Probing of conformational changes in human *O*⁶-alkylguanine-DNA alkyl transferase protein in its alkylated and DNA-bound states by limited proteolysis

Sreenivas KANUGULA, Karina GOODTZOVA and Anthony E. PEGG¹

Department of Cellular and Molecular Physiology, H-166, M. S. Hershey Medical Center, Pennsylvania State University College of Medicine, Hershey, PA 17033-0850, U.S.A.

Human *O*⁶-alkylguanine-DNA alkyl transferase (hAGT) is a DNA repair protein that protects cells from alkylation damage by transferring an alkyl group from the *O*⁶-position of guanine to a cysteine residue in the active site (-PCHR-) of the protein. The structure of the hAGT protein (23 kDa) has been probed by limited proteolysis with trypsin and Glu-C endoproteases and analysis of the polypeptide fragments by SDS/PAGE. The native hAGT protein had limited accessibility to digestion with trypsin and Glu-C in spite of a number of potential cleavage sites. Initial cleavage by trypsin occurred at residue Lys-193 to give a 21 kDa polypeptide fragment, and this polypeptide underwent further cleavage at residues Arg-128 and Lys-165. These trypsin-cleavage sites became more accessible to digestion in the presence of double-stranded DNA (dsDNA), indicating that hAGT undergoes a change in its conformation on binding to DNA. However, the trypsin cutting site at the Arg-128 position was less available

for digestion in the presence of single-stranded DNA (ssDNA), suggesting that the hAGT protein has a different conformation when bound to ssDNA compared with dsDNA. When protease digestion was carried out on wild-type protein, preincubated with the low-molecular-mass pseudosubstrate *O*⁶-benzylguanine, increased susceptibility to proteases was observed. A mutant C145A hAGT protein, which cannot repair *O*⁶-alkylguanine because the Cys-145 acceptor site in the active site of the protein is changed to Ala, showed identical trypsin cleavage to the wild type, but its digestion was not affected by *O*⁶-benzylguanine. These results suggest that alkylation of hAGT leads to an altered conformation. The acquisition of increased susceptibility to proteases upon DNA binding and alkylation demonstrates that hAGT undergoes considerable conformational changes in its structure upon binding to DNA and after repair of alkylation damage.

INTRODUCTION

Alkylation at the *O*⁶-position of guanine and the *O*⁴-position of thymine in DNA makes these bases highly mutagenic, causing G·C → A·T and A·T → G·C transitions [1–3]. Organisms have developed an effective repair mechanism to remove these mutagenic lesions from DNA through the action of *O*⁶-alkylguanine-DNA alkyltransferase (AGT) (EC 2.1.1.63) [4–7]. Alkyltransferase proteins have been identified from many different species. All have a -PCHR- sequence containing the cysteine acceptor site and repair the alkylation damage in DNA by transferring the alkyl group from the *O*⁶-position of guanine to this cysteine residue, thus restoring the DNA to its unmodified form. The cysteine is not regenerated, but the mechanism of the removal of the alkylated form of the protein is not well understood. Investigations using immunoblots of extracts from cells treated with *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine or *O*⁶-benzylguanine (BG) indicated that the human AGT (hAGT) was rapidly degraded [8] and, recently, evidence has been presented that alkylated hAGT may be ubiquitinated [9,10]. This may indicate that it is removed from the cell by the ATP-dependent, ubiquitin-mediated, proteosomal degradation pathway.

Only a very limited amount of information is available concerning the structure of hAGT. A recent study using CD and fluorescence spectroscopy suggests that the protein is highly constrained and is approximately spherical [11]. The X-ray structure of the *Escherichia coli* C-terminal domain of Ada (Ada-

C) alkyltransferase (amino acids 178–354 of the 39 kDa *E. coli* *O*⁶-alkylguanine-DNA alkyltransferase (Ada) protein) indicates a buried active site and a helix–turn–helix DNA-binding motif [12,13]. Although the structure of hAGT is yet to be solved, the structure of Ada-C could be viewed as a model for hAGT in view of the considerable amino acid sequence similarity between these two proteins [7,14]. It has been suggested that Ada-C needs to undergo substantial conformational change in its structure on binding to DNA in order to expose the active site to the alkyl group on the DNA substrate [13], although a substantial change in the DNA structure may also contribute to this juxtaposition [15–17]. AGT has been shown to bind to both double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA) [16,18–20], and more tightly to DNA containing an *O*⁶-methylguanine adduct [21]. Binding to non-adducted DNA appears to involve interaction with from 4–9 nucleotides [18–20], while footprinting studies with a dsDNA substrate containing a single *O*⁶-methylguanine indicate interaction with up to 18 nucleotides on the adducted strand and 14 in the complementary strand [21].

DNA-dependent conformational changes in AGT proteins were observed for Ada-C protein [18] and hAGT [16,19]. The DNA-dependent stimulation of the rate of alkyl transfer from the low-molecular-mass substrate, BG, which has been observed for both hAGT and Ada-C [17,22], is consistent with the conformational change that activates the protein [13].

Probing protein structure with protease sensitivity has been

Abbreviations used: Ada, *E. coli* *O*⁶-alkylguanine-DNA alkyltransferase; Ada-C, the C-terminal domain of Ada; hAGT, human *O*⁶-alkylguanine-DNA alkyl transferase; DTT, dithiothreitol; dsDNA, double-stranded DNA; ssDNA, single-stranded DNA; BG, *O*⁶-benzylguanine.

¹ To whom correspondence should be addressed.

widely used to understand the structures of a variety of proteins. Hubbard et al. [23] suggested that the proteases generally cleave peptide sequences only in the regions that can adopt rather extended conformations. Thus, proteolytic cleavages occur easily at the less structured and more flexible regions, and the pattern and sequence of the proteolytic cuts of a protein can be useful in understanding its domain structure and conformation. Some examples of this are: identification of the domain organization in the alpha subunit of *E. coli* RNA polymerase [24], demonstration of nucleotide- and DNA-induced conformational changes in phage T7 gene 4 protein [25] and probing of DNA-induced conformational changes in vaccinia topoisomerase [26] or *Sulfolobus solfataricus* DNA polymerase [27].

The possibility that protease digestion can be used to evaluate the structure of hAGT was supported by reports that V8 protease preferentially cleaves the alkylated form of the protein at residues Glu-30, Glu-172 and in a region around Lys-107 [28,29]. We have carried out a more detailed study of purified hAGT and of mutant hAGT molecules generated by site-directed mutagenesis using proteolytic enzymes as probes of the protein structure. The sites of proteolytic cleavage were identified and changes in protease sensitivity in response to alkylation of the active site or to binding of DNA were observed. Native hAGT was found to have very limited accessibility to proteases in its native state, but the sensitivity to digestion was altered in response to DNA binding or alkylation, indicating that the hAGT undergoes considerable conformational change under these conditions.

EXPERIMENTAL PROCEDURES

Materials

All chemicals used were reagent grade. BG was generously provided by Dr. R. C. Moschel (ABL-Basic Research Program, National Cancer Institute-Frederick Research and Development Center, Frederick, MD, U.S.A.) [30]. Trypsin and Glu-C and other biochemical reagents were purchased from Sigma Chemical Co. (St. Louis, MO, U.S.A.). The Sequenase version 2.0 kit and M13 mp18 ssDNA was procured from Amersham Corporation (Arlington Heights, IL, U.S.A.). Plasmid pGEM, T7 RNA polymerase and the TnT rabbit reticulocyte lysate were purchased from Promega (Madison, WI, U.S.A.) and Pfu DNA polymerase was from Stratagene (La Jolla, CA, U.S.A.). [³⁵S]Methionine (1000 Ci/mmol) was purchased from NEN Life Science Products (Boston, MA, U.S.A.).

Mutant AGT proteins

The construction of AGT mutants C145A, Y114E, R128A and G177stop has been described previously [31–33]. Mutant K165A was made by PCR using 5'-GGACTGGCCGTGGCGGAATGGCTTCTGGCC-3' as the mutagenic primer and the conditions for PCR previously described [31]. The entire coding sequence of all mutant AGT proteins was verified by DNA sequencing.

Purification of wild-type and mutant AGT proteins

Wild-type and mutant hAGT proteins were expressed in *E. coli* and purified to homogeneity [34]. The wild-type (207 amino acids) and C145A mutant proteins were expressed from pCX35 hAGT plasmid [35]. Mutants Y114E, R128A, K165A and G177stop were expressed from pInhAGT plasmid, and insertion into these vectors adds a five amino acid extension (Met-Lys-Gly-Gly-Ile-Leu- in place of Met-) at the N-terminus [34]. The purified proteins were analysed by SDS/PAGE and the protein concentrations were estimated by using the protein assay reagent kit from Bio-Rad (Richmond, CA, U.S.A.).

Protease digestion of hAGT

Wild-type and mutant hAGT proteins in 50 mM Tris/HCl, pH 7.6/5 mM dithiothreitol (DTT)/0.1 mM EDTA were subjected to protease digestion at 25 °C with either trypsin or Glu-C at different enzyme concentrations [indicated as enzyme/substrate (E/S) ratio, w/w] or for different time intervals, as indicated in the Figure legends. The digestion reaction was stopped by adding the SDS/PAGE loading buffer and immediately boiling the samples at 100 °C for 5 min. Alkylation of the hAGT protein at the Cys-145 acceptor site was accomplished by incubating the proteins with 100 μM BG at 37 °C for 30 min. This reaction mixture was then subjected to protease digestion as described above. The hAGT-DNA complexes were made with M13 ssDNA for ssDNA, and with pGEM plasmid DNA digested with *Eco*RI for dsDNA. The complexes were formed by incubating the protein and the DNA at 37 °C for 30 min and these pre-formed hAGT-DNA complexes were then subjected to trypsin digestion as described above. The polypeptide fragment products of digestion were analysed by SDS/PAGE.

Analysis of products by gel electrophoresis and immunoblotting

The protease digestion products were separated by SDS/PAGE on 15 % polyacrylamide gels [36] and the polypeptide fragments were visualized by silver staining [37] or were transferred onto 0.45 μm nitrocellulose membranes and probed with antibodies raised to peptides corresponding to N-terminal (amino acid residues 8–20) or C-terminal (amino acid residues 197–207) sequences from hAGT [38]. The nitrocellulose membranes were developed using the enhanced chemiluminescence detection system (ECL; Amersham). The stripping and re-probing of the nitrocellulose membranes, wherever necessary, was done according to the manufacturer's instructions.

In vitro transcription and translation

The ³⁵S-labelled AGT protein was synthesized using a pGEMAGT plasmid, which contains the entire hAGT coding sequence downstream from the T7 RNA polymerase promoter. pGEMAGT plasmid DNA (4 μg) was incubated in the TnT transcription/translation system for 60 min at 30 °C in the presence of 160 μCi of [³⁵S]methionine and T7 RNA polymerase in a final volume of 0.2 ml. The radiolabelled protein was precipitated by adding ammonium sulphate solution to give a final concentration of 1.6 M, and was incubated on ice for 30 min. The protein precipitate was then collected by centrifugation at 16000 g for 30 min, dissolved in 2.0 ml of 50 mM Tris/HCl, pH 7.6/5 mM DTT/0.1 mM EDTA and concentrated to a volume of 0.1 ml using a Centricon-10 microconcentrator (Amicon, Beverly, MA, U.S.A.). The concentrated protein solution was diluted to 2.0 ml with the same buffer and the concentration step was repeated once more to a final volume of 0.1 ml. An aliquot of this solution was used for trypsin digestion and the reaction products were separated by SDS/PAGE. The gels were then dried and exposed to photographic film or the radioactivity in the bands was quantified using a Molecular Dynamics 425E-120 PhosphorImager.

RESULTS

Trypsin digestion of wild-type and mutant hAGT proteins

The primary structure of hAGT (23 kDa) has 19 potential trypsin-cleavage sites (Figure 1), but only a limited number of

```

1 MDKDCEMKRTTLDSPLGKLELSGCEQGLHEIKLLGKGTSA 40
41 ADAVEVPAPAAVLGGPEPLMQCTAWLNAYFHQPEAIEEFP 80
81 VPALHHPVFQQESFTTRQVLWKLLKVVKFGEVISYQQLAAL 120
    ↓
121 AGNPKAARAVGGAMRGNPVPILIPCHRVVCSSGAVGNYSG 160
    ↓
161 GLAVKEWLLAHEGHRLGKPGLGGSSGLAGAWLKGAGATSG 200
201 SPPAGRN                                     207

```

Figure 1 Amino acid sequence of hAGT protein

Potential trypsin-cleavage sites (19) are indicated in bold and the actual sites as determined are shown by arrows. The sequences to which antibodies were derived are underlined.

Figure 2 Effect of trypsin concentration on digestion of hAGT

Aliquots of 2 μ g of hAGT protein in 50 mM Tris/HCl, pH 7.6/5 mM DTT/0.1 mM EDTA were digested with different concentrations of trypsin at 25 °C for 30 min (A) and 15 min (B and C). The reaction was terminated by boiling with SDS/PAGE loading buffer at 100 °C for 5 min, and 1 μ g of protein from the digestion reaction mixture was loaded into each lane and the polypeptide fragments were resolved by SDS/15% PAGE. The resulting bands were detected by either silver staining (A) or Western blotting after transfer to nitrocellulose membranes (B and C). The E/S ratio for lanes 1–5 was 0:100, 1:100, 1:50, 1:20 and 1:10 respectively. Results for membranes probed with N-terminal specific anti-hAGT antibody are shown in (B). The membrane was then stripped and re-probed with C-terminal-specific anti-hAGT antibody and the results are shown in (C).

these sites were available for digestion with trypsin as only four polypeptide fragments of sizes approximately 21, 18, 14 and 11–12 kDa were formed on trypsin digestion (Figure 2A). The 21 kDa fragment appeared even at very low trypsin concentrations (1:100; lane 2, Figure 2A), whereas the 12, 14 and 18 kDa fragments were seen only after treatment with relatively high concentrations (lanes 3–5, Figure 2A). The possible positions of these cleavage sites were mapped by checking the cross-reactivity of these polypeptide fragments with N-terminal- and C-terminal-specific antibodies to hAGT protein. (Figure 1 shows the sequences used to generate the antibodies). The N-terminal antibody reacted with all three polypeptide fragments formed

Figure 3 Trypsin digestion of different hAGT mutant proteins

Aliquots of 1 μ g of hAGT wild-type or mutant proteins were incubated with trypsin at a 1:10 ratio for 30 min and analysed as in Figure 2(B) by Western blotting with an antibody specific for the N-terminal region to detect the protein bands. Lanes 1, 3, 5 and 7 contain wild-type, G177stop, K165A and R128A proteins respectively, incubated without trypsin, whereas lanes 2, 4, 6 and 8 contain wild-type, G177stop, K165A or R128A protein samples digested with trypsin.

(Figure 2B), whereas the C-terminal antibody failed to react with any of the digestion products (Figure 2C). Therefore trypsin cleavage of hAGT must be occurring at a site(s) present in the C-terminal and central part of the native protein.

The full-length G177stop mutant protein, which has a calculated molecular mass of 19418 Da, (lane 3, Figure 3), migrated more rapidly on SDS/PAGE than the 21 kDa fragment produced from wild-type protein by trypsin (Figure 3, lane 2). This strongly suggests Lys-193 as the cleavage site for trypsin in hAGT, since other potential sites (Arg-175 or Lys-178) would give fragments of almost the same size as the full-length G177stop mutant protein. The hAGT cut at Lys-193 would have a calculated molecular mass of 20451 Da.

Digestion of the G177stop protein gives the same 14 and 18 kDa fragments as the wild-type protein (lane 4, Figure 3). The 18 kDa fragment suggests attack at Lys-165, as there were no other potential trypsin-sensitive sites that could generate a fragment of 18 kDa from the G177stop mutant. This was tested by digesting the K165A mutant hAGT with trypsin. This mutant protein failed to give the 18 kDa fragment (Figure 3, lane 6), confirming Lys-165 as the cleavage site for trypsin. Cleavage at Lys-165 would give a fragment with a calculated molecular mass of 17573 Da.

There are three potential cleavage sites (Lys-125, Arg-128 and Arg-135; see Figure 1) in the central part of hAGT that could generate the 14 kDa fragment which retains the N-terminal sequence, but the digestion of the R128A mutant protein failed to produce the 14 kDa fragment (Figure 3, lane 8), strongly suggesting that Arg-128 is the third cleavage site for trypsin, giving a fragment with a calculated size of 13985 Da.

The 11–12 kDa product was seen in only small amounts at high trypsin concentrations and is likely to result from further degradation of the 14 kDa fragment. There are several sites in the AGT sequence at which cleavage would convert the 14 kDa fragment into this polypeptide, including Arg-96, Lys-101, Lys-104 and Lys-107.

Trypsin digestion of hAGT–DNA complexes

DNA-induced conformational changes in hAGT were probed by digestion of the pre-formed hAGT–DNA complexes with trypsin (Figure 4A). At very low trypsin concentration (E/S ratio, 1:200), only a small fraction of the hAGT molecules were digested to form the 21 kDa fragment within 5 min (Figure 4A, lane 2), but, after pre-incubation with dsDNA or ssDNA, a much more extensive conversion into this product was seen (Figure 4A, lanes 3 and 4). At an E/S ratio of 1:100, hAGT

Figure 4 Effect of DNA on trypsin digestion of hAGT

Aliquots of 1 μ g of hAGT protein were incubated with 2 μ g of either M13 ssDNA or linearized pGEM plasmid dsDNA in 50 mM Tris/HCl, pH 7.6/5 mM DTT/0.1 mM EDTA at 37 °C for 30 min. The protein–DNA complexes were then subjected to trypsin digestion at different E/S ratios and the digestion was stopped after 5 min (A) or 15 min (B) and the products were analysed as described in Figure 2. The bands were detected using Western blotting with the N-terminal specific anti-hAGT antibody. Results shown in (A) are for trypsin digestion (1:200) of hAGT alone (lane 2), hAGT incubated with dsDNA (lane 3), hAGT incubated with ssDNA (lane 4), and for trypsin digestion (1:100) of hAGT alone (lane 5), hAGT-dsDNA (lane 6) and hAGT-ssDNA (lane 7). Lane 1 contained hAGT protein incubated without trypsin. Results shown in (B) are for trypsin digestion of hAGT and mutant Y114E at an E/S ratio of 1:10. Lane 1, hAGT without trypsin treatment; lanes 2, 3 and 4, effects of trypsin on hAGT alone, hAGT-dsDNA and hAGT-ssDNA respectively; lane 5, Y114E without trypsin treatment; lanes 6, 7 and 8, effects of trypsin on Y114E alone, Y114E-dsDNA and Y114E-ssDNA respectively. WT, wild type,

protein incubated with trypsin generated only the 21 kDa fragment and there was still a substantial amount of the intact full-length protein remaining (Figure 4A, lane 5), whereas hAGT preincubated with dsDNA and ssDNA (Figure 4A, lanes 6 and 7) did not show any full-length protein and the 18 kDa fragment was seen in addition to the 21 kDa product. These results show clearly that enhanced digestion of hAGT at Lys-193 and Lys-165 occurs on binding to DNA.

Trypsin digestion of hAGT–DNA complexes was also carried out at an E/S ratio of 1:10 (Figure 4B). At this concentration, the 14 kDa degradation product was produced from hAGT. The amount of this band was similar when hAGT incubated with dsDNA was used as a substrate (Figure 4B, lanes 2 and 3), but this band did not occur when hAGT combined with ssDNA was used (Figure 4B, lane 4). Alteration of the Tyr-114 residue in hAGT has been shown to affect DNA binding [31]. As shown in Figure 4(B) (lanes 5–7), the presence of ssDNA did not affect the cleavage of the Y114E mutant hAGT to form the 14 kDa product. These results indicate that binding of ssDNA prevents the tryptic cleavage of hAGT at Arg-128.

Protease digestion of alkylated hAGT protein

hAGT was allowed to react with BG to produce *S*-benzylcysteine at the active site. The alkylated form of the protein was much more sensitive to digestion by trypsin (Figure 5A) than the native protein (Figure 2A). The amount of the 18 kDa fragment detected was less than in the case of the unalkylated protein, but this could be due to the rapid conversion of the alkylated 18 kDa fragment into the 14 kDa species.

A time course of digestion of hAGT by trypsin confirmed that the alkylated form was degraded more rapidly (Figure 5B). Similar results were obtained for experiments done with 35 S-

Figure 5 Trypsin digestion of hAGT preincubated with BG

hAGT protein (1 μ g per 20 μ l of 50 mM Tris/HCl, pH 7.6/5 mM DTT/0.1 mM EDTA) was preincubated with 100 μ M BG at 37 °C for 30 min, followed by trypsin digestion of 1 μ g aliquots of hAGT at different E/S ratios. Samples were analysed as in Figure 2 and bands were detected by silver staining. Results in (A) are for hAGT without trypsin treatment (lane 1) and samples treated with trypsin at 1:100, 1:50, 1:20 and 1:10 ratios (lanes 2–5 respectively). Results in (B) are for samples of control hAGT protein: proteins were preincubated with BG digested with trypsin (1:10) for the times shown. Lanes 1–5, protein incubated without BG; lanes 6–10, protein incubated with BG. Samples were taken at 0 min (lanes 1 and 6), 5 min (lanes 2 and 7), 15 min (lanes 3 and 8), 30 min (lanes 4 and 9) and 60 min (lanes 5 and 10).

labelled hAGT protein (Figure 6). In this experiment, more than 50 % of the benzylated hAGT protein was digested within 15 min of incubation with trypsin, whereas less than 20 % of the unalkylated hAGT was degraded in this time period (Figure 6, square symbols).

Tryptic digestion of the C145A mutant hAGT protein (which cannot accept alkyl groups, as the Cys-145 acceptor residue is changed to Ala) showed no difference with respect to BG treatment (Figure 7). The digestion pattern of this mutant C145A protein was like that of hAGT in the absence of BG, both with regard to trypsin concentration (Figure 7A) and rate of degradation (Figure 7B). These results confirm that the increased

Figure 6 Time course of trypsin digestion of 35 S-labelled hAGT

A 22 μ l portion of the concentrated [35 S]AGT protein solution was diluted to a volume of 55 μ l with 50 mM Tris/HCl, pH 7.6/5 mM DTT/0.1 mM EDTA and incubated with (●, ■) or without (control: ○, □) 100 μ M BG at 37 °C for 30 min, mixed with 10 μ l of trypsin solution (16 ng/ μ l) and incubated at 25 °C for the times indicated. Aliquots of 12 μ l were removed at zero time and at the times shown and analysed by SDS/PAGE. The radioactivity present in the full-length AGT protein (■, □) and in the full length AGT protein plus the 21 kDa fragment (●, ○) was then expressed as a percentage of the radioactivity present in the full-length protein at zero time.

Figure 7 Trypsin digestion of C145A hAGT mutant protein

In the experiment shown in (A), C145A hAGT was subjected to trypsin digestion at the E/S ratios shown, with (+) and without (–) preincubation with 100 μ M BG at 37 °C for 30 min (using the same conditions as described in Figure 5). Samples were analysed as described in Figures 2 and 5 and bands were detected by silver staining. In the experiment shown in (B), samples were treated with trypsin for the times indicated, as described in Figure 5(B), and bands were detected by silver staining.

Figure 8 Glu-C digestion of hAGT preincubated with BG

hAGT protein was preincubated without (lanes 1–5) or with (lanes 6–10) 100 μ M BG at 37 °C for 30 min. Aliquots of 1 μ g of hAGT were then digested with Glu-C at different E/S ratios and analysed as described in Figure 2 with silver staining to detect the bands. hAGT without Glu-C treatment was loaded in lanes 1 and 6. Lanes 2 and 7, samples treated at a 1:100 ratio of hAGT:Glu-C; lanes 3 and 8, results for a 1:50 ratio; lanes 4 and 9, results for a 1:20 ratio; and lanes 5 and 10, results for a 1:10 ratio.

sensitivity to trypsin in hAGT exposed to BG is due to alkylation of the Cys-145 acceptor site. Using another protease, Glu-C, we also observed increased digestion of hAGT protein preincubated with BG (Figure 8).

DISCUSSION

Pure hAGT has limited accessibility to the proteases trypsin and Glu-C in its native state. By using R128A, K165A and G177stop hAGT mutant proteins, in addition to the wild-type protein, for trypsin digestion experiments, we have identified Lys-193, Arg-128 and Lys-165 as the sites of cleavage by trypsin. Of these sites, Lys-193 is by far the most sensitive. The site at Arg-128 was attacked slightly more readily than that at Lys-165 (see Figure 2B). Further cleavage of the fragment representing residues 1–128 is likely to occur at the region covering amino acids 96–107 to yield the smaller fragment of about 11–12 kDa seen in Figures 2, 4, 5 and 7.

Addition of dsDNA clearly increased the sensitivity of hAGT to trypsin at all three sites. This provides evidence that the structure of the protein is altered by binding to DNA. This concept is also supported by the observations that the presence of DNA enhances the rate of reaction of hAGT or Ada-C with BG [17,22], and by an alteration in the fluorescence and CD spectra of the hAGT protein [16,18,19]. These observations support the hypothesis, suggested on the basis of the crystal structure of the Ada-C protein [13], that a conformational change in the protein structure occurs and that this could break the interaction between His-146 and Glu-172 to stimulate the production of a thiolate anion from Cys-145.

However, it should be noted that the buried site of the Cys acceptor in the Ada-C structure [12,13] does not necessarily imply that a major conformational change must occur in the protein, as the reaction could be permitted after a protein-induced alteration in the DNA structure leading to the ‘flipping out’ of the *O*⁶-methylguanine base. Such base-flipping mechanisms have now been identified in numerous DNA methylases and DNA repair enzymes [39–43], and it has been suggested that the AGT reaction may also proceed in this way [15,41,44]. Studies with the Y114E and R128A mutant hAGT proteins, which have impaired abilities to bind DNA and to repair *O*⁶-methylguanine in DNA, are consistent with a base-flipping mechanism [31]. These experiments also revealed that the mutation at Arg-128 impaired binding of ssDNA without preventing the binding of dsDNA. This difference is also reflected in the tryptic digestion pattern, since, as shown in Figure 4, the cutting site at Arg-128 was masked from digestion in the presence of ssDNA but is made more available for cleavage in the presence of dsDNA. Glu-C digestion of hAGT–ssDNA complexes also showed a different digestion pattern compared with hAGT protein alone or hAGT–dsDNA complexes (results not shown). The finding that Arg-128 is not accessible to trypsin when bound to ssDNA is consistent with results suggesting that this residue is involved in binding the single-stranded region of the DNA in a way that allows reaction to occur.

Except for the immediate C-terminal domain, our results indicate that hAGT has a compact structure, since only a few of the potential tryptic-cleavage sites are actually sensitive to the protease. This conclusion is consistent with the results reported using physical chemical techniques or model building [11,14]. The very facile cleavage at Lys-193 indicates that the C-terminal region is likely to be only loosely associated with the rest of the protein, and experiments in which the protein is truncated by site-directed mutagenesis and comparisons of alkyltransferase sequences from different species are in agreement with this. The last totally conserved residue is that equivalent to Glu-172 in hAGT, and many of the known alkyltransferases contain only a few residues after this position [7]. These include the rabbit AGT, which lacks the residues past position 182 that are present in all other mammalian AGTs [45]. All of the residues after Leu-176 can be removed from hAGT without affecting activity [21].

The compact structure of hAGT may contribute to its stability in the cell. Many studies suggest that AGT does not turn over rapidly (reviewed in [7,46]) and this is consistent with its physiological role. The hAGT protein is able to act only once, since the protein is irreversibly inactivated after repairing alkylation damage, and therefore DNA repair involves synthesis of AGT protein. However, since AGT acts alone to restore the alkylated DNA to its original state, this provides an efficient means of dealing with damage that occurs only rarely, provided that each AGT molecule survives long enough to provide an adequate capacity for repair when needed.

Conversely, the alkylated form of the AGT protein is not

required further and its presence could interfere with repair if it still has the capacity to bind to DNA containing *O*⁶-methyl-guanine lesions. Thus, a conformational change favouring degradation may allow for the rapid removal of the inactivated alkylated form. There is good evidence that the degradation of AGT is increased in cells treated with alkylating agents or BG [8,9]. Our results provide clear evidence that a conformational change in the hAGT protein, which renders the protein more sensitive to trypsin, occurs after reaction with BG. Since the C145A mutant hAGT, which is known to bind to substrates in the same way as the wild-type protein [21], was not sensitized to tryptic degradation by incubation with BG, this conformational change is due to the formation of *S*-benzylcysteine at the active site.

These results are in agreement with the recent demonstration by CD spectroscopy that there was a small alteration in the α -helical content of hAGT on benzylation and a release of some constrained Trp residues [11]. They are also broadly consistent with the reports from Li and co-workers [28,29] that the alkylated inactive form of hAGT differs in structure from the active form. They have reported that a monoclonal antibody, which recognizes a region centred around Lys-107, reacts exclusively with the alkylated form of AGT [29] and that the alkylated protein is cleaved specifically by protease V8 [28,29]. However, in their experiments, this cleavage, which was tentatively assigned to occur at residues Glu-30 and Glu-172, did not occur with active AGT protein. In our studies with trypsin, the degradation pattern of hAGT was not altered with respect to sites attacked but was rendered much more rapid. Although not analysed in detail, the same results were obtained by us when hAGT was digested with Glu-C (V8 protease). The pure hAGT was sensitive to Glu-C degradation, although degradation was facilitated after reaction of the protein with BG (Figure 8). It should be noted that the completely alkylation-dependent V8 cleavage reported by Li and co-workers [28,29] was observed only with crude extracts where the binding of other proteins or of DNA could have protected the non-alkylated form of the protein from the protease. They noted that cleavage of pure recombinant protein did occur, although they attributed this to incorrect folding of the protein. This explanation is unlikely, since the recombinant protein used in our studies is fully active in DNA repair [34].

Convincing evidence that the methylated form of hAGT is conjugated with ubiquitin has been presented [9,10]. It is possible that the conformational change identified in the experiments mentioned above provides a signal for ubiquitination, which then leads to proteasomal degradation of the hAGT protein. Since the sites for ubiquitin conjugation are at lysine residues, the possibility that this reaction occurs at either Lys-165 or Lys-193, which are rendered more accessible after alkylation, should be considered.

This research was supported by grants CA-18138, CA-57725 and CA-71976 from the National Cancer Institute.

REFERENCES

- Loveless, A. (1969) *Nature* (London) **223**, 206–207
- Saffhill, R., Margison, G. P. and O'Connor, P. J. (1985) *Biochim. Biophys. Acta* **823**, 111–145
- Singer, B. and Dosanjh, M. K. (1990) *Mutation Res.* **233**, 45–51
- Samson, L. (1992) *Mol. Microbiol.* **6**, 825–831
- Lindahl, T., Sedgwick, B., Sekiguchi, M. and Nakabeppu, Y. (1988) *Annu. Rev. Biochem.* **57**, 133–157
- Pegg, A. E. (1990) *Cancer Res.* **50**, 6119–6129
- Pegg, A. E., Dolan, M. E. and Moschel, R. C. (1995) *Prog. Nucleic Acid Res. Mol. Biol.* **51**, 167–223
- Pegg, A. E., Wiest, L., Mummert, C., Stine, L., Moschel, R. C. and Dolan, M. E. (1991) *Carcinogenesis* **12**, 1679–1683
- Srivenugopal, K. S., Yuan, X. H., Bigner, D. D., Friedman, H. S. and Ali-Osman, F. (1996) *Biochemistry* **35**, 1328–1334
- Major, G. J., Brady, M., Notarianni, G. B., Collier, J. D. and Douglas, M. S. (1997) *Biochem. Soc. Trans.* **25**, 3595
- Federwisch, M., Hassiepen, U., Bender, K., Dewor, M., Rajewsky, M. F. and Wollmer, A. (1997) *Biochem. J.* **324**, 321–328
- Moody, P. C. E. and Moore, M. E. (1995) in *Novel Approaches in Anticancer Drug Design Molecular Modelling: New Treatment Strategies*, vol. 49 (Zeller, W. J., D'Incalci, M. and Newell, D. R., eds.), pp. 16–24, Karger, Basel
- Moore, M. H., Gulbus, J. M., Dodson, E. J., Demple, B. and Moody, P. C. E. (1994) *EMBO J.* **13**, 1495–1501
- Wibley, J. E. A., McKie, J. H., Embrey, K., Marks, D. S., Douglas, K. T., Moore, M. H. and Moody, P. C. E. (1995) *Anti-Cancer Drug Design* **10**, 75–95
- Roberts, R. J. (1995) *Cell* **82**, 9–12
- Federwisch, M., Hassiepen, U., Bender, K., Rajewsky, M. F. and Wollmer, A. (1997) *FEBS Lett.* **407**, 333–336
- Goodtzova, K., Kanugula, S., Edara, S., Pauly, G. T., Moschel, R. C. and Pegg, A. E. (1997) *J. Biol. Chem.* **272**, 8332–8339
- Takahashi, M., Sakumi, K. and Sekiguchi, M. (1990) *Biochemistry* **29**, 3431–3436
- Chan, C. L., Wu, Z., Ciardelli, T., Eastman, A. and Bresnick, E. (1993) *Arch. Biochem. Biophys.* **300**, 193–200
- Fried, M. G., Kanugula, S., Bromberg, J. L. and Pegg, A. E. (1996) *Biochemistry* **35**, 15295–15301
- Hazra, T. K., Roy, R., Biswas, T., Grabowski, D. T., Pegg, A. E. and Mitra, S. (1997) *Biochemistry* **36**, 5769–5776
- Goodtzova, K., Crone, T. and Pegg, A. E. (1994) *Biochemistry* **33**, 8385–8390
- Hubbard, S. J., Campbell, S. F. and Thornton, J. M. (1991) *J. Mol. Biol.* **220**, 507–530
- Negishi, T., Fujita, N. and Ishihama, A. (1995) *J. Mol. Biol.* **248**, 723–728
- Yong, Y. and Romano, L. J. (1995) *J. Biol. Chem.* **270**, 24509–24517
- Sekiguchi, J. and Shuman, S. (1995) *J. Biol. Chem.* **270**, 11636–11645
- Pisani, F. M., Manco, G., Carratore, V. and Rossi, M. (1996) *Biochemistry* **35**, 9158–9166
- Ayi, T. C., Oh, H. K., Lee, T. K. Y. and Li, B. F. L. (1994) *Cancer Res.* **54**, 3726–3731
- Oh, H.-K., Teo, A. K.-C., Ali, R. B., Lim, A., Ayi, T.-C., Yarosh, D. B. and Li, B. F.-L. (1996) *Biochemistry* **35**, 12259–12266
- Dolan, M. E., Moschel, R. C. and Pegg, A. E. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5368–5372
- Kanugula, S., Goodtzova, K., Edara, S. and Pegg, A. E. (1995) *Biochemistry* **34**, 7113–7119
- Crone, T. M., Goodtzova, K. and Pegg, A. E. (1996) *Mutation Res.* **363**, 15–25
- Crone, T. M. and Pegg, A. E. (1993) *Cancer Res.* **53**, 4750–4753
- Pegg, A. E., Boosalis, M., Samson, L., Moschel, R. C., Byers, T. L., Swenn, K. and Dolan, M. E. (1993) *Biochemistry* **32**, 11998–12006
- Potter, P. M., Lassiter, A. and Brent, T. P. (1993) *Methods Mol. Cell. Biol.* **4**, 139–146
- Laemmli, U. K. (1970) *Nature* (London) **227**, 680–685
- Blum, H., Beier, H. and Gross, H. J. (1987) *Electrophoresis* **8**, 93–99
- Pegg, A. E., Wiest, L., Mummert, C. and Dolan, M. E. (1991) *Carcinogenesis* **12**, 1671–1677
- Mol, C. D., Arvai, A. S., Slupphaug, G., Kavli, B., Alseth, I., Krokan, H. E. and Tainer, J. A. (1995) *Cell* **80**, 869–878
- Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E. and Tainer, J. A. (1996) *Nature* (London) **384**, 87–91
- Demple, B. (1995) *Curr. Biol.* **5**, 719–721
- Cheng, X. and Blumenthal, R. M. (1996) *Structure* **4**, 639–645
- Pearl, L. H. and Savva, R. (1995) *Trends Biochem. Sci.* **20**, 421–426
- Verdine, G. L. (1994) *Cell* **78**, 197–200
- Iyama, A., Sakumi, K., Nakabeppu, Y. and Sekiguchi, M. (1994) *Carcinogenesis* **15**, 627–633
- Mitra, S. and Kaina, B. (1993) *Prog. Nucleic Acid Res. Mol. Biol.* **44**, 109–142