Nucleic Acids Research, 1994, Vol. 22, No. 9 1613–1619

Specificities of human, rat and *E.coli* O^6-methylguanine-DNA methyltransferases towards the repair of O^6-methyl and O^6-ethylguanine in DNA

Lie-Kuo Liem, Allan Lim and Benjamin F.L.Li*
Chemical Carcinogenesis Laboratory, Institute of Molecular and Cell Biology, National University of Singapore, Kent Ridge, Singapore 0511, Republic of Singapore

Received February 3, 1994; Revised and Accepted March 22, 1994

ABSTRACT

The behaviour of highly purified bacterial expressed rat O^6-methylguanine-DNA methyltransferase (MGMT) towards the repair of CGCm6GAGCTCGCG and CGCe6GAGCTCGCG (k_{m6G}/k_{e6G} = 1.45, where k is the second order repair rate constant determined, m6G and e6G are O^6-methyl and O^6-ethylguanine) is similar to that of *E.coli* 39kD Ada protein (k_{m6G}/k_{e6G} = 1.6). However, the human MGMT is very different (k_{m6G}/k_{e6G} = 163). The preferential repair of O^6-ethylguanine lesion by the rat MGMT appears not to be related to the lack of the initiator methionine in the expressed protein since similar results were obtained from N-terminal Glutathione-S-transferase (GST) fused protein (GSTMGMT) which retains the methionine. The possible relationship between these findings and the differences observed in the primary amino acid sequence of these proteins is discussed. In addition the preferential repair of O^6-ethylguanine substrate by the 39kD Ada protein as compared to the catalytic C-terminus alone (different by 134 times) suggests that the N-terminus plays a crucial role in the repair of O^6-ethylguanine. This is in constrast to the minor effects of the GST domain when fused to the N-terminus of mammalian MGMT.

INTRODUCTION

Mutagenic (or carcinogenic) potency of alkylating agents reside in their intrinsic reactivity with cellular DNA to produce high proportions of mutagenic lesions, O^6-alkylguanine and O^4-alkylthymine (1). However, the DNA repair enzyme, O^6-methylguanine-DNA methyltransferase (MGMT), protects cells against the mutagenic effect of these alkylating agents (2). This is due to its ability to restore these damaged bases to their normal counterparts by transferring the alkyl groups of these damaged bases to a cysteine residue in the active site of the protein (3). The repair protein is inactivated through this process, therefore, the concentrations of this protein are critical in determining the susceptibility of respective cells towards mutagenesis by these chemicals (4). Because of the importance of MGMT for cellular protection, we have begun a series of experiments to study the repair mechanism of this protein.

The two fundamental factors related to the effectiveness of DNA repair are addressed in these studies; namely the substrate and the repair protein. The former factor has been widely studied because there is evidence suggesting that the formation and repair of O^6-alkylguanine residues in DNA are sequence specific (5–8). It is possible that the observed 'mutation hot spots' induced by SN1 alkylating agents could be a composite effect of these two parameters. Interestingly, recent observations indicate that there is a link between the overall structure of the substrate and the repair protein. This arises from the similarity between the regional distortion of DNA helix by the O^6-methylguanine residue (9) and the optimal substrate determined for the *E.coli* 39kD Ada and human MGMT: this is a pentanucleotide motif containing the lesion (ref 10, which is close to a 7-mer reported, ref 11). It is possible that regions of the repair protein contacting the DNA might have evolved around the structure of this regionally distorted DNA: if these domain(s) can be identified, it would improve our understanding of the repair mechanism. At present, very little is known about the molecular basis of recognition of the DNA lesions by the repair protein. However, the importance of the amino acid residues within the highly conserved active site for the catalytic event has been studied by mutagenesis (12). As the primary amino acid sequences of the repair proteins from various species are now available, determining the intrinsic properties of these different repair proteins should allow us to define regions of the protein that are important in the repair.

In this paper, we report the use of two oligonucleotides, CGCm6GAGCTCGCG and CGCe6GAGCTCGCG [which contain O^6-methylguanine (m6G) and O^6-ethylguanine (e6G) residues in identical flanking sequence with solution structures that are well characterised (13, 14)] to study the intrinsic repair properties of MGMTs from human, rat and *E.coli* (39 kD Ada).

*To whom correspondence should be addressed

This study shows that the human and rat proteins have intrinsically different specificities towards the repair of these two lesions despite serving the same function in DNA repair. Furthermore, the preferential repair of O^6-ethylguanine substrate by the full length *E.coli* 39kD Ada protein (134 times that of its catalytic C-terminus alone) suggests that the N-terminus may play a modulatory role in the repair of O^6-ethylguanine. This effect is not seen in the case of the GST domain fused to the N-terminus of mammalian MGMTs.

MATERIALS AND METHODS

Source of oligonucleotides

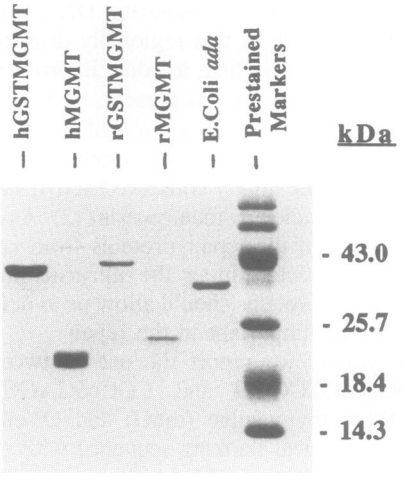
Oligodeoxynucleotides containing the O^6-methylguanine and O^6-ethylguanine residues were synthesised and deprotected as described previously (15). They were purified and quantified according to published procedures (16).

Source of recombinant proteins and specific activities

The bacteria stock (pSM41) containing the plasmid encoding the *E.coli* 39kD Ada protein was a generous gift from Dr. S. Mitra, (Oak Ridge National Laboratory, Tennessee, USA). The procedures for growth and purification of the protein from the bacterial culture are essentially the same as reported (17), except that two cycles of AcA54 and ssDNA-cellulose column chromatography were required to obtain a homogeneous preparation. The human MGMT and Glutathione-S-transferase fused MGMT (GSTMGMT) were obtained by cloning of the cDNA obtained from CEM cells (acute lymphoblast leukemia, ATCC, USA). The procedures for cloning and purification of these proteins were reported previously (18). The human repair proteins (i.e. MGMT and GSTMGMT) differ from the published sequences in codon 84, i.e. phenylalanine instead of leucine (19–21) while codon 127 is alanine (19, 20) rather than threonine (21). The rat MGMT and GSTMGMT were obtained from screening of rat cDNA library using the human MGMT cDNA probe (22). The rat MGMT cDNA was then cloned into the expression vectors and purified exactly as previously described for the human MGMT (18). However, N-terminal sequencing of the purified rat MGMT, by Edman's degradation, shows the absence of the first methionine. Despite careful manipulation of the culture conditions, purification and bacterial hosts, it is not possible to obtain rat MGMT with the first methionine residue. This is similar to the reported mouse MGMT. The N-terminal methionine is cleaved by methionine aminopeptidase (MAP) which is specific for methioine residues adjacent to alanine (and also glycine and proline) found in the rat MGMT (23). However, this first methionine should be in place in the rat GSTMGMT since it is protected by the fused GST protein. The purified proteins were subjected to amino acid analysis (using the Pico-Tag system from Waters, Millipore, USA) for quantification. The concentrations of proteins were then calculated by taking the average concentrations of 10 well resolved amino acid derivatives determined and their compositions deduced from our own cDNA sequence. The activities of the proteins were then titrated by incubating 1μg of the proteins (determined by amino acid analysis) with various known concentrations of ^{32}P TATA-Cm6GTATA in an assay buffer of 100μl at 37°C for 1 hr. Under this condition the repair would have completed more than $100 \times t_{1/2}$. The activities of the proteins were determined by the amount of the product ^{32}P TATACGTATA formed (because the repairis suicidal). The final specific activities of the proteins and the 13.5% SDS-PAGE analysis were shown in Figure 1. These proteins represent the combined fractions of the final ssDNA chromatography during purification and were used directly for the kinetic experiments. They were shown to be at least >50% active. The proteins are reasonably stable at 37°C, with observable loss of 10–15% activities over 120 min but not at high dilution.

Kinetic experiments

a. 5′ ^{32}P-labeling and purification of substrates. Oligonucleotide substrates (0.1 OD at 260nm) were 5′ ^{32}P labelled by using the blunt-end phosphorylation protocol with γ^{32}P-ATP (10 μl, 222



	Specific Activities pm/mg	
	calculated	determined
hGSTMGMT	23256	13570
hMGMT	47619	28505
rGSTMGMT	23256	12325
rMGMT	47619	23809
E.Coli *ada*	25641	13405

Figure 1. SDS/PAGE (13.5%) analysis and specific activities of recombinant proteins; hMGMT = human O^6-methylguanine-DNA methyltransferase, hGSTMGMT = Glutathione-S-Transferase(GST) fused to the N-terminus of human MGMT, rGSTMGMT = GST fused to the N-terminus of rat MGMT, rMGMT = rat O^6-methylguanine-DNA methyltransferase, *E.coli* Ada = *E.coli* 39kD Ada protein, M = prestained molecular weight markers. The specificity was determined by amino acid analysis and assay using a known concentration of ^{32}P TATACm6GTATA as a substrate (see experimental).

TBq/ml from Amersham, UK) as described (24). After labelling for 30 min the reaction was chased with cold ATP (2 μl of 0.1M ATP) for another 15 min. The phosphorylated oligonucleotides were then purified by reversed phase chromatopgraphy using a linear gradient of acetonitrile in triethylammonium-acetate buffer (TEAA, 0.1M, pH 7 as buffer A and buffer A containing 40% acetonitrile as buffer B). The phosphorylated oligonucleotides were well separated from the non phosphorylated starting material under this condition (25). The purified ^{32}P labelled substrates were then freeze-dried and use directly after quantification.

b. Experimental conditions. The kinetics was followed by using two fold excesses of ^{32}P end labelled substrate to one fold of enzyme. The use of this condition enables us to determine accurately the concentrations of the reactant and product at $t_\infty(10\times t_{1/2})$ and good approximation to the second order reaction. The volume of the assay buffers (50mM Tris.HCl pH 8.0 containing 1mM EDTA and 5mM DTT) used are; 30ml for CGCm6GAGCTCGC, 20ml for CGCe6TAGCTCGCG and 15ml for CGCm6GCG. Initially, the reaction kinetics were tested using 40 pmol of DNA substrates and 20 pmol of repair proteins. The concentrations of the substrates and the repair proteins (maintained at a ratio of 2:1) were then adjusted until the kinetics can be followed at managable time scale. All reactions were carried out at 37°C except the rat MGMT and GSTMGMT were also studied at 10°C. The reaction mixtures containing the substrate DNA and buffer were incubated at 37°C for 15 min before the addition of the repair protein. Kinetic time points were obtained by quenching aliquots of the reaction mixture (1.0ml) with 0.02 OD at 260nm TATACm6GTATA (in 100 μl assay buffer) at required intervals. The reaction mixtures (1.0ml) were analysed using an HPLC system equipped with an autosampler (Waters, USA) linked directly to a radioactive flow detector (Radiometic, USA). The conditions for HPLC analysis are as follow: 1. stationary phase, 8mm diameter Nova Pak C-phenyl cartridges on a Z module (Millipore, USA), and 2. mobile phase, a linear gradient (1%B/min, 3ml/min) of buffer A (0.35M KH_2PO_4 pH 6.4) and B (30% acetonitrile buffer A). Conditions for the radioactive flow detector were as follow: scintillant (Escosin USA, 1ml/min), 0.5ml flow cell, 6 second update time. The concentrations of products and reactants were obtained from the corresponding radioactive peak areas determined by the radioactive detector. The second order rate constants were calculated, from 3 independent experiments, using the experimental infinity values (the amount of product and reactant determined after > ten $t^{1/2}$). A typical kinetic measurement and calculation were shown in Figure 2.

RESULTS

Specific activities of the repair proteins

Figure 1 summarises the SDS-PAGE analysis and specific activities of purified recombinant proteins (they represent the combined fractions of the final ssDNA chromatography) used in these kinetic studies. These analyses show that the proteins are active and pure.

Repair kinetics of CGCm6GAGCTCGCG and CGCe6G-AGCTCGCG by *E.coli* 39kD Ada protein

It is apparent from Table 1 that the repair kinetics of the two substrates by *E.coli* 39kD Ada protein is quite similar, i.e. k_{m6G}/k_{e6G} = 1.6. Surprisingly, this differs substantially from our previous report on the catalytic activity of the 19kD C-terminus, with k_{m6G}/k_{e6G} = 1000 assayed under the same condition (25). The substantial difference (620 fold) observed between these two proteins in their relative repair rates of O^6-methyl- and O^6-ethylguanine lesions in DNA cannot be ascribed to contaminating factors because of the purity of protein and substrates (15, 25) used in these experiments. In addition, comparisons of the repair rate of methyl and ethyl substrates within the same sequence background eliminates sequence related effects [also note that similar rates were observed between

Table 1.

	Second order rate constants 10^5 $M^{-1}S^{-1}$				
MGMT species	CGC**m6**GCG	CGC**m6**GAGCTCGCG	CGC**e6**GAGCTCGCG	k_{m6G}/k_{e6G}	*Ref.*
Human MGMT	3.44 ± 0.13	490 ± 34.2	3.00 ± 0.10	**163**	Present study
Rat MGMT	8.43 ± 0.59	160 ± 9.6 (30)	110 ± 5.5 (13)	**1.45**	Present study
Human GSTMGMT	1.76 ± 0.05	970 ± 87.3	9.20 ± 0.37	**105**	Present study
Rat GSTMGMT	4.84 ± 0.24	69.0 ± 3.4 (22)	130 ± 9.17 (31)	**0.53**	Present study
E.Coli 39kD Ada	3.30 ± 0.13	56.0 ± 2.3	35.0 ± 1.4	**1.6**	Present study
19 kD Ada C-terminus	N.A.	260	0.26	**1000**	*ref. 25*
Ogt	N.A.	290	45	**6.4**	*ref. 37*

Second order repair rate constants of O^6-methyl and O^6-ethylguanine containing substrates obtained from various MGMTs at 37°C; *E.coli* 39kD Ada protein and its 19kD independent C-terminus, *E.coli* Ogt; human 21 kD O^6-methylguanine-DNA methyltransferase (human MGMT); human 43 kD Glutathione-S-Transferase fused MGMT, (human GSTMGMT); Rat 23kD O^6-methylguanine-DNA methyltransferase (rat MGMT); Rat 43 kD Glutathione-S-Transferase fused MGMT (rat GSTMGMT); N.A., not available. Data in brackets are obtained at 10°C (from a single measurement). All other rate constants are obtained from the average of 3 independent experiments.

reported data on the repair of O^6-methylguanine residues in poly(dC.dG) substrate (180×10^5 $M^{-1}S^{-1}$, this substrate also contains other lesions such as phosphotriesters and depurinated sites, ref 17) and the present CGCm6GAGCTCGCG substrate (56×10^5 $M^{-1}S^{-1}$, Table 1) by the 39kD Ada protein]. These data demonstrate that the N-terminal domain in the 39kD Ada protein substantially alters its C-terminus activity (by comparing activity of the catalytic C-terminus alone), as a result the 39kD Ada protein barely discriminates between the repair of m6G and e6G residues.

Repair kinetics of CGCm6GAGCTCGCG and CGCe6G-AGCTCGCG by human and rat MGMTs and their glutathione-S-transferase (GST) fused derivatives (GSTMGMTs)

Data in Table 1 show that in all cases human and rat MGMTs repair the 12mer substrate much faster than the 6mer [k_{12mer}/k_{6mer} for human MGMT is 140 and the rat MGMT is 20]. However, while the human MGMT repairs the 12mer substrate 3 times faster than the rat MGMT, it is two-fold slower

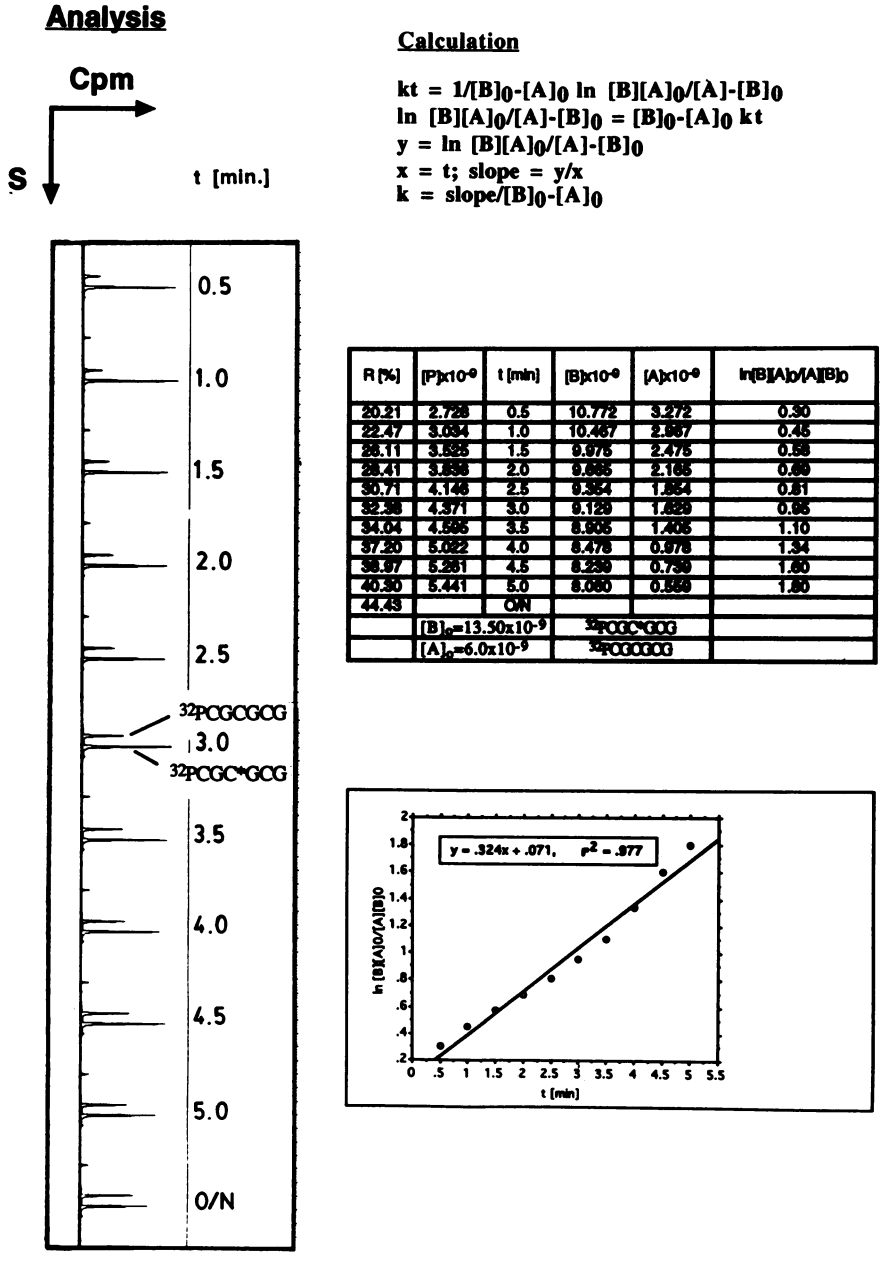


Figure 2. Example of kinetic measurement of the repair of the ^{32}P CGCm6GCG substrate by the Rat fusion protein rGSTMGMT, abbreviations in HPLC analysis are; ^{32}P CGCGCG = product and ^{32}P CGCm6GCG = reactant, cpm = radioactivity detected by the on-line flow radioactive detector, s = sample at different time points, abbreivations for kinetic= calculation are; $[A]_0$ and $[B]_0$ in molar = concentration of enzyme and DNA at time zero, [A] and [B] in molar = concentration of enzyme and DNA at time t, k = second order rate constant in $M^{-1}S^{-1}$, R = % of repair, P in molar = amount of DNA at time t.

for the 6mer. Whether these subtle differences are due to the preference of the rat MGMT towards the repair of shorter or CG rich substrates requires further study (10). Apparently these data also suggest that it would be inappropiate to use m6G and e6G containing 6mer as substrates for comparative study of the repair efficiency of these lesions by different repair proteins, since recognition of short substrates by the human MGMT is the rate-limiting step in the repair (10). However, the 12mer would be the ideal substrate for the purpose because the two mammalian proteins repair the m6G containing 12mer with an extreme but similar efficiency, see Table 1. Therefore, the substantial difference in the repair rate constants observed (37 fold) between the two proteins in the repair of O^6-ethylguanine lesion (see in Table 1) is most likely due to the intrinsic specificity of the two proteins towards the repair of the e6G lesion. The absence of the first methionine residue from the rat MGMT (see material and methods) does not appear to contribute to the difference observed between the human and rat proteins, since comparison of methyl and ethyl data obtained from the rat MGMT (k_{m6G}/k_{e6G} = 1.45) clearly shows that the rat MGMT is less discriminatory in the repair of the two lesions. The human MGMT is intrinsically less efficient in the repair of O^6-ethylguanine lesion (k_{m6G}/k_{e6G} = 163). As the human and rat proteins share substantial homology in their primary amino acid sequences, these data are unexpected. Similar trends in specificity were observed with the mammalian GSTMGMTs fusion proteins (which contain the complete sequence of the rat MGMT). In fact the preferential repair of O^6-ethylguanine lesion by rat GSTMGMT is enhanced (see Table 1). Although the presence GST protein at the N-terminus of human MGMT increases the rate of repair of m6G and e6G containing substrates, these do not parallel the effects on the rat MGMT. Nevertheless, it would be extremely difficult to study the contribution of the first methionine of rat MGMT to the repair kinetic since there is no protease at present that can completely cleave the GST away from the rat GSTMGMT, which could release the rat MGMT with the first methionine exposed (thrombin cleaves the GSTMGMT fusion protein and releases the MGMT with extra 7 amino acid residues, due to the location of the thrombin site, fused to the first methionine). Therefore, one would envisage that the comparison of the GSTMGMT and thrombin cleaved GSTMGMT would be inappropiate for an understanding of the importance of the first methionine since the size of the fused proteins are now compared but not the presence and absence of the first methioine. However, these changes in specificity are not significant as compared to our previous reported data on the effects of the length of phosphodiesters (10).

DISCUSSION

The role of the N-terminus phosphotriester repair domain and structure of O^6-methyl- and O^6-ethylguanine containing oligonucleotide substrates in the repair of O^6-alkylguanine by *E.coli* 39kD Ada protein

Our previous structural study by NMR indicates that the phosphodiester backbone in the O^6-ethylguanine, but not O^6-methylguanine, containing oligonucleotide is distorted (13,14). Could this altered conformation of the phosphodiester backbone be an important determining factor in the repair of O^6-ethylguaniune lesion by MGMT? The observation that the presence of the N-terminal domain in the 39 kD Ada protein increases the repair activity of its C-terminus towards the repair of O^6-ethylguanine by 135 fold, but decreases the O^6-methylguanine by 4.6 fold (as compared to the activity of the independent C-terminus, Table 1), has some interesting implications. The different specificity observed between the two proteins could be the result of the following: 1. catalysis due the preferential stabilisation of the transition state of the O^6-ethylguanine substrate and protein complex, 2. inefficiency of the independent C-terminus in the recognition of O^6-ethylguanine substrate or 3. the N-terminus may be involved in scanning of DNA for damage sites (i.e. recognition event). These observations suggest that points 2 and 3 are responsible because: a. the rate of repair of O^6-methylguanine lesion by the full length protein is similar to that of the O^6-ethylguanine, b. O^6-methyl and O^6-ethylguanine are repaired through the same mechanism (26) and c. the two proteins have identical active sites. Thus the N-terminus of the protein may function in the overall recognition event. As a result, the repair activities of the C-terminus of the 39kD Ada protein would be dependent upon the efficiency of the N-terminus in recognising damaged DNA. The observation that methyl and ethyl lesions are repaired at similar rate would imply that the N-terminus might be preferentially associated with the O^6-ethylguanine substrate. This suggests a basis for the recognition of phosphotriester lesions in DNA by the N-terminus phophotriester recognition domain involving modified phosphodiester backbone which might resemble to that found in O^6-ethylguanine containing DNA. This result shows that the recognition of O^6-alkylguanine containing substrate by the 39kD Ada protein is not only present in the catalytic domain of the C-terminus. However, if the increase in repair rate of the O^6-ethylguanine containing substrate is not due to preferential recognition of this substrate, i.e. the N-terminus recognises O^6-methyl and O^6-ethylguanine substrates with same efficiency, one can infer that the removal of methyl and ethyl groups from O^6-alkylguanine residues by the catalytic domain are energetically similar (inferred from k_{m6G}/k_{e6G} = 1.6). This is important for an understanding of the repair mechanism. Furthermore, it will be interesting to discover whether the Zn binding motif (27) in the N-terminus of the 39kD Ada protein can recognise other damaged DNA in an analogous fashion to the Zn binding motifs of mammalian poly(ADP-ribose) polymerase (28).

The primary protein sequence of MGMTs involved in the repair of O^6-ethylguanine

Although it is unlikely that the loss of the first methionine in the rat MGMT alters its repair activity, more substantial N-terminal

	Codon no.			
Human	126	AARAVGGAMR	GNPVPILI**P***C*	HRVVCSSGAV
Rat	130	AARAVGGAMR	SNPVPILI**P***C*	HRVIRSDGAI
Ogt	120	AARAVGAANG	SNPISIVV**P***C*	HRVIGRNGTM
Ada (I)	302	AVRAVASACA	ANKLAIVI**P***C*	HRVVRGDGSL
Ada (II)	302	AVRAVASACA	ANKLAIII**P***C*	HRVVRGDGTL

Figure 3. Comparison of the predicted 30 amino acid residues around the active site domains (-PCHRV-) of human O^6-methylguanine-DNA methyltransferase, rat O^6-methylguanine-DNA methyltransferase, *E.coli* Ogt, *E.coli* Ada C-terminal (version I) and *E.coli* Ada C-terminal (version II). Data obtained from Santibanez-Koref *et al.*, 1992 (ref. 30). Bold characters indicate the arginine (R) residue which may be involved in the repair of O^6-ethylguanine lesion. The methyl acceptor cysteine in the underlined -PCHRV active site peptide is in italic.

deletions of human and rat MGMTs (lacking codon 2 to 11) do give rise to inactive proteins (unpublished results). By contrast, it has been shown that deletion of the 30 amino acid residues at the C-terminus of the human MGMT does not affect its repair activity (29). The combination of the data from both MGMT and GSTMGMT is sufficient to show that there is an intrinsic specificity of rat MGMT towards the repair of O^6-ethylguanine lesion. Since Ada, Ogt and rat MGMTs efficiently repair O^6-ethylguanine while their primary amino acid sequences are significantly less homologous (30), it is unlikely that the -PEGV- peptide domain in the mouse and rat MGMTs (31) is responsible for the difference observed between rat and human MGMTs. It is more likely that the slow repair of O^6-ethylguanine by human MGMT is the result of difference(s) in amino acid residues around its active-site as compared to the other proteins. Careful inspection of the primary amino acid sequences among these proteins reveals the absence of a basic arginine residue immediately after the highly conserved active-site peptide domain, (-PCHRV-), of the human protein (see Figure 3). As the Ada, Ogt and rat MGMTs efficiently repair O^6-ethylguanine and contain this arginine (R) residue, which is 5−6 amino acid away (on the C-terminal side) from their active-site cysteine, it might contribute to the repair or recognition of the O^6-ethylguanine lesion. This could involve in the stabilisation of altered phosphodiester backbone of the O^6-ethylguanine containing substrate by the basic arginine residue. Whether the precise location of this arginine residue relative to the alkyl acceptor cysteine would be an important factor for the catalytic event is unclear (note that the arginine residue of the Ogt protein is not in alignment, see Figure 3). In line with this hypothesis, the active-site, i.e. -FRPCKRC- (32), of phosphotriester repair domain at the N-terminus of the 39kD Ada protein is extremely basic. In addition, the differences in the predicted primary amino-acid sequence of human MGMTs derived from different cDNA sources (obtained from various tumor cell lines, ref. 19, 20 and 21, including ours) also needs to be addressed. Although we are not in a position to comment on what is the 'normal' human cDNA sequence, we believe that the differences at codons 84 and 127 [which are 61 and 18 amino acid residues away (on the N-terminal side) from the active-site cysteine at codon 145] in our human MGMT are unlikely to be reponsible for the observed slow repair of the O^6-ethylguanine substrate. For examples, 1. Crone and Pegg 1993 (ref 38) reported that mutation at codon W100A (45 amino acid away from the active-site cysteine at codon 145) has no effect on the repair activity, while mutation at P140A [which is 5 amino acids (on the N-terminal side) away from the active-site cysteine] does have an inhibitory effect on the repair of O^6-benzylguanine (but not O^6-methylguanine containing DNA), 2. The human MGMT used in this study repairs the CGCm6GAGCTCGCG substrate at such high efficiency (second order rate constant of 4.9×107 $M^{-1}S^{-1}$, see Table 1) which is close to the encountered control rate indicates that our protein is fully functional and 3. Chueh *et al*, 1992 (ref 12) had shown that only substitution of amino acid residues within the active-site (ie., LIPCHRV) can have significant inhibitory effect on the repair activity of human MGMT.

Relevance of kinetic data to the carcinogenicity of methyl and ethylating agents in man and rodent

It is important to address the question of whether *in vitro* data (repair kinetic data obtained from short synthetic oligonucleotide substrates) can be usefully extrapolated to *in vivo* repair behaviour. The results described here show the efficient *in vitro* repair of these synthetic substrates by MGMTs from diverse species in all cases. As the complexities of DNA repair *in vivo* are only beginning to emerge (strand bias repair, ref. 33; interplay between different repair pathways, ref. 34; cell cycle control, ref. 35), these *in vitro* data are useful for comparison. Furthermore, the data reported here may have some use in extrapolating carcinogenicity tests using rodents to the human situation. If repair efficiency of O^6-alkylguanine residues in DNA by MGMT is an important factor in deciding the fate of mutations induced by S_N1 type alkylating agents, the above data would suggest that human cells are more susceptible to ethylating agents. Thus the rat may be a poor model for studying *in vivo* carcinogenicity of methyl- and ethylating agents in relation to human cancers. Studies have already shown that rats are an order of magnitude less susceptible to bifunctional ethylating agents than humans (36).

CONCLUSION

These data show the potential of O^6-methyl and O^6-ethylguanine containing synthetic substrates in characterising MGMTs from different species and elucidating the repair mechanism. The two novel findings in this study, that the N-terminus of *E.coli* 39kD Ada protein can substantially influence the recognition event in the repair pathway and there is different substrate selectivity exhibited by homologous human and rat MGMTs, should initiate further experiments to understand their relevance *in vivo*. For example, the relationship between rate of repair of O^6-alkylguanine lesions by MGMT *in vivo* and mutagenesis can be investigated by comparing N-ethylnitrosourea induced mutations in cell lines transfected with human or rat MGMT.

ACKNOWLEDGEMENTS

We thank Dr Lin Wen-Shu for protein sequencing and Dr Edward Manser for reading of the manuscript. This research is funded by the National University of Singapore.

REFERENCES

1. Lawley, P.D. (1979) Chemical Carcinogens and DNA, P.L. Golver ed., CRC Press, Boca Raton, Florida, USA.
2. Pegg, A.E. (1990) *Cancer Res.* 50, 6119−6129.
3. Lindahl, T., Sedgwick, B., Sekiguchi, M. and Nakabeppu, Y (1988) *Annu. Rev Biochem.* 57, 133−157.
4. Dumenco, L.L., Allay, E., Norton, K. and Gerson, K.L. (1993) *Science* 259, 219−222.
5. Buckley, N. (1987) *J. Amer. Chem. Soc.* 109, 7918−7920.
6. Richardson, F.C., Boucheron, J.A., Skopek, T.R. and Swenberg, J.A. (1989) *J. Biol. Chem.* 264, 838−841.
7. Topal, M.D., Eadie, J.A. and Conrad, M. (1986) *J. Biol. Chem.* 261, 9879−9885.
8. Georgiadis, P., Smith, C.A. and Swann, P.F. (1991) *Cancer Res.* 51, 5943−5850.
9. Wong, C.W., Tan, N.W. and Li, B.F.L. (1992) *J. Mol. Biol.* 228, 1137−1146.
10. Liem, L.K., Wong, C.W., Lim, A. and Li, B.F.L.(1993) *J. Mol. Biol.* 231, 950−959.
11. Chan, C-L., Wu. Z., Ciardelli, T., Eastman, A. and Bresnick, E. (1993) *Arch. Biochem. Biophys.* 300, 193−200.
12. Chueh, L.L., Nakamura, T., Nakatsu, Y., Sakumi, K., Hayakawa, H. and Sekiguchi, M. (1992) *Carcinogenesis* 13, 837−843.
13. Kalnik, M.W., Li, B.F.L., Swann, R.F. and Patel, D.J. (1989) *Biochemistry* 28, 6182−6192.
14. Swann, P.F. (1990) *Mut. Res.* 233, 81−94.

15. Li, B.F.L. and Swann, P.F. (1989) *Biochemistry* 28, 5779–5786.
16. Tan, N.W. and Li, B.F.L. (1990) *Biochemistry* 29, 9234–9240.
17. Bhattacharyya, D., Tano, K., Bunick, G.J., Uberbacher, E.C., Behnke, W.D. and Mitra, S. (1988) *Nucl. Acid. Res.* 16, 6397–6410.
18. Ayi, T.C., Loh, K.C., Rahmen, B. Ali and Li, B.F.L. (1992) *Cancer Res.* 52, 6423- 6430.
19. Tano, K., Shiota, S., Collier, J., Foote, R.S. and Mitra, S. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 686–690.
20. Hayakawa, K., Koike, G. and Sekiguchi, M. (1990) *J. Mol. Biol.* 213, 739–747.
21. Rydberg, B., Spurr, N. and Karran, P. (1990) *J. Biol. Chem.* 265, 9563–9569.
22. Staron, I.R. and Laval, F. (1991) *Biochem. Biophys. Res. Comun.* 177, 597–602.
23. Ben-Bassat, A., Bauer, K., Chang, S-Y., Myambo, K., Boosman, A. and Chang S. (1987) *J. Bacteriol.* 169, 751–757.
24. Maniatis, T., Fritsch, E.F. and Sambrok, J. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Sping Habour Laboratory Press, Cold Spring Harbor, NY.
25. Graves, R.J., Li, B.F.L. and Swann, P.F. (1989) *Carcinogenesis* (Lond.) 10, 661–666.
26. Sedgwick, B. and Lindahl, T. (1982) *J. Mol. Biol.* 154, 169–175.
27. Myers, L.C., Terranova, M.P., Nash, H.M., Markus, M.A. and Verdine, G.L. (1992) *Biochemistry* 31, 4541–4547.
28. Molinete, M., Vermeulen, W., Burkle, A., Menissier-de Murcia, J., Kupper, J.H., Hoeijmakers, Jan H.J. and de Murcia, G. (1993) *EMBO J.* 12, 2109–2117.
29. Elder R.H., Tumelty, J., Douglas, K.T., Margison, G.P. and Refferty, J.A. (1992) *Biochem. J.* 285, 707–709.
30. Santibanez-Koref, M., Elder, R.H., Fan, C.Y., Cawkwell, L., McKie, J.H., Douglas, K.T., Margison, G.P. and Rafferty, J.A. (1992) *Mol. Carcinogenesis* 5, 161–169.
31. Shiota, S. von Wronski, M.A., Tano, K., Bigner, D.D., Brent, T.P., and Mitra, S. (1992) *Biochemistry* 31, 1897–1903.
32. Pegg, A.E. and Byers, T.L. (1992) *FASEB* . 6, 2302–2309.
33. Palombo, F., Kohfeldt, E., Calcagnile, A., Nehls, P. and Dogliotti, E. (1992) *J. Mol. Biol.* 223, 587–594.
34. Bronstein, S.M., Skopek, T.P. and Swenberg, J.A. (1992) *Cancer Res.* 52, 2008- 2011.
35. Kastan, M.B., Zhan, Q., El-Deiry, W.S., Carrier, F., Jacks, T., Walsh W.V., Plunkett, B.S., Vogelstein, B. and Fornace, Jr. J. (1992) *Cell* 71, 587–597.
36. Dedrick, R.L. and Morrison, P.F. (1992) *Cancer Res.* 52, 2464–2467.
37. Wilkinson, M.C., Potter, P.M., Georgiadis, P., Patel, D.J., Swann, P.F. and Margison, G.P. (1989) *Nucl. Acids Res.* 17, 8475–8484.
38. Crone, T.M. and Pegg, A.E. (1993) *Cancer Res.* 53, 4750–4753.