (MMS), 0-1 μg/mL MNNG or 0-400 μM 3-morpholinosydnonimine (SIN-1)) while in other cases cells were allowed to grow to mid-log phase before dosing. Optical density measurements were taken at 600 nm over the course of several hours in order to detect cell killing and growth inhibition caused by the presence of damaging agent.

β-Galactosidase Assays. β-Galactosidase assays were performed as previously described (*28*). Saturated overnight cultures were diluted (4 drops into 5 mL) in either LB medium or M9 minimal salts supplemented with 50 μg/mL kanamycin and the dose of agent to be analyzed. Cells were grown to an OD_{600} of about 0.4, at which point they were placed on ice and subjected to the assay, allowing the expression from the ada promoter to be quantified. Briefly, the assay involves the addition of 0.1 mL culture to 0.9 mL Z buffer (60 mM $Na_2HPO_4 \cdot 7H_2O$, 40 mM $NaH_2PO_4 \cdot H_2O$, 10 mM KCl, 1 mM $MgSO_4 \cdot 7H_2O$, 50 mM β-mercaptoethanol). Cells are opened by the addition of 2 drops of chloroform and 1 drop 0.1% SDS followed by vortexing briefly. After incubation at 28 °C for 5 min, reactions are initiated by the addition of 0.2 mL ONPG (4 mg/mL in Z buffer). Upon completion (appearance of yellow color), the reaction is stopped by the addition of 0.5 mL 1 M Na_2CO_3 and the time taken for color to develop is noted. The OD_{420} and OD_{550} are measured and the enzyme units are calculated as $[1000 \times (OD_{420} - 1.75 \times OD_{550})]/[t \times v \times OD_{600}]$, where t = the reaction time in minutes and v = mL of culture used in the reaction.

Results

Sensitivity of AlkB-proficient and –deficient *E. coli* to several classes of DNA damaging agent. In order to determine the optimal concentration of drug that would produce the most differential response between the AlkB-proficient and AlkB-deficient strains, growth curves were constructed for a range of doses of several alkylating and oxidative agents.

Exocyclic ring formation. Because the etheno- lesions were of major interest, the ability of the HK81 and HK82 strains to grow in the presence of chloroacetaldehyde (CAA), a metabolite of vinyl chloride that produces cyclic etheno-lesions in DNA (*36*), was monitored over a range of drug concentrations spanning three orders of magnitude. Figures A1.2A and A1.2B show the impaired growth of HK81 and HK82, respectively, in liquid cultures to which drug was added immediately upon cellular dilution from the 10 mL overnight culture (at t=0). In both strains, the 0.0001% dose was almost indistinguishable from the untreated control, while the 0.005 and 0.01% doses produced identical curves not rising above baseline. The similarity in the curve sets demonstrates that the AlkB status of the cells does not greatly influence their ability to replicate in the presence of CAA at any of the doses examined. Characterization in this manner of the cellular response to exposure to CAA also revealed a non-linear relationship between dose and growth ability that features a threshold range in which the majority of growth inhibition or cell death occurs. A very narrow dose range is observed over which the majority of cells die. While this experiment shows similar responses of the HK81 and HK82 strains to concentrations of CAA spanning several orders of magnitude, further studies were conducted in an attempt to optimize the dose range to focus on the threshold area to detect differential behavior between AlkB-proficient and AlkB-deficient cells.

Because such a large gap in growth was observed in both strains between the 0.001% and the 0.0025% doses, further experiments were conducted employing concentrations in this range administered immediately and monitoring growth for longer periods of time in order to more precisely determine the dose that would produce the largest difference in survival between the two strains (Figure A1.3). The 0.002% dose produced the largest differential growth response between AlkB-proficient and AlkB-deficient strains (Figure A1.3, squares), although some convergence was observed at later time-points. The 0.0025% dose also induced slower growth in the AlkB-deficient strain than in the AlkB-proficient, although the divergence occurred later and was of smaller magnitude. Treatment with 0.003% CAA delayed growth substantially in both strains and did not produce a differential response until the last time point measured.

Normalization of the data to the appropriate strain's untreated control highlights the large growth advantage possessed by HK81 under treatment with 0.002% CAA (i.e, the treated AlkB-containing cells lag behind the untreated AlkB-containing cells less than the treated AlkB-deficient cultures lag behind their untreated equivalent). The differential phenotypes induced by the two higher doses are more easily shown to be insignificant when the data are presented this way, and both the raw and the normalized data will be presented for subsequent studies.

One way in which such experiments have been conducted in the past involves the addition of damaging agent not immediately upon dilution, but rather once the cells have reached their mid-log phase of growth. Consequently, the sensitivity of cells treated in this manner was assayed as well. Interestingly, none of the concentrations of CAA that caused considerable growth inhibition when administered immediately were effective at killing or slowing growth of mid-log cultures (Figure A1.4). Additionally, the doses used (0.002% - 0.003% CAA) did not cause divergence of the growth patterns between HK81 and HK82. The normalized data underscores the very small differences in growth when cells are treated mid-log and shows an unexpected slight growth advantage of HK82 under these conditions.

S_N2 methylation. In addition to CAA, the S_N2 alkylating agent MMS, which produces the AlkB substrates 1-methyladenine and 3-methylcytosine in DNA, was investigated. Cells were treated at mid-log phase with 0% - 0.15% MMS and subsequent growth was monitored as above (Figure A1.5). Again a threshold effect is observed, as the differences between untreated, 0.025% MMS and 0.05% MMS are relatively small compared with the large decrease in growth induced in both strains by the 0.15% MMS dose. Normalization of the growth curves to the untreated controls (Figure A1.5, bottom) indicates that, surprisingly, the HK82 strain again exhibits a very slightly increased ability to grow in the presence of the drug.

S_N1 methylation. The S_N1 alkylating agent MNNG was similarly characterized and found to possess only a weak ability to alter the growth of cells treated at mid-log. The

top panel of Figure A1.6 shows the comparable growth of both strains treated with two doses of MNNG, both of which are sufficient to cause robust induction of the adaptive response. Normalization of the data (treated/untreated for each strain, Figure A1.6, bottom) shows that these doses produce only minor changes in the growth of cultures of either strain.

Oxidation. Finally, SIN-1, a compound that decomposes through a three-step mechanism to release nitric oxide and superoxide, was used to study the effects of growth under oxidative stress. Figure A1.7 shows that mid-log cells are again resistant to the toxic effects of this oxidizing agent. The lower panel of Figure A1.7 reveals that any observed differences are miniscule and that if any growth advantage exists, it counterintuitively lies in favor of the AlkB-deficient HK82.

Ability of various damaging agents to induce the adaptive response. Cultures of HK81 and HK82 containing pGW2622 (plasmid containing the ada-LacZ fusion; designated HK81/22 and HK82/22) or pGW2620 (control plasmid containing ada+; designated HK81/20 and HK82/20) were grown in the presence or absence of the chemical agents examined above. Analysis by the standard β-galactosidase method was conducted to determine each agent's ability to induce the adaptive response.

Comparison of untreated control-plasmid-containing strains (HK81/20 and HK82/20) with untreated fusion-plasmid-containing strains (HK81/22 and HK82/22) (Figure A1.8, black bars represent HK81 derivatives, gray bars correspond to HK82 strains) reveals that, while no activity is observed in control strains, expression of the gene fusion contained in the pGW2622 plasmid takes place even without cellular exposure to an alkylating agent in both strains (HK81/22 and HK82/22). Consequently, all data are presented as induction relative to the baseline expression observed in untreated HK81/22 cells to correct for this background signal (although the background expression in the AlkB+ and AlkB- strains are not significantly different and either could have served as the normalization factor). As expected, given pGW2620's lack of a lacZ gene, treatment of HK81/20 and HK82/20 with 0.001% CAA does not alter the lack of β-

galactosidase activity. The background expression of functional protein in the HK81/22 strain is similar to that observed for HK82/22, and these levels do not change significantly upon exposure to CAA, either at the 0.001% or the 0.002% dose.

MMS exposure was compared with CAA treatment and observed to be even less efficient at inducing expression from the ada promoter in either strain (Figure A1.9), despite the fact that DNA damage levels are expected to differ greatly between the two strains due to their differential abilities to conduct AlkB-mediated repair of MMS-induced lesions.

Even at completely non-toxic doses, the positive control compound MNNG was able to dramatically increase the amount of β-galactosidase produced from the ada promoter in both the HK81- and HK82-derived strains (Figure A1.10). Again, the levels of expression between the two strains under any of the conditions studied were not significantly different, as expected due to the fact that both possess functional Ada.

Finally, the ability of oxidative stress to induce the adaptive response was examined (Figure A1.11). Doses of SIN-1 up to 400 μM were unable to increase expression while the low, non-toxic dose of MNNG induced a large up-regulation in both strains (Figure A1.11). Concentrations of SIN-1 up to 800 μM also did not increase LacZ activity above background, although growth was slightly inhibited at these higher concentrations (OD_{600} = 0.320 (600 μM) and 0.293 (800 μM) after 2 h growth as compared to 0.423 for the untreated culture; data not shown).

Discussion

The HK81 and HK82 strains of *E. coli* were used here in an attempt to characterize further the adaptive response to alkylating agents. The phenomenon was originally discovered to induce the expression of genes that encode proteins that are responsible for removing alkylated DNA damage, specifically those lesions formed by the inducing agent itself. Further studies of the response revealed that the up-regulated proteins removed additional lesions from DNA beyond those formed by MNNG. After the discovery of novel substrates for one of these proteins, AlkB, and speculation that another, AidB, might be involved in the repair of as-yet-unidentified DNA lesions, it was

hypothesized that perhaps agents which produce these other lesions could also be capable of inducing the up-regulation of this set of genes. Perhaps *E. coli* protect themselves from genomic degradation through an efficient system in which exposure to a variety of damage-inducing stresses or agents could induce expression of precisely the proteins required to mitigate the negative consequences of each particular agent.

While immediate addition of damaging agent to diluted cultures did produce substantial differences in the growth and perhaps the survival of AlkB-proficient as compared to AlkB-deficient strains, supplementation of mid-log cultures did not change the strains' growth differentially. Perhaps these agents affect cells exiting stationary phase more dramatically than those in mid-log phase due to intrinsic differences in replication rates at each stage. Cells already in mid-log phase are actively transcribing and translating the genome, possibly resulting in higher baseline levels of the relevant repair proteins or in decreased time required to synthesize those proteins. Cells just exiting stationary phase were not actively dividing and could possess fewer of the proteins required to protect the genome from damage by these agents, thus resulting in slowed growth while those enzymes are synthesized. Previous studies have examined cellular survival, rather than growth, and it is possible that cells that are already replicating in liquid culture (i.e. in mid-log phase) when treated continue to do so regardless of the presence of AlkB over the time-scale monitored here, but would exhibit different abilities to survive long-term and form colonies according to AlkB status.

Comparison of untreated control-plasmid-containing strains (HK81/20 and HK82/20) with untreated fusion-plasmid-containing strains (HK81/22 and HK82/22) (Figure A1.8) reveals that expression of the fusion protein takes place even without exposure to an alkylating agent, indicating either that the plasmid is leaky or that a background level of environmental damage is sufficient to induce detectable levels of protein production. This low level of baseline expression is, however, far below the induction caused by the positive control and was accounted for by normalization.

All of the experiments described here were conducted in cells uninduced for the adaptive response, meaning that observed differences in growth ability or induction of ß-galactosidase activity after exposure to a damaging agent were created by baseline levels

of repair protein. Perhaps pre-induction of the strains would produce larger phenotypic differences between AlkB+ and AlkB- strains in these growth assays.

Although it would seem useful and efficient for *E. coli* to extend the phenomenon of alkylation-induced expression of alkylated-base repair genes to the arena of oxidative damage, this does not appear to be the case, as none of the damaging agents examined except for the canonical inducing agent, MNNG, were able to significantly induce expression of lacZ from the ada promoter. Although substrates for adaptive response proteins are created by these other agents, the presence of these types of damage does not signal the cell to increase the baseline protein expression. In the case of AlkA, the 3-alkyladenine glycosylase, induction could be rendered unnecessary due to the presence in *E. coli* of a constitutively-expressed analog, Tag (*37*), which could be capable of handling the damage of this type. *E. coli* also produce a pair of methyltransferases, one inducible (Ada) and one constitutive (Ogt). It is conceivable that the duality of these two repair systems serves to avoid a need for adaptive response induction by all types of substrate damage.

However, the other two induced proteins, AlkB and AidB, do not have identified constitutively-expressed partners, and their expression depends wholly (AlkB) or partially (AidB, which is also induced by growth in anaerobic conditions) on induction by methylated Ada protein. Perhaps counterintuitive is the observation that the formation of non-substrate lesions (i.e, MPTEs) causes production of these certain (AlkB) or putative (AidB) repair proteins, yet the presence of substrate lesions does not. The repair of substrate lesions is thus tied to the presence of a separate class of damage, which could indicate that under certain circumstances, AlkB substrates are simply not repaired, or are removed from the genome by an as-yet-unidentified functional homolog.

Figure A1.1: Reactions catalyzed by β-galactosidase. **A.** The biologically-relevant cleavage of lactose into galactose and glucose. **B.** Cleavage of the colorless 5-bromo-4-chloro-3-indolyl-β-galactose (X-gal) into galactose and 5-bromo-4-chloro-3-hydroxyindole, a colorless compound that upon spontaneous dimerization and oxidation forms the blue 5,5'-dibromo-4,4'-dichloro-indigo, the color of which can be exploited to demonstrate enzymatic activity. **C.** Cleavage of the colorless o-nitrophenyl-β-D-galactopyranoside (ONPG) into galactose and o-nitrophenol, the presence of which can be detected by its yellow color.

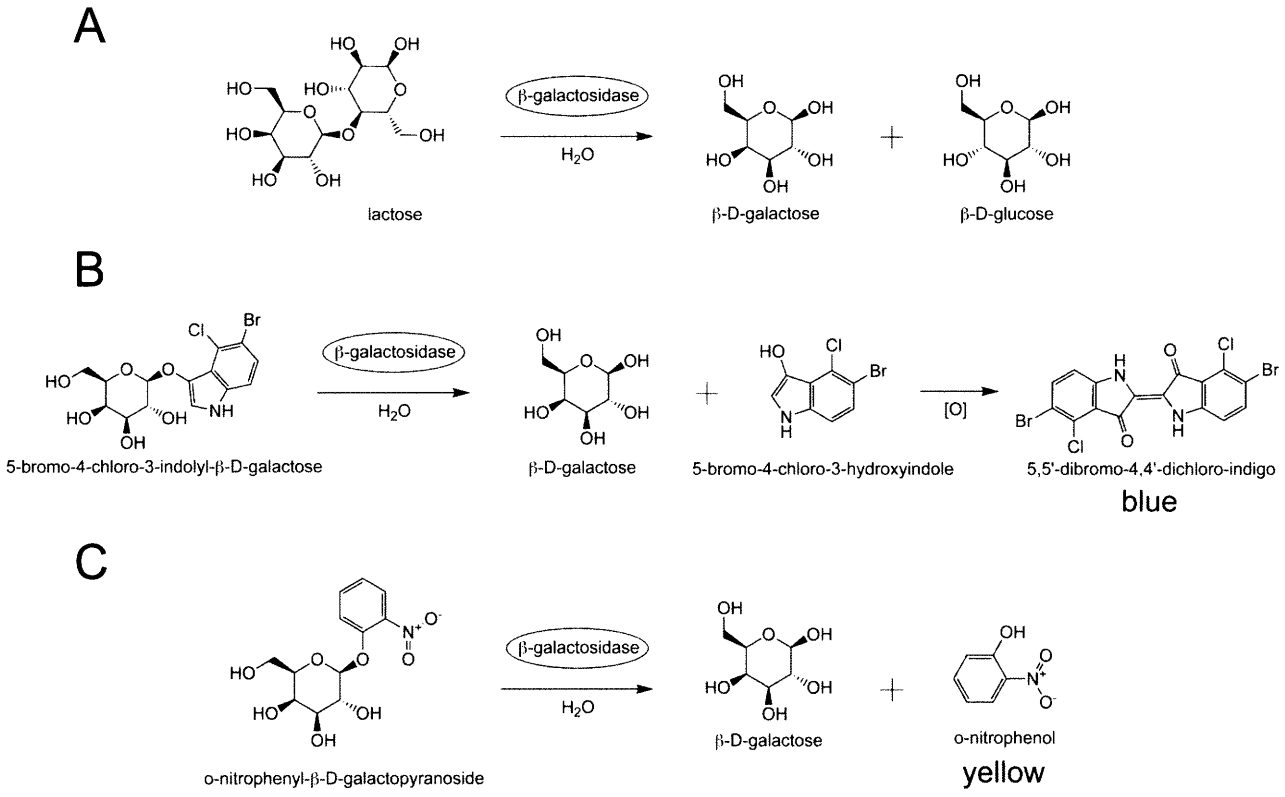
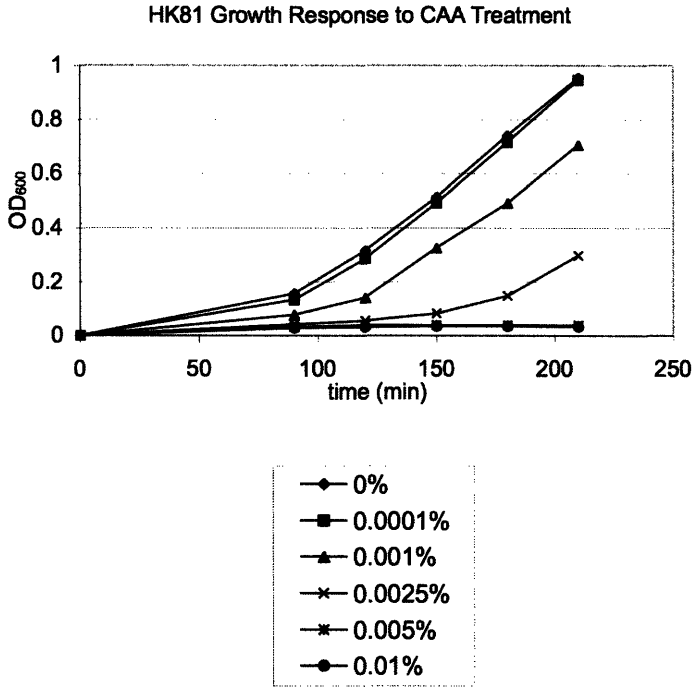


Figure A1.2: Growth curves showing similar inhibition of growth of AlkB+ and AlkB- strains containing pGW2622 (ada-LacZ fusion) upon treatment with a large range of chloroacetaldehyde (CAA) concentrations **A.** HK81/22 (AlkB+) **B.** HK82/22 (AlkB -)

A.



B.

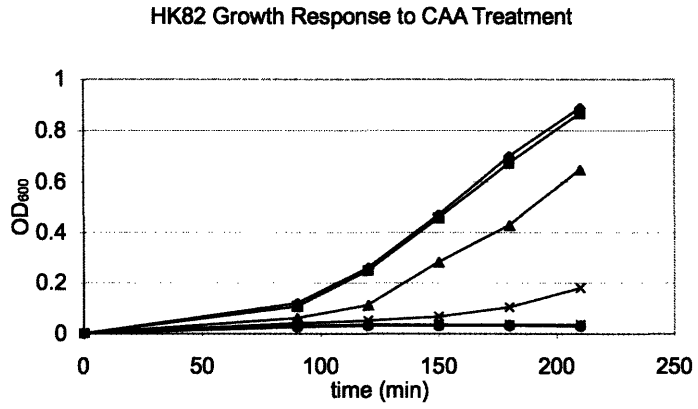


Figure A1.3: Growth curves showing differential growth between the AlkB-proficient (HK81/22, shown in black) and the AlkB-deficient (HK82/22, shown in gray) strains upon treatment with a narrow dose range of CAA. The lower panel shows the same data normalized to the untreated control of each strain.

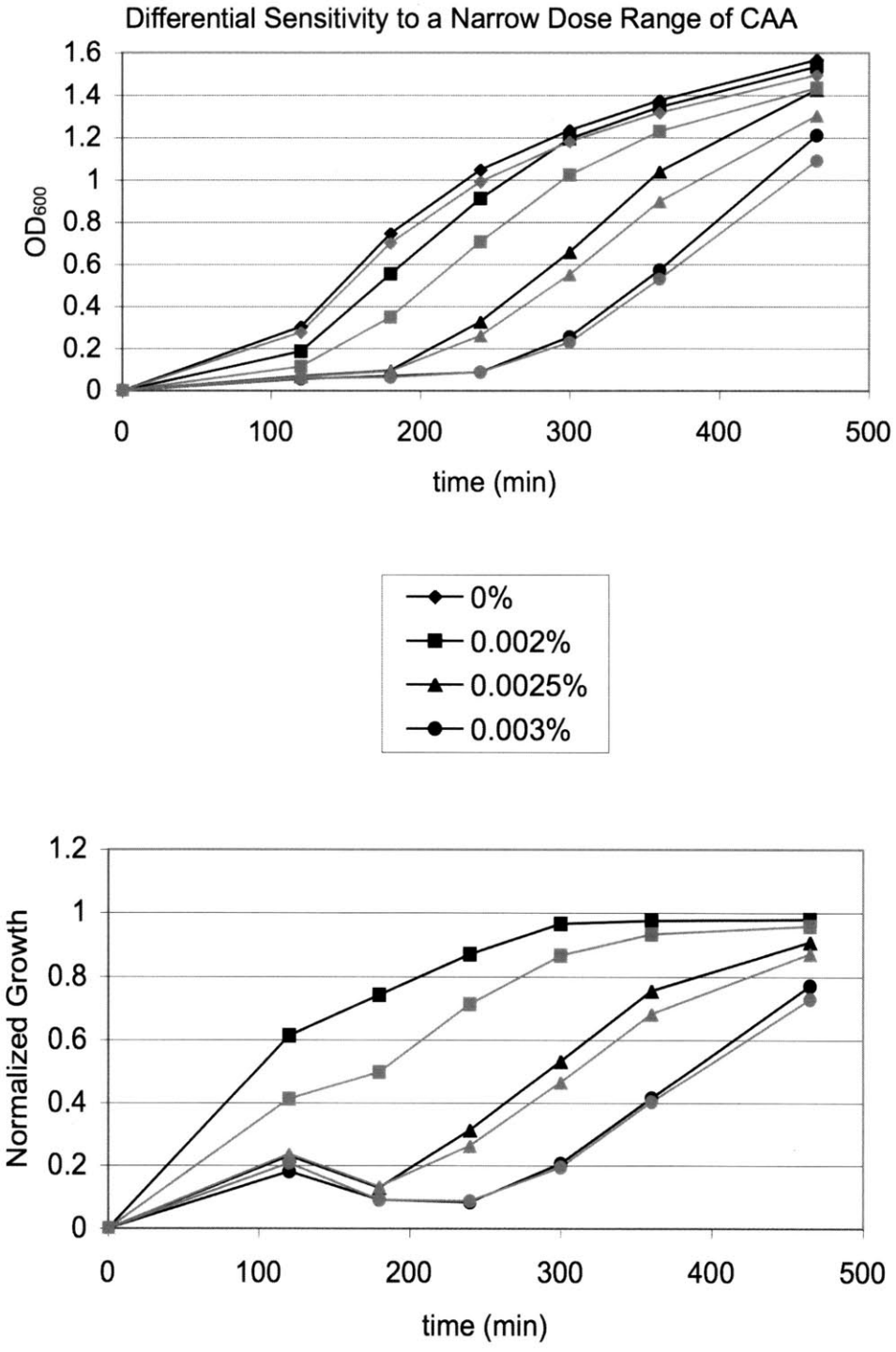
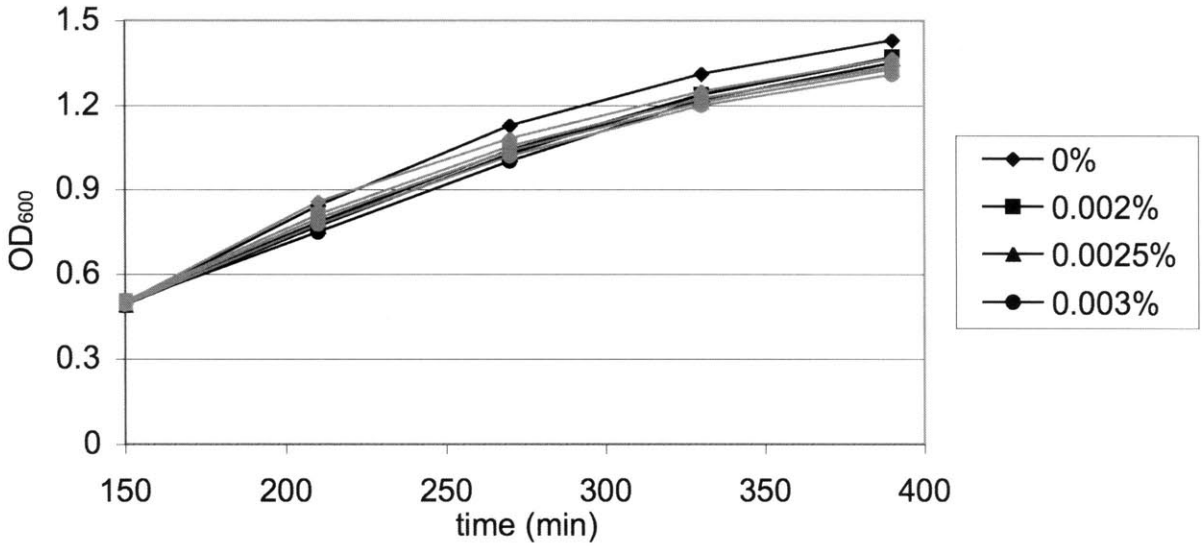


Figure A1.4: When treated at mid-log phase, HK81/22 (AlkB-proficient, black) and HK82/22 (AlkB-deficient, gray) are similarly resistant to doses of CAA that were toxic when administered immediately upon cultural dilution. Top panel: Optical density at 600 nm over time. Bottom panel: Relative growth of treated cultures as compared to each strain's untreated control.



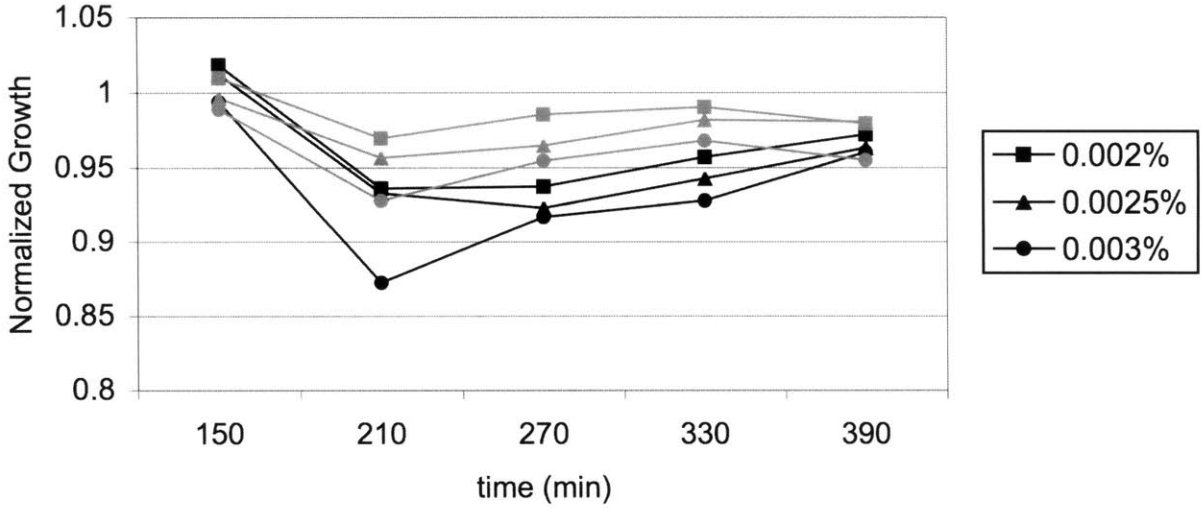
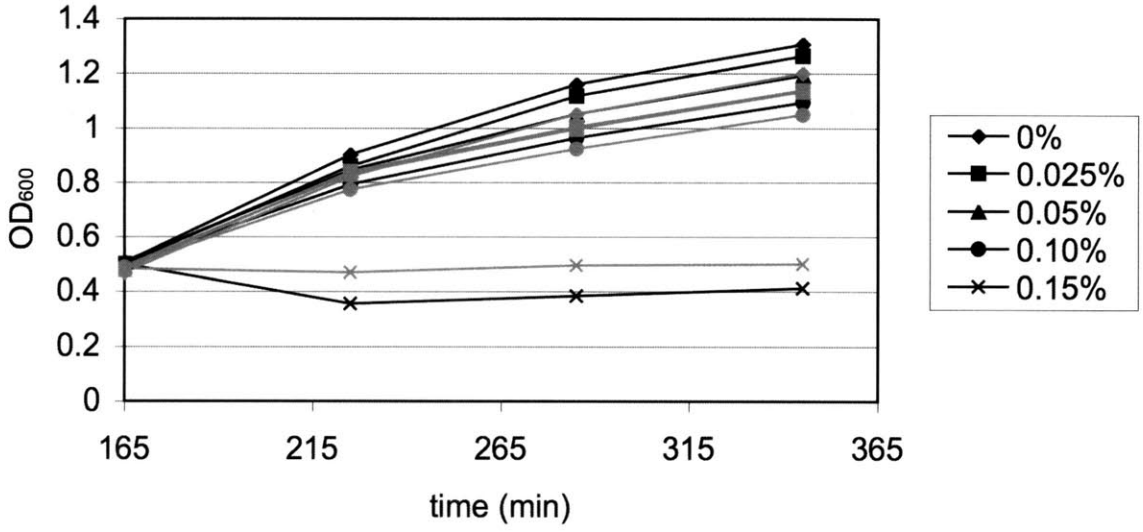


Figure A1.5: When treated at mid-log phase with MMS, HK81/22 (AlkB-proficient, black) and HK82/22 (AlkB-deficient, gray) show similar responses of resistance at low doses followed by a sharp drop in growth over a small increase in MMS concentration. Top panel: Optical density at 600 nm over time. Bottom panel: Relative growth of treated cultures as compared to each strain's untreated control.



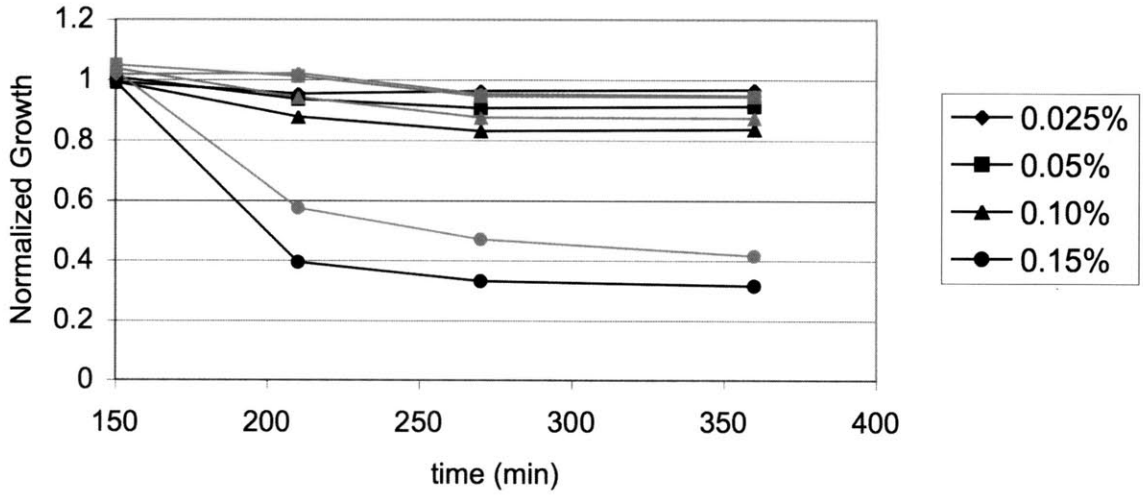


Figure A1.6: When treated at mid-log phase with MNNG, HK81/22 (AlkB-proficient, black) and HK82/22 (AlkB-deficient, gray) are similarly resistant to doses that generated robust expression of the adaptive response. Top panel: Optical density at 600 nm over time. Bottom panel: Relative growth of treated cultures as compared to each strain's untreated control.

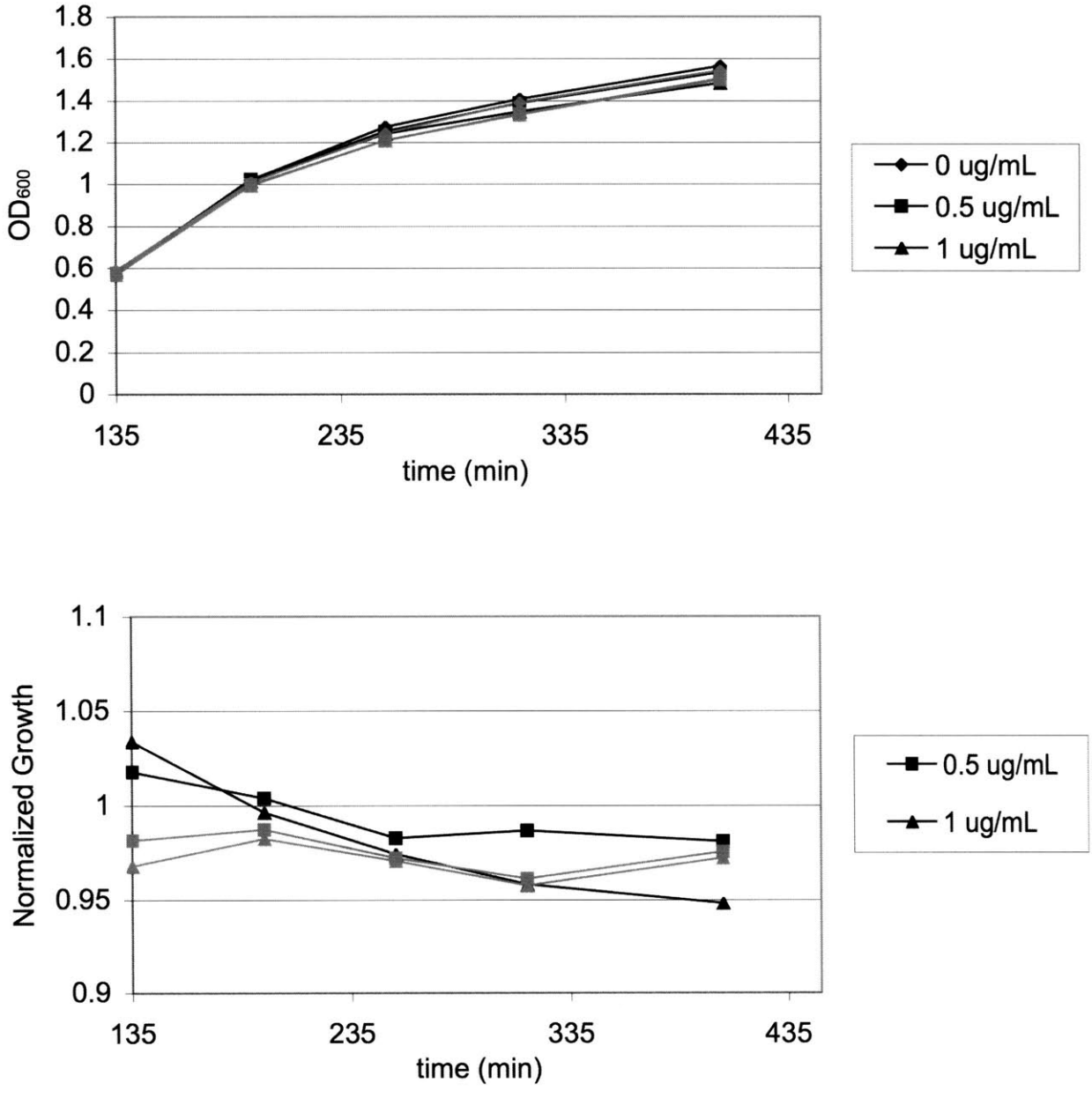


Figure A1.7: Treatment with SIN-1 at mid-log phase did not inhibit the growth of either HK81/22 (AlkB-proficient, black) or HK82/22 (AlkB-deficient, gray). Administration of higher doses (600 and 800 μM) was able to delay growth slightly (data not shown). Top panel: Optical density at 600 nm over time. Bottom panel: Relative growth of treated cultures as compared to each strain's untreated control.

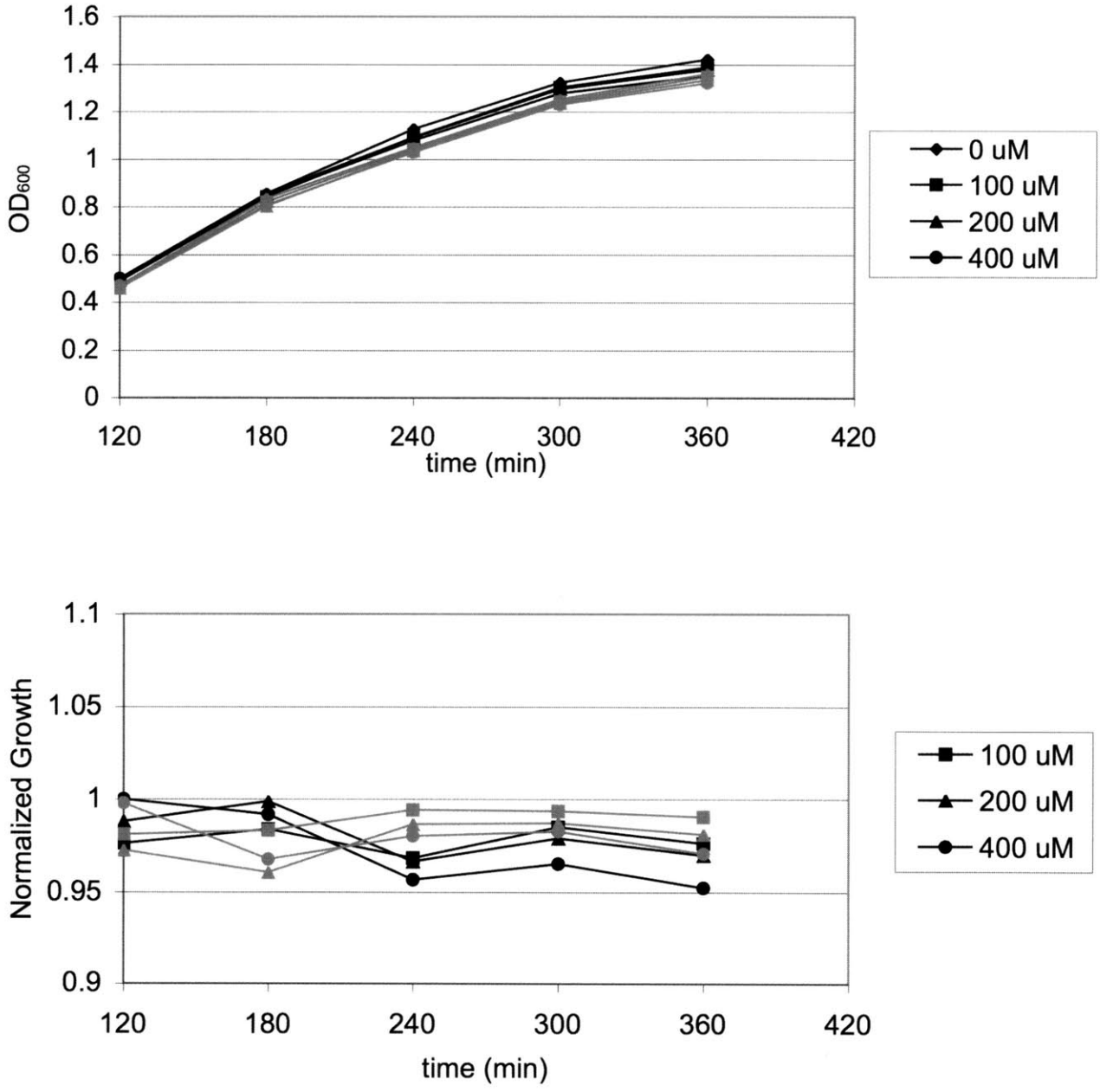


Figure A1.8: Comparison of the LacZ activity of untreated strains containing the control plasmid (pGW2620, "untreated 20") with untreated strains containing the ada-lacZ fusion (pGW2622, "untreated 22") shows a high level of background expression from the fusion-containing plasmid. Treatement with CAA does not substantially increase expression in either set of strains. HK81 derivatives (wild-type) shown in black, HK82 derivatives (AlkB-deficient) shown in gray.

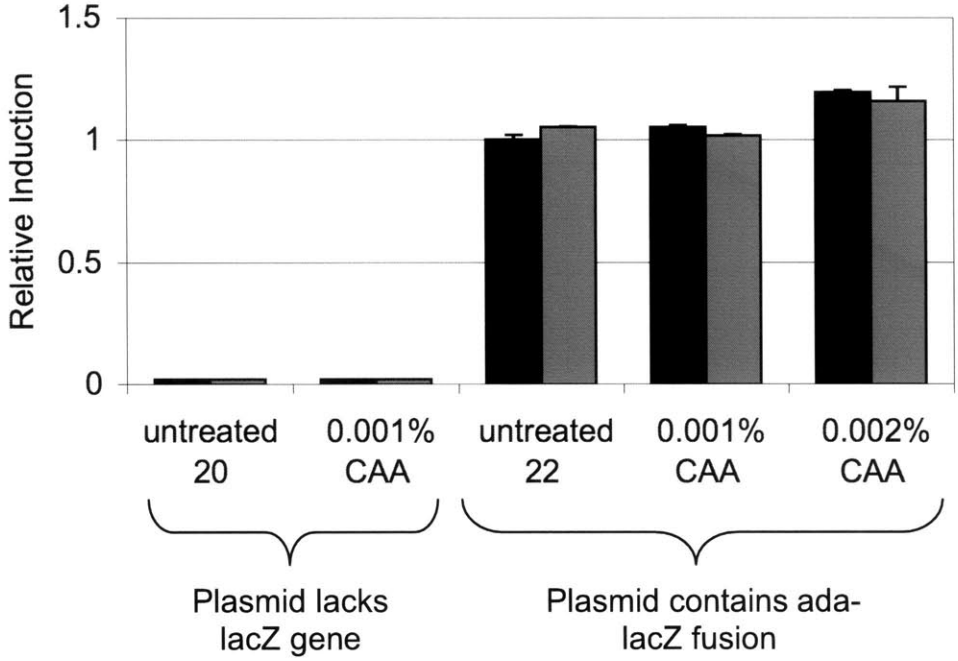


Figure A1.9: Treatment of fusion-containing plasmid strains (HK81/22 and HK82/22) with either CAA or MMS does not substantially increase the β-galactosidase activity above the untreated background levels.

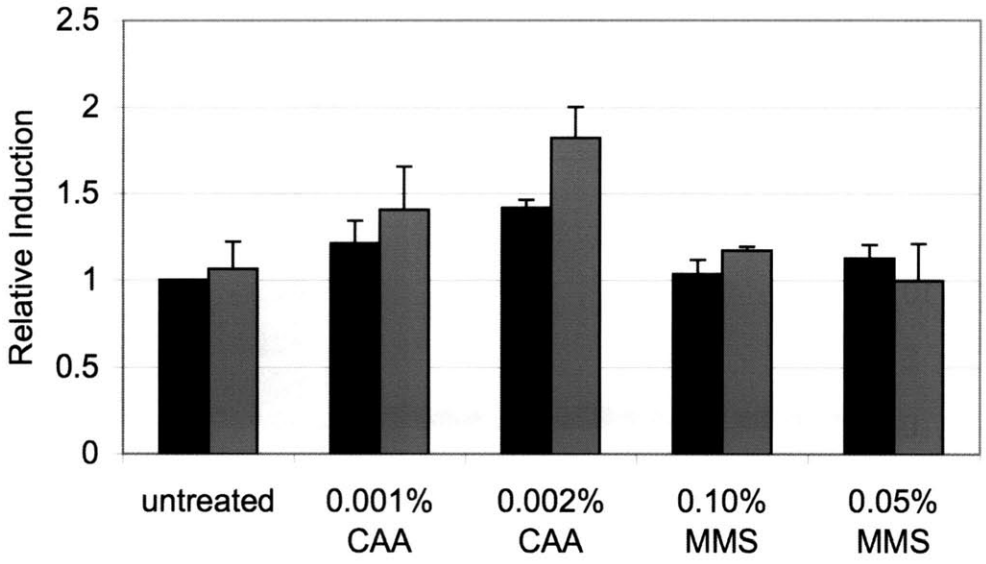


Figure A1.10: Treatment of ada-lacZ fusion-plasmid-containing strains (HK81/22, in black and HK82/22 , in gray) with MNNG caused a large induction of β-galactosidase activity in both the AlkB+ (in black) and the AlkB- (in gray) strains.

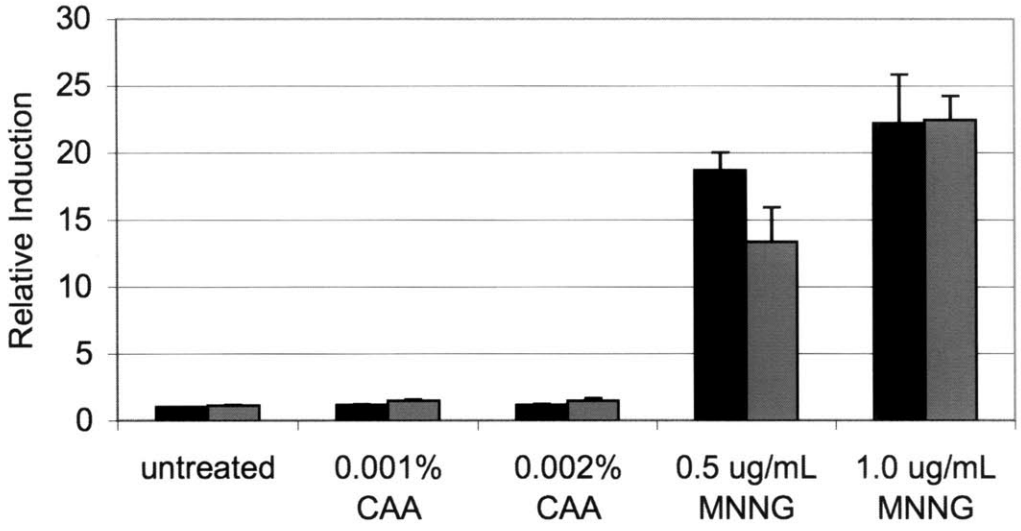
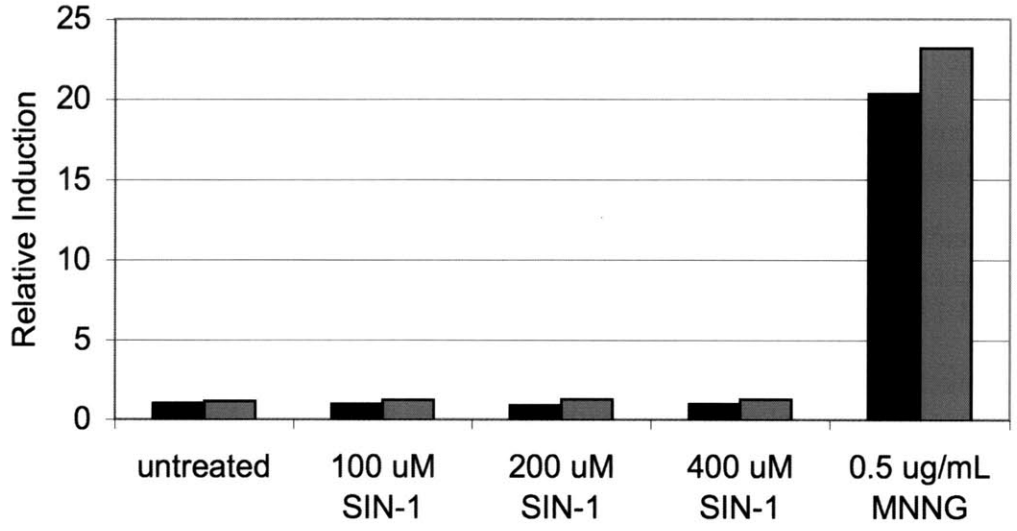


Figure A1.11: Treatment of fusion-plasmid-contianing strains (HK81/22, in black and HK82/22 , in gray) with the oxidizing agent SIN-1 did not induce β-galactosidase activity.



References

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

2. Schendel, P. F. and P. E. Robins. 1978. Repair of O6-methylguanine in adapted Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A* **75**:6017-6020.

3. Yamamoto, Y., M. Katsuki, M. Sekiguchi, and N. Otsuji. 1978. Escherichia coli gene that controls sensitivity to alkylating agents. *J. Bacteriol.* **135**:144-152.

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

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

6. Volkert, M. R. and D. C. Nguyen. 1984. Induction of specific Escherichia coli genes by sublethal treatments with alkylating agents. *Proc. Natl. Acad. Sci. U. S. A* **81**:4110-4114.

7. Sedgwick, B., P. Robins, and T. Lindahl. 2006. Direct removal of alkylation damage from DNA by AlkB and related DNA dioxygenases. *Methods Enzymol.* **408**:108-120.

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

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

10. Delaney, J. C. and J. M. Essigmann. 2004. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A* **101**:14051-14056.

11. Falnes, P. O. 2004. Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. *Nucleic Acids Res.* **32**:6260-6267.

12. Koivisto, P., P. Robins, T. Lindahl, and B. Sedgwick. 2004. Demethylation of 3-methylthymine in DNA by bacterial and human DNA dioxygenases. *J. Biol. Chem.* **279**:40470-40474.

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

14. Koivisto, P., T. Duncan, T. Lindahl, and B. Sedgwick. 2003. Minimal methylated substrate and extended substrate range of Escherichia coli AlkB protein, a 1-methyladenine-DNA dioxygenase. *J. Biol. Chem.* **278**:44348-44354.

15. Delaney, J. C., L. Smeester, C. Wong, L. E. Frick, K. Taghizadeh, J. S. Wishnok, C. L. Drennan, L. D. Samson, and J. M. Essigmann. 2005. AlkB reverses etheno DNA lesions caused by lipid oxidation in vitro and in vivo. *Nat. Struct. Mol. Biol.* **12**:855-860.

16. Mishina, Y., C. G. Yang, and C. He. 2005. Direct repair of the exocyclic DNA adduct 1,N6-ethenoadenine by the DNA repair AlkB proteins. *J. Am. Chem. Soc.* **127**:14594-14595.

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

18. Thomas, L., C. H. Yang, and D. A. Goldthwait. 1982. Two DNA glycosylases in Escherichia coli which release primarily 3-methyladenine. *Biochemistry* **21**:1162-1169.

19. Lindahl, T. 1976. New class of enzymes acting on damaged DNA. *Nature* **259**:64-66.

20. Laval, J. 1977. Two enzymes are required from strand incision in repair of alkylated DNA. *Nature* **269**:829-832.

21. McCarthy, T. V., P. Karran, and T. Lindahl. 1984. Inducible repair of O-alkylated DNA pyrimidines in Escherichia coli. *EMBO J.* **3**:545-550.

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

23. Privezentzev, C. V., M. Saparbaev, A. Sambandam, M. M. Greenberg, and J. Laval. 2000. AlkA protein is the third Escherichia coli DNA repair protein excising a ring fragmentation product of thymine. *Biochemistry* **39**:14263-14268.

24. Matijasevic, Z., M. Sekiguchi, and D. B. Ludlum. 1992. Release of N2,3-ethenoguanine from chloroacetaldehyde-treated DNA by Escherichia coli 3-methyladenine DNA glycosylase II. *Proc. Natl. Acad. Sci. U. S. A* **89**:9331-9334.

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

26. Rohankhedkar, M. S., S. B. Mulrooney, W. J. Wedemeyer, and R. P. Hausinger. 2006. The AidB component of the Escherichia coli adaptive response to alkylating agents is a flavin-containing, DNA-binding protein. *J. Bacteriol.* **188**:223-230.

27. Voet D.and Voet J.G. 1995. Biochemistry. Wiley.

28. Miller, J. 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

29. Lemotte, P. K. and G. C. Walker. 1985. Induction and autoregulation of ada, a positively acting element regulating the response of Escherichia coli K-12 to methylating agents. *J. Bacteriol.* **161**:888-895.

30. Shevell, D. E. and G. C. Walker. 1991. A region of the Ada DNA-repair protein required for the activation of ada transcription is not necessary for activation of alkA. *Proc. Natl. Acad. Sci. U. S. A* **88**:9001-9005.

31. Volkert, M. R., F. H. Gately, and L. I. Hajec. 1989. Expression of DNA damage-inducible genes of Escherichia coli upon treatment with methylating, ethylating and propylating agents. *Mutat. Res.* **217**:109-115.

32. Volkert, M. R., D. C. Nguyen, and K. C. Beard. 1986. Escherichia coli gene induction by alkylation treatment. *Genetics* **112**:11-26.

33. Nakabeppu, Y. and M. Sekiguchi. 1986. Regulatory mechanisms for induction of synthesis of repair enzymes in response to alkylating agents: ada protein acts as a transcriptional regulator. *Proc. Natl. Acad. Sci. U. S. A* **83**:6297-6301.

34. Volkert, M. R., L. I. Hajec, Z. Matijasevic, F. C. Fang, and R. Prince. 1994. Induction of the Escherichia coli aidB gene under oxygen-limiting conditions requires a functional rpoS (katF) gene. *J. Bacteriol.* **176**:7638-7645.

35. Landini, P., L. I. Hajec, L. H. Nguyen, R. R. Burgess, and M. R. Volkert. 1996. The leucine-responsive regulatory protein (Lrp) acts as a specific repressor for sigma s-dependent transcription of the Escherichia coli aidB gene. *Mol. Microbiol.* **20**:947-955.

36. Barbin, A. 2000. Etheno-adduct-forming chemicals: from mutagenicity testing to tumor mutation spectra. *Mutat. Res.* **462**:55-69.

37. Friedberg E.C. 2006. DNA Repair and Mutagenesis. AMS Press, Washington, DC.

APPENDIX 2:

Investigation of the Possible Combinatorial Function of AidB with AlkB in the Repair of Methylated DNA Bases

Studies of MNNG sensitivity of *aidB E. coli* as well as mutagenicity and polymerase bypass analysis of candidate substrate lesions in wild-type, *aidB, alkB,* and *aidB alkB E. coli.*

Introduction

The AlkB protein of the *Escherichia coli* adaptive response is an α-ketoglutarate/iron (II)-dependent dioxygenase that demethylates adducted bases through a mechanism of direct reversal involving hydroxylation of the unnatural methyl group, which falls off as formaldehyde, leaving the intact base (*1,2*). The versatility of this protein has been demonstrated by the addition of 3-ethylcytosine (*3*), 1-methylguanine (*3,4*), 3-methylthymine (*3-5*), 1-ethyladenine (*6*), and 1,N^6-ethanoadenine (*7*) to its substrate pool. Beyond its ability to repair alkylated bases, AlkB possesses a second mechanism through which it can restore another class of DNA lesions, the exocyclic ethenobases formed at high levels at sites of inflammation or by reaction of DNA with vinyl chloride or its metabolites (*8,9*). In this pathway, AlkB epoxidizes the double bond of the etheno-group, creating an intermediate that, with the addition of water, opens into a glycol-containing intermediate that decomposes to the repaired base and a glyoxal molecule (*8,9*). The enzyme is also capable of repairing many substrate bases in RNA (*10,11*), underscoring the flexibility of repair of alkylation damage by this oxidative strategy.

As the adaptable nature of AlkB has become clearer, additional possible substrates have been suggested for the enzyme. The list includes a variety of methylated lesions, some of which have other established repair pathways and some of which do not. The N^2-methylguanine lesion, which forms from the reaction of formaldehyde with DNA, is both toxic and mutagenic upon replication with the Klenow fragment of *E. coli* DNA polymerase I as well as various human polymerases *in vitro* (*12-14*). Although this lesion has been hypothesized to be an AlkB substrate (*15,16*), no repair pathway in DNA has yet been established. The analogous N^2-ethylguanine is formed upon exposure of DNA to acetaldehyde, the primary metabolite of ethanol, and its presence is thought to be a possible initiating event in alcohol-induced cancer (*17*). It has also been shown to be miscoding when replicated *in vitro* (*13,18,19*). Repair of these substituted guanine adducts by the AlkB protein would be an efficient means of removing deleterious lesions from the genome, and the work done here was intended to investigate whether the

methylated or ethylated guanine derivatives can be repaired by AlkB as well as to determine their toxic and mutagenic properties *in vivo*.

AidB is one of four proteins encoded by the alkylation-inducible adaptive response of *E. coli*. Despite decades of study, the function of the protein remains elusive. The gene has been shown to be expressed not only in response to the presence of methylated Ada protein (the canonical inducer of all four adaptive response genes), but also by growth in anaerobic conditions or in the presence of acetate at a slightly acidic pH (*20*). Induction by anaerobiosis is independent of cellular ada status (*20*), but requires a functional rpoS gene (*21*). RpoS encodes an alternative RNA polymerase sigma factor, σ^s, that is normally expressed at low levels in exponentially-growing cultures and at higher levels as *E. coli* enter the late-logarithmic and stationary phases of growth. Sigma factors function to modulate the affinity of the RNA polymerase holoenzyme toward particular promoters, as each sigma factor displays a different order of affinities for various promoters. Expression of the alternative σ^s factor in later growth phases enables replacement of the traditional σ^{70} factor, inducing expression of stationary phase genes and halting expression of genes involved in rapid growth. The alternative factor is additionally involved in cellular responses to carbon starvation, oxidative damage, and osmotic shock (*22*). The σ^s-dependent expression of aidB is, however, not induced merely by the reduction in growth rate that accompanies entry into stationary phase, but rather requires conditions of limited oxygen (*21*). It seems that the accumulation of organic acids and the resulting decrease in cellular pH play some role in the anaerobic expression of the gene, as artificial decrease of the pH to 6.8 by the addition of acetate causes production of AidB at similar levels to that seen in cells grown without aeration (*23*).

Studies of an rpoS knockout have revealed that the more common σ^{70} factor is also able to transcribe the aidB gene, albeit much less efficiently than the σ^s factor. Furthermore, in the absence of transcription factors, aidB is transcribed efficiently *in vitro* by the σ^s RNA polymerase holoenzyme but at only a very low level by the σ^{70} form (*24*). Levels of transcription by holoenzymes containing either sigma factor are

significantly increased in the presence of methylated Ada, which interacts with a negatively-charged patch that is conserved between the two sigma factors (*22*).

Strains lacking functional AidB have shown inconsistent phenotypes, with *aidB:lacZ* insertion mutants falling into two phenotypic categories: those with alkylation sensitivity identical to that of wild-type cells, and those that demonstrate resistance to both the lethal and mutagenic effects of the S_N1 alkylating agent *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG), but no alteration in sensitivity to other alkylating agents (*25*). This observation led to the hypothesis that perhaps AidB is involved in the metabolic activation of MNNG (*26*) and that loss of its function, which occurred in only a subset of insertion mutants, renders the drug innocuous. Paradoxically, however, over-expression of the AidB protein in *E. coli* also produced cells that were resistant to mutagenesis by MNNG, prompting the opposing hypothesis that the protein facilitates the metabolic inactivation of the methylating agent. Landini et al. (1994) examined the data in an attempt to resolve the apparent contradiction, and found that the *aidB:lacZ* insertion mutants studied all affected only the terminal third of the gene. The observed phenotypes would be explained if expression of the first two-thirds of the gene produced a functional protein that is either stabilized or activated in the resistant strains and similar to wild-type in the non-resistant strains. However, a complete knockout of the aidB gene has also been shown to provide MNNG resistance (M. Volkert, unpublished observations), a paradox that has yet to be explained. Here, a systematic comparison of several sets of isogenic "wild-type" and complete *aidB* knockouts and one inducible over-expressing strain was conducted in an attempt to elucidate the effect conferred by complete lack of AidB as well as whether the previously-reported phenotypes are reproducible and consistent among laboratories.

In an attempt to determine the possible function of this protein, the predicted amino acid sequence of the 60.5 kDa AidB was entered into a homology search of the Swiss Protein and GenBank-EMBL databases. While no homology was found with any known DNA repair protein, significant homology with several mammalian acyl coenzyme A (CoA) and isovaleryl-CoA dehydrogenases was uncovered. Because the protein sequence indicates 24.6% homology with human isovaleryl-CoA dehydrogenase

(IVD) precursor (*25*), an enzyme involved in leucine metabolism (Figure A2.1) in mammalian cells, extracts containing AidB were tested for IVD activity *in vitro*. In an isovaleryl-CoA reduction assay, IVD function was detected in crude extracts of cells over-expressing AidB, but not in extracts of cells that were not over-expressing the protein. Discovery of the IVD activity of AidB led to the hypothesis that the protein, like other acyl-CoA dehydrogenases, plays a role in metabolism and energy generation during times of limiting oxygen availability (*25*). This hypothesized involvement in metabolism would explain the observed induction of the protein in an ada-independent manner in response to anaerobiosis, however it does not provide a simple explanation for the ada-dependent induction of the protein in response to methyl damage.

The homology that AidB shares with human isovaleryl-CoA dehydrogenase has led to investigations into the possible role of AidB in leucine metabolism. The ability of leucine to stimulate AidB production *in vivo* has been established (*24*) and shown to be specific to the AidB promoter, and not the alternative sigma factor (σ^S). Leucine responsive protein (Lrp), which functions as a transcriptional regulator of multiple genes and most operons involving leucine in *E. coli*, binds to its target promoters and activates or inhibits transcription of the target gene. The consequence of Lrp binding depends on other factors, and leucine itself often modulates Lrp activity. Binding of pure Lrp protein to the aidB promoter was detected in a DNAseI-protection assay, but vanished with the addition of leucine or methylated Ada to the reaction due to competitive binding at the promoter region (*24*). *In vitro* transcription experiments revealed that Lrp binding functions to prevent AidB transcription and that the presence of either leucine or methylated Ada allows transcription to take place.

Additionally, *E. coli* lacking a functional lrp gene were found to express AidB during aerobic growth in a phase-dependent manner, an expression pattern characteristic of other rpoS-dependent genes. This evidence implicates lrp in a regulatory role, keeping AidB expression low during times of rapid growth or leucine starvation, and allowing levels to increase when cells are not dividing quickly and excess leucine is present. Such regulation is characteristic of proteins involved in the breakdown of certain amino acids, and the fact that AidB is subject to this type of negative regulation lends support to its

hypothesized role in leucine metabolism. Other mechanisms of AidB regulation are clearly at work, however, as the lack of a functional lrp gene does not affect induction by methylating agents, and growth of lrp-deficient cells in anaerobic conditions resulted in significantly higher levels of AidB expression than growth in oxygenated conditions, indicating that some method of AidB induction beyond simple de-repression is involved. Much evidence indicates that AidB regulation is complex, and without understanding of the protein's function, the reasons for such intricate control are unclear.

Recent studies of purified AidB protein revealed the presence of stoichiometric amounts of FAD and an ability to bind double-stranded DNA (*27*). It was hypothesized that perhaps the protein is involved in the repair of methylated bases (Figure A2.5) through a dehydrogenase-based mechanism (Figure A2.6). The 3-methylcytosine, N^2-methylguanine, and N^2-ethylguanine lesions are example adducts that could be repaired if this hypothesis were to hold true. Consequently, the bypass and mutational characteristics of these three lesions were examined in four strains of *E. coli*, wild-type, *alkB*, *aidB*, and *alkB aidB* to determine whether evidence to support this hypothetical role for AidB could be uncovered.

In addition to a possible role in dehydrogenation of methylated bases, a second idea has been put forward regarding the putative function of AidB in *E. coli*. The 1,N^6-ethanoadenine adduct is formed in DNA upon reaction with the bi-functional alkylating agent 1,3-bis(2-chloroethyl)-1-nitrosourea and is excised only inefficiently by the 3-alkyladenine DNA glycosylases of humans or *E. coli* (*28*). The co-regulation of AidB and AlkB paired with the newly-discovered activity of AlkB in repairing the analogous 1,N^6-ethenoadenine adduct led to the proposal that perhaps AidB is able to dehydrogenate EA, turning it into εA, which is an efficient substrate for both the AlkB and AlkA enzymes of the adaptive response (Figure A2.6). Accordingly, the ability of *E. coli* to replicate EA in the presence or absence of the AlkB and/or AidB proteins was assessed, as was the mutational spectrum generated in each context.

The work described here was intended to shed light on the role of AidB in DNA repair in *E. coli*, possibly providing an explanation for the multiple layers of regulation to

which it is subject as well as an understanding of the function of the protein in defending the human genome against DNA damage and mutation.

Materials and Methods

Bacterial Strains. *E. coli* strains used for MNNG toxicity assays included HK81 (as AB1157, but *nalA*; wild-type), MV5933 (as HK81, but Δ*aidB35::Tet*; AidB-deficient), MV5934 (as HK81, but Δ*aidB35::Tet alkB22*; AlkB- and AidB-deficient), MV1161 (AB1157 derivative; wild-type), MV5925 (as MV1161, but Δ*aidB::TetR*; AidB-deficient), MG1655 (K-12 derivative; wild-type), MV5924 (as MG1655, but Δ*aidB::TetR*; AidB-deficient), and MV3590 (K-12 derivative carrying pMV435 (pTrc99a-aidB); AidB-overexpressing). Strains MV5933, MV5934, MV5925, MV5924, and MV3590 were generous gifts of Michael Volkert. *E. coli* strains used for the lesion bypass and mutagenicity assays were HK81 (wild-type) and HK82 (as HK81, but *alkB22*; AlkB-deficient), MV5933 (AidB-deficient), and MV5934 (AidB/AlkB deficient).

MNNG toxicity assays. Cultures of each strain to be analyzed were grown in 10 mL LB overnight at 37 °C to saturation with antibiotic as appropriate (6 μg/mL tetracycline for MV5933, MV5934, MV5924, and MV5925, and 50 μg/mL carbenicillin for MV3590). One hundred μL of each saturated culture were diluted into 10 mL fresh LB again with selection and with the addition of 10 μL 0.1 M IPTG to the MV3590 culture. Cells were grown with aeration at 37 °C for approximately 2 h (to an OD_{600} of ~0.4) at which point they were pelleted by centrifugation. Isolated cells were resuspended in 5 mL M9 minimal salts and 1 mL aliquots were distributed into 1.5 mL Eppendorf tubes. The appropriate quantity of MNNG was added and cells were incubated for 30 min at 37 °C. Drug exposure was halted by dilution of cultures 1:100 into fresh M9 medium on ice. Subsequent serial dilutions were made (10^{-3} to 10^{-6}) and 100 μL of the appropriate dilutions were plated in triplicate on LB agar and allowed to grow overnight. The next day, surviving colonies were counted and relative survival calculated as a percentage of each strain's untreated control.

Oligonucleotides. A set of sixteen-mer oligonucleotide inserts of sequence 5'-GAAGACCTXGGCGTCC-3' were synthesized using phenoxyacetal-protected phosphoramidites and capping reagents from Glen Research (*3*). Approximately 50 mg of EA phosphoramidite (Chemgenes), which has previously been site-specifically incorporated into an oligonucleotide (*29*), were used during a 30 min coupling with the solid support. Following synthesis, the resin was incubated with concentrated ammonium hydroxide for one hour prior to lyophilization. The 16-mer insert oligonucleotide 5'-GAAGACCTXGGCGTCC-3' (X=EA) was purified by anion exchange and reversed phase HPLC, and the molecular weight was verified by MALDI-TOF mass spectrometry (4916.21 observed, 4916.22 calculated). Sixteen-mer oligonucleotides of the same sequence, but where X = εA, T, m2G, e2G, and m3C were synthesized and purified as described (*8*). The m2G and e2G oligonucleotides were made using a 2-fluoro-O^6-(trimethylsilylethyl)-2'-deoxyinosine phosphoramidite (ChemGenes, Wilmington, MA). After the oligonucleotide was reacted with methylamine or ethylamine, this convertible nucleoside approach (*30*) yielded the m2G or e2G oligonucleotide, respectively. Scaffold oligonucleotides (of sequence 5'-GGTCTTCCAC TGAATCATGGTCATAGC-3' and 5'-AAAACGACGGCCAGTGA ATTGGACGC-3') used in genome construction were obtained from Integrated DNA Technologies.

Genome construction and concentration normalization. Single-stranded M13 genomes containing a site-specific lesion of interest were constructed as previously described (*7,31*). Briefly, wild-type M13mp7L2 genome was digested with EcoRI, and annealed to two scaffolds to correctly position the phosphorylated 16-mer oligonucleotide insert, which was then ligated into the resulting gap. After digestion of the scaffolds by T4 DNA polymerase, genomes were purified by PCI extraction, desalted using Centricon-100 spin-dialysis columns (Millipore), and stored at -20 °C. Concentrations were normalized across the six genomes by annealing the scaffolds used for construction, digesting with *HinFI*, and dephosphorylating and radiolabeling the fragment. Following a final digestion with *HaeIII*, products were separated by denaturing PAGE and

quantified using phosphorimagery (Molecular Dynamics). Concentrations were equalized by the addition of water.

Preparation of electrocompetent cells. Three baffled flasks each containing 150 mL of LB media were inoculated with 1.5 mL saturated overnight culture of the strain to be transformed (HK81, wild-type; HK82, AlkB deficient; MV5933, AidB deficient; or MV5934, AlkB and AidB deficient) and grown with shaking at 37 °C to an OD_{600} of about 0.5 (~2.5 h). The three cultures were then pelleted and the cells pooled, followed by 3 washes at 4 °C each using 175 mL water. The final resuspension of cells was in 6 mL 10% glycerol, after which the cells were transformed within 2 h.

Assessment of lesion bypass. Normalized, lesion-containing genomes were mixed to a ratio of ~10:1 with an internal standard genome containing a T instead of a lesion in a 19-mer insert. One hundred μL of electrocompetent cells were mixed with 4.8 μL of genome mixture, transferred to a 2 mm gap electroporation cuvette on ice, and exposed to ~2.5 kV and ~125 Ω. After electroporation, cells were immediately transferred to 10 mL fresh LB and grown for 6 h at 37 °C. Cells were then pelleted, and the progeny-containing supernatant was removed and stored at 4 °C. The number of successful transformation events was measured by immediately plating 10 μL of the 10 mL culture onto a lawn of NR9050 indicator bacteria. Each independent event resulted in the formation of a plaque, enabling quantification of initial events and ensuring a robust progeny population for analysis. All electroporation procedures reported here produced $>10^4$ events, with the vast majority resulting in $>10^5$.

One hundred μL progeny phage (~10^9) were re-grown for 6 h in 10 mL LB inoculated with 10 μL of an overnight culture of SCS110 *E. coli* (Stratagene). Only viable progeny phage, i.e. those previously replicated within a cell, would be amplified, while any unincorporated genomes remaining in the medium from the electorporation step would not be affected. This selective amplification effectively increases the signal to noise ratio by several orders of magnitude. Cells were again pelleted and progeny-containing supernatant was isolated and stored at 4 °C.

Single-stranded phage DNA was isolated using a QIAprep Spin M13 kit (Qiagen). PCR amplification was performed using primers (5'-YCAGCTATGACCATGATTC AGTGGAAGAC-3' and 5'-YCAGGGTTTTCCCAGTCACGACGTTGTAA-3' where Y signifies amino modification of the 5' nucleotide) as described (*3,8,11,31*). After PCI extraction and removal of salt and organic residue with Sephadex G50 fine resin (Amersham), the PCR product was digested with *BbsI* and dephosphorylated with shrimp alkaline phosphatase. T4 polynucleotide kinase was used to 5'-radiolabel the resulting fragment, which was further digested by *HaeIII*. Radiolabeled products separated by PAGE were quantified by phosphorimagery. Comparison of the intensity of the 18-mer band (corresponding to the lesion-containing genome) to that of the 21-mer band (representing the internal standard) gave a measure of lesion bypass.

Mutational spectrum analysis. The mutational specificities and frequencies of the six lesions were determined by following the lesion bypass assay up to the PCR step. A different set of primers (5'-YCAGCTATGACCATGATTCAGTGGAAGAC-3' and 5'YTGTAAAACGACGGCCAGTGAATTGGACG-3') were used to obtain specific amplification of progeny from genomes that originally contained the lesion. The second primer is mismatched when annealed to the internal standard, which is thus not amplified. Samples were processed as in the lesion bypass assay through the PAGE step, at which point the 18-mer bands were excised from the gel, and the DNA extracted by crushing and soaking. Notably, the restriction enzyme incises the DNA to expose the base that corresponds to the original site of the lesion at the 5' end, which is then radiolabeled. This site-specific labeling allows for precise inquiry into the base present at the lesion site after replication. The extracted DNA is desalted on Sephadex G50 fine resin (GE Healthcare), lyophilized, and digested to free 5'-nucleotide monophosphates by nuclease P1 (United States Biological). The digested nucleotides are separated on a polyethyleneimine thin layer chromatography (TLC) plate, and the resulting spots, each representing a normal base, are quantified by phosphorimagery. Because only the 5'-base was radiolabeled, the TLC represents the base composition at the lesion site after replication by *E. coli*, enabling calculation of both mutation frequency and specificity.

Results

MNNG toxicity. The phenotypes of several sets of isogenic wild-type/*aidB* pairs were examined to determine what effect a lack of functional AidB protein has on sensitivity to MNNG. The HK81 set, consisting of HK81, MV5933 (an *aidB* derivative of HK81), and MV5934 (the *aidB alkB* double mutant derived from HK82), showed identical sensitivity among all strains to treatment with MNNG (Figure A2.2). All three strains were resistant to the toxic effects of the drug. In contrast, two other wild-type/mutant pairs MG1655 and MV5924 (an *aidB* strain derived from MG1655) and MV1161 and MV5925 (an *aidB* derivative of MV1161) showed some increased resistance in the AidB mutant (Figure A2.3). The magnitudes of observed resistance differed, however, with MV5924 being only slightly better able than MG1655 to survive, while MV5925 exhibited half a log of increased survival. Finally, over-expression of AidB from plasmid pMV435 increased survival still more than either of the two knock-out strains (Figure A2.4). As seen previously (*25,26*), both too little and too much AidB protein seems to render *E. coli* more resistant to the toxicity induced by MNNG.

Bypass of cyclic adenine adducts *in vivo* in *aidB E. coli.* In wild-type *E. coli* (black bars), both the EA and the εA adducts are replicated as easily as a control thymine (Figure A2.7). In cells lacking AlkB, however (dark gray bars), EA reveals itself to be potently toxic, as it is replicated only ~14% as well as T. In cells lacking AidB (light gray bars), the bypass efficiency of EA rises to ~83%, suggesting that AidB plays little, if any, role in mitigating the block that unrepaired EA poses to replication. In *E. coli* lacking both AlkB and AidB (white bars), EA produces a bypass efficiency of just ~11%, which is not significantly different from that seen in *alkB* cells. These data implicate AlkB, but not AidB, in the reduction of toxicity of EA, most likely through repair or metabolism of the adduct. The εA lesion produces similar results, showing very low rates of bypass in *alkB* (~10%, dark gray bars) and *alkB aidB* (5%, white bars) strains. The replication efficiency of εA in the AidB single mutant was again intermediate, as bypass was observed ~55% of the time. AlkB is again shown to be critical to the

alleviation of adduct toxicity, but the role of AidB in the reduction of toxicity is unclear for this lesion as well, as a slightly lower rate of bypass is observed, but the magnitude of the reduction is minute.

Mutagenicity of cyclic adenine adducts *in vivo* in *aidB E. coli.* EA is only weakly mutagenic in any of the cell strains examined, as >96% A was seen at the lesion site in *alkB* strains (in dark gray), >93% A was observed in *aidB* strains (light gray), and >86% A in the *alkB aidB* double mutant (white; Figure A2.7). Low levels of A to C (~2%), A to T (~1%), and A to G (~1%) mutations were observed in the *alkB* cells. While the *aidB* strain produced mutations of ~1% A to C, ~4% A to T, and ~1% A to G, the *alkB aidB* cells generated a mutational spectrum of ~6% A to C, ~5% A to T, and ~3% A to G. The mutagenic frequencies in any of the four strains do not speak to the miscoding nature of EA, but the small differences observed could indicate a weak ability for the AlkB and AidB proteins to somehow moderate the coding specificity of the adduct.

Replication of the structural analog εA in the four *E. coli* strains produced more variation in the observed mutation pattern. In wild-type cells, εA was poorly mutagenic, coding as >97% A. The remaining 3% was distributed equally among A to C, A to T, and A to G mutations. In the absence of AlkB (dark gray), εA is much more miscoding, generating ~10% A to C, ~20% A to T, and ~8% A to G mutations at the lesion site. Loss of AidB is much less deleterious (light gray), as >94% A remains at the site, along with ~1% C, ~4% T, and ~1% G. The double mutant is very similar to the *alkB* strain, with ~55% A, ~13% C, ~21% T, and ~11% G seen after replication of the lesion. Here, AlkB is implicated in combating the mutagenic nature of εA *in vivo*, while no evidence for a similar role by AidB is observed.

Bypass of methylated bases *in vivo* in *aidB*, *alkB*, and *aidB alkB E. coli.* The 3-methylcytosine adduct, which is formed at high levels in single-stranded nucleic acids upon exposure to methylating agents, is non-toxic when replicated in wild-type cells (Figure A2.8). In strains lacking AlkB (dark gray), however, m3C is a strong block to replication, as it is bypassed only ~20% as often as T. In *aidB* cells (in light gray), a

more moderate decrease in replication efficiency is observed (~50%). Bypass in the double mutant is inefficient (white bars, about 12%) and of the same magnitude as that seen in the *alkB* strain. While AlkB is clearly implicated in reducing the toxicity of m3C *in vivo*, the data are suggestive of a possible role played by AidB as well.

The substituted guanine adduct, m2G, is not efficiently replicated in *E. coli* of any of the genotypes examined, as bypass was observed ~35, 32, 32, and 25% as often as bypass of T in wild-type (black), *alkB* (dark gray), *aidB* (light gray), and *alkB aidB* (white bars), respectively. Poor bypass regardless of cell strain indicates that neither protein is involved in the repair of m2G.

The analogous e2G is also replicated poorly in all four strains used. Efficiencies were ~45, 35, 39, and 36% that of T in wild-type (black), *alkB* (dark gray), *aidB* (light gray), and *alkB aidB* (white bars) strains, respectively. Again, no support for AlkB- or AidB-mediated repair of this lesion was uncovered.

Mutational frequency and specificity of methylated bases *in vivo* in wild-type, *aidB*, *alkB*, and *aidB alkB E. coli.* In wild-type cells, m3C is weakly mutagenic (Figure A2.8), coding mainly as C (~95%), but also occasionally as T (~3%), G (~1%), and A (~1%). The mutagenicity of m3C increases in the absence of AlkB (dark gray bars) to ~74% C, 13% T, 5% G, and 9% A. The absence of AidB (light gray bars) creates a mutational spectrum almost identical to that generated in wild-type cells, specifically ~93% C, 3% T, 2% G, and 2% A. The mutational specificity of m3C in cells lacking both AlkB and AidB (white bars) is very similar to that seen in the AlkB single mutant (~67% C, 17% T, 5% G, and 11% A), indicating that although AlkB is capable of rendering m3C less mutagenic, the presence of AidB does not alter greatly the mutagenicity of the adduct.

The mutational spectrum generated by replication of m2G in wild-type cells is generally non-mutagenic, with ~94% G, 2% C, 0% T, and 4% A appearing at the lesion site. The spectra produced in *alkB* and *aidB* cells show a slight increase in mutagenicity and are very similar to one another (*alkB*: ~89% G, 3% C, 1% T, and 7% A; *aidB*: ~89 %

G, 3% C, 2% T, and 6% A). When m2G is replicated in cells lacking both proteins, another minor increase in mutagenicity is observed (~86% G, %4 C, 1% T, and 9% A).

Upon replication, the e2G adduct codes almost exactly as m2G, showing poor mutagenicity in wild-type cells (~94% G, 2% C, 0% T, and 4% A), a very slight increase in miscoding in *alkB* cells (~92% G, 2% C, 0% T, and 5% A), a slightly larger increase in *aidB* cells (~90% G, 3% C, 2% T, and 5% A) and a more substantial increase in the doubly-deficient *alkB aidB* strain (~84% G, 6% C, 2% T, and 8% A).

Discussion

Studies of the MNNG sensitivity of *E. coli* strains lacking the AidB protein have been conducted in the past in an attempt to determine an *in vivo* function for the protein (*25,26*). In those studies, strains lacking AidB were either equally or less sensitive than their wild-type counterparts to killing by MNNG. Paradoxically, however, strains deficient in or over-expressing AidB both exhibited enhanced resistance to MNNG toxicity, complicating hypothetical explanation of the protein's function. Here, again, results were seemingly contradictory. One deletion strain (MV5933) was exactly as sensitive as its wild-type analog, showing no phenotype due to the lack of AidB expression. Others (MV5925 and MV5924), however, derived from alternative wild-type parental strains, showed increased resistance, as has been seen in the past. Further, over-expression of the protein from the inducible MV3590 strain caused even greater resistance to MNNG. Yet again, no clear role for AidB in enhancement or mitigation of methylating agent-induced toxicity has emerged from examination of the consequences of gene deletion. Perhaps further examination of the purified protein will reveal a mechanism of action possibly involved in the demethylation of damaged DNA, RNA, or protein.

The hypothesis (outlined in Figure A2.6) that AidB serves to convert a lesion that is refractory to repair (EA) to one that is removed from DNA by at least two mechanisms (εA) was explored by monitoring the ability of cells possessing or lacking the AlkB and AidB proteins to replicate a single-stranded M13 genome containing a single site-specific lesion and to read that base correctly. *E. coli* lacking either protein were less able to

replicate either exocyclic lesion, although the phenotype conferred by loss of AlkB was much more drastic than that produced by loss of AidB in both cases. Perhaps the AidB protein is involved in some way in assisting the cell during replication of toxic bases, but AidB is not clearly implicated in the repair of either base lesion.

The mutagenic spectrum generated by replication of the exocyclic adenine adducts is similarly inconclusive. The EA adduct is poorly mutagenic in any of the contexts examined, coding mainly as A regardless of the presence of the AlkB or AidB proteins. The slight increase in mutagenicity seen in the double mutant is most likely artifactual, as the responsible increase in C is seen in all samples, including the T control. The identical coding of EA in all four strains provides no evidence that either protein is involved in repair of the adduct. The coding specificity of εA, on the other hand, does differ by strain. Although poorly-mutagenic in wild-type cells, εA is ~35% mutagenic in strains lacking AlkB, a phenomenon also seen a previous study (*8*). Lack of the aidB gene, however, does not seem to increase the mutagenicity of εA substantially, as the spectrum is almost identical to that created in wild-type cells. The spectrum corresponding to replication in the double mutant is extremely similar to the AlkB single mutant, suggesting that while the AlkB protein is important to the repair of εA, the presence of AidB is inconsequential. Thus, the bypass and mutational data shown here do not support the conversion hypothesis of AidB function. Additionally, the observed alleviation of EA toxicity through direct metabolism by AlkB represents a mechanism of removal of this toxic lesion from DNA despite the formation of a secondary adduct (*7*), and renders elimination of EA by conversion into εA unnecessary.

The proposal that AidB could be a dehydrogenase that repairs methylated bases through the mechanism outlined in Figure A2.6 was tested similarly to the conversion hypothesis, by monitoring the bypass and mutagenicity of candidate substrate lesions in *E. coli* with and without native protein. Three alkylated bases that would fit this hypothesis (proposed in (*27*)) were examined in four strains of *E. coli* differing in their AidB and AlkB statuses. 3-Methylcytosine was replicated as easily as thymine in wild-type cells, but at only ~20% the rate of thymine in cells lacking AlkB, suggesting that AlkB removes the methyl group that acts as a block to DNA polymerase. This result is

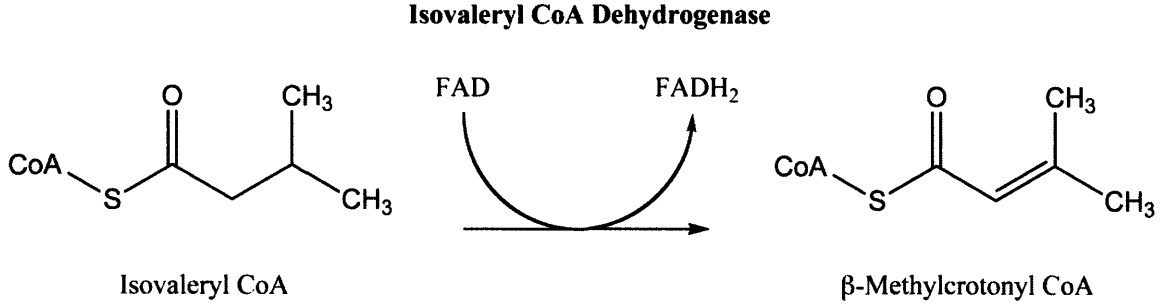
consistent with the establishment of m3C as one of the canonical substrates for the AlkB protein (*1,2*). In cells lacking only AidB, replication of m3C proceeds ~50% of the time, suggesting that AidB is to some extent capable of alleviating the toxicity of m3C, but that it is not as efficient as AlkB. The double mutant renders m3C slightly more toxic than the *alkB* strain, further supporting a possible minor role for AidB in reducing the block to replication represented by unrepaired m3C. N^2-Methylguanine was a poor substrate for replication in all of the strains examined, suggesting that neither protein was capable of reducing the potent toxicity of the intact lesion. Replication of an N^2-ethylguanine-containing genome produced very similar results, with consistently low bypass efficiencies observed in all four strains. These results agree with the previous observation (*5*) that N^2-ethylguanine is not a substrate for the AlkB protein, and suggest that e2G is not demethylated by AidB, either. Intriguingly, the only indication of AidB function in the reduction of adduct toxicity was observed in the case of m3C, which is efficiently repaired by the AlkB protein. Neither of the guanine adducts was affected by the presence of AidB or AlkB, and they consequently remain without an established repair pathway.

The mutational spectrum generated by replication of m3C in the four strains shows clearly that AlkB is involved in suppressing the adduct's mutagenicity but was less obvious as to the possible role of AidB. Substantial mutagenicity was observed in the *alkB* and the *alkB aidB* strains, while the *aidB* strain appeared similar to wild-type, although with a slightly higher mutation rate (of only a few percent). Analysis of the mutational spectra produced by m2G and e2G support the conclusions of the bypass assay in that no evidence for repair by either protein is observed. Neither lesion is drastically miscoding, as >84% G is seen at the site of investigation in all cases.

Neither of the hypotheses regarding AidB function tested here was supported convincingly by the MNNG sensitivity, lesion bypass, or lesion mutagenicity assay results. The small changes in bypass and mutagenicity of certain candidate substrate lesions were insufficient to draw strong conclusions about protein activity. Similarly, the resistance to the methylating agent MNNG caused in some cases by loss of the AidB gene as well as by the over-production of AidB protein confounds the search for function.

Clearly, further studies of the protein will be required to elucidate how or even whether the protein participates in the process of policing the genome and maintaining genetic information in an undamaged form. It is possible that the protein does not function in DNA repair, but rather is involved in demethylation of another macromolecule through its proposed mechanism of dehydrogenation. Future genetic and biochemical experiments will be required to uncover the answer.

Figure A2.1: The reaction carried out by human isovaleryl coenzymeA dehydrogenase, a protein to which AidB is homologous.



Isovaleryl CoA

β-Methylcrotonyl CoA

Figure A2.2: Identical resistance to MNNG is observed in the isogenic set HK81 (wild-type; black), MV5933 (AidB-deficient; dark gray) and MV5934 (AlkB and AidB-deficient; light gray).

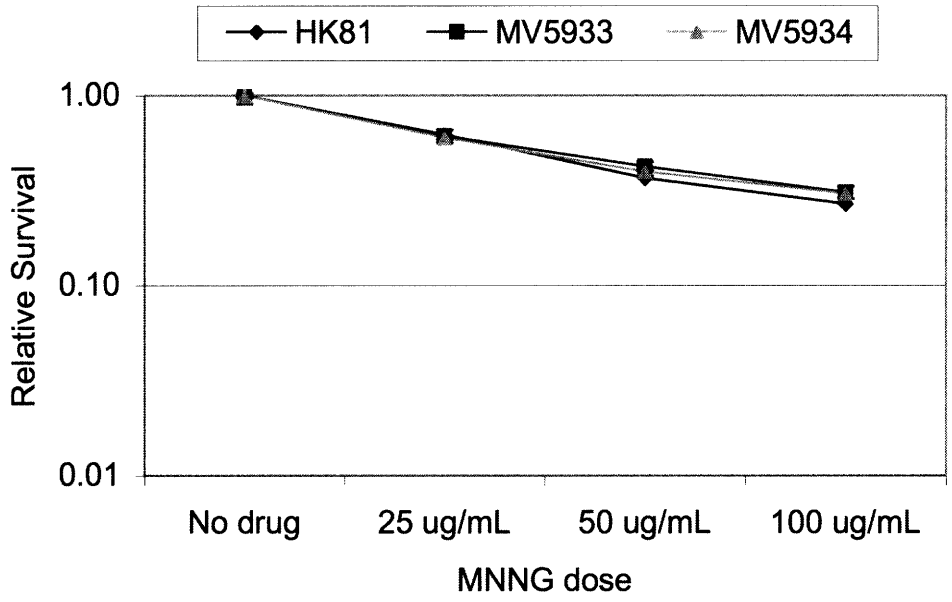
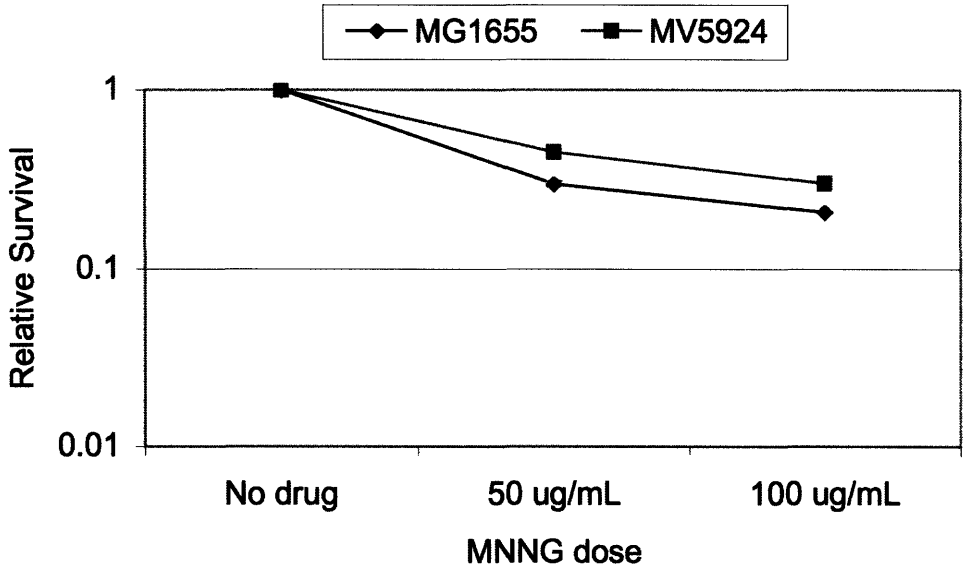


Figure A2.3: Resistance to killing by exposure to MNNG is observed in two knock-out strains in which the aidB gene has been replaced with a tetracycline-resistance marker. Top panel: MV5924 (gray) is the AidB-deficient derivative of MG1655 (black). Bottom panel: MV5925 (gray) is the AidB-deficient derivative of MV1161.



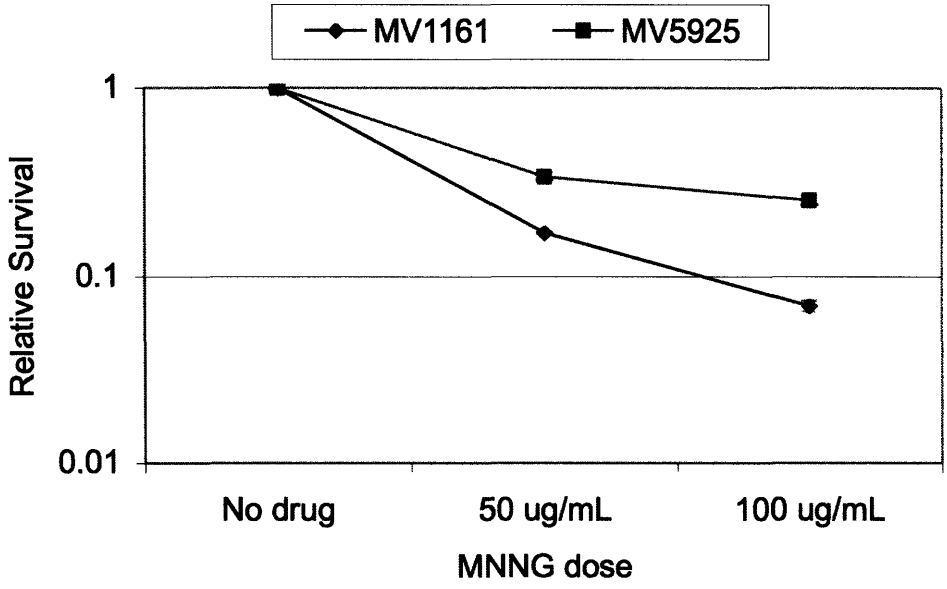


Figure A2.4: Over-expression of AidB from an inducible plasmid renders cells even more resistant to the toxic effects of MNNG. The magnitude of resistance of MV3590 (over-expressing, gray circles) is larger than that observed for either of the knock-out strains shown in Figure A2.3 (MV1161 and MV5925 shown in black and gray diamonds, respectively, and MG1655 and MV5924 shown in black and gray squares, respectively).

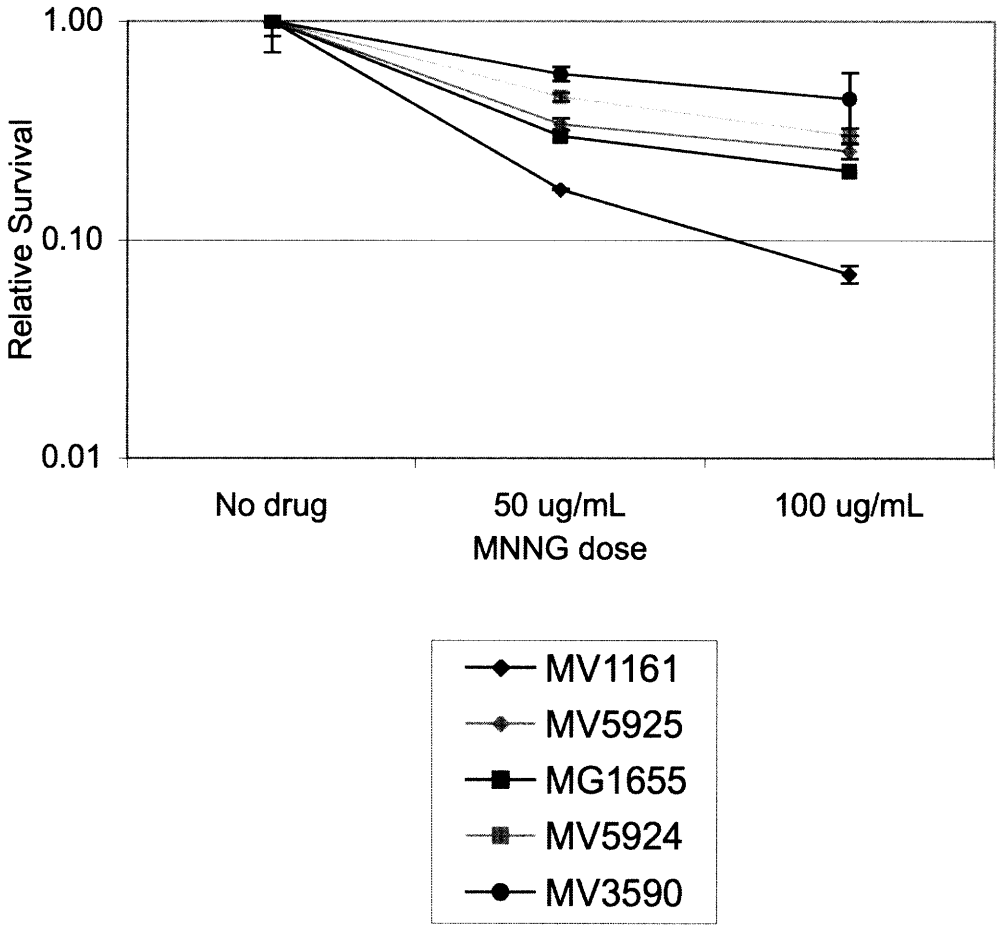
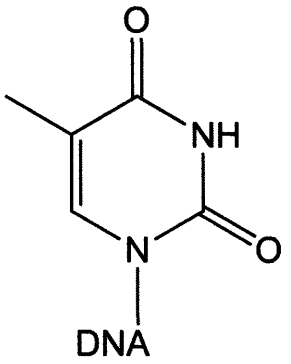
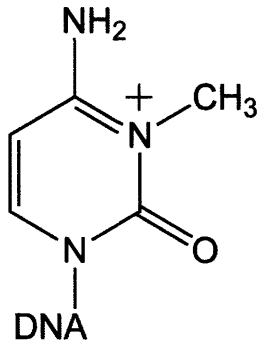


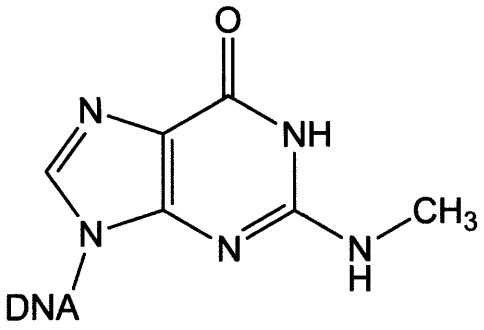
Figure A2.5: Structures of the adducted bases used in the lesion bypass and mutagenesis assays. The latter five were considered to be candidate substrates for repair by the AidB and/or the AlkB protein.



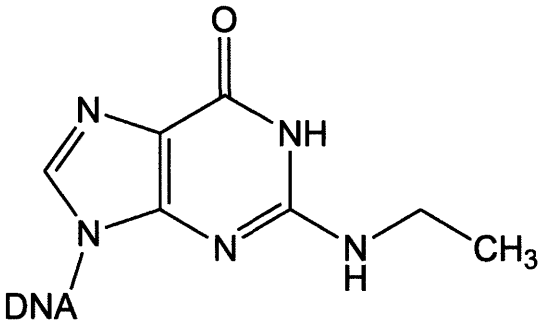
thymine



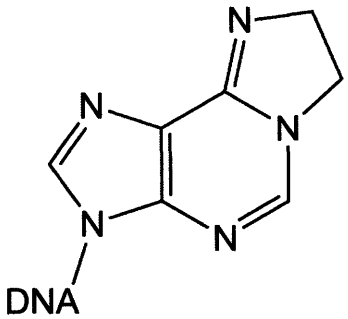
3-methylcytosine



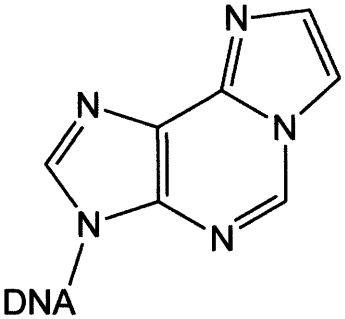
N^2-methylguanine



N^2-ethylguanine

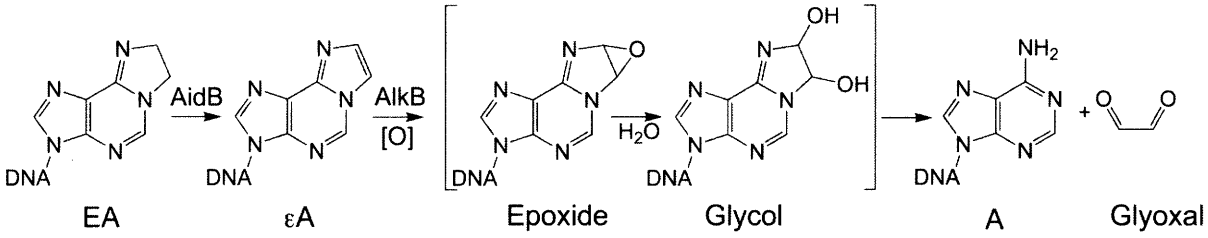


1,N^6-ethanoadenine



1,N^6-ethenoadenine

Figure A2.6: Outlines of the two hypotheses tested. Top panel: Hypothesis that AidB serves to convert the poorly-repaired EA adduct into the efficiently-repaired εA. Bottom panel: Hypothesis proposed in ref. (27) that AidB is a methylated-base dehydrogenase (Figure adapted from ref (27)).



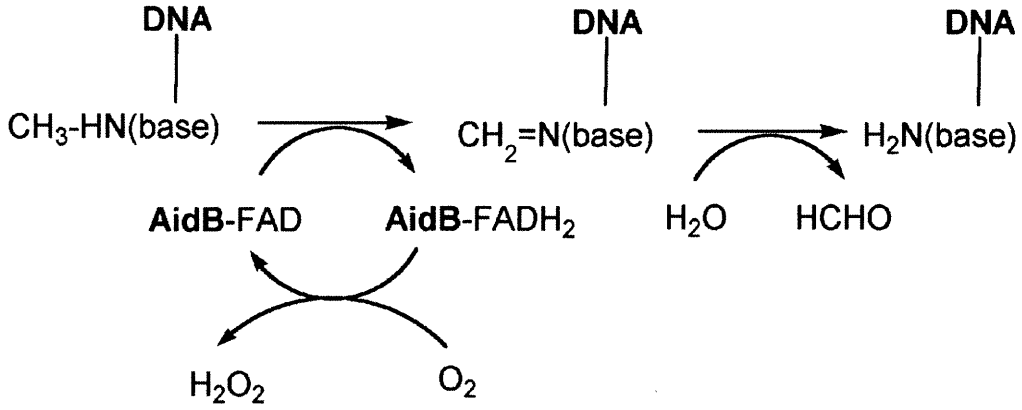


Figure A2.7: Bypass and mutagenicity analysis of the exocyclic EA and εA adducts in wild-type (black), *alkB* (dark gray), *aidB* (light gray), and *alkB aidB* (white) strains of *E. coli.* Top panel: Replication efficiency of EA and εA as compared to T. Bottom panel: Base composition at the lesion site after replication of an M13 genome containing a single EA or εA adduct.

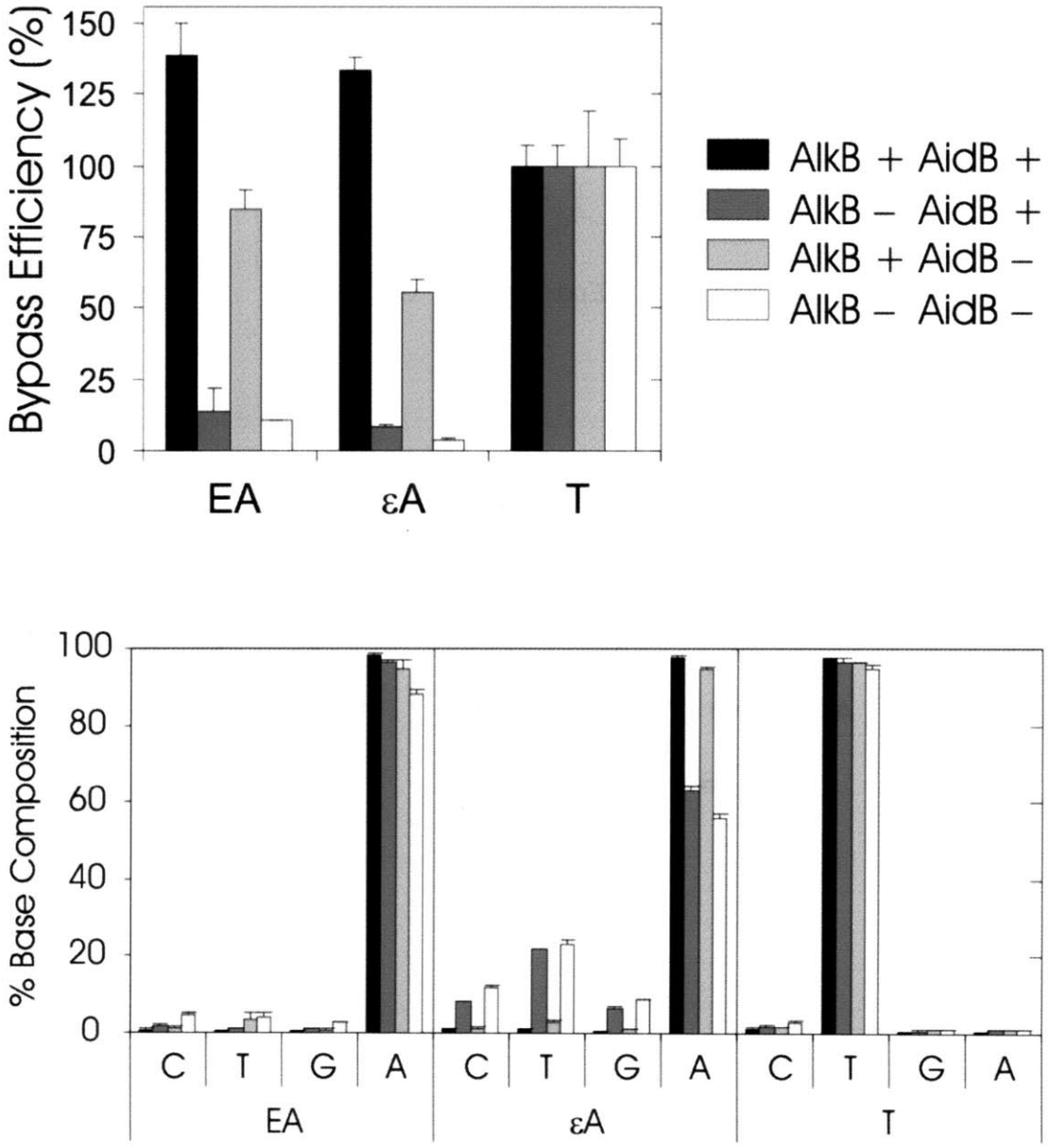
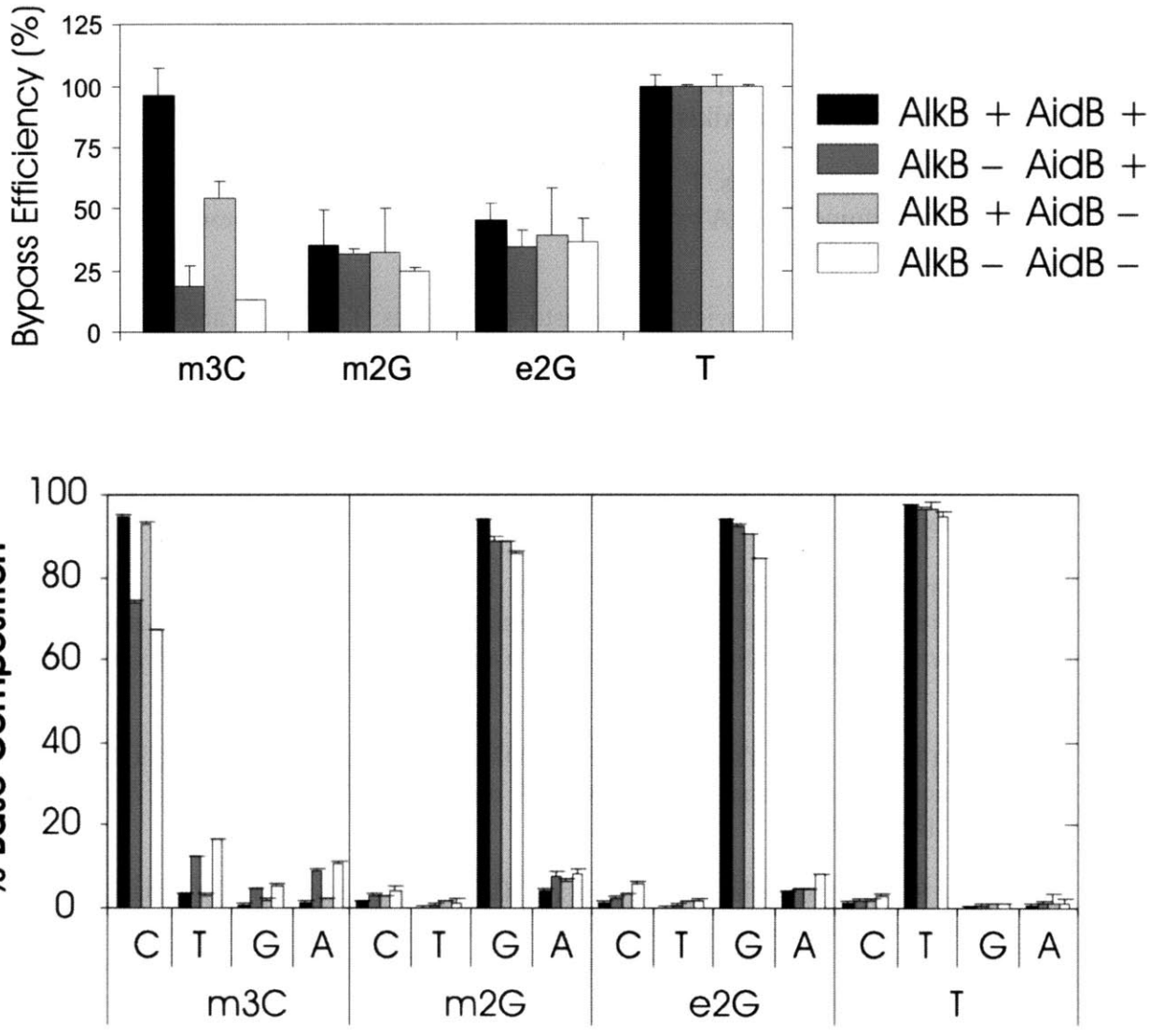


Figure A2.8: Replication efficiency and coding specificity of the methylated bases proposed to be substrates for a dehydrogenase mechanism of AidB. Top panel: Relative replication efficiency of m3C, m2G, and e2G in wild-type (black), *alkB* (dark gray), *aidB* (light gray), and *alkB aidB* (white) cells. Bottom panel: Mutational frequency and specificity of each lesion when replicated *in vivo* in the presence or absence of putative repair proteins.



References

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

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

3. Delaney, J. C. and J. M. Essigmann. 2004. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A* **101**:14051-14056.

4. Falnes, P. O. 2004. Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. *Nucleic Acids Res.* **32**:6260-6267.

5. Koivisto, P., P. Robins, T. Lindahl, and B. Sedgwick. 2004. Demethylation of 3-methylthymine in DNA by bacterial and human DNA dioxygenases. *J. Biol. Chem.* **279**:40470-40474.

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

7. Frick, L. E., J. C. Delaney, C. Wong, C. L. Drennan, and J. M. Essigmann. 2007. Alleviation of 1,N6-ethanoadenine genotoxicity by the Escherichia coli adaptive response protein AlkB. *Proc. Natl. Acad. Sci. U. S. A* **104**:755-760.

8. Delaney, J. C., L. Smeester, C. Wong, L. E. Frick, K. Taghizadeh, J. S. Wishnok, C. L. Drennan, L. D. Samson, and J. M. Essigmann. 2005. AlkB reverses etheno DNA lesions caused by lipid oxidation in vitro and in vivo. *Nat. Struct. Mol. Biol.* **12**:855-860.

9. Mishina, Y., C. G. Yang, and C. He. 2005. Direct repair of the exocyclic DNA adduct 1,N6-ethenoadenine by the DNA repair AlkB proteins. *J. Am. Chem. Soc.* **127**:14594-14595.

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

11. Ougland, R., C. M. Zhang, A. Liiv, R. F. Johansen, E. Seeberg, Y. M. Hou, J. Remme, and P. O. Falnes. 2004. AlkB restores the biological function of mRNA and tRNA inactivated by chemical methylation. *Mol. Cell* **16**:107-116.

12. Yasui, M., S. Matsui, M. Ihara, Y. R. Laxmi, S. Shibutani, and T. Matsuda. 2001. Translesional synthesis on a DNA template containing N2-methyl-2'-deoxyguanosine catalyzed by the Klenow fragment of Escherichia coli DNA polymerase I. *Nucleic Acids Res.* **29**:1994-2001.

13. Choi, J. Y. and F. P. Guengerich. 2006. Kinetic evidence for inefficient and error-prone bypass across bulky N2-guanine DNA adducts by human DNA polymerase iota. *J. Biol Chem.* **281**:12315-12324.

14. Choi, J. Y., K. C. Angel, and F. P. Guengerich. 2006. Translesion synthesis across bulky N2-alkyl guanine DNA adducts by human DNA polymerase kappa. *J. Biol Chem.* **281**:21062-21072.

15. Mishina, Y., E. M. Duguid, and C. He. 2006. Direct reversal of DNA alkylation damage. *Chem. Rev.* **106**:215-232.

16. Mishina, Y. and C. He. 2006. Oxidative dealkylation DNA repair mediated by the mononuclear non-heme iron AlkB proteins. *J. Inorg. Biochem.* **100**:670-678.

17. Fang, J. L. and C. E. Vaca. 1997. Detection of DNA adducts of acetaldehyde in peripheral white blood cells of alcohol abusers. *Carcinogenesis* **18**:627-632.

18. Terashima, I., T. Matsuda, T. W. Fang, N. Suzuki, J. Kobayashi, K. Kohda, and S. Shibutani. 2001. Miscoding potential of the N2-ethyl-2'-deoxyguanosine DNA adduct by the exonuclease-free Klenow fragment of Escherichia coli DNA polymerase I. *Biochemistry* **40**:4106-4114.

19. Choi, J. Y. and F. P. Guengerich. 2004. Analysis of the effect of bulk at N2-alkylguanine DNA adducts on catalytic efficiency and fidelity of the processive DNA polymerases bacteriophage T7 exonuclease- and HIV-1 reverse transcriptase. *J. Biol Chem.* **279**:19217-19229.

20. Volkert, M. R., L. I. Hajec, and D. C. Nguyen. 1989. Induction of the alkylation-inducible aidB gene of Escherichia coli by anaerobiosis. *J. Bacteriol.* **171**:1196-1198.

21. Volkert, M. R., L. I. Hajec, Z. Matijasevic, F. C. Fang, and R. Prince. 1994. Induction of the Escherichia coli aidB gene under oxygen-limiting conditions requires a functional rpoS (katF) gene. *J. Bacteriol.* **176**:7638-7645.

22. Landini, P. and M. R. Volkert. 2000. Regulatory responses of the adaptive response to alkylation damage: a simple regulon with complex regulatory features. *J. Bacteriol.* **182**:6543-6549.

23. Smirnova, G. V., O. N. Oktyabrsky, E. V. Moshonkina, and N. V. Zakirova. 1994. Induction of the alkylation-inducible aidB gene of Escherichia coli by cytoplasmic acidification and N-ethylmaleimide. *Mutat. Res.* **314**:51-56.

24. Landini, P., L. I. Hajec, L. H. Nguyen, R. R. Burgess, and M. R. Volkert. 1996. The leucine-responsive regulatory protein (Lrp) acts as a specific repressor for sigma s-dependent transcription of the Escherichia coli aidB gene. *Mol. Microbiol.* **20**:947-955.

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

26. Volkert, M. R., D. C. Nguyen, and K. C. Beard. 1986. Escherichia coli gene induction by alkylation treatment. *Genetics* **112**:11-26.

27. Rohankhedkar, M. S., S. B. Mulrooney, W. J. Wedemeyer, and R. P. Hausinger. 2006. The AidB component of the Escherichia coli adaptive response to alkylating agents is a flavin-containing, DNA-binding protein. *J. Bacteriol.* **188**:223-230.

28. Hang, B., A. Chenna, A. B. Guliaev, and B. Singer. 2003. Miscoding properties of 1,N6-ethanoadenine, a DNA adduct derived from reaction with the antitumor agent 1,3-bis(2-chloroethyl)-1-nitrosourea. *Mutat. Res.* **531**:191-203.

29. Chenna, A., H. Maruenda, and B. Singer. 1999. Synthesis of para-benzoquinone and 1,3-bis(2-chloroethyl)nitrosourea adducts and their incorporation into oligonucleotides. *IARC Sci. Publ.*89-101.

30. DeCorte, B. L., D. Tsarouhtsis, S. Kuchimanchi, M. D. Cooper, P. Horton, C. M. Harris, and T. M. Harris. 1996. Improved strategies for postoligomerization synthesis of oligodeoxynucleotides bearing structurally defined adducts at the N2 position of deoxyguanosine. *Chem. Res. Toxicol.* **9**:630-637.

31. Delaney, J. C. and J. M. Essigmann. 2006. Assays for Determining Lesion Bypass Efficiency and Mutagenicity of Site-Specific DNA Lesions *In Vivo*. *Methods in Enzymology* **408A**:1-15.

LAUREN E. FRICK
lfrick@mit.edu
(617) 253-6207

Massachusetts Institute of Technology
32 Vassar St. 56-686
Cambridge, MA 02139

EDUCATION

Massachusetts Institute of Technology Cambridge, MA
Candidate for PhD in Biological Engineering, Sept 2007 GPA: 5.0/5.0
Research focuses on the *E. coli* DNA repair protein AlkB and its function in preventing toxicity and mutation caused by substrate DNA lesions.

Massachusetts Institute of Technology Cambridge, MA
S.B. in Materials Science and Engineering, June 2003 GPA: 4.9/5.0
Minor in German

RESEARCH EXPERIENCE

MIT, Department of Biological Engineering Cambridge, MA
Advisor: John M. Essigmann
Research focuses on the genetic consequences of alkylative DNA damage in *Escherichia coli* and the ability of the repair protein AlkB to mitigate the genotoxicity and mutagenicity of novel substrate lesions.

MIT, Department of Materials Science and Engineering Cambridge, MA
Advisor: Christine Ortiz
Characterization of nanostructure and mechanical properties of tough natural materials to aid in development of biomimetic materials that offer superior impact protection to be employed in new energy-absorbing body armor materials technologies. AFM and three-point bend tests. Funded by the MIT Institute for Soldier Nanotechnologies. (Jan – Sept 2003)

BASF Aktiengesellschaft Ludwigshafen, Germany
Advisor: Ingolf Hennig
Conducted research on the electrical properties of various electrode, separator, and electrolyte materials in an attempt to optimize electrochemical double-layer capacitors. Work was in German, using techniques such as impedance spectroscopy, contact angle measurements, and differential scanning calorimetry. (June – August 2002)

NASA Langley Research Center Hampton, VA
Advisor: Lisa Jones
Devised more accurate method for characterization of polyurethane foam for application in improving crashworthiness of aircraft seating. Re-designed hydrostatic test method and wrote code to analyze data. Helped modify and prepare a Lancair aircraft for impact dynamic testing. (June – August 2001)

Westinghouse Electric Corporation Madison, PA
Advisor: William Soule
Wrote an online Electronic Circuit Analysis course, which is currently being taught through a local college, and compiled a course in Nuclear Detector Maintenance, which was sold to customers. (June – August 2000)

TEACHING EXPERIENCE

MIT, Department of Biological Engineering Cambridge, MA
Analysis of Biological Networks - Teaching Assistant
Assisted in creating problem sets and exams, ran tutorials, aided in critical analysis of students' proposal ideas, and graded all problem sets, exams, and final grant proposals. (Aug – Dec 2004)

Media and Technology Charter High School Brookline, MA
Taught math and science to freshmen. (Jan – May 2001)

MIT, Department of Physics Cambridge, MA
Physics 1 - Teaching Assistant
Ran a weekly tutorial teaching Newtonian Mechanics. Developed problems for online physics tutorial program owned by MIT. (Feb – May 2000)

PUBLICATIONS

"Alleviation of 1,N6-ethanoadenine genotoxicity by the Escherichia coli adaptive response protein AlkB." Frick LE, Delaney JC, Wong C, Drennan CL, Essigmann JM, *Proceedings of the National Academy of Sciences,* USA, 2007, 104(3), 755.

"AlkB reverses etheno DNA lesions caused by lipid oxidation in vitro and in vivo." Delaney JC, Smeester L, Wong C, Frick LE, Taghizadeh K, Wishnok JS, Drennan CL, Samson LD, Essigmann JM., *Nature Structural and Molecular Biology*, 2005, 12(10), 855.

"Nanoscale morphology and indentation of individual nacre tablets from the gastropod mollusc *Trochus niloticus."* B.J.F. Bruet , H.J. Qi, M.C. Boyce, R. Panas, K. Tai, L. Frick, C. Ortiz, *Journal of Materials Research*, 2005, 20(9), 2400.

AWARDS

22nd Aspen Cancer Conference Fellow *July 2007*
MIT Center for Environmental Health Sciences 4th Annual Poster Session – Outstanding Poster 2nd Place *May 2007*
Department of Materials Science and Engineering Outstanding Student *June 2003*
NEWMAC All-Academic Honors *November 2001, November 2002*
Award for Outstanding Service to the Department of Materials Science Community *June 2001, June 2002*
Robert T. Haslam Scholarship *May 2001, Jan. 2002*
Kenneth Tray Memorial Scholar-Athlete *June 1999*
National Merit Finalist *May 1999*
Westinghouse Family Scholarship *March 1999*

ACTIVITIES

MIT Women's Volleyball Club – Treasurer *2006-2007*, Member *2003-2007*
SUMS (Society of Undergraduate Materials Students) – President *2002*, Vice-President *2001*
MIT Women's Varsity Volleyball – *1999-2002* - elected co-captain *2002*
McCormick House Government – President *2002*, Room Reservation Chair *2001*