Alkylation damage repair protein *O*⁶-alkylguanine-DNA alkyltransferase from the hyperthermophiles *Aquifex aeolicus* and *Archaeoglobus fulgidus*

Sreenivas KANUGULA and Anthony E. PEGG¹

Department of Cellular and Molecular Physiology, Pennsylvania State University College of Medicine, The Milton S. Hershey Medical Center, P.O. Box 850, 500 University Drive, Hershey, PA 17033, U.S.A.

AGT (*O*⁶-alkylguanine DNA alkyltransferase) is an important DNA-repair protein that protects cells from killing and mutagenesis by alkylating agents. The AGT genes from two extremely thermophilic organisms, the bacterium *Aquifex aeolicus* and the archaeon *Archaeoglobus fulgidus* were PCR-derived and cloned into an expression vector. The nucleotide sequence of the *Aq. aeolicus* AGT encodes a 201-amino-acid protein with a molecular mass of 23 000 Da and *Ar. fulgidus* AGT codes for a 147-amino-acid protein with a molecular mass of 16 718 Da. The *Aq. aeolicus* and *Ar. fulgidus* AGTs were expressed at high levels in *Escherichia coli* fused to an N-terminal polyhistidine tag that allowed single-step isolation and purification by metal-affinity chromatography. Both AGTs formed inclusion bodies and were not soluble under native purification conditions. Therefore AGT

isolation was performed under protein-denaturation conditions in the presence of 8.0 M urea. Soluble AGT was obtained by refolding the AGT in the presence of calf thymus DNA. Both AGTs were active in repairing *O*⁶-methylguanine and, at a lower rate, *O*⁴-methylthymine in DNA. They exhibited thermostability and optimum activity at high temperature. The thermostable AGTs, particularly that from *Aq. aeolicus*, were readily inactivated by the low-molecular-mass inhibitor *O*⁶-benzylguanine, which is currently in clinical trials to enhance cancer chemotherapy.

Key words: *O*⁶-alkylguanine-DNA alkyltransferase, *O*⁶-benzylguanine, alkylating agent, DNA repair, *O*⁶-methylguanine, thermophile.

INTRODUCTION

AGT (*O*⁶-alkylguanine-DNA alkyltransferase) is a DNA-repair protein that removes alkyl adducts from the *O*⁶ position of guanine and the *O*⁴ position of thymine in DNA [1–4]. These lesions, if not removed, are highly mutagenic and carcinogenic [5–7]. AGT repairs these DNA adducts by transferring the alkyl group on to a cysteine amino acid residue at the active site of the protein and restores DNA to its unmodified form in a single step. This reaction permanently inactivates the AGT and synthesis of new protein is required for repair of additional DNA damage. The presence of a characterized AGT gene product or a putative gene coding for AGT protein has been reported so far in more than 100 different species from the three domains of life, Archaea, Bacteria and Eukaryota, although it is not ubiquitous, apparently being absent from plants, *Schizosaccharomyces pombe* and *Deinococcus radiodurans*. All known AGTs (except a variant from *Caenorhabditis elegans* that has Pro-Cys-His-Pro sequence) share a highly conserved Pro-Cys-His-Arg amino acid sequence in their active site and contain a number of other conserved amino acids in the active site and DNA-binding domain of the protein [4,8]. The widespread presence of AGTs indicates the occurrence of alkylation damage to DNA under a variety of living conditions and the need to repair such damage.

The presence of AGT protects cells from killing by methylating and chloroethylating agents, which are widely used in cancer chemotherapy. AGT is therefore an important source of the resistance of many tumour cells to these agents. Attempts are being made to circumvent this resistance by using inactivators of the protein such as *O*⁶-benzylguanine (BG) [9,10]. A full understanding

of the structure/mechanism of AGT will be useful in the further development of drugs that can specifically inactivate AGT.

Crystal structures for AGTs from three species (*Escherichia coli*, *Homo sapiens* and *Pyrococcus kodakaraensis*) have now been determined [11–14]. These provide a general understanding of the active site and the role of certain conserved residues that allow the modelling of the repair reaction, but at present, no structure of the complex between AGT and its DNA substrate is available, and many details of the DNA binding, substrate recognition and mechanism of alkyl transfer remain unclear.

Archaea represent the third domain of living organisms and are considered to be the most primitive organisms on earth. Recent identification of DNA-repair processes in thermophilic members of the Archaea and Bacteria have revealed that they are a valuable source of proteins involved in genome maintenance that aid in our understanding of the process of DNA repair in eukaryotes and in the identification of novel repair pathways, which have diverged from common ancestors [15–17].

Proteins from thermophilic organisms are particularly useful for the study of structure–function relationships at the molecular level because they are often stable and easily crystallized. In the present study, we report cloning, expression, purification and characterization of the AGT genes from two hyperthermophiles *Aquifex aeolicus* and *Archaeoglobus fulgidus*. *Aq. aeolicus* was one of the earliest diverging, most thermophilic bacteria known to date and grows at temperatures in excess of 80 °C [18]. The purified *Aq. aeolicus* AGT exhibited optimum activity at 70 °C and, unlike other bacterial AGTs known so far, was highly sensitive to inactivation with BG. *Ar. fulgidus* belongs to sulphate-reducing hyperthermophilic Archaea and grows at 83 °C [19]. The

Abbreviations used: Ada-C, C-terminal domain of the *Escherichia coli* Ada gene product; AGT, *O*⁶-alkylguanine-DNA alkyltransferase; hAGT, human AGT; BG, *O*⁶-benzylguanine; DTT, dithiothreitol; IPTG, isopropyl β -D-thiogalactoside; Ogt, *E. coli* AGT derived from the *Ogt* gene.

¹ To whom correspondence should be addressed (e-mail aep1@psu.edu).

Ar. fulgidus protein with only 147 amino acids is the smallest AGT to be purified to date and showed optimum DNA repair at 80 °C.

MATERIALS AND METHODS

Materials

All oligodeoxyribonucleotides were purchased from Life Technologies (Gaithersburg, MD, U.S.A.). *E. coli* XL1-Blue bacterial strain and *Pfu* polymerase enzyme were purchased from Stratagene (La Jolla, CA, U.S.A.). pQE-30 plasmid and the DNA-isolation kits were obtained from Qiagen (Chatsworth, CA, U.S.A.). Restriction enzymes were purchased from New England Biolabs (Beverly, MA, U.S.A.). Talon Metal Affinity IMAC Resin and T4 ligase with the Ligation Express kit were purchased from Clontech (Palo Alto, CA, U.S.A.). Ampicillin, IPTG (isopropyl β -D-thiogalactoside), haemocyanin, calf thymus DNA and most other biochemical reagents were purchased from Sigma Chemical Co. (St. Louis, MO, U.S.A.). Urea was purchased from Boehringer Mannheim (Mannheim, Germany). Nitrocellulose filters (0.45 μ m) were obtained from Millipore (Bedford, MA, U.S.A.). BG was generously provided by Dr R. C. Moschel (ABL-Basic Research program, NCI-Frederick Cancer Research and Development Center, Frederick, MD, U.S.A.).

Construction of pQEAGT plasmids for expression of *Aq. aeolicus* and *Ar. fulgidus* AGT

Based on the reported DNA sequences of the putative *Aq. aeolicus* (GenBank® accession number AE000743) and *Ar. fulgidus* (GenBank® accession number AE000944) AGT genes, specific primers were synthesized and used for their amplification from DNA of these two organisms. The primers used for *Aq. aeolicus* were 5'-CCATCACGGATCCATGAGTACAGAAAGAAGAAGG-3' (sense) and 5'-GACCCGGGGTACCTCA-AAGGATTAAACCCTCGTG-3' (antisense). The primers used for *Ar. fulgidus* were 5'-ACCATCACGGATCCATGTTCTCG-GTTAAGTGGGGAG-3' (sense) and 5'-CGACCCGGGGTACC-TCAAAGACGCCCTCAAGCCTG-3' (antisense). [In both cases, the *Bam*HI (sense) and *Kpn*I (antisense) restriction endonuclease sites are indicated in bold, and the start and stop codons are underlined.] PCR was carried out using *Pfu* polymerase with 50 ng of genomic DNA from *Aq. aeolicus* (a gift from Dr Robert Huber and Dr Karl O. Stetter, Lehrstuhl für Mikrobiologie, Universität Regensburg, Regensburg D-93053, Germany) or *Ar. fulgidus* (a gift from Dr David E. Graham, Department of Microbiology, B103 Chemical and Life Sciences Laboratory, MC-110, University of Illinois at Urbana-Champaign, Urbana, IL, U.S.A.) as templates under the following conditions: initial denaturation for 2 min at 92 °C, followed by 25 cycles of denaturation at 92 °C for 30 s, annealing at 52 °C for 30 s and extension at 72 °C for 30 s, with a final extension at 72 °C for 5 min. The PCR products (665 bp for *Aq. aeolicus* AGT DNA and 503 bp for *Ar. fulgidus* AGT DNA) were digested with *Bam*HI and *Kpn*I restriction endonucleases, purified on a 1% (w/v) agarose gel using Qiagen gel-extraction kit and ligated into pQE30 vector digested with the same enzymes. The entire coding region of both AGT plasmids was sequenced to ensure that no secondary mutations were introduced during the constructions.

Protein purification

The pQEAGT plasmids were transformed into *E. coli* XL1-B cells for protein expression and purification. The proteins expressed

from these plasmids carry a 12-amino-acid (MRGSH₆GS-) tag at their N-terminus. The AGTs were purified in a single step of immobilized metal-affinity chromatography by the following modification of the method described for human AGT (hAGT) [20]. Bacteria harbouring the pQEAGT plasmids were grown overnight in 10 ml of LB (Luria-Bertani) medium at 37 °C with 50 μ g/ml ampicillin. The overnight grown culture was used to inoculate 500 ml of LB medium supplemented with 50 μ g/ml ampicillin. The 500 ml cultures were grown at 37 °C with shaking until the culture reached a D_{595} of 0.6–0.8. Protein expression was then induced by the addition of IPTG to a final concentration of 150 μ M and the cells were harvested 4 h later. The cell pellets were suspended in 20 ml of 20 mM Tris/HCl (pH 8.0) and 250 mM NaCl buffer and were lysed by sonication. The soluble protein extract and pellets were collected by centrifugation at 16 000 *g*. Analysis by SDS/PAGE indicated that two AGTs were insoluble and recovered in the inclusion bodies. These protein pellets were suspended in 20 mM Tris/HCl (pH 8.0), 250 mM NaCl and 8.0 M urea and stirred gently at room temperature (25 °C) for 3 h to extract the proteins. The suspension was then centrifuged at 16 000 *g* for 30 min at 20 °C. The supernatant fraction was collected and applied to a column containing 1.0 ml of Talon IMAC resin equilibrated with the above buffer. The column was washed extensively with above buffer and the AGT was eluted with 200 mM imidazole in the same buffer. The fractions containing AGT were pooled and 3 mg of calf thymus DNA was added. The AGT/DNA mixture was dialysed first against 50 mM Tris/HCl (pH 7.6), 250 mM NaCl, 5 mM DTT (dithiothreitol), 0.1 mM EDTA and 5% (v/v) glycerol buffer containing 5 M urea for 3 h. This was then followed by dialysis for 3 h against 2.5 M, 1.0 M and no urea in the above buffer. Both AGTs remained in solution under these refolding conditions.

Pure soluble *Aq. aeolicus* AGT was also obtained in the absence of DNA by refolding the protein under slightly acidic conditions. The protein isolated from the inclusion bodies as described above was then dialysed against 20 mM sodium phosphate (pH 4.5), 250 mM NaCl, 5 mM DTT, 0.1 mM EDTA and 5% (v/v) glycerol buffer containing 8.0 M urea followed by dialysis against 5.0 M, 2.5 M, 1.0 M and no urea in the above buffer.

The purified refolded AGTs were pre-incubated at 37 °C for 12 h and were stored at 4 °C. The purified proteins were analysed by SDS/PAGE and were visualized by Coomassie Blue staining. The hAGT was purified as described previously [20].

AGT activity assay

AGT activity was measured by determining the transfer of [³H]methyl groups to AGT from *O*⁶-[³H]methylguanine in [³H]methylated DNA [8]. Purified AGT preparations were incubated with 15 μ g of [³H]methylated DNA substrate and 50 μ g of haemocyanin in 1.0 ml of 50 mM Tris/HCl (pH 7.6), 5 mM DTT and 0.1 mM EDTA (Buffer A) for 30 min at different temperatures. The AGT reaction was terminated by addition of 7 ml of denaturation buffer [20 mM Tris/HCl (pH 7.6), 600 mM NaCl and 8 M urea] to the samples and incubation further at room temperature for 60 min. The samples were then passed through 0.45 μ m (2.5 cm) pore-size nitrocellulose filters on a Millipore 12-well holder unit. The reaction tubes were washed with 5 ml of denaturation buffer, and the wash was applied on to the filters. The filters were washed twice further with fresh denaturation buffer (5.0 ml), followed by two washes with 10% (v/v) ethanol in water (5.0 ml). The filters were dried and the radioactivity was scintillation-counted in 10 ml of Econofluor-2 (Packard Instrument Co., Downers Grove, IL, U.S.A.). AGT specific activities were determined by incubating substrate DNA with

0.01 µg of hAGT at 37 °C for 3 h, 0.1 µg of *Aq. aeolicus* AGT at 70 °C for 3 h and 0.1 µg of *Ar. fulgidus* AGT at 80 °C for 6 h. The kinetics of the methyl transfer were measured as previously described [21]. The reaction mixture (1 ml) for each time point contained 9.4 nM of *Aq. aeolicus* AGT (at 70 °C), 13 nM of *Ar. fulgidus* AGT (at 80 °C) and 0.26 nM O⁶-[³H]methylguanine in [³H]methylated DNA substrate in Buffer A.

A similar assay was used to determine O⁴-methylthymine repair but in this case the [³H]methylated DNA substrate was replaced by [³H]methylated poly(dA · dT).

Heat-stability of the AGT proteins

Aliquots of different AGTs in the presence of 10 µg of calf thymus DNA were incubated for different time intervals at different temperatures in 1 ml of Buffer A. The AGT activity remaining was then measured as described above by incubation of the samples for 30 min with [³H]methylated DNA at 37 °C for hAGT, 70 °C for *Aq. aeolicus* AGT and 80 °C for *Ar. fulgidus* AGT.

AGT inactivation by BG

Purified AGTs were incubated with different concentrations of BG in the presence of 10 µg of calf thymus DNA at 37 °C (hAGT) and 60 °C (*Aq. aeolicus* AGT and *Ar. fulgidus* AGT) for 30 min in 0.5 ml of Buffer A. The residual AGT activity was then determined as described above by incubation of the BG-incubated protein samples with [³H]methylated DNA (15 µg) for 30 min at 37 °C (hAGT), 70 °C (*Aq. aeolicus* AGT) and 80 °C (*Ar. fulgidus* AGT). The results were expressed as the percentage of the AGT activity remaining, and the graphs of AGT activity remaining against inhibitor concentration were used to calculate an ED₅₀ value, representing the amount of inhibitor needed to produce a 50 % loss of activity.

Measurements of guanine formation from BG were carried out with *Aq. aeolicus* AGT purified under denaturing conditions and renatured at pH 7.6 in the presence of calf thymus DNA or renatured at pH 4.5 in the absence of DNA in Buffer A containing 0.7 µM O⁶-benzyl[8-³H]guanine (which was prepared by tritium exchange from BG [22]), in an assay volume of 0.25 ml. After incubation for 20 min at different temperatures, the reaction was stopped by the addition of 0.5 ml of the same buffer containing 0.2 mM guanine and 0.2 mM BG. The extent of formation of [8-³H]guanine was determined by HPLC as described previously [21].

RESULTS

The nucleotide sequence of the AGT gene amplified from *Ar. fulgidus* and *Aq. aeolicus* genomic DNA matched accurately with the AGT gene sequence identified from complete genome sequence analysis. The AGT gene from *Aq. aeolicus* codes for a protein with 201 amino acids, whereas the AGT gene from the archaeon *Ar. fulgidus* produces a protein with only 147 amino acids. Comparison of the primary amino acid sequence of these two proteins with human, *E. coli* and *P. kodakaraensis* AGTs shows 17 conserved amino acids (Figure 1).

Aq. aeolicus and *Ar. fulgidus* AGTs were overproduced in *E. coli* for protein purification, but the proteins were insoluble and formed inclusion bodies under all conditions tested. The proteins could be purified to homogeneity by immobilized metal-affinity chromatography in the presence of 8 M urea. When dialysed against decreasing concentration of urea in a step-wise manner to

Figure 1 Amino acid sequence and alignment of AGTs from *Aq. aeolicus*, *Ar. fulgidus*, *P. kodakaraensis* (also described as *Thermococcus kodakaraensis*), *E. coli* Ada-C and human

Identical amino acids in all five sequences are indicated by black shading and the active-site cysteine residue is shown with an asterisk. The Pro¹⁴⁰ and Gly¹⁵⁶ residues in hAGT and the equivalent residues in *Aq. aeolicus*, *Ar. fulgidus* and *P. kodakaraensis* are underlined and in italics, and are indicated by a closed diamond. The tryptophan residue causing steric hindrance to the binding of BG in Ada-C is shown in bold and is indicated by #. The acidic and basic residues thought to stabilize *P. kodakaraensis* AGT are indicated by italics and an open triangle (acidic) and underlining and a closed circle (basic). The equivalent residues in *Aq. aeolicus*, and *Ar. fulgidus* are indicated similarly.

Figure 2 Purified *Aq. aeolicus* and *Ar. fulgidus* AGTs were analysed by SDS/PAGE stained with Coomassie Blue

The left lane shows *Aq. aeolicus* AGT (5 µg) and the right lane shows *Ar. fulgidus* AGT (5 µg). The positions of molecular-mass markers are shown by arrows (sizes in kDa). The gel was intentionally overloaded to assess protein purity. The minor 14 kDa contaminating band in the *Aq. aeolicus* AGT amounted to < 3 % of the total protein.

refold the proteins, AGTs aggregated out of the solution when the concentration of urea in the dialysis buffer was lowered to 2.5 M. Addition of DNA to the protein before renaturation by step-wise decreasing of the urea allowed their isolation in a soluble form. Both AGT preparations were more than 97 % pure as analysed by SDS/PAGE and Coomassie Blue staining (Figure 2). AGTs were stored at 4 °C. Storage at either -20 °C or -80 °C resulted in protein precipitation.

Both thermophilic AGTs showed increased transferase activity as the temperature of incubation of the proteins with DNA substrate was increased (Figure 3). *Aq. aeolicus* AGT gave optimum

Figure 3 Activity assay of purified AGTs at different temperatures

AGT activity with the purified proteins was carried out as described in the Materials and methods section by incubation of the pure proteins for 30 min with the [3 H]methylated DNA substrate at different temperatures. Results are shown for hAGT (circles), *Aq. aeolicus* AGT (squares) and *Ar. fulgidus* AGT (triangles).

repair of O^6 -methylguanine in DNA at 70 °C and at temperatures above this, the DNA-repair capacity of this protein disappeared rather quickly. *Ar. fulgidus* AGT exhibited maximum DNA repair at 80 °C and, unlike the bacterial *Aq. aeolicus* protein, showed substantial repair activity even at 100 °C. hAGT, which exhibits optimum DNA repair at about 37 °C, lost its DNA repair activity completely at temperatures above 60 °C, presumably due to denaturation of the hAGT at these temperatures.

Even at the optimal temperature, the thermophilic AGTs had much lower rate constants for repair of DNA than hAGT ($1.6 \times 10^5 \text{ M}^{-1} \cdot \text{min}^{-1}$ for *Ar. fulgidus* at 80 °C and $1.6 \times 10^6 \text{ M}^{-1} \cdot \text{min}^{-1}$ for *Aq. aeolicus* at 70 °C compared with $3.8 \times 10^7 \text{ M}^{-1} \cdot \text{min}^{-1}$ for hAGT at 37 °C). To some extent, this difference may reflect the melting of the DNA substrate at high temperature, since AGTs are known to have lower repair rates with single-stranded substrates.

The thermal stability of these proteins was tested by pre-incubation of aliquots of these proteins at different temperatures and then assaying the remaining AGT activity by incubation for 30 min with substrate DNA at 37 °C for hAGT, 70 °C for *Aq. aeolicus* AGT and 80 °C for *Ar. fulgidus* AGT. The pre-incubation for 48 h at 37 °C did not alter the activity of the two thermostable proteins, but hAGT lost 50 % of its activity after 6 h at this temperature. More than 90 % of hAGT activity was lost after pre-incubation for only 30 min at 60 °C. In contrast, *Aq. aeolicus* AGT had a half-life of 12 h at 60 °C and that from *Ar. fulgidus* was even more stable with a 24 h half-life at 60 °C. The half-lives of these proteins at temperatures where they exhibited optimum activity were 6 h for hAGT (37 °C), 3 h for *Aq. aeolicus* AGT (70 °C) and 8 h for *Ar. fulgidus* AGT (80 °C).

All of the activity studies described above measured the repair of O^6 -methylguanine in DNA. AGTs are also able to repair O^4 -methylthymine, although the rate of such repair varies significantly with the source of AGT [23–27]. Mammalian AGTs are particularly poor at carrying out this reaction [25,28]. Both *Aq. aeolicus* and *Ar. fulgidus* AGTs were able to repair O^4 -methylthymine in a methylated poly(dA · dT) substrate, but they were less active than in the repair of O^6 -methylguanine. The *Aq. aeolicus* AGT was about 14 times less active and the *Ar. fulgidus* AGT was 48 times less active. These assays were carried out at the optimal temperature for repair of O^6 -methylguanine (70 °C and 80 °C respectively). At these temperatures, it is likely that the methylated poly(dA · dT) substrate used for assay of O^4 -methylthymine repair is not double-stranded, but the ratio of activities was not greatly different when both assays were carried out at 37 °C (15-fold for *Aq. aeolicus* and 42-fold for *Ar. fulgidus*). Mammalian AGTs have been reported to have

Figure 4 Inactivation of purified AGTs by BG

(A) Inactivation experiments were carried out as described in the Materials and methods section by incubation of the pure proteins with the BG concentrations shown, for 30 min in the presence of 10 μg of calf thymus DNA. The residual activity was then determined by incubation with [3 H]methylated DNA for 30 min at 37 °C for hAGT (circles), at 70 °C for *Aq. aeolicus* AGT (squares) and at 80 °C for *Ar. fulgidus* AGT (triangles). (B) [3 H]guanine formation from O^6 -benzyl[8- 3 H]guanine by *Aq. aeolicus* AGT as a function of the temperature examined. Results are shown for the protein renatured in the presence of DNA (squares) and renatured at pH 4.5 (circles).

30–10000 fold better activity on O^6 -methylguanine than on O^4 -methylthymine [23–27]. This wide variation probably relates to the assays and substrates employed. hAGT at 37 °C showed approx. 100-fold greater activity towards O^6 -methylguanine with the substrates we employed, whereas the *E. coli* Ogt (AGT derived from the *Ogt* gene), which is known to be much less discriminatory, showed only a 3-fold difference.

Aq. aeolicus AGT and *Ar. fulgidus* AGT were both inactivated by BG as shown in Figure 4(A). *Aq. aeolicus* AGT was more sensitive to reaction with BG with an ED_{50} value of only 7.0 μM , whereas the archaeal *Ar. fulgidus* AGT had an ED_{50} value of 200 μM . Although both proteins were less sensitive than hAGT, they are much more sensitive than the *E. coli* Ada-C (C-terminal domain of the *Ada* gene product) and that from *Saccharomyces cerevisiae*, which are almost totally refractory to BG, and the *E. coli* Ogt, which had an ED_{50} of about 600 μM under these conditions [22,29].

The ability of the *Aq. aeolicus* AGT to react with BG was also demonstrated by showing that [8- 3 H]guanine was formed from O^6 -benzyl[8- 3 H]guanine by this protein (Figure 4B). This finding confirms that inactivation occurs in the same way as originally documented for the hAGT [22] by an alkyl transfer reaction at the active site of the protein generating *S*-benzylcysteine at the cysteine acceptor residue. The reaction with BG had a slightly lower optimal temperature (60 °C) than the repair of DNA (70 °C). *Aq. aeolicus* AGT isolated by refolding at pH 4.6 was less active than that renatured at pH 7.6 in the presence of DNA and had a lower optimal pH, suggesting that refolding under acidic conditions does not restore the structure as efficiently as renaturation in the presence of DNA.

DISCUSSION

Although many probable AGT sequences have been identified in genome data bases based on the presence of the PCHR (Pro-Cys-His-Arg) motif around the active site, a much smaller number of such proteins have been purified and characterized, and there have been only very limited reports of the characterization of AGTs from acute thermophiles. There have been no previous reports of the properties of AGTs from thermophilic bacteria. Studies of crude extracts of several Archaea, including *Sulfolobus acidocaldarius*, *Pyrococcus furiosus*, *Pyrobaculum islandicum* and *Thermococcus litoralis* have been described [30]. These indicate the presence of AGT activity in crude extracts from these organisms and that the activities had a high optimum temperature and stability against thermodenaturation, but the extent to which other proteins present in these crude extracts contribute to these properties is unknown. So far, the only archaeal AGT that has been studied as purified protein is that from *P. kodakaraensis* [12,31]. This purified protein was shown to have the capacity to repair methylated DNA at 50 °C, but was not compared in rate at other temperatures or to other AGTs. The *P. kodakaraensis* AGT was shown to be thermostable in a very limited study in which, after incubation for 1 h, it lost 25 % of its activity at 75 °C and 66 % at 90 °C [31]. In a more detailed biophysical analysis of its conformational stability, it was found to unfold at 98.6 °C, which is more than 50 °C greater than the melting temperature for *E. coli* Ada-C AGT [32].

The recombinant thermostable AGTs from *Aq. aeolicus* and *Ar. fulgidus* were purified under denaturing conditions because both proteins formed inclusion bodies in *E. coli*. Soluble AGT was obtained in the form of an AGT–DNA complex, as renaturation of the protein without DNA resulted in protein precipitation. Expression of these two thermophilic proteins at 37 °C, purification and refolding at low temperature along with their high pI values (pH 9.7 for *Aq. aeolicus* and pH 9.8 for *Ar. fulgidus* compared with pH 8.1 for hAGT) could have contributed to protein–protein interactions resulting in AGT aggregation. The presence of DNA during protein refolding may interfere with these protein–protein interactions, resulting in the yield of soluble AGT. Although the *Aq. aeolicus* AGT could also be isolated in a soluble form after removal of urea at low pH in the absence of DNA, this protein was less active than that isolated in the presence of DNA, suggesting that it is not fully or correctly renatured.

The observed optimum temperature (70 °C) for DNA repair by *Aq. aeolicus* AGT was well below the reported growth temperature (85 °C) for this organism [18]. The protein was highly stable only up to 60 °C, suggesting that chaperone-like proteins or other stabilizing factors present in this organism are needed to keep it in an active state at the growth temperature of *Aq. aeolicus*. On the other hand, the optimum temperature for activity (80 °C) of the *Ar. fulgidus* AGT was very close to its optimal growth temperature of 83 °C [19].

The presence of AGTs and methylpurine glycosylases [33,34] in hyperthermophilic organisms suggests they are naturally exposed to endogenous methylating agents. Several such agents have been proposed, including S-adenosylmethionine (which is known to be present in hyperthermophiles) and N-nitroso compounds [4,35,36]. Although such agents are chemically unstable at the high temperatures at which these organisms grow, such chemical decomposition may actually increase the extent to which they form DNA alkylation products.

Previous studies [22,29] have indicated that AGTs from microbial sources react poorly, if at all, with the irreversible inhibitor BG, which is a powerful inactivator of mammalian AGTs [9,22] and is currently undergoing clinical trials as an adjunct to cancer

chemotherapy [10,37,38]. The *Aq. aeolicus* AGT resembles the mammalian AGTs in its sensitivity to BG and is the first bacterial AGT reported to exhibit this property. The original concept leading to the synthesis of BG was that it would serve as a substrate for the protein because of the ability of the benzyl group to enter into a bimolecular displacement reaction with the active-site cysteine residue [9]. Subsequent studies by biochemical, site-directed mutagenesis and structural approaches [13,14,22,39,40] suggest that its ability to benzylate this cysteine requires its binding in the active-site pocket and that a major component of this binding is an interaction with the side chain of residue Pro¹⁴⁰ (residue numbers in AGT referred to in the text are based on the hAGT sequence in Figure 1) in hAGT (Figure 1). This proline residue is absent in many microbial AGTs, but is present in both the *Aq. aeolicus* and the *Ar. fulgidus* AGTs, which is consistent with the hypothesis for the activity of BG described above and the sensitivity to BG of these AGTs shown in Figure 4.

Another factor that determines sensitivity to BG is the size of the active-site pocket, which, in some cases, is too small to readily accommodate the inhibitor [13,14,39,41]. *Aq. aeolicus* and *Ar. fulgidus* AGTs resemble hAGT in the presence of a glycine residue at position Gly¹⁵⁶ (Figure 4). This orients the surface of the substrate-binding pocket allowing adequate space to bind BG. An additional steric factor contributing to the resistance of Ada-C and some other microbial AGTs to BG is the bulky side chain of the tryptophan residue that occurs in a RWG (Arg-Trp-Gly) sequence equivalent to SGG (Ser-Gly-Gly) at residues 159–161 in hAGT [39]. This tryptophan residue is absent in the *Aq. aeolicus* and the *Ar. fulgidus* proteins. However, *Ar. fulgidus* AGT, which was less sensitive to BG than that from *Aq. aeolicus*, has an insert of a stretch of six extra amino acids in this region (Figure 1). This is a unique feature not found in any other AGT (including the other archaeal AGTs), and may account for its diminished reaction with BG compared with *Aq. aeolicus*.

Crystal structures are available for three AGTs (*E. coli* Ada-C, human and *P. kodakaraensis*) [11–14]. These structural studies have indicated that the protein contains two domains. The role of the N-terminal domain, which is represented by residues 1–85 in hAGT, is unclear, but it is essential for AGT stability and for the correct orientation of the C-terminal domain. There is no obvious similarity in the residues in this domain between the five AGTs shown in Figure 1, although the structural studies show that its overall structure is quite similar between the Ada-C, hAGT and *P. kodakaraensis* AGTs. It has been proposed that the thermostability of the latter protein is due to the presence of intra- and inter-helix ion pairs. The residues forming these ion pairs, which are indicated in Figure 1, are frequently, although not invariantly, conserved between the *P. kodakaraensis* AGT and those from *Aq. aeolicus* and the *Ar. fulgidus*. In several cases where they are not conserved, a minor change in the alignment (which is not very definitive in the N-terminal domain in which there are so few common residues) indicates other charged residues that could fulfil this requirement.

The C-terminal domain of the AGT contains the active-site binding pocket and the DNA binding domain which is a winged helix–turn–helix motif [3,11]. All of the 17 fully conserved residues of the *Aq. aeolicus* and *Ar. fulgidus* AGTs are present in this domain (Figure 1), including the key structural residue Asn¹³⁷ that stabilizes a tight turn, which is a critical part of the hinge region between the two domains [13]. The cysteine acceptor site of AGTs is buried in the protein and, in order for reaction with a DNA substrate to take place, the substrate O⁶-alkylguanine must be displaced from the DNA helix [3]. This nucleoside flipping requires the insertion of an amino acid residue into the DNA to displace the base. This role is filled by Arg¹²⁸ [13,42], which is one of the fully conserved residues.

The active-site acceptor cysteine residue in AGTs has a low pK_a consistent with a thiolate anion [43] and is highly reactive with electrophiles. This reactivity is likely to result from a concerted action of a network of amino acids and a bound water linking residues His¹⁴⁶, Arg¹⁴⁷ and Glu¹⁷² to Cys¹⁴⁵ [13]. All of these residues are conserved, as is Tyr¹¹⁴, which is involved in both DNA binding and in the facilitation of the alkyl transfer reaction probably by acting to protonate the guanine substrate at the N3 position [44].

The features of the AGT active site allowing good reaction with *O*⁴-methylthymine are not well understood, although some attempts have been made to mutate hAGT to facilitate its very poor reaction with this pyrimidine [27]. The *E. coli* Ada and Ogt AGTs are much better than hAGT in the repair of *O*⁴-methylthymine [24,26]. The thermostable AGTs showed intermediate activity, but these proteins (especially that from *Aq. aeolicus*) are more like the *E. coli* AGTs in being competent to repair *O*⁴-methylthymine. The residues in the substrate-binding pocket located at positions 144–164 in hAGT are those which influence the reaction with *O*⁴-methylthymine, since altering these residues to those present in Ogt changes the specificity to that of Ogt (K. Goodtzova, S. Kanugula and A. E. Pegg, unpublished work). The sequences of the thermostable AGTs differ significantly from both hAGT and the *E. coli* AGTs in this region (Figure 1). Of particular interest is the replacement of Ser¹⁵² by a basic residue since Ser¹⁵² → Arg was one of the mutations in the active site of hAGT that was found to improve reaction with *O*⁴-methylthymine by Encell and Loeb [27].

In summary, two novel thermostable AGTs from Archaea and Bacteria were purified and characterized. Apart from their thermal stability, each of these proteins have novel features that may advance the study of direct repair of alkylation damage and its role in protecting cells from environmental and therapeutic alkylating agents. It is noteworthy that *Aq. aeolicus*, which was identified as the most primitive bacterium [18], has an AGT that resembles the mammalian AGTs more closely than other bacterial equivalents with respect to sensitivity to BG. Further study of this thermostable protein with respect to inactivation by such pseudo-substrates may provide a useful model system for this reaction. The archaeal *Ar. fulgidus* AGT is the smallest well-characterized AGT yet described with only 147 amino acids, and its remarkable optimum temperature and stability also offers advantages for further study of the AGT reaction.

This research was supported in part by grant CA-18137 from the National Cancer Institute, National Institutes of Health, Bethesda, MD, U.S.A.

REFERENCES

- Pegg, A. E., Dolan, M. E. and Moschel, R. C. (1995) Structure, function and inhibition of *O*⁶-alkylguanine-DNA alkyltransferase. *Progr. Nucleic Acid Res. Mol. Biol.* **51**, 167–223
- Pieper, R. O. (1997) Understanding and manipulating *O*⁶-methylguanine-DNA methyltransferase expression. *Pharmacol. Ther.* **74**, 285–297
- Daniels, D. S. and Tainer, J. A. (2000) Conserved structural motifs governing the stoichiometric repair of alkylated DNA by *O*⁶-alkylguanine-DNA alkyltransferase. *Mutation Res.* **460**, 151–163
- Margison, G., Povey, A. C., Kaina, B. and Santibáñez-Koref, M. F. (2003) Variability and regulation of *O*⁶-alkylguanine-DNA alkyltransferase. *Carcinogenesis* **24**, 625–635
- Singer, B. and Essigmann, J. M. (1991) Site-specific mutagenesis: retrospective and prospective. *Carcinogenesis* **12**, 949–955
- Margison, G. P., Santibáñez-Koref, M. F. and Povey, A. C. (2002) Mechanisms of carcinogenicity/chemotherapy by *O*⁶-methylguanine. *Mutagenesis* **17**, 483–487
- Pauly, G. T. and Moschel, R. C. (2001) Mutagenesis by *O*⁶-methyl-, *O*⁶-ethyl-, and *O*⁶-benzylguanine and *O*⁴-methylthymine in human cells: effects of *O*⁶-alkylguanine-DNA alkyltransferase and mismatch repair. *Chem. Res. Toxicol.* **14**, 894–900
- Kanugula, S. and Pegg, A. E. (2001) A novel DNA repair alkyltransferase from *Caenorhabditis elegans*. *Environ. Mol. Mutagen.* **38**, 235–243
- Dolan, M. E., Moschel, R. C. and Pegg, A. E. (1990) Depletion of mammalian *O*⁶-alkylguanine-DNA alkyltransferase activity by *O*⁶-benzylguanine provides a means to evaluate the role of this protein in protection against carcinogenic and therapeutic alkylating agents. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5368–5372
- Friedman, H. S. (2000) Can *O*⁶-alkylguanine-DNA alkyltransferase depletion enhance alkylator activity in the clinic? *Clin. Cancer Res.* **6**, 2967–2968
- Moore, M. H., Gulbus, J. M., Dodson, E. J., Demple, B. and Moody, P. C. E. (1994) Crystal structure of a suicidal DNA repair protein: the Ada *O*⁶-methylguanine-DNA methyltransferase from *E. coli*. *EMBO J.* **13**, 1495–1501
- Hashimoto, H., Inoue, T., Nishioka, M., Fujiwara, S., Takagi, M., Imanaka, T. and Kai, Y. (1999) Hyperthermostable protein structure maintained by intra- and inter-helix ion pairs in archaeal *O*⁶-methylguanine-DNA methyltransferase. *J. Mol. Biol.* **292**, 707–716
- Daniels, D. S., Mol, C. D., Arval, A. S., Kanugula, S., Pegg, A. E. and Tainer, J. A. (2000) Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical base binding. *EMBO J.* **19**, 1719–1730
- Wibley, J. E. A., Pegg, A. E. and Moody, P. C. E. (2000) Crystal structure of the human *O*⁶-alkylguanine-DNA alkyltransferase. *Nucleic Acid Res.* **28**, 393–401
- DiRuggiero, J., Brown, J. R., Bogert, A. P. and Robb, F. T. (1999) DNA repair systems in archaea: mementos from the last universal common ancestor? *J. Mol. Evol.* **49**, 474–484
- Makarova, K. S., Aravind, L., Grishin, N. V., Rogozin, I. B. and Koonin, E. V. (2002) A DNA repair system specific for thermophilic archaea and bacteria predicted by genomic analysis. *Nucleic Acids Res.* **30**, 482–496
- White, M. F. (2003) Archaeal DNA repair: paradigms and puzzles. *Biochem. Soc. Trans.* **31**, 690–693
- Deckert, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E. and Overbeek, R. (1998) The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature (London)* **392**, 353–358
- Klenk, H. P., Clayton, R. A., Tomb, J. F., White, O., Nelson, K. E., Ketchum, K. A., Dodson, R. J., Gwinn, M., Hickey, E. K., Peterson, J. D. et al. (1997) The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature (London)* **390**, 364–370
- Edara, S., Kanugula, S., Goodtzova, K. and Pegg, A. E. (1996) Resistance of the human *O*⁶-alkylguanine-DNA alkyltransferase containing arginine at codon 160 to inactivation by *O*⁶-benzylguanine. *Cancer Res.* **56**, 5571–5575
- Goodtzova, K., Crone, T. and Pegg, A. E. (1994) Activation of human *O*⁶-alkylguanine-DNA alkyltransferase by DNA. *Biochemistry* **33**, 8385–8390
- Pegg, A. E., Boosalis, M., Samson, L., Moschel, R. C., Byers, T. L., Swenn, K. and Dolan, M. E. (1993) Mechanism of inactivation of human *O*⁶-alkylguanine-DNA alkyltransferase by *O*⁶-benzylguanine. *Biochemistry* **32**, 11998–12006
- Graves, R. J., Li, B. F. L. and Swann, P. F. (1989) Repair of *O*⁶-methylguanine, *O*⁶-ethylguanine, *O*⁶-isopropylguanine and *O*⁴-methylthymine in synthetic oligodeoxynucleotides by *Escherichia coli* ada gene *O*⁶-alkylguanine-DNA-alkyltransferase. *Carcinogenesis* **10**, 661–666
- Sassanfar, M., Dosanjh, M. K., Essigmann, J. M. and Samson, L. (1991) Relative efficiencies of the bacterial, yeast, and human DNA methyltransferases for the repair of *O*⁶-methylguanine and *O*⁴-methylthymine. *J. Biol. Chem.* **266**, 2767–2771
- Zak, P., Kleib, K. and Laval, F. (1994) Repair of *O*⁶-methylguanine and *O*⁴-methylthymine by the human and rat *O*⁶-methylguanine-DNA methyltransferase. *J. Biol. Chem.* **269**, 730–733
- Paalman, S. R., Sung, C. and Clarke, N. D. (1997) Specificity of DNA repair methyltransferases determined by competitive inactivation with oligonucleotide substrates: evidence that *Escherichia coli* Ada repairs *O*⁶-methylguanine and *O*⁴-methylthymine with similar efficiency. *Biochemistry* **36**, 11118–11124
- Encell, L. P. and Loeb, L. A. (1999) Redesigning the substrate specificity of human *O*⁶-alkylguanine-DNA alkyltransferase: mutants with enhanced repair of *O*⁴-methylthymine. *Biochemistry* **38**, 12097–12103
- Samson, L., Han, S., Marquis, J. C. and Rasmussen, L. J. (1997) Mammalian DNA repair methyltransferases shield *O*⁴-methylthymine from nucleotide excision repair. *Carcinogenesis* **18**, 919–924
- Elder, R. H., Margison, G. P. and Rafferty, J. A. (1994) Differential inactivation of mammalian and *Escherichia coli* *O*⁶-alkylguanine-DNA alkyltransferase by *O*⁶-benzylguanine. *Biochem. J.* **298**, 231–235
- Skorvaga, M., Raven, N. D. H. and Margison, G. P. (1998) Thermostable archaeal *O*⁶-alkylguanine-DNA alkyltransferases. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6711–6715
- Leclerc, M. M., Nishioka, M., Yuasa, T., Fujiwara, S., Takagi, M. and Imanaka, T. (1998) The *O*⁶-methylguanine-DNA methyltransferase from the hyperthermophilic archaeon *Pyrococcus* sp. KOD1: a thermostable repair enzyme. *Mol. Gen. Genet.* **258**, 69–77

- 32 Shiraki, K., Nishikori, S., Fujiwara, S., Hashimoto, H., Kai, Y., Takagi, M. and Imanaka, T. (2002) Comparative analyses of the conformation stability of a hyperthermophilic protein and its mesophilic counterpart. *Eur. J. Biochem.* **268**, 4144–4150
- 33 Begley, T. J., Haas, B. J., Noel, J., Shekhtman, A., Williams, W. A. and Cunningham, R. P. (1999) A new member of the endonuclease III family of DNA repair enzymes that removes methylated purines from DNA. *Curr. Biol.* **17**, 653–656
- 34 Birkeland, N. -K., Ånensen, H., Knaevelsrud, I., Kristoffersen, W., Bjørås, M., Robb, F. T., Klungland, A. and Bjelland, S. (2002) Methylpurine DNA glycosylase of the hyperthermophilic archaeon *Archaeoglobus fulgidus*. *Biochemistry* **41**, 12697–12705
- 35 Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature (London)* **362**, 709–715
- 36 Taverna, P. and Sedgwick, B. (1996) Generation of an endogenous DNA-methylating agent by nitrosation in *Escherichia coli*. *J. Bacteriol.* **178**, 5105–5111
- 37 Quinn, J. A., Pluda, J., Dolan, M. E., Delaney, S., Kaplan, R., Rich, J. N., Friedman, A. H., Reardon, D. A., Sampson, J. H., Colvin, O. M. et al. (2002) Phase II trial of carmustine plus *O*⁶-benzylguanine for patients with nitrosourea-resistant recurrent or progressive malignant glioma. *J. Clin. Oncol.* **2**, 2277–2283
- 38 Margison, G. P. and Santibáñez-Koref, M. F. (2002) *O*⁶-Alkylguanine-DNA alkyltransferase: role in carcinogenesis and chemotherapy. *BioEssays* **24**, 255–266
- 39 Goodtzova, K., Kanugula, S., Edara, S., Pauly, G. T., Moschel, R. C. and Pegg, A. E. (1997) Repair of *O*⁶-benzylguanine by the *Escherichia coli* Ada and Ogt and the human *O*⁶-alkylguanine-DNA alkyltransferase. *J. Biol. Chem.* **272**, 8332–8339
- 40 Xu-Welliver, M., Kanugula, S. and Pegg, A. E. (1998) Isolation of human *O*⁶-alkylguanine-DNA alkyltransferase mutants highly resistant to inactivation by *O*⁶-benzylguanine. *Cancer Res.* **58**, 1936–1945
- 41 Crone, T. M., Kanugula, S. and Pegg, A. E. (1995) Mutations in the Ada *O*⁶-alkylguanine-DNA alkyltransferase conferring sensitivity to inactivation by *O*⁶-benzylguanine and 2,4-diamino-6-benzoyloxy-5-nitrosopyrimidine. *Carcinogenesis* **16**, 1687–1692
- 42 Kanugula, S., Goodtzova, K., Edara, S. and Pegg, A. E. (1995) Alteration of arginine-128 to alanine abolishes the ability of human *O*⁶-alkylguanine-DNA alkyltransferase to repair methylated DNA but has no effect on its reaction with *O*⁶-benzylguanine. *Biochemistry* **34**, 7113–7119
- 43 Guengerich, F. P., Fang, Q., Liu, L. and Pegg, A. E. (2003) *O*⁶-Alkylguanine-DNA alkyltransferase: low p*K*_a and reactivity of cysteine **145**. *Biochemistry*, in the press
- 44 Spratt, T. E., Wu, J. D., Levy, D. E., Kanugula, S. and Pegg, A. E. (1999) Reaction and binding of oligodeoxynucleotides containing analogs of *O*⁶-methylguanine with wild type and mutant human *O*⁶-alkylguanine-DNA alkyltransferase. *Biochemistry* **38**, 6801–6806

Received 2 June 2003/24 July 2003; accepted 1 August 2003

Published as BJ Immediate Publication 1 August 2003, DOI 10.1042/BJ20030809