Mol Gen Genet (1998) 258: 69–77

ORIGINAL PAPER

M. M. Leclere · M. Nishioka · T. Yuasa
S. Fujiwara · M. Takagi · T. Imanaka

The O^6-methylguanine-DNA methyltransferase from the hyperthermophilic archaeon *Pyrococcus* sp. KOD1: a thermostable repair enzyme

Received: 2 October 1997 / Accepted: 28 November 1997

Abstract The enzyme O^6-methylguanine-DNA methyltransferase (MGMT) is the most common form of cellular defense against the biological effects of O^6-methylguanine (O^6-MeG) in DNA. Based on PCR amplification using primers derived from conserved amino acid sequences of MGMTs from 11 species, we isolated the DNA region coding for MGMT from the hyperthermophilic archaeon *Pyrococcus* sp. KOD1. The MGMT gene from KOD1 (*mgtk*) comprises 522 nucleotides, encoding 174 amino acid residues; its product shows considerable similarity to the corresponding mammalian, yeast and bacterial enzymes, especially around putative methyl acceptor sites. Phylogenetic analysis of MGMTs showed that archaeal MGMTs were grouped with their bacterial counterparts. The location of the MGMT gene on the KOD1 chromosome was also determined. The cloned KOD1 MGMT gene was overexpressed using the T7 RNA polymerase expression system, and the recombinant protein was purified by ammonium sulfate fractionation, heat treatment, ion-exchange chromatography and gel filtration chromatography. The purified recombinant protein was assayed for its enzyme activity by monitoring transfer of [^{3}H]methyl groups from the substrate DNA to the MGMT protein; the activity was found to be stable at 90° C for at least 30 min. When the *mgtk* gene was placed under the control of the *lac* promoter and expressed in the methyltransferase-deficient *Escherichia coli* strain KT233 (Δ*ada*, Δ*ogt*) cells, a MGMT was produced. The enzyme was functional in vivo and complemented the mutant phenotype, making the cells resistant to the cytotoxic properties of the alkylating agent *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine.

Key words DNA repair · O^6-methylguanine-DNA methyltransferase · Archaea · Mutagenesis · Alkylation

Communicated by W. Goebel

M. M. Leclere · M. Nishioka · T. Yuasa
S. Fujiwara · M. Takagi
Department of Biotechnology, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0891, Japan

T. Imanaka (✉)
Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University, Kyoto 606-8501, Japan
Fax: +81-75-753-4703; e-mail: imanaka@sbchem.kyoto-u.ac.jp

Introduction

The alkylation of guanine bases at the O^6 position is considered to be a major source of induced mutations and an important contributory factor in the development of cancer in higher organisms. Simple methylating agents, such as *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine (MNNG) and *N*-methyl-*N*-nitrosourea (MNU), are potent mutagens and carcinogens that can cause cell death, acting through covalent modification of the genome to generate miscoding alkylated base derivatives and lesions that block DNA replication. Among the many methylated bases formed, O^6-methylguanine (O^6-MeG) appears to be a major premutagenic lesion in organisms (Loveless 1969; Strauss et al. 1975; Sukumar et al. 1983). Mutations involving the replacement of G-C by A-T basepairs in DNA are often found in tumors induced by alkylating agents. These mutations arise when O^6-MeG is formed and pairs with thymine during DNA replication (Coulondre and Miller 1977; Ito et al. 1994). O^6-MeG is efficiently repaired by specialized DNA repair enzymes, the O^6-MeG DNA methyltransferases (MGMT), which transfer methyl groups from O^6-MeG, as well as from the minor methylated base O^4-methylthymine (O^4-MeG), in the DNA to an internal cysteine residue, repairing DNA lesions in a single-step reaction (Karran and Lindahl 1979; Olsson and Lindahl 1980; Koike et al. 1990; Zak et al. 1994; Kawate et al. 1995). Since this reaction irreversibly inactivates the enzyme, the capacity for repair of O^6-MeG depends on the number of active enzyme molecules per cell. The enzyme is present in organisms from bacteria to humans, and the characteristic suicidal reaction mechanism is

conserved in all bacterial and eukaryotic cells (Demple 1990). The hyperthermophilic archaeon *Pyrococcus* sp. strain KOD1 was isolated from a solfataric hot spring on Kodakara Island, Kagoshima, Japan (Morikawa et al. 1994). Enzymes derived from this strain have been shown to be extremely thermostable, and amino acid sequences of those enzymes are closely related to their eukaryotic homologs (Fujiwara et al. 1996). Studies on hyperthermophilic archaea have begun only very recently and our understanding of their biochemistry, genetics, and molecular biology is still at an early stage. Archaea constitute a third domain of living organisms, sharing characteristics with both bacterial and eukaryotic cells (Woese et al. 1990). On the molecular level, Archaea are related to Eukarya and they can be used to understand the nature and evolution of eukaryotic cells. So far, no information on MGMT has been available from Archaea, with the exception of the sequence of an MGMT homolog from *Methanococcus jannaschii* (Genbank accession number U67593). Thus, we set out to characterize MGMT from KOD1 with the aim of understanding the DNA repair mechanism in Archaea and its evolution. This paper describes the cloning, sequencing and expression of the gene, and the purification and basic characterization of KOD1 MGMT. The amino acid sequence exhibits extensive homology to those of the mammalian, yeast and bacterial enzymes, and the recombinant methyltransferase is functional both in vitro and in vivo.

Materials and methods

Bacterial strains, plasmids, and phages

E. coli JM109 was used as a host for plasmid DNA. *E. coli* TG-1 was used to subclone gene fragments for nucleotide sequencing. *E. coli* strain KT233 (Δ*ada* Δ*ogt*; Takano et al. 1991) was used to express the KOD1 MGMT gene and test for complementation of the mutant phenotype of the host. *E. coli* HMS174(DE3)pLysS [F-*recA* hsd^R (r_{k12}-m_{k12}^+) Rif^R (DE3) pLysS(Cm^R)] was used as a host for pET-8c (Studier et al. 1986) to overexpress the MGMT gene from strain KOD1. Bacteriophages M13 mp18 and mp19 were used for preparation of single-stranded DNA.

Media

L broth (10 g tryptone, 5 g yeast extract, and 5 g NaCl in 1 l of deionized water, pH 7.2) was used for the culture of *E. coli*. 2 × YT broth (16 g tryptone, 10 g yeast extract, and 5 g NaCl in 1 liter of deionized water) was used for the preparation of phage DNA. NZCYM medium (10 g NZ amine, 5 g NaCl, 5 g yeast extract, 1 g casamino acids, and 2 g $MgSO_4.7H_2O$ in 1 l of deionized water, pH 7.0) was used for enzyme preparation. Ampicillin was used at a final concentration of 100 μg/ml.

Chemicals

N-[^{3}H]methyl-*N*-nitrosourea (*N*-[^{3}H]MNU) (0.2 mCi), was purchased from Amersham Japan (Tokyo, Japan). *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine (MNNG) was from Nacalai-Tesque (Kyoto, Japan), and restriction endonucleases, DNA polymerase, and T4 DNA ligase were purchased from Takara Shuzo (Kyoto, Japan). Enzymes were used according to the methods recommended by the suppliers. Chromosomal DNA of strain KOD1 was prepared from a Sarkosyl lysate of cells as previously described (Imanaka et al. 1981). The GeneClean kit (Bio101, La Jolla, Calif., USA) was used to recover DNA fragments from agarose gels.

Gene amplification, nucleotide sequence determination and analysis

All the O^6-MeG-DNA methyltransferases show significant homology within the core motif I (region I) and very high homology within the putative active site (region II). In order to isolate the MGMT gene from KOD1, PCR was carried out using primers with sequences based on the conserved regions I and II (forward primer: 5′-A GGGCAGT(G/A/T/C)GG(G/A/T/C)G(G/C)(G/A/T/C)GC(G/A/T/C)-3′, reverse primer: 5′-GACCCT(A/G)TG(A/G)CA(G/A/T/C)AC-3′). The amplification reaction was performed using genomic DNA (1 μg) as a template in a 100-μl reaction containing 200 pmol of each primer; 10 μl of 10 × Vent DNA polymerase buffer (New England Biolabs, Beverly, Mass., USA); 10 μl of 2 mM dNTP mixture (Takara Shuzo) and 1 U of Vent DNA polymerase. The mixture was subjected to 30 cycles of heat treatment in a thermal cycler (Gene Amp PCR System 2400, Perkin Elmer, Foster City, Calif., USA) at 94° C, 55° C and 72° C for 1, 1 and 2 min, respectively. The PCR product was cloned into the pUC19 vector and the sequence analysis confirmed that the amplified DNA fragment was derived from a methyltransferase gene. The entire DNA region encoding MGMT was cloned using the PCR fragment labeled with digoxigenin-dUTP (DIG DNA labeling system, Boehringer Mannheim GmbH, Mannheim, Germany) as a probe for Southern hybridization. The plasmid library was constructed from 2- to 2.5-kb *Sac*I fragments of KOD1 chromosomal DNA in the pUC18 vector, and a positive clone harboring a 2.3-kb insert was selected. DNA sequencing analysis was performed by the dideoxy-chain termination method with fluorescent primers and T7 DNA polymerase (A.L.F DNA Sequencer, Pharmacia Biochemicals, Uppsala, Sweden) M13 mp18 and mp19 phages were used for preparation of the single-stranded DNA. Nucleotide and amino acid sequence analyses, including an open reading frame search, molecular weight calculation, and sequence alignment, were performed using DNASIS software (Hitachi Software Engineering, Yokohama, Japan). Homology searches were done with the BLAST program from the DDBJ E-mail Server.

Phylogenetic analysis

Alignment of amino acid sequences and estimation of the number of amino acid substitutions per site (evolutionary distances between sequences of extant species) was measured by calculating the degrees of amino acid difference between the sequences compared (Kimura 1983). Positions where gaps are present in any of the aligned sequences were excluded from the analysis. Based on the resulting evolutionary distance matrix, a phylogenetic tree was inferred using the Neighbor Joining (NJ) method (Saitou and Nei 1987). To assess the reliability of the inferred phylogenetic tree, the bootstrap method (Felsenstein 1985) was also applied. The bootstrap resamplings were repeated 1000 times, and for each resampling a tree was inferred. The bootstrap probability that a particular tree topology occurs during the resampling was evaluated. Amino acid sequences in the conserved central domain was used for the calculation using the ODEN (Molecular Evolutionary Analysis System for DNA and Amino Acid Sequences) software version 1.1.1 (National Institute of Genetics, Mishima, Japan) and Bioresearch/Sinca software (Institute of Microbial Diseases, Osaka University, Osaka, Japan).

Overexpression of the KOD1 O^6-MeG-DNA methyltransferase gene in *E. coli* and purification of the enzyme product

To overproduce the KOD1 MGMT protein, the fragment encoding the protein was cloned into an expression vector, pET-8c,

containing the T7 promoter upstream of the *Nco*I site (the upstream cloning site) and the T7 terminator downstream of *Bam*HI (the downstream cloning site). The following pair of primers were used: forward primer: 5′-CCGCAGCTGAGCCGTGGAGAAGTTTAG-3′; reverse primer: 5′-AGGGATCCTTTCAGCTTGTCCATT-3′. The fragment was amplified by using KOD1 chromosomal DNA (1μg) as a template. The expected product of 525 bp was digested with *Pvu*II and *Bam*HI and inserted into pET-8c after filling in the *Nco*I cohesive ends using T4 DNA polymerase. The structure of the construct was verified by DNA sequencing and the resultant plasmid was designated pKM-1m. *E. coli* HM174(DE3)pLysS cells, harboring pKM-1m were grown overnight at 37° C in NZCYM medium containing ampicillin (100 μg/ml). The overnight culture was inoculated (1% inoculation) into fresh NZCYM medium containing ampicillin (100 μg/ml), and incubation at 37° C was continued until the optical density at 660 nm reached approximately 0.4. The inducer IPTG (final concentration, 1 mM) was added and culture was continued for 6 h at 37° C. Cells were harvested by centrifugation (8000 × *g* for 10 min) and then washed with buffer A (50 mM TRIS-HCl, 0.1 mM EDTA pH 8.0). The cell pellet was resuspended in the same buffer. The washed cells were disrupted by sonication, centrifuged (8000 × *g* for 60 min) and the supernatant was taken as a crude extract (Fraction I). Fraction I was incubated at 90° C for 10 min, and then centrifuged (8000 × *g* for 60 min) to remove denatured proteins. The clear supernatant (Fraction II) was brought to 80% saturation by adding ammonium sulfate, and kept at 4 °C for 2 h. The precipitate was collected by centrifugation (8000 × *g* for 60 min), dissolved in buffer A, and dialyzed overnight against the same buffer (Fraction III). Fraction III was applied to a Sepharose Q column (1.6 cm × 2.5 cm, Hi-Trap Q) (Pharmacia), equilibrated with buffer A and eluted with a linear gradient of NaCl (0 to 1 M) at a flow rate of 1 ml/min. Flow-through fractions (Fraction IV) containing MGMT activity were pooled, dialyzed overnight against buffer B (50 mM potassium phosphate pH 6.5, 0.1 mM EDTA, 50 mM KCl) and applied to a Mono S column (Pharmacia, Uppsala, Sweden), previously equilibrated with buffer B. The homogeneous enzyme was eluted with a linear gradient of KCl (50 mM to 1 M) at a flow rate of 1 ml/min. Eluted MGMT fractions (Fraction V) were pooled, dialyzed against buffer A containing 0.15 M KCl, and concentrated with Centriprep-3 (Amicon, Beverly, Mass., USA). Concentrated MGMT fractions were applied to a Superose 6 column (Pharmacia) equilibrated with buffer A containing 0.15 M KCl and eluted at a flow rate of 0.5 ml/min. The molecular mass of recombinant KOD1 MGMT estimated by gel filtration chromatography showed that the protein exists in monomeric form (data not shown).

Gel electrophoresis of protein

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was carried out by the method of Laemmli (1970). Protein bands were detected by staining with 0.1% Coomassie brilliant blue R-250. The molecular weight standards (Pharmacia) and their molecular masses in kDa were: α-lactalbumin, 14.0; trypsin inhibitor, 20.1; bovine carbonic anhydrase, 30.0; ovalbumin, 43.0; albumin, 67.0; phosphorylase b, 94.0.

Enzyme assay

DNA substrate containing O^6-MeG residues was prepared by alkylation of calf thymus DNA with [^{3}H]MNU by a modification of the method previously described (Bogden et al. 1981). A 1-ml aliquot of calf thymus DNA (2 mg/ml in 50 mM TRIS-HCl pH 9.0) was incubated with 40 ml of [^{3}H]MNU (0.2 mCi) at 37° C for 4 h. The DNA was precipitated by the addition of 0.1 volume of 3 M sodium acetate and two volumes of cold ethanol. After centrifugation, the DNA pellet was washed several times with cold 95% ethanol, dried in a desiccator and redissolved overnight at 4° C in 0.15 N sodium chloride containing 0.015 N sodium citrate pH 7.0. The DNA was reprecipitated, washed, dried as above and stored at –20° C. The methylated substrate prepared in this manner had a specific activity of 2520 cpm/mg.

The enzymatic activity of MGMT was measured by detection of [^{3}H]methyl transfer from the substrate DNA to the cysteine residue at the active site of the MGMT protein (Koike et al. 1990). In the standard assay, the reaction mixture (100 μl) containing 70 mM HEPES-KOH (pH 7.5), 1 mM DTT, 5 mM EDTA, 10 mg of BSA, 13.5 mg [^{3}H]MNU-treated calf thymus DNA (34,000 cpm) and the enzyme, was incubated at 50° C for 1 h, and the reaction was terminated by adding 500 μl of 5% trichloroacetic acid and placing the sample on ice for 5 min. The mixture was heated at 90° C for 30 min to hydrolyze the DNA, placed on ice for 5 min, and then 100 μl of BSA solution (1 mg/ml) were added. The hydrolyzed DNA was transferred onto a circular Whatman GF/C filter (25 mm in diameter) under suction, and the filter was washed several times with 5% trichloroacetic acid and ethanol. Radioactivity of enzymes trapped on the filter was measured in a liquid scintillation counter.

Thermostability of the KOD1 O^6-MeG-DNA methyltransferase

The homogeneous enzyme was incubated at 75° C and 90° C in 50 mM TRIS-HCl buffer (pH 8.0), for 10, 30, and 60 min, and the remaining activity was assayed at 50° C in the standard assay described above.

Analysis of the thermal denaturation profile of the KOD1 methyltransferase by monitoring circular dichroism

Circular dichroism (CD) spectra were measured on a J-720 W automatic spectropolarimeter (Japan Spectroscopic Company, Tokyo, Japan) based on the published procedure (Kanaya and Itaya 1992). The far UV spectra (200–260 nm) were obtained using solutions containing 61.3 μg protein/ml in 0.5 mM TRIS-HCl (pH 7.5), 0.5 mM NaCl in a 1-cm optical path cell. The mean residue ellipticity, θ (measured in deg/cm^2 per dmol), was calculated using an average amino acid molecular weight of 110.

Thermal denaturation curves of KOD1 methyltransferase were obtained by monitoring the change in the CD value at 222 nm of a solution containing 61.3 μg protein/ml in 0.5 mM TRIS-HCl (pH 7.5) in a 1-cm optical path cell as the temperature was increased at a rate of 0.5° C/min and the absorbance was monitored at 222 nm. The temperature at the midpoint of transition (*Tm*) was estimated from curve fitting of the resultant CD values plotted against temperature.

Expression of KOD1 MGMT in the *E. coli* mutant KT233

The plasmid pKM-1m containing the MGMT gene was digested with *Xba*I and *Bam*HI. The resultant fragment, containing the RBS and the translational start and stop codons, was subcloned into pUC19 previously digested with the same restriction enzymes. The resulting plasmid, named pKM-2m, was used to transform *E. coli* strain KT233, which lacks both *ogt* and *ada* genes (Takano et al. 1991). A 3-ml aliquot of the culture was added to 30 ml of NZCYM broth containing 50 μg/ml of ampicillin. Exponentially growing cells (OD_{660} = 0.3~0.4) were induced with 1 mM IPTG at 37° C. After 3 h of incubation, the cells were harvested and suspended in the same volume of M9 salt medium. A 1-ml aliquot of the suspension was added to 1 ml of M9 salt medium containing various concentrations of MNNG.

After incubation at 37° C for 10 min, 2 ml of LB broth were added and, after appropriate dilution, the cells were plated on LB (50 μg/ml ampicillin) plates. The number of viable cells (colonies) was counted after overnight incubation at 37° C.

Nucleotide sequence accession number

The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL and GenBank nucleotide sequence databases under the accession number D86335.

Results

Cloning, nucleotide sequence analysis and chromosomal location of the O^6-MeG-DNA methyltransferase gene from *Pyrococcus* sp. KOD1

The probe DNA used to clone the complete gene was made by PCR. PCR primer sequences were designed after due consideration of the O^6-MeG-DNA methyltransferases from 11 different species. After confirming that the PCR product represented part of the nucleotide sequence of an MGMT gene, this fragment was used as a probe to isolate the whole gene from a library containing *Xho*I fragments of defined size. The DNA sequences of both strands were determined by the dideoxy-chain termination method with M13 phages. Figure 1 shows the nucleotide sequence of the KOD1 MGMT gene and its upstream region.

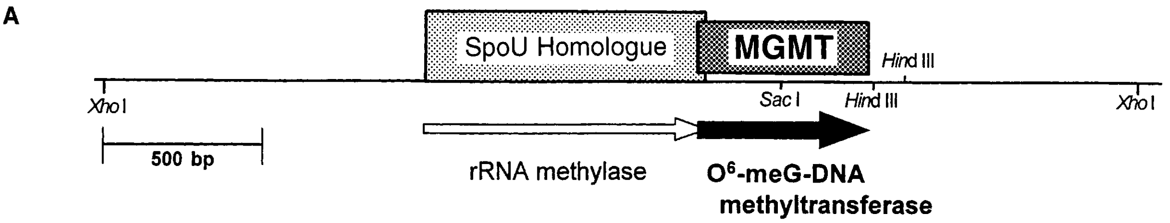
The open reading frame contains 522 nucleotides, corresponding to a protein product of 174 amino acid residues. The molecular weight of the protein calculated from the predicted amino acid sequence was 19 515. Based on the amino acid composition, KOD1 MGMT is a highly positively charged protein, with a calculated isoelectric point (pI) of 9.49. The putative ribosomal binding site (GGAGGA), was located 6 bp upstream from the start codon (Morikawa et al. 1994; Dennis 1997). A comparison of the amino acid sequences of the KOD1 MGMT with those of mammalian, yeast and bacterial homologs shows regions of sequence identity scattered along the entire molecule. Remarkable sequence similarity is observed in the C-terminal region, especially around the putative active site of methyltransferase (Fig. 2).

A physical map of the KOD1 chromosome has been constructed in our laboratory using pulsed-field gel electrophoresis of *Asc*I, *Pme*I, and *Pac*I digests (Fujiwara et al. 1996), and the location of the gene for MGMT was determined. Southern hybridization shows that the MGMT gene is located on the fragments *Asc*I D, *Pme*I A, and *Pac*I D of the KOD1 chromosome (Fig. 3).

Another ORF, encoding a protein with 236 residues, was found to show some overlap with the N-terminal region of the MGMT gene (Fig. 1). Its product showed high homology to the proteins SpoU and LAST, which belong to a new family of putative rRNA methylases (Koonin et al. 1993). The KOD1 rRNA methylase homolog showed homology especially in the region α (underlined in Fig. 1), which contains two conserved glycines preceded by hydrophobic amino acid residues, and resembles the putative S-adenosylmethionine

Fig. 1A The O^6-MeG-DNA methyltransferase gene of KOD1 (*mgtk*) overlaps a gene encoding an rRNA methyltransferase homolog (SpoU homologue) gene located upstream. **B** Nucleotide and deduced amino acid sequences of the MGMT gene and the overlapping gene for the SpoU homologue. The putative ribosomal binding sites (RBS) are underlined, and the first amino acid residue of each protein is numbered +1. Conserved regions 1 and 2 of MGMT are *boxed*. Region 2 contains the methyl acceptor cysteine residue, which is marked by a *circle*. Region α, which is also *boxed*, is a region conserved in rRNA methyltransferases. In both sequences *asterisks* mark the stop codons

A



B

```
1                                     SpoU homologue -->     M  I  S  V      4
1    AGGCCTCATCCAACCACCACAACCGTTTAAAGCATCTCCCTTAAAGACCTACCGATGATATCGGTC      66
                                               RBS
5    V  L  V  E  P  E  G  P  A  N  I  G  M  I  A  R  T  M  K  N  F  G       26
67   GTTCTGGTCGAGCCCGAGGGGCCGGCGAACATCGGCATGATAGCGAGGACGATGAAGAACTTCGGC     132

27   F  S  R  L  V  L  I  N  P  N  L  T  E  E  S  Y  A  Y  A  V  H  A       48
133  TTCTCAAGGTTAGTGCTCATCAACCCGAACCTGACCGAAGAGAGCTACGCCTACGCCGTCCACGCG     198

49   R  D  V  L  E  N  A  L  I  L  D  S  F  E  K  A  L  E  L  F  D  F       70
199  CGGGACGTCCTTGAAAACGCCCTGATTCTGGACTCATTTGAAAAGGCCCTTGAACTCTTCGACTTC     264

71   T  V  G  T  T  G  K  P  G  K  R  F  I  P  H  R  A  P  L  M  P  W       92
265  ACAGTTGGGACGACTGGAAAGCCCGGAAAGCGCTTTATCCCGCATAGAGCCCCGCTGATGCCCTGG     330

93   E  L  V  G  L  I  K  D  Y  P  G  E  V  G  I  F  F  G  R  E  S  I      114
331  GAGCTGGTTGGGCTTATCAAGGACTACCCTGGAGAGGTAGGAATATTCTTTGGGCGTGAGAGCATC     396
                                                          Region α
115  G  L  K  N  E  E  L  D  A  M  D  V  T  L  T  I  P  T  S  E  E  Y      136
397  GGCCTGAAGAACGAGGAGCTTGATGCAATGGACGTCACCCTCACGATACCCACGAGCGAGGAGTAC     462

137  P  V  M  N  L  A  Q  A  A  A  V  V  L  Y  E  L  S  K  K  R  P  E      158
463  CCCGTTATGAACCTCGCACAGGCCGCAGCGGTAGTCTTGTACGAGCTGTCGAAGAAGAGGCCAGAG     528

159  P  Y  V  E  A  L  K  P  A  T  R  E  E  K  E  A  L  V  R  T  W  E      180
529  CCATACGTTGAGGCATTAAAGCCAGCTACCAGGGAGGAGAAGGAAGCGCTTGTAAGAACCTGGGAA     594

181  K  L  L  N  T  L  D  Y  P  K  D  A  E  R  R  D  V  F  V  K  V  F      202
595  AAACTTCTCAATACTCTCGACTATCCAAAAGATGCTGAGAGAAGGGACGTCTTTGTTAAGGTCTTT     660

203  Q  R  A  V  G  R  A  F  L  Y  G  R  E  V  H  T  L  I  G  P  L  R      224
661  CAGAGGGCCGTTGGCAGGGCGTTCCTCTACGGGAGGGAAGTCCACACGCTCATCGGGCCTCTAAGA     726

225  K  A  L  K  R  L  E  E  C  R  C  *                                    235
727  AAGGCCCTTAAAAGGCTGGAGGAATGCCGATGCTGAGCGTGGAGAAGTTTAGAGTCGGGGAGAGAG     792
1      MGMT -->      RBS            M  L  S  V  E  K  F  R  V  G  E  R      12

793  TGGTTTGGATCGGTGTAATCTTCTCTGGGAGGGTGCAGGGGATAGCCTTCGCCTTCGACAGGGGCA     858
13   V  V  W  I  G  V  I  F  S  G  R  V  Q  G  I  A  F  A  F  D  R  G       34

859  CGCTGATGAAAAGAATCCACGACCTTGCCGAGCACCTGGGGAAGAGAGGAGTCAGCATATCACTCG     924
35   T  L  M  K  R  I  H  D  L  A  E  H  L  G  K  R  G  V  S  I  S  L       56

925  ACGTCCAGCCGAGCGATTACCCTGAGAAGGTCTTCAAAGTCCTGATTGGGGAGCTCGACAATGCCA     990
57   D  V  Q  P  S  D  Y  P  E  K  V  F  K  V  L  I  G  E  L  D  N  A       78

991  GCTTTCTCCGGGAACTTTCCTTTGAGGGCGTTACACCCTTTGAGAAGAAGGTTTACGAGTGGCTCA    1056
79   S  F  L  R  E  L  S  F  E  G  V  T  P  F  E  K  K  V  Y  E  W  L      100

1057 CAAAAAACGTTAAAAGAGGGAGCGTTATAACCTACGGTGACCTTGCAAAGGCGCTCAACACGTCCC    1122
101  T  K  N  V  K  R  G  S  V  I  T  Y  G  D  L  A  K  A  L  N  T  S      122

1123 CGCGGGCAGTTGGTGGGGCAATGAAGAGAAACCCGTATCCAATCGTTGTTCCCTGTCACAGGGTCG    1188
123  P  R  A  V  G  G  A  M  K  R  N  P  Y  P  I  V  V  P  C  H  R  V      144
       Region 1                                            Region 2
1189 TTGCACACGATGGTATTGGTTACTACAGCTCCGGGATTGAGGAAAAGAAGTTCCTGCTGGAAATTG    1254
145  V  A  H  D  G  I  G  Y  Y  S  S  G  I  E  E  K  K  F  L  L  E  I      166

1255 AGGGGGTGAAAGAATGGACAAGCTGAAAGCTTAC                                    1288
16   E  G  V  K  E  W  T  S  *
```

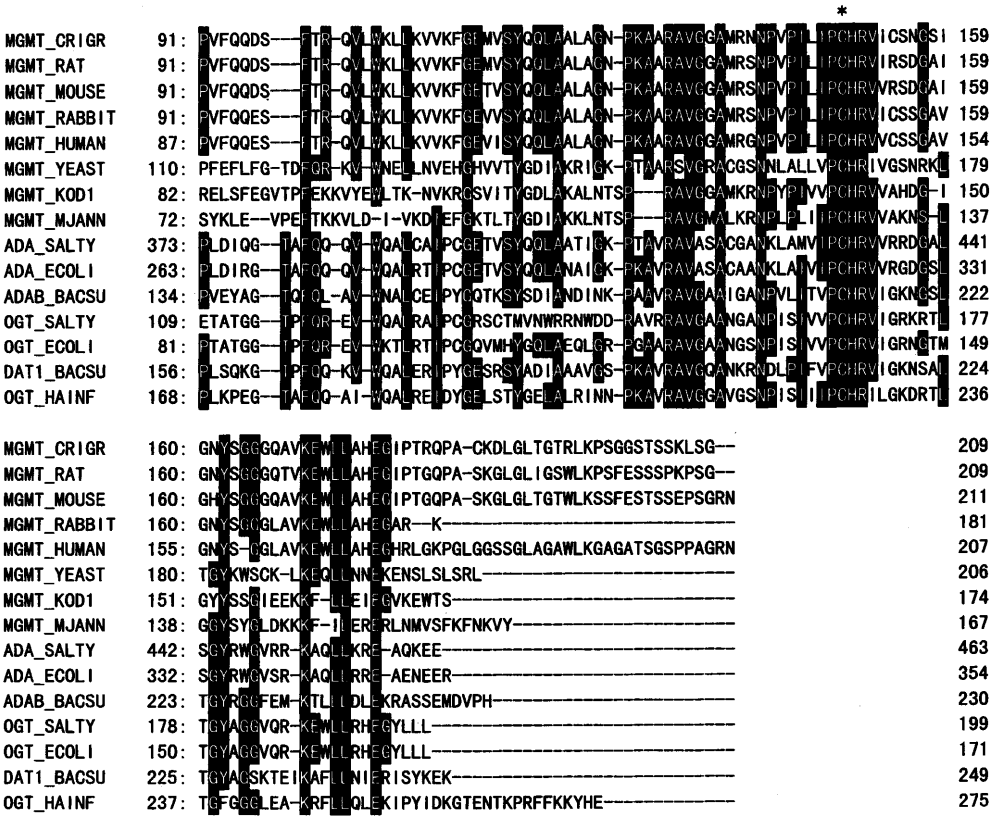


Fig. 2 Comparison of the amino acid sequences of the C-terminal halves of O^6-MeG-DNA methyltransferases from various sources. Gaps, indicated by *dashes*, were introduced to optimize the alignment. *White letters on a black background* indicate amino acid residues conserved in at least nine of the 15 species included. The methyl acceptor cysteine is indicated by the *asterisk*. The abbreviations are as follows and the entry names or accession numbers in protein or DNA databanks are given in *parentheses*. PIR: Protein Information Resource, SP: Swiss-Prot, GB: GenBank. MGMT_CRIGR, *Cricetulus griseus* (Chinese hamster) (SP P26186); MGMT_RAT, (SP P24528); MGMT_MOUSE (PIR A41809); MGMT_RABBIT (GB 177047); MGMT_HUMAN (SP P16455); MGMT_YEAST, *Saccharomyces cerevisiae* (SP P26188); MGMT_KOD1, *Pyrococcus* sp. KOD1 (This paper, GB D86335), MGMT_MJANN *Methanococcus jannaschii* (PIR H64490) ADA_SALTY, Ada *Salmonella typhimurium* (PIR XYEBOT); ADA_ECOLI, Ada *Escherichia coli* (SP P06134); OGT_ECOLI, Ogt *E. coli* (SP P09168); OGT_SALTY, Ogt *S. typhimurium* (SP P37429); ADAB_BACSU, AdaB *Bacillus subtilis* (SP P19220); DAT1_BACSU, Dat1 *B. subtilis* (SP P11742); OGT_HAINF, Dat1 *Haemophilus influenza* (SP P44687)

(SAM)-binding motif that is conserved in a variety of methyltransferases (Ingrosso et al. 1989; Koonin 1993).

Phylogenetic analysis

A phylogenetic tree was constructed by comparing the amino acid sequences of known MGMTs from Eukarya, Archaea, and Bacteria. As shown in Fig. 4, the phylogenetic tree shows a different profile from that of the universal tree based on 16S rRNA sequences (Woese et al. 1990). Archaeal MGMTs are grouped on the bacterial branch rather than with Eukarya.

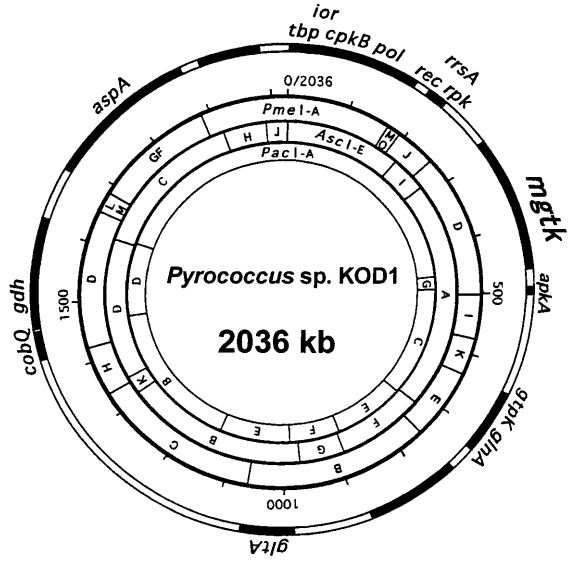


Fig. 3 Physical map of the KOD1 chromosome and location of the O^6-MeG-DNA methyltransferase-encoding gene (*mgtk*)

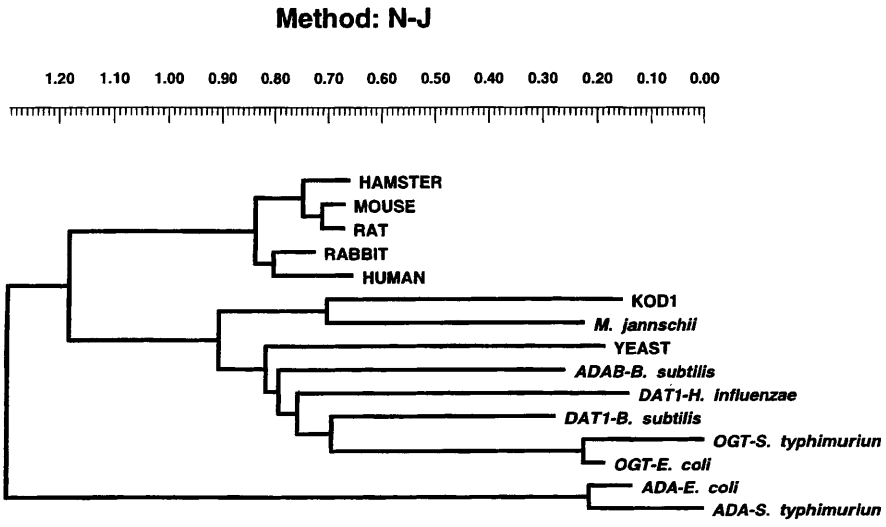


Fig. 4 Phylogenetic tree for O^6-MeG-DNA methyltransferases. Evolutionary distances were computed by the Neighbor Joining method and the tree was constructed using ODEN and Biosearch/Sinca programs. Segments corresponding to an evolutionary distance of 0.1 are shown. Each name at the terminus refers the species from which the sequence originated

Overexpression and purification of KOD1 methyltransferase from recombinant *E. coli*

E. coli HM174(DE3)pLysS harboring pKM-1m was grown and gene expression was induced by IPTG. Cells were disrupted by sonication, and the crude extract was used for further enzyme purification. Taking advantage of the overproduction of the methyltransferase protein in cells harboring pKM-1m, MGMT was purified to apparent physical homogeneity. The purity of the final preparation after the FPLC step on Superose 6 was examined by electrophoresis on 16% polyacrylamide gels containing 0.1% SDS (Fig. 5). A single protein band of 20 kDa was observed, in excellent agreement with the value calculated from the nucleotide sequence. The purified recombinant protein was assayed for its methyl transferase activity and exhibited an activity of 1100 cpm/mg.

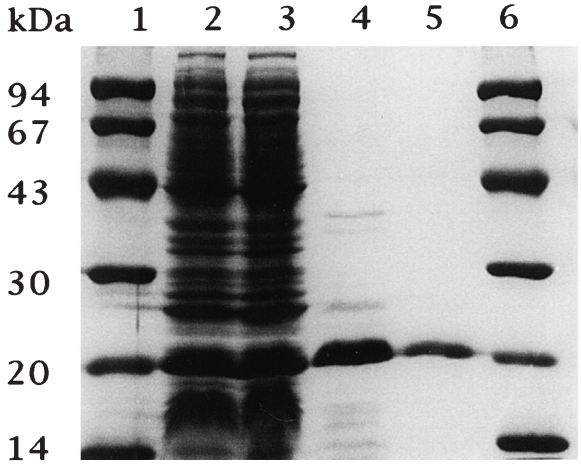


Fig. 5 SDS-PAGE analysis of MGMT purification. Lanes 1 and 6, molecular markers; lane 2, total cell protein; lane 3, Fraction I (crude extract); lane 4, Fraction II (heat treated); lane 5, purified MGMT. The purified protein has an estimated molecular mass of 20 kDa, which agrees with the value deduced by sequence analysis

Thermostability of KOD1 O^6-MeG-DNA methyltransferase

The thermostability of recombinant KOD1 MGMT was estimated from the activity remaining after heat treatment at 75° C and 90° C. After incubation at both 75° C and 90° C for 30 min, KOD1 MGMT showed no decrease in enzymatic activity. When KOD1 MGMT was incubated for 60 min at both temperatures, the remaining activities were 77% at 75° C and 34% at 90° C, respectively.

Thermal denaturation of MGMT was examined by monitoring CD spectra at 30° C and 90° C. The results indicated that the spectrum of the enzyme does not show any drastic change at 90° C (Fig. 6A). The thermal denaturation curve was then obtained by monitoring the change in the CD value at 222 nm. The curve in Fig. 6B shows that KOD1 MGMT is a very thermostable repair enzyme which is still not fully denatured even at 98° C.

Phenotypic complementation assay of *E. coli* mutant

In order to confirm that the cloned fragment encodes a O^6-MeG-DNA methyltransferase that is functional in vivo, the gene was expressed in *E. coli* (Δ*ada* Δ*ogt*) cells. The plasmid pKM-2m, containing the *lac* promoter upstream of the MGMT coding sequence, was introduced into the methyltransferase-deficient *E. coli* strain KT233 (Δ*ada* Δ*ogt*), which was then exposed to different concentrations of MNNG. We determined the survival ratios of KT233 (Δ*ada* Δ*ogt*) cells with and without the KOD1 MGMT gene, after MNNG treatment. When the protein was overproduced by expression of the KOD1 MGMT gene from a multicopy plasmid vector, the sensitivity of KT233 (Δ*ada* Δ*ogt*) cells to the alkylating agent was clearly decreased, as shown in Fig. 7. The data indicate that the methyltransferase from KOD1 encoded by the MGMT gene in pKM-2m is

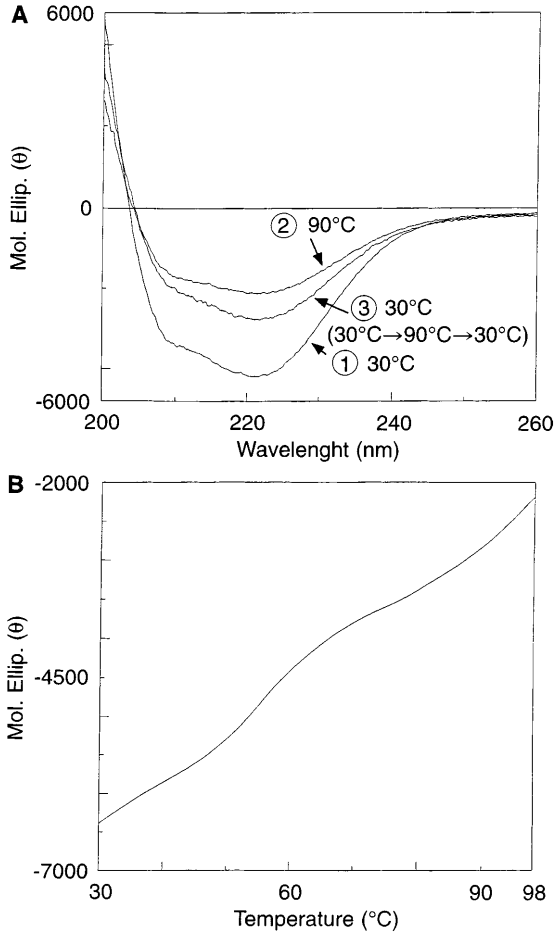


Fig. 6A, B Analysis of the thermal denaturation profile of KOD1 methyltransferase. **A** CD spectra. **B** Thermal denaturation curve monitored by CD analysis at 222 nm

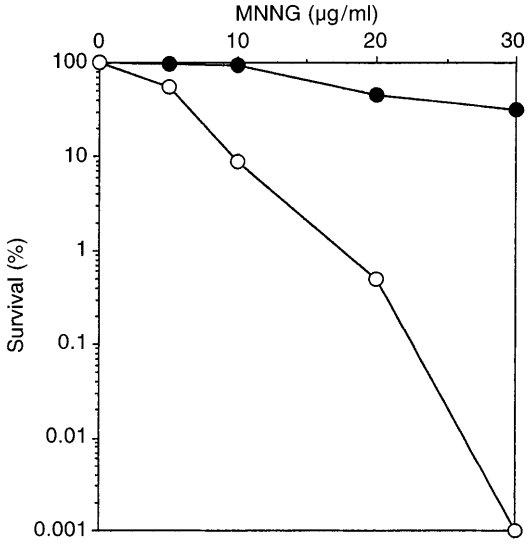


Fig. 7 Survival of cells after treatment with *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG). *E. coli* KT233 (Δ*ada*, Δ*ogt*) cells carrying one of the following plasmids were treated with various concentrations of MNNG at 37° C for 10 min and the survival frequencies were determined. *Filled circles*, pKM-2m; *open circles*, pUC-19

biologically functional and can recognize and demethylate O^6-MeG residues in *E. coli* cells at 37° C

Discussion

O^6-MeG-DNA methyltransferase (MGMT) is present in organisms from bacteria to humans. The structure and function of MGMTs have been studied extensively in some organisms, particularly in *E. coli* (Demple et al. 1985; Le Motte and Walker 1985; Marginson et al. 1985; Nakabeppu et al. 1985; Sekiguchi et al. 1996). However, no report of MGMT has been available from Archaea, except for the sequence of an MGMT homolog from *M. jannaschii* (Genbank accession number U67593). Ours is thus the first report to document the in vivo and in vitro enzymatic characteristics, as well as the thermostability profile, of an MGMT from a hyperthermophilic archaeon, *Pyrococcus* sp. strain KOD1.

We designed the PCR primers based on known MGMT sequences from 11 different species and obtained part of the gene corresponding to the KOD1 MGMT gene. The entire gene was cloned using the PCR product as a probe for Southern and colony hybridizations. The KOD1 MGMT gene has an ORF composed of 522 bases and preceded by a putative ribosomal binding site (GGAGGA). The deduced isoelectric point of KOD1 MGMT (pI = 9.49) showed that this protein was positively charged, suggesting that it can efficiently bind DNA (Graying et al. 1996). When the amino acid sequences of the KOD1, mammalian, yeast and bacterial MGMTs were aligned (Fig. 2), regions of high sequence conservation was observed. Of these the most notable region is the sequence PCHRV (Region 2 in Fig. 1), which is found near the C-terminus and contains the cysteine residue that serves as the methyl group acceptor (Olsson and Lindahl 1980; Demple et al. 1985). The sequence RAV (at positions 110–112 of the KOD1 sequence) is also conserved as the core of a sequence composed of ten amino acids, PXA(A/V)RAV(G/A)XA (Region 1 in Fig. 1), which may possibly constitute the active site of MGMT (Sakumi et al. 1991). A feature which differentiates archaeal MGMTs from bacterial and eukaryotic ones is also found within the latter region. The MGMTs of KOD1 and *Methanococcus jannaschii* do not have the motif XA(A/V), which are conserved in bacterial and eukaryotic MGMTs. At present it is not clear what kinds of role these residues might play. Bacterial and eukaryotic MGMTs might have picked up the tripeptide in the course of evolution.

No phylogenetic analysis has previously been performed on the available MGMT sequences. We therefore constructed a MGMT tree, including the newly determined KOD1 sequence, in the hope of better understanding the general evolutionary relationships among the three domains, Bacteria, Eukarya and Archaea. The phylogenetic tree was constructed by the NJ method as shown in Fig. 4. Archaeal MGMTs were

grouped with their bacterial counterparts. The profile of the tree does not agree with that of the universal tree of life based on 16S rRNA sequences (Woese et al. 1990). The MGMT tree is quite unique, because archaeal proteins which are involved in genetic information processing generally share characteristics with their eukaryotic counterparts. More sequence information on MGMTs from various organisms including other archaea should clarify the phylogenetic tree. KOD1 may possess another, unknown MGMT gene which is more closely related to eukaryotic MGMTs.

KOD1 MGMT was overproduced using the pET expression system and the recombinant protein was purified to homogeneity by ammonium sulfate fractionation, heat treatment, ion-exchange chromatography and gel filtration chromatography. The molecular mass of the recombinant KOD1 MGMT, estimated by gel filtration chromatography, shows that the protein exists in a monomeric form, as has also been shown for other MGMTs. KOD1 MGMT possesses methyl transferase activity and it was found to be stable at 90° C for 30 min. The thermal denaturation curves showed that KOD1 methyltransferase is a very thermostable protein, which is not fully denatured even at 98° C. However, the enzymatic activity of KOD1 MGMT could not be measured precisely at temperatures higher than 50° C probably due to instability of the substrate DNA. High temperatures may cause melting of the substrate DNA, resulting in inhibition of the interaction between substrate DNA and MGMT. This idea is compatible with the evidence that MGMT shows a strong preference for duplex over single-stranded DNA (Demple 1990). The DNA-protein complex model based on the three-dimensional structure of Ada O^6-methylguanine-DNA methyltransferase from *E. coli* also supports this idea (Moore et al. 1994). The thermostable DNA polymerase from KOD1 also shows a lower optimum for activity in vitro than the optimal temperature for hyperthermophile cell growth (Takagi et al. 1997). In cells of hyperthermophiles, some accessory proteins, such as a histone-like protein, may stabilize DNA and facilitate the catalytic reactions of DNA-related enzymes (Graying et al. 1996). Further in vitro study is necessary for better understanding of reaction mechanisms at high- temperature environments.

The KOD1 MGMT can be expressed in *E. coli* as an active form which is functional even at 37° C. The repair system in Archaea seems to be quite similar to that of bacteria, as revealed by the fact that a repair-deficient *E.coli* mutant could be rescued by expression of the archaeal protein in the cell.

Sequencing also uncovered an ORF encoding a protein of 236 residues which overlaps the N-terminal region of the MGMT gene. The deduced product showed high sequence similarity to the protein SpoU, which belongs to a new family of rRNA methylases (Koonin and Rudd 1993). It is noteworthy that two genes encoding different kinds of methylation-related enzyme are arranged in tandem. At present, the biological function of the SpoU homolog is unclear. Further analysis is required to obtain detailed information about the biological roles of this gene.

Hyperthermophilic archaea live in an environment that is very similar to that of the primitive earth. Early organisms must also have been exposed to deleterious, high-temperature environments. It was therefore essential to develop effective DNA repair systems. Hence, the study of archaeal MGMT is expected to provide interesting clues concerning primitive systems of DNA repair. Crystallization and detailed structural analysis are now in progress. The three-dimensional structure will allow us to elucidate the relationship between structure and function of their enzymes.

Acknowledgment We extend special thanks to Prof. Dr. M. Sekiguchi for providing the *E. coli* strain KT233 (Δ*ada*, Δ*ogt*).

References

Bogden JM, Eastman A, Bresnick E (1981) A system in mouse liver for the repair of O^6-methylguanine lesions in methylated DNA. Nucleic Acids Res 13:3089–3103

Coulondre C, Miller JH (1977) Genetic studies of the *lac* repressor. IV. Mutagenic specificity in the *lacI* gene of *Escherichia coli*. J Mol Biol 117:577–606

Demple B (1990) In: Paik WK, Kim S (eds) Protein methylation. CRC Press, Boca Raton, Florida, pp 285–304

Demple B, Sedgwick B, Robins P, Totty N, Waterfield MD, Lindahl T (1985) Active site and complete sequence of the suicidal methyltransferase that counters alkylation mutagenesis. Proc Natl Acad Sci USA 82:2688–2692

Dennis PP (1997) Ancient cipher: translation in Archaea. Cell 89:1007–1010

Felsenstein J (1985) Confidence limits on phylogenies: an approach using bootstrap. Evolution 39:783–791

Graying RA, Sandman K, Reeve JN (1996) DNA stability and binding proteins. Adv Protein Chem 48:437–467

Fujiwara S, Okuyama S, Imanaka T (1996) The world of archaea: genome analysis, evolution and thermostable enzymes. Gene 179:165–170

Imanaka T, Tanaka T, Tsunekawa H, Aiba S (1981) Cloning of the genes for penicillinase, *penP* and *penJ*, of *Bacillus licheniformis* in some vector plasmids and their expression in *Escherichia coli*, *Bacillus subtilis*, and *Bacillus licheniformis*. J Bacteriol 147:776–786

Ingrosso D, Fowler AV, Bleibaum J, Clarke S (1989) Sequence of the D-aspartyl/L-isoaspartyl protein methyltransferase from human erythrocytes. J Biol Chem 264:20131–20139

Ito T, Nakamura T, Maki H, Sekiguchi M (1994) Roles of transcription and repair in alkylation mutagenesis. Mutant Res. 314:273–285

Kanaya S, Itaya M (1992) Expression, purification, and characterization of a recombinant ribonuclease H from *Thermus thermophilus* HB8. J Biol Chem 267:10184–10192

Karran P, Lindahl T (1979) Adaptive response to alkylating agents involves alteration in situ of O^6-methylguanine residues in DNA. Nature 280:76–77

Kawate H, Ihara K, Kohda K, Sakumi K, Sekiguchi M (1995) Mouse methyltransferase for repair of O^6-methylguanine and O^4-methylthymine in DNA. Carcinogenesis 16:1595–1602

Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge, UK

Koike G, Maki H, Takeya H, Sekiguchi M (1990) Purification, structure and biochemical properties of human O^6-methylguanine-DNA methyltransferase. J Biol Chem 265:14754–14762

Koonin EV (1993) Computer-assisted identification of a putative methyltransferase domain in NS5 protein of flavivirus and λ2 protein of reovirus. J Gen Virol 73:733–740

Koonin EV, Rudd KE (1993) SpoU protein of *Escherichia coli* belongs to a new family of putative rRNA methylases. Nucleic Acid Res 23:5519

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 277:680–685

Le Motte PK, Walker GC (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

Loveless A (1969) Possible relevance of O^6-alkylation of deoxyguanosine to the mutagenicity and carcinogenicity of nitrosamines and nitrosamides. Nature 223:206–208

Marginson GP, Cooper DP, Brennand J (1985) Cloning of the *E. coli* O^6-methylguanine and methylphosphotriester methyltransferase gene using a functional DNA repair assay. Nucleic Acids Res 13:1939–1952

Moore MH, Gulbis JM, Dodson EJ, Demple B, Moody PCE (1994) Crystal structure of a suicidal DNA repair protein: the Ada O^6-methylguanine-DNA methyltransferase from *E. coli*. EMBO J 13:1495–1501

Morikawa M, Izawa Y, Rashid N, Hoaki T, Imanaka T (1994) Purification and characterization of a thermostable thiol protease from a newly isolated hyperthermophilic *Pyrococcus* sp. Appl Environ Microbiol 60:4559–4566

Nakabeppu Y, Kondo H, Kawabata S, Iwanaga S, Sekiguchi M (1985) Purification and structure of the intact Ada regulatory protein of *Escherichia coli* K12, O^6-methylguanine-DNA methyltransferase. J Biol Chem 260:7281–7288

Olsson M, Lindahl T (1980) Repair of alkylated DNA in *Escherichia coli*. J Biol Chem 255:10569–10571

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sakumi K, Shiraishi A, Hayakawa H, Sekiguchi M (1991) Cloning and expression of cDNA for rat O^6-methylguanine-DNA methyltransferase. Nucleic Acids Res 19:5597–5601

Sekiguchi M, Nakabeppu Y, Sakumi K, Tuzuki T (1996) DNA-repair methyltransferase as a molecular device for preventing mutation and cancer. J Cancer Res Clin Oncol 122: 199–206

Strauss B, Scudiero D, Henderson E (1975) The nature of the alkylation lesion in mammalian cells. In: Hanawalt PC, Setlow RB (eds) Molecular mechanisms for repair of DNA, Part A. Plenum, New York, pp 13–24

Studier FW, Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol 189:113–130

Sukumar S, Natario V, Martin-Zanca D, Barbacid M (1983) Induction of mammary carcinomas in rats by nitroso-methylurea involves malignant activation of H-*ras*–1 locus by single point mutations. Nature 306:658–661

Takagi M, Nishioka M, Kakihara H, Kitabayashi M, Inoue H, Kawakami B, Oka M, Imanaka T (1997) Characterization of DNA polymerase from *Pyrococcus* sp. strain KOD1 and its application to PCR. Appl Environ Microbiol 63:4504–4510

Takano K, Nakamura T, Sekiguchi M (1991) Roles of two types of O^6-methylguanine-DNA methyltransferases in DNA repair. Mutat Res 254:37–44

Woese CR, Kandler O, Wheelis ML (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eukarya. Proc Natl Acad Sci USA 87:4576–4579

Zak P, Kleibl K, Laval F (1994) Repair of O^6-methylguanine-DNA methyltransferase and O^4-methylthymine by the human and rat O^6-methylguanine-DNA methyltransferases. J Biol Chem 269:730–733