The EMBO Journal vol.13 no.7 pp.1495–1501, 1994

Crystal structure of a suicidal DNA repair protein: the Ada O^6-methylguanine-DNA methyltransferase from *E.coli*

Madeleine H.Moore, Jacqueline M.Gulbis[1], Eleanor J.Dodson, Bruce Demple[2] and Peter C.E.Moody[3]

Department of Chemistry, University of York, York, YO1 5DD, UK and [2]Department of Molecular and Cellular Toxicology, Harvard School of Public Health, Boston MA 02115 USA

[1]Present address: Laboratory of Molecular Biophysics, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA
[3]Corresponding author

Communicated by T.Lindahl

The mutagenic and carcinogenic effects of simple alkylating agents are mainly due to methylation at the O^6 position of guanine in DNA. O^6-methylguanine directs the incorporation of either thymine or cytosine without blocking DNA replication, resulting in GC to AT transition mutations. In prokaryotic and eukaryotic cells antimutagenic repair is effected by direct reversal of this DNA damage. A suicidal methyltransferase repair protein removes the methyl group from DNA to one of its own cysteine residues. The resulting self-methylation of the active site cysteine renders the protein inactive. Here we report the X-ray structure of the 19 kDa C-terminal domain of the *Escherichia coli ada* gene product, the prototype of these suicidal methyltransferases. In the crystal structure the active site cysteine is buried. We propose a model for the significant conformational change that the protein must undergo in order to bind DNA and effect methyl transfer.

Key words: alkyltransferase/conformational change/DNA–protein interactions/DNA repair/nucleophilic attack

Introduction

Cellular DNA is subject to attack by a variety of environmental and cellular mutagens, which can be carcinogenic. Organisms have developed various repair mechanisms to recognize and counteract chemical modification of their DNA. Recently, X-ray structures of two DNA repair enzymes were reported: the T4*denV* endonuclease, which recognizes UV-induced photo dimers in DNA (Morikawa *et al.*, 1992) and *Escherichia coli* endonuclease III, which removes oxidized pyrimidines from DNA (Kuo *et al.*, 1992). However, the mechanisms employed by these enzymes to cleave the phosphodiester backbone and achieve glycosidic bond breakage in DNA are unknown.

The principal mutagenic effect of monofunctional alkylating agents, O^6-methylation of guanine (Schendel and Robins, 1978; Gordon *et al.*, 1990) inducing G to A transitions (Coulondre and Miller, 1977), is avoided through the action of suicidal O^6-methylguanine methyltransferases. These proteins serve as alkyl group acceptors participating in stoichiometric transfer of the O^6-methyl group to a thiol on the protein to form a thioether. The transfer returns the damaged DNA base to an unmodified state and permanently inactivates the protein (Olson and Lindahl, 1980; Lindahl *et al.*, 1982; Teo *et al.*, 1984; Demple *et al.*, 1985; Brent *et al.*, 1988). The chemistry of this reaction involves ether bond breakage associated with a large energy barrier (Lindahl *et al.*, 1988). The paradigm for this type of repair is the product of the *ada* gene of *Escherichia coli*, whose induction substantially increases the resistance of bacteria to methylation mutagenesis (Lindahl *et al.*, 1988). The *ada* gene codes for a 354 amino acid protein that possesses both O^6-methylguanine-DNA methyltransferase activity and acts as a positive regulator both for its own synthesis and for the expression of other repair genes involved in the adaptive response (Demple *et al.*, 1982, 1985; Teo *et al.*, 1984; Lindahl *et al.*, 1988). This Ada protein contains two cysteine residues active in repair: Cys^{321} for O^6-guanine and O^4-thymine adducts (Demple *et al.*, 1985) and Cys^{69} for phosphotriester adducts in DNA (Sedgwick *et al.*, 1988). The O^6-methylguanine-DNA methyltransferase was initially isolated as a stable 19 kDa C-terminal fragment of Ada (Demple *et al.*, 1982, 1985; Teo *et al.*, 1984) arising from cleavage by *OmpT* protease in cell extracts at Lys^{178} (Sedgwick, 1989). The known prokaryotic and eukaryotic O^6-alkyltransferase polypeptides in which a single cysteine is the only alkyl acceptor are homologous to the 19 kDa C-terminal Ada fragment (Brent *et al.*, 1988; Lindhal *et al.*, 1988). Secondary structure predictions (Rost and Sander, 1992) indicate a high degree of structural homology within this family of proteins (Figure 1). Thus, knowledge of the structure of the Ada O^6-alkylguanine alkyltransferase would illuminate the properties of this entire family of proteins.

Results and discussion

Protein isolation

The protein used for crystallization was the product of the *E.coli* B *ada* gene, engineered to express the C-terminal Met^{175}–Arg^{354} methyltransferase, Ada-C protein, fragment. A high level expression vector for the production of Ada-C under *tac* control was constructed (personal communication and gift from G.Verdine). Soluble Ada-C was produced in *E.coli* JM101, the cells were then harvested and lysed by sonication. Ada-C protein was purified by a combination of anion exchange and gel filtration chromatography. The native protein exhibited a pI of 7.5 upon isoelectric focusing and N-terminal sequencing carried out on the purified protein revealed loss of the methionine. The sequence was therefore numbered $Thr^{1}C$–$Arg^{179}C$, corresponding to residues Thr^{176}–Arg^{354} of full-length Ada protein.

Overall structure

The X-ray crystal structure of this 178 amino acid C-terminal fragment of Ada (Ada-C) which contains the O^6-methyl-

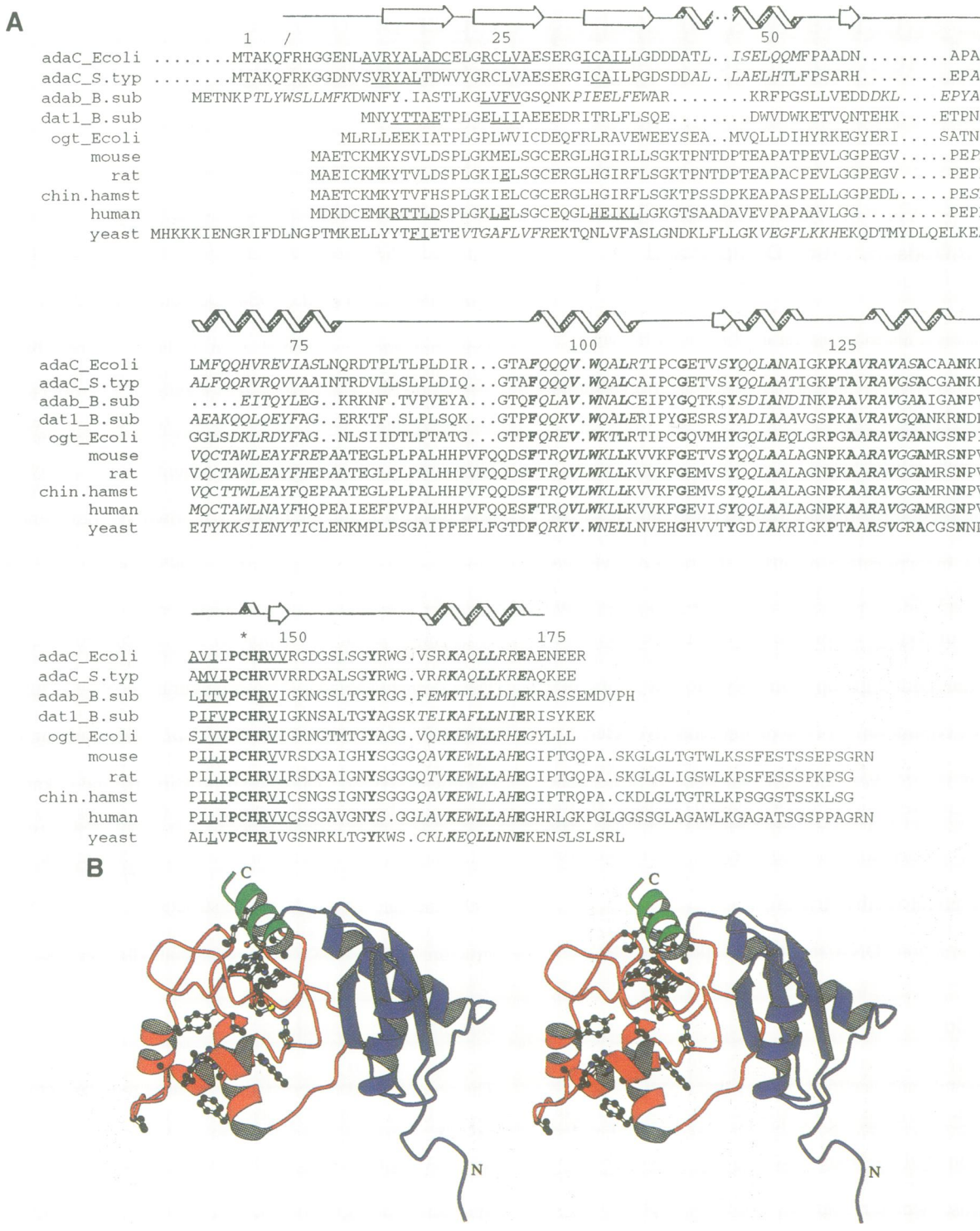

Fig. 1. (**A**) A GCG (Genetics Computer Group) amino acid sequence alignment of all O^6-alkylguanine DNA alkyltransferase (EC 2.1.1.63) sequences found in the genEMBL database. To highlight the similarity between all known alkyltransferases, absolutely conserved residues are typed in bold. In addition, secondary structure elements predicted (Rost and Sander, 1992) with 82% or greater confidence are indicated by different typeface (α-helix in italics, β-strand underlined). To compare the secondary structure predictions with the secondary structure found for the AdaC protein, the observed secondary structural elements are drawn across the top of the alignments for the following species: adaC__Ecoli, C-terminal 180 amino acid residues of *E.coli* K12 Ada protein (the structure presented is of the protein from *E.coli* B, which has residue 142 as I and 143 as V); adaC__S.typ, C-terminal 179 amino acid residues of *Salmonella typhimurium* Ada protein; adab__B.sub, sequence of AdaB protein from *Bacillus subtilis*; dat1__B.sub, the dat protein from *B.subtilis*; ogt__Ecoli, the ogt protein from *E.coli*; chin.hamst, chinese hamster alkyltransferase; yeast, alkyltransferase in *Saccharomyces cerevisiae*. (**B**) A schematic stereo drawing of the 178 amino acid C-terminal fragment of Ada with O^6-alkylguanine DNA alkyltransferase acitvity. The predominantly alpha helical and random coil nature of this protein is illustrated by the cartoon representation of secondary structural elements. The different secondary stuctural elements of domains one (N-terminal) and two (C-terminal) are coloured in blue and red, respectively. Residues $^{163}C-^{174}C$ (338–349 in full-length Ada protein) which comprise the C-terminal helix are illustrated in green. To illustrate the location in the C-terminal domain of all residues identical in all the alkyltransferase sequences listed in Figure 1A, the side-chains of these residues are shown in ball and stick and colour coded according to atom type. The figure was generated by the computer program MOLSCRIPT (Kraulis, 1990).

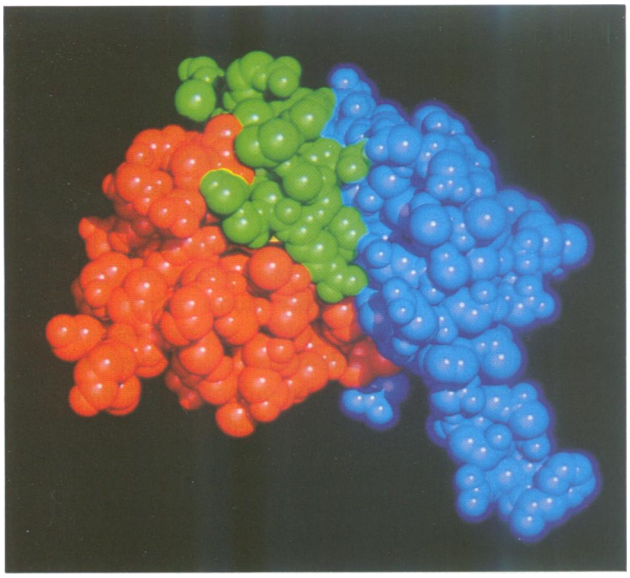

Fig. 2. A space filling drawing of the alkyltransferase structure with all solvent atoms removed. This representation of the protein illustrates the almost completely buried environment of the active site Cys^{146}C (Cys^{321} in Ada); the barely visible yellow speck in the centre of the protein is its sulfur side-chain atom. For easy comparison, the orientation of this display is identical to Figure 1B) and the same colour code has been used. The figure was generated by the computer program QUANTA (Molecular Simulations, Burlington, MA).

guanine-DNA methyltransferase activity has revealed two distinct domains with novel topology. The N-terminal domain comprises the first 84–88 residues and the C-terminal domain is made up of the remaining residues (Figure 1). The N-terminal domain has a fold reminiscent of part of ribonuclease H (Katayanagi *et al.*, 1990; Yang *et al.*, 1990), whereas the C-terminal residues fold to form an exclusively helix and coil domain with a previously unobserved organization. The most highly conserved methyltransferase sequences correspond well to the secondary structural elements predicted for this structurally novel C-terminal domain of Ada-C (Figure 1). The disposition of the three helices coloured red in Figure 1B suggests a 'helix–turn–helix (HTH) variant' DNA-binding motif (Schwabe and Travers, 1993). However, for Ada-C the residues in the putative DNA binding helix of this HTH region are far from the active site of the protein. Perhaps the HTH motif mediates a generalized binding to DNA substrate rather than specifically to the O^6-alkylguanine target (Demple, 1990). The C-terminal domain contains not only the Pro-Cys-His-Arg active site sequence found in all known O^6-alkylguanine-DNA alkyltransferases but also all the other identical and chemically similar residues found in this family (Figure 1).

Requirement for a conformational change

The most surprising feature of this structure is the buried position of the active site thiol of Cys^{146}C which

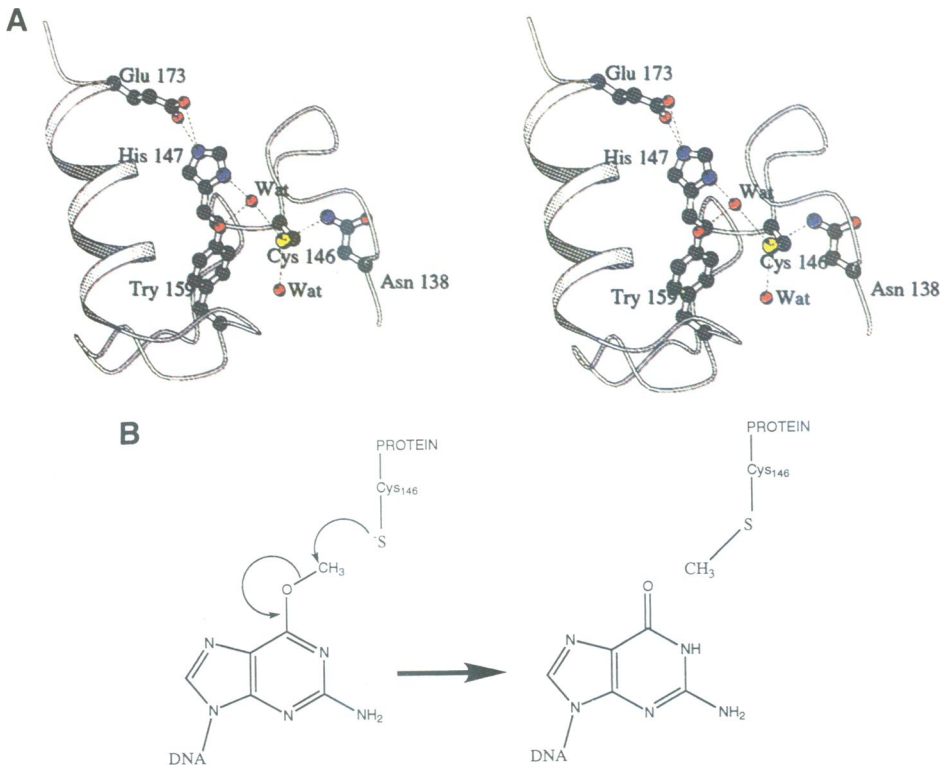

Fig. 3. (**A**) Stereoview of the environment of the active site thiol of Cys^{146}C. The spatial arrangement of all key residues and the two solvent molecules in close proximity to Cys^{146}C is illustrated. Hydrogen bonds are drawn as dashed lines and a smoothed protein backbone is traced. One of the ordered water molecules is involved in a hydrogen-bonded network being situated 3.2 Å from Sγ of Cys^{146}C, 3.0 Å from Nδ of His^{147}C and 3.0 Å from Oη of Tyr^{159}C. The latter two groups are perhaps involved in stabilizing the bound substrate. The side-chain of conserved Glu^{173}C is 3.2 Å from Nϵ of His^{147}C. Rotation of the His^{147}C side-chain would place its Nδ in the position occupied by the water, directly hydrogen bonded to Sγ of Cys^{146}C. A methyl group could easily replace the other ordered water molecule interacting with this thiol. The 150$_C$-165$_C$ residue loop leading into the C-terminal helix can be seen in the bottom left hand corner of the figure. The figure was generated by the computer program MOLSCRIPT (Kraulis, 1990). (**B**) A line drawing of the overall methyltransferase suicide repair reaction. This shows the methyl group being transferred from the O^6 postion on a guanine in DNA to the sulfur side-chain of the active site thiol, Sγ of Cys^{146}C.

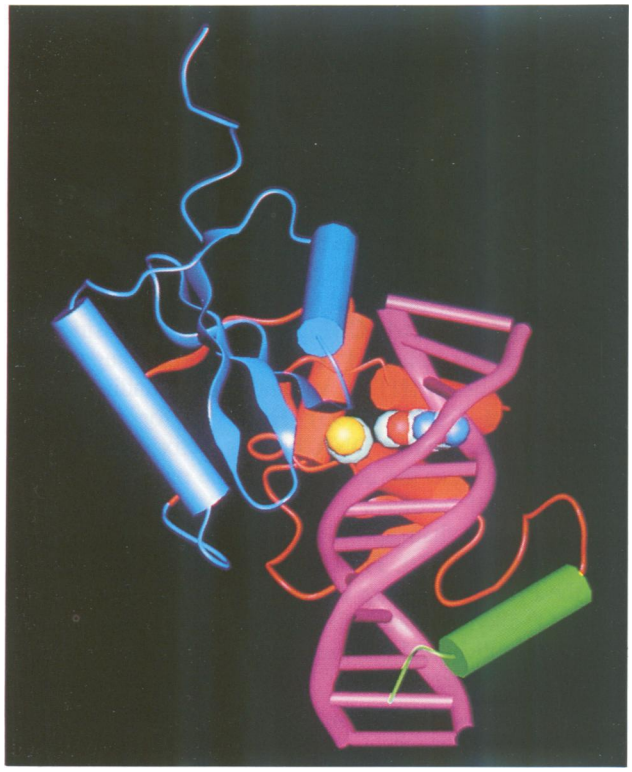

Fig. 4. A schematic cartoon drawing of our simple DNA–protein complex model. The protein has been coloured with the same colour scheme as used in Figures 1B and 2. The van der Waal's surface of the active site cysteine side-chain has been colour coded according to atom type. A cartoon representation of the duplex DNA substrate was generated from coordinates of d[CGC(**m**^{6}**G**)AATTTGCG]$_2$ obtained from Brookhaven protein data bank (Leonard *et al.*, 1990). The van der Waal's surface of one of the O^6-methylguanine bases has been colour coded according to atom type in order to illustrate its position relative to Cys^{146}C in our model. This figure highlights the extent of conformational change in the alkyltransferase required to bind its preferred substrate, duplex DNA. In our model the active site thiol is positioned directly in line for nucleophilic attack (*S*-methyl–O bond angle of 180°) and the C-terminal helix (green) of the protein lies in the major groove of bound DNA. In this model the absolutely conserved lysines and arginine located in the C-terminal helix make good contacts with the phosphate backbone. The figure was generated by the computer program QUANTA (Molecular Simulations, Burlington, MA).

corresponds to Cys^{321} in Ada (see Materials and methods) (Figure 2). Consequently, a conformational change must take place to allow this cysteine to carry out nucleophilic attack on the target methyl group either in duplex DNA or even as a simple O^6-methylguanine base. The rate determining step in the reaction of the homologous (Figure 1) human O^6-alkylguanine-DNA alkyltransferase with DNA appears to be DNA binding (Takahashi *et al.*, 1990; Chan *et al.*, 1993). In addition, the interaction of the human methyltransferase with DNA induces a conformational change in the protein, as monitored by circular dichroism and fluorescence analysis (Takahashi *et al.*, 1990; Chan *et al.*, 1993). It is noteworthy that, were the active site cysteine not already known, this structure might suggest that it does not play a role in the active site chemistry. Since the active site residues have not been determined unequivocally for the other repair enzymes subjected to structural analysis (Kuo *et al.*, 1992; Morikawa *et al.*, 1992), it is not known whether these structures represent the substrate bound enzyme conformations.

The methyl transfer reaction

The mechanism for alkyl transfer from O^6-methylguanine to Cys^{146}C (Demple, 1990; Spratt and de los Santos, 1992) involves classical acid catalysed S_N2 chemistry with a cysteine thiolate acting as the nucleophile. A concomitant positive charge transfer onto O^6 of guanine and nucleophilic attack at the carbon centre has been proposed (Spratt and de los Santos, 1992). The only atoms within hydrogen bonding distance of Sγ of Cys^{146}C are two well-ordered water molecules and the side-chain of conserved Asn^{138}C (Figures 1 and 3). One of these ordered water molecules may provide a clue to the position that will be occupied by O^6 of the bound substrate. One of these waters is also hydrogen bonded to the Nδ of conserved His^{147}C and to the Oη of conserved Tyr^{159}C. The Nϵ of His^{147}C is, in turn, hydrogen bonded to the side-chain of conserved Glu^{173}C. Serine proteases and lipases use this type of acid and histidine stereochemistry to generate the reactive nucleophile (Blow *et al.*, 1969). However, it is not known whether this arrangement is preserved in the methyltransferase structure during methyl group transfer.

A DNA binding model

This enzyme's favoured substrate is duplex DNA, however, the structure we present here does not allow the active site cysteine to contact the target O^6-methyl group in a DNA molecule. Nevertheless, a simple movement about the loop residues 150$_C$–160$_C$ would swivel the C-terminal helix (residues 165$_C$–175$_C$) to expose a potential DNA binding surface and simultaneously render the thiol of Cys^{146}C accessible to duplex DNA. In this scheme, the C-terminal helix can be accommodated in the major groove (Figure 4) and the entire protein would cover ~8 bp, consistent with fluorescence and circular dichroism studies (Takahashi *et al.*, 1990; Chan *et al.*, 1993). This movement would also break the His^{147}C–Glu^{173}C interaction and could allow His^{147}C to rotate about χ1 and χ2. Such a rotation would allow the Nδ of His^{147}C may replace one of the ordered water molecules, only 3.2 Å from the Sγ of Cys^{146}C, and well situated to generate the reactive thiolate anion (Figure 3). Such a complex would be poised to carry out the methyl transfer reaction, and does not require a conformational change in the DNA. Nevertheless, a conformational change in the substrate DNA during its repair certainly cannot be excluded.

Mutant Ada proteins with strongly reduced methyl transfer rates have been identified (Demple, 1986). These proteins nevertheless form *S*-methylcysteine stoichiometrically during their reaction with O^6-methylguanine in DNA, which indicates that the fundamental chemistry of the methyl transfer is preserved. For one of the mutants, Ada3 protein, the mutational change has been determined, a GC to AT transition that alters Ala^{136}C to a valine residue (A.Abou-Zamzam and B.Demple, unpublished data). This is a surface residue which, if altered to a valine, may bury itself in the active site pocket and in so doing alter the ability of the methyltransferase to undergo the conformational change required for methyl transfer. Indeed the corresponding residue in other methyltransferases is hydrophilic (glycine, lysine or arginine; Figure 1A).

Table I. Data collection and heavy atom refinement statistics for both crystal forms of O^6-methylguanine-DNA methyltransferase

Crystal form	Monoclinic	Orthorhombic
Spacegroup	$P2_1$	$P2_12_12_1$
Cell dimensions	a = 42.6 Å b = 45.9 Å c = 46.4 Å β = 114.0°	a = 38.0 Å b = 55.5 Å c = 76.5 Å
Native data		
Resolution (Å)	40–2.1	35–2.1
Total No. of reflections	19437	39698
No. of unique reflections	9017	9926
$R_{merge} = \Sigma\|I - <I>\| / \Sigma I$	0.058	0.078
Completeness (%)	86	99
Derivative data		
Resolution (Å)	35–3.5	35–3.5
No. of reflections phased (acentric; centric)	2301; 206	2279; 675
Overall figure of merit[a] (acentric; centric)	0.41; 0.60	0.33; 0.55
Ethyl mercury phosphate		
Fractional coordinates x,y,z, occupancy, anomalous occupancy	0.141, 0.000, 0.383, 0.54, 0.45	0.125, 0.171, 0.082, 0.36, 0.31
No. of reflections phased (acentric, centric)	2211; 198	2277; 670
Phasing power[b] (acentric; centric)	1.4; 1.6	0.9; 0.8
Cullis R[c] (acentric; centric)	0.75; 0.68	0.82; 0.70
Lead (II) acetate		
Fractional coordinates x,y,z, occupancy, anomalous occupancy	0.011, 0.165, 0.894, 0.25, 0.09	0.505, 0.394, 0.259, 1.00, 0.28 0.079, 0.160, 0.078, 0.62, 0.00
No. of reflections phased (acentric; centric)	1370; 122	1373; 429
Phasing power[b] (acentric; centric)	0.6; 0.7	0.5; 0.5
Cullis R[c] (acentric; centric)	0.93; 0.82	0.94; 0.85

[a]Overall figure of merit = $\int P(\theta)\exp(i\theta)d\theta / \int P(\theta)d\theta$, where P is the probability distribution of θ, the phase angle and $\int$ is from $\theta = 0$ to 2π.
[b]Phasing power = heavy-atom structure factor amplitude/residual lack of closure = $\Sigma|F_hc| / \Sigma[||F_po|\exp(i\theta) + F_hc| - |F_{ph}o||]$, where $|F_hc|$ is the calculated diffraction amplitude of the heavy atom, $|F_po|$ and $|F_{ph}o|$ are the observed amplitudes for the protein and heavy atom derivative, respectively, θ_c is the calculated phase and the sum, Σ, is over all observations.
[c]Cullis $R = \Sigma||F_{ph}o \pm F_po| - F_hc| / \Sigma|F_{ph}o - F_po|$

Substrate specificity

Alkylation at the O^6 position of guanine slightly distorts the phosphodiester backbone in double-stranded DNA (Patel *et al.*, 1986; Kalnik *et al.*, 1988; Leonard *et al.*, 1990), an effect that may be important for recognition by the methyltransferase. The overall structure of DNA must play a role in the recognition and transfer process, because the alkyltransferase shows a strong preference for duplex over single-stranded DNA and the rate of repair of synthetic oligonucleotides increases with increasing length of substrate (Demple, 1990). For comparison, the free O^6-methylguanine base is an $\sim 10^7$-fold poorer substrate than O^6-methylguanine in duplex DNA (Demple, 1990). In addition, the *E.coli* Ada methyltransferase does not repair O^6-methylguanine residues present in Z-DNA (Boiteux *et al.*, 1985). The necessity for a conformational change to produce an active protein may explain these large differences in reaction rates, as the DNA binding step may optimize the positions of functional groups essential for promoting the suicide reaction.

Suicide enzymes

The methyltransferase overcomes a significant energy barrier in cleaving the ether bond in O^6-methylguanine (Lindahl *et al.*, 1988). The protein undergoes suicide inactivation (Lindahl *et al.*, 1982); apparently no mechanism exists to dealkylate the resulting *S*-alkylcysteine moiety (Demple *et al.*, 1982; Lindahl *et al.*, 1988). Other methyltransferring systems employ cofactors or side-chains whose chemistry facilitates reversibility (e.g. glutamate methylation in chemotaxis; Kleene *et al.*, 1977). The chemical stability of *S*-methylcysteine and the lack of cofactors as transfer sites, and perhaps the reburial of *S*-alkylcysteine itself could, after the transfer has taken place, limit further reactions.

Obtaining a suitable substrate-bound complex for X-ray structural studies will be difficult because in the S_N2 methyl transfer reaction there are no chemical intermediates and only a single transition state. In addition the repair reaction with O^6-methylguanine in DNA as substrate is extremely fast for a DNA-processing enzyme, $k > 10^8\ M^{-1}\ sec^{-1}$ (Demple, 1990), which indicates that pre-transfer complexes would not be obtained with the wild-type protein. Even the Ada3 mutant protein mentioned above has a sufficient rate of methyl transfer to preclude formation of a stable complex with O^6-methylguanine-containing DNA. Mutant proteins which bind O^6-methylguanine in DNA but do not carry out the transfer reaction and non-transferable substrate analogues that bind to the protein have not yet been identified. The construction of such substrates or the isolation of mutant methyltransferases that bind without methyl transfer will greatly facilitate structural analysis of this fundamental repair process.

Table II. Refinement parameters of the monoclinic crystal form of O^6-methylguanine-DNA methyltransferase

Final refinement parameters	
No. of atoms (non-hydrogen)	1465
No. of water molecules	157
Range of spacings (Å)	10.0 - 2.1
R-factor = $\Sigma\|F_{obs}-F_{calc}\| / \Sigma\|F_{obs}\|$ (all data $F > 0$)	19.0%
R.m.s. deviations from ideal geometry of the final model distances (target values) (Å)	
Bonds (1−2 neighbour)	0.014 (0.020)
Angles (1−3 neighbour)	0.057 (0.040)
Interplaner (1−4 neighbour)	0.067 (0.050)
Planer groups (Å)	0.008 (0.020)
Chiral volumes (Å^3)	0.171 (0.150)

Materials and methods

The structure was solved by a combination of multiple isomorphous heavy-atom replacement, averaging between two crystal forms and density modification. The structure has been refined in space group $P2_1$ to a crystallographic *R*-value of 0.19 using reflections from 10 to 2.1 Å spacing (Table I).

Crystallization

The first crystals of a $Gln^{179}-Arg^{354}$ C-terminal fragment were obtained in space group $P2_1$ (Moody and Demple, 1988). Crystals of the $Thr^{176}-Arg^{354}$ C-terminal protein, Ada-C, have since been obtained in space groups $P2_1$ and $P2_12_12_1$. Two crystal forms were grown by microdialysis. Protein (5−50 mg/ml) in 200−500 mM sodium acetate, 10 mM dithiothreitol, 10−20 mM Tris−acetate buffer pH 7.8−8.0 was dialysed against 10 mM dithiothreitol, 300 mM glucose or maltose, 10 mM Tris−acetate, pH 7.8−8.0. The monoclinic crystal form is isomorphous with the original $P2_1$ crystals (Moody and Demple, 1988).

Data collection

Native data for the monoclinic crystal form were collected at a wavelength of 0.89 Å on BL6 at the Photon Factory, using a Sakabe/Weissenberg camera and merged with data from a MAR image plate and RU300HB molybdenum rotating anode in the laboratory of Robert Liddington at Harvard Medical School. Data were processed using DENZO and merged in SCALEPACK (Z.Otwinoski, Yale). The native orthorhombic data was collected with an RAXIS IIC image plate system and an RU200HB copper rotating anode and processed using MOSFLM (V4.3, A.G.W.Leslie, MRC-LMB Cambridge) and the CCP4 program suite. All derivative data (Table I) were also collected at York with a Xentronics multi-wire proportional counter and RU200HB copper rotating anode and processed using XDS (Kabsch, 1988).

Crystallographic analysis

An initial model comprising 834 atoms was built into a 3.5 Å MIR phased orthorhombic map (Table I), however, further confident interpretation was impossible. This model was then used as a molecular replacement search object against the monoclinic native data using a rotation search (Rossmann and Blow, 1962). The phased translation function solution (Reed and Schierbek, 1988) resulted in the superimposition of the orthorhombic and monoclinic mercury sites. Rigid body refinement [as implemented in X-plor (Brunger *et al.*, 1987)] of the monoclinic model allowed the matrix relating the two crystal forms to be defined more precisely. The monoclinic and orthorhombic maps where then averaged (Bricogne, 1976), the solvent boundary defined using O and the phases extended to 2.5 Å using SQUASH (Cowtan and Main, 1993). The 2.5 Å maps where then averaged again. At this stage the SIRAS monoclinic map was greatly improved and phases extended to 2.1 Å by the use of the DENMOD program by Z.Otwinoski (unpublished). A combination of averaged and DENMOD maps was used to extend and modify the model to include 1187 atoms using O (Jones *et al.*, 1991). The *R* factor at this stage with the monoclinic data was 48.73% against all data in the range 10−2.1 Å. This model was subjected to Powell minimization and simulated annealing in X-plor to give an *R* of 34.1% for the same data. Several cycles of rebuilding using FRODO (Jones, 1978), Hendrickson−Konnert least-squares refinement (Konnert and Hendrickson, 1981) and X-plor (Brunger *et al.*, 1988) allowed further correction and the insertion of 157 solvent molecules. The quality of the monoclinic structure is shown in Table II.

Acknowledgements

This work was funded by the Yorkshire Cancer Research Campaign. A NATO collaborative research grant funded travel and the Dreyfus Foundation supported work done in Boston. The authors acknowledge the help of A.M.Brozowski, S.M.Cutfield, J.F.Cutfield, G.G.Dodson, W.N.Hunter, J.Jeffery, H.J.Kemmish, R.Liddington, Z.Otwinowski, H.F.J.Savage, M.Suzuki, S.P.Tolley, R.Ward and the SERC and MRC supported Protein Structure Group at York. Coordinates for the refined model of the *E.coli* O^6-methylguanine-DNA methyltransferase will be submitted to the Brookhaven Protein Data Bank upon publication.

References

Boiteux,S., de Oliveira,R.C. and Laval,J. (1985) *J. Biol. Chem.*, **260**, 8711−8715.
Blow,D.M., Birktoft.J.J. and Hartley,B.S. (1969) *Nature*, **221**, 337−340.
Brent,T.P. *et al.* (1988) *Proc. Natl Acad. Sci. USA*, **85**, 1759−1762.
Bricogne,G. (1976) *Acta Crystallogr.*, **A32**, 832−847.
Brunger,A.T., Kurian,J. and Karplus,M. (1987) *Science*, **235**, 458−460.
Chan,C.-L., Wu,Z., Ciardelli,T., Eastman,A. and Bresnick,E. (1993) *Arch. Biochem. Biophys.*, **300**, 193−200.
Coulondre,C. and Miller,J.H. (1977) *J. Mol. Biol.*, **117**, 577−606.
Cowtan,K.D. and Main.,P. (1993) *Acta Crystallogr.*, **D49**, 148−157.
Demple,B. (1986) *Nucleic Acids Res.*, **14**, 5575−5589.
Demple,B. (1990) In Paik,W.K. and Kim,K. (eds), *Protein Methylation*. CRC Press, Boca Raton, Florida.
Demple,B., Jacobsson,A., Olsson,M., Robins,P. and Lindahl,T. (1982) *J. Biol. Chem.*, **257**, 13776−13780.
Demple,B., Sedgewick,B., Robins,P., Totty,N., Waterfield,M.D. and Lindahl,T. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 2688−2692.
Gordon,A.J.E., Burns,P.A. and Glickman,B.W. (1990) *Mutat. Res.*, **233**, 92−103.
Jones,T.A. (1978) *J. Appl. Crystallogr.*, **11**, 268−272.
Jones,T.A., Zou,J.-Y., Cowan,S.W. and Kjeldgaard,M. (1991) *Acta Crystallogr.*, **A47**, 110−119.
Kabsch,W. (1988) *J. Appl. Crystallogr.*, **21**, 916−924.
Kalnik,M.W., Kouchakdjian,M., Li,B.F.L., Swann,P.F. and Patel,D.J. (1988a) *Biochemistry*, **27**, 100−108.
Kalnik,M.W., Kouchakdjian,M., Li,B.F.L., Swann,P.F. and Patel,D.J. (1988b) *Biochemistry*, **27**, 108−122.
Katayanagi,K., Miyagawa,M., Matsushima,M., Ishikawa,M., Kanaya,S., Ikehara,M., Matsuzaki,T. and Morikawa,K. (1990) *Nature*, **347**, 306−309.
Kleene,S.J., Toews,M.L. and Adler,J. (1977) *J. Biol. Chem.*, **252**, 3214−3218.
Konnert,J.H. and Hendrickson,W.A. (1981) In Srinivasen,R. (ed.), *Biomolecular Structure. Conformation, Function and Evolution, Vol. I.* Pergamon Press, Oxford, pp. 43−57.
Kuo,C.-F., McRee,D.E., Fisher,C.L., O'Handley,S.F., Cunningham,R.P. and Tainer,J.A. (1992) *Science*, **258**, 434−440.
Kraulis,P.J. (1990) *J. Appl. Crystallogr.*, **24**, 946−950.
Leonard,G.A., Thomson,J., Watson,W.P. and Brown,T. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 9573−9576.
Lindahl,T., Demple,B. and Robins,P. (1982) *EMBO J.*, **1**, 1359−1363.
Lindahl,T., Sedgewick,B., Sekiguchi,M. and Nakabeppu,Y. (1988) *Annu. Rev. Biochem.*, **57**, 133−157.
Moody,P.C.E. and Demple,B. (1988) *J. Mol. Biol.*, **200**, 751−752.
Morikawa,K., Matsumoto,O., Tsujimoto,M., Katayanagi,K., Ariyoshi,M., Doi,T., Ikehara,M., Inaoka,T. and Ohtsuka,E. (1992) *Science*, **256**, 523−526.
Olson,M. and Lindahl,T. (1980) *J. Biol. Chem.*, **255**, 10569−10571.
Patel,D.J., Shapiro,L., Kozlowski,S.A., Gaffney,B.L. and Jones,R.A. (1986) *Biochemistry*, **25**, 1027−1036.
Patel,D.J., Shapiro,L., Kozlowski,S.A., Gaffney,B.L. and Jones,R.A. (1986) *Biochemistry*, **25**, 1036−1042.
Read,R.J. and Schierbek,A.J. (1988) *J. Appl. Crystallogr.*, **21**, 490−495.
Rossmann,M.G. and Blow,D.M. (1962) *Acta Crystallogr.*, **15**, 24−31.
Rost,B. and Sander,C. (1992) *Nature*, **360**, 540.
Schendel,P.F. and Robins,P.E. (1978) *Proc. Natl Acad. Sci. USA*, **75**, 6017−6020.

Schwabe,J.W.R. and Travers,A. (1993) *Curr. Biol.*, **3**, 628–630
Sedgwick,B. (1989) *J. Bacteriol.*, **171**, 2249–2251.
Sedgwick,B., Robins,P., Totty,N. and Lindhal,T. (1988) *J. Biol. Chem.*, **263**, 4430–4433.
Spratt,T.E. and de los Santos,H. (1992) *Biochemistry*, **31**, 3688–3694.
Takahashi,M., Sakumi,K. and Sekiguchi.M. (1990) *Biochemistry*, **29**, 3431–3436.
Teo,I., Sedgwick,B., Demple,B., Li,B. and Lindahl,T. (1984) *EMBO J.*, **3**, 2151–2157.
Yang,W., Hendrickson,W.A., Crouch., R.J. and Satow,Y. (1990) *Science*, **249**, 1398–1405.

Received on November 11, 1993; revised on January 24, 1994