Point mutations at multiple sites including highly conserved amino acids maintain activity, but render *O*⁶-alkylguanine–DNA alkyltransferase insensitive to *O*⁶-benzylguanine

Meng XU-WELLIVER 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 17033-0850, U.S.A.

The DNA repair protein, *O*⁶-alkylguanine–DNA alkyltransferase (AGT), is inactivated by reaction with the pseudosubstrate, *O*⁶-benzylguanine (BG). This inactivation sensitizes tumour cells to chemotherapeutic alkylating agents, and BG is aimed at enhancing cancer treatment in clinical trials. Point mutations in a 24 amino acid sequence likely to form the BG-binding pocket were identified using a screening method designed to identify BG-resistant mutants. It was found that alterations in 21 of these residues were able to render AGT resistant to BG. These included mutations at the highly conserved residues Lys¹⁶⁵, Leu¹⁶⁸ and Leu¹⁶⁹. The two positions at which changes led to the largest increase in resistance to BG were Gly¹⁵⁶ and Lys¹⁶⁵. Eleven mutants at Gly¹⁵⁶ were identified, with increases in resistance ranging from 190-fold (G156V) to 4400-fold (G156P). Two mutants at Lys¹⁶⁵ found in the screen (K165S and K165A)

showed 620-fold and 100-fold increases in resistance to BG. Two mutants at the Ser¹⁵⁹ position (S159I and S159V) were > 80-fold more resistant than wild-type AGT. Eleven active mutants at Leu¹⁶⁹ were also resistant to BG, but with lower increases (5–86-fold). Fourteen BG-resistant mutants were found for position Cys¹⁵⁰, with 3–26-fold increases in the amount of inhibitor needed to produce a 50 % loss of activity in a 30 min incubation. Six BG-resistant mutants at Asn¹⁵⁷ were found with increases of 4–13-fold. These results show that many changes can render human AGT resistant to BG without preventing the ability to protect tumour cells from therapeutic alkylating agents.

Key words: alkylating agents, chemotherapy, DNA repair, *O*⁶-methylguanine.

INTRODUCTION

*O*⁶-Alkylguanine–DNA alkyltransferase (AGT) is an important DNA repair protein which is able to protect cells from the mutagenic and toxic effects of monofunctional alkylating agents and some chloroethylating drugs [1–5]. It acts to repair *O*⁶-alkylguanine residues formed in DNA by these agents. Repair is carried out by the transfer of the alkyl group from the DNA to an active-site cysteine residue of the AGT protein itself resulting in the restoration of the DNA structure in a single step. The *S*-alkylcysteine formed at the active site of AGT is not converted back into cysteine; the protein, therefore, can act only once. The alkylated form of the AGT protein undergoes a conformational change, which leads to its rapid degradation, possibly by facilitating its ubiquitinylation [6–8].

AGT genes have been found in many living systems, including Eubacteria, Archaea and Eucarya. At least 28 DNA sequences coding for AGT genes representing 25 different species have now been reported. There are approx. 30 residues in these sequences that are either absolutely invariant or have only a small number of very conservative changes. These highly conserved amino acids are located in the carboxyl half of the protein, which includes the DNA-binding domain and the cysteine acceptor site [9–11] (Figure 1). Based on the alkyltransferase crystal structure [12–14], models of the binding of the DNA substrate to the protein, and studies of the reaction mechanism [9,15,16], it is likely that the *O*⁶-alkylguanine residue is flipped out of the DNA helix and bound in an active-site pocket containing the Cys-145 acceptor.

Human AGT is rapidly inactivated by reaction with *O*⁶-benzylguanine (BG) [17]. The mechanism of this inactivation is well understood and involves the recognition of BG as a substrate by the AGT. BG is bound at the active site, where the benzyl group is transferred, releasing guanine and forming *S*-benzylcysteine in the AGT protein [18]. In animal models, BG has been shown to enhance the cancer chemotherapeutic effects of chloroethylating agents, such as 1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU), and methylating agents, such as temozolomide [4,19,20]. BG is currently in Phase II trials for use as an anti-tumour agent [21,22].

However, BG is not effective as an inactivator of the AGT proteins from *Escherichia coli* and yeast [18,23,24]. This is probably due to the steric exclusion of BG from the active-site pocket [14], consisting of the residues shown in Figure 1. Mutation of two residues in the *E. coli* Ada-C AGT sequence (equivalent to the replacement of alanine with proline at position 140 and the substitution of the tryptophan at position 156 with glycine or alanine) increases the space available at the active site and renders this C-terminal domain of the *Ada* gene product capable of reacting with BG [25,26]. These results suggest that alterations in human AGT are possible which reduce the ability of the protein to react with BG, without eliminating activity towards alkylated DNA substrates. Several mutants have been produced that support this hypothesis. Single point mutations at Pro¹³⁸, Pro¹⁴⁰, Gly¹⁵⁶, Tyr¹⁵⁸ and Gly¹⁶⁰ have all been shown to be associated with resistance to BG [27–30]. A detailed study of mutants of the region Pro¹³⁸Val¹³⁹Pro¹⁴⁰, on the N-terminal side of the active site, has determined that any replacement of

Abbreviations used: AGT, *O*⁶-alkylguanine–DNA alkyltransferase; BCNU, 1,3-bis(2-chloroethyl)-1-nitrosourea; BG, *O*⁶-benzylguanine; DTT, dithiothreitol; LB, Luria–Bertani; MNNG, *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine; WT, wild-type; ED₅₀, the amount of inhibitor needed to produce a 50 % loss of activity in a 30 min incubation.

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

- NPVP	LI	PC	HR	VV	CS	SG	AV	GN	VS	G	---	G	---	L	AV	K	EW	L	A	H	G	H	Human (137-174)
- NPVP	LI	PC	HR	VV	CS	SG	AV	GN	VS	G	---	G	---	G	L	AV	K	EW	L	A	H	G	Rabbit
- NPVP	LI	PC	HR	VV	RS	SD	GA	IG	VS	G	---	G	---	G	Q	AV	K	EW	L	A	H	G	Mouse
- NPVP	LI	PC	HR	VV	RS	SD	GA	IG	VS	G	---	G	---	G	Q	TV	K	EW	L	A	H	G	Rat
- NPVP	LI	PC	HR	VV	CS	NG	S	IG	VS	G	---	G	---	G	Q	AV	K	EW	L	A	H	G	Chinese Hamster
- NELA	LI	PC	HR	VV	SQ	NG	A	---	VS	G	---	G	---	A	AL	K	Q	L	L	A	D	K	Drosophila
- NNLA	LI	PC	HR	VV	SN	R	K	L	T	G	---	S	---	C	K	L	K	E	L	L	N	N	S. cerevisiae
- NPVL	LI	PC	HR	VV	G	K	N	G	S	L	T	G	---	F	E	M	A	T	L	L	D	L	B. subtilis AdA8
- NDLP	LI	PC	HR	VV	G	K	N	S	A	L	T	G	---	S	K	T	E	I	A	L	L	I	B. subtilis Dac1
- NKLA	LI	PC	HR	VV	R	G	D	G	S	L	S	G	---	V	S	R	A	A	L	L	R	R	E. coli AdA8
- NPIS	LI	PC	HR	VV	G	R	N	G	T	M	T	G	---	G	V	Q	R	K	E	L	L	R	E. coli Ogt
- NPIS	LI	PC	HR	VV	G	K	D	R	T	L	T	G	---	G	L	E	A	K	R	L	L	Q	H. influenzae
- NPIS	LI	PC	HR	VV	R	K	N	G	T	L	G	---	G	I	E	V	K	K	L	L	E	F	H. pylori
- NPIS	LI	PC	HR	VV	G	A	S	G	Q	L	I	G	---	G	L	N	R	K	L	L	L	E	M. leprae
- NPIS	LI	PC	HR	VV	G	A	S	G	K	L	T	G	---	G	I	N	R	K	R	L	L	E	M. tuberculosis
- NPIS	LI	PC	HR	VV	G	K	D	G	S	L	T	G	---	G	L	E	R	K	A	L	L	R	P. algeriense sp. s9
- NKLA	LI	PC	HR	VV	R	D	G	A	L	S	G	---	G	V	R	R	E	A	L	L	K	R	S. typhimurium AdA
- NPIS	LI	PC	HR	VV	G	R	N	G	T	L	T	G	---	G	V	Q	R	K	E	L	L	R	S. typhimurium Ogt
- NPIS	LI	PC	HR	VV	G	R	N	G	T	L	T	G	---	G	V	Q	R	K	E	L	L	R	S. muenster
- NPIS	LI	PC	HR	VV	G	A	D	G	S	L	T	G	---	G	L	E	R	K	T	R	L	S	S. coelicolor
- NPIS	LI	PC	HR	VV	S	A	R	A	S	I	G	G	---	G	S	---	A	L	K	H	F	L	T. pallidum
- NPIS	LI	PC	HR	VV	S	---	---	---	---	---	---	---	---	G	E	K	L	Y	K	---	A	I	C. trachomatis
- NPIS	LI	PC	HR	VV	S	K	R	D	---	L	G	G	---	G	---	E	I	X	S	E	L	L	A. aeolicus
- NPIS	LI	PC	HR	VV	S	K	R	D	---	L	G	G	---	G	---	E	I	X	S	E	L	L	P. sp. K001
- NPIS	LI	PC	HR	VV	A	K	N	S	---	L	G	G	---	G	---	L	D	K	K	E	L	L	M. jannaschii
- NPIS	LI	PC	HR	VV	R	S	D	G	N	A	G	G	---	G	---	V	K	L	K	R	L	L	M. thermotrophicum
- NPIS	LI	PC	HR	VV	G	K	N	N	---	P	W	L	---	K	---	P	S	Y	K	F	L	L	P. horikoshii
- NPIS	LI	PC	HR	VV	G	K	K	E	---	I	G	G	---	G	---	K	S	L	K	R	L	L	A. fulgidus

Figure 1 AGT sequence surrounding the active site

The residues corresponding to the human AGT sequence from Asn¹³⁷ to His¹⁷⁴, which contains the cysteine acceptor site at residue 145, are shown for the 28 known AGT sequences. Highly conserved consensus residues and the residues with conservative alterations at these consensus positions are shaded in black.

Pro¹⁴⁰ renders AGT resistant to BG, although insertion of the basic residues, such as lysine and arginine, is most effective [31]. Human AGT mutants with multiple changes, some involving these sites, have also been reported to be strongly BG-resistant [32,33].

In the present study, we have carried out a detailed analysis of the residues from 150 to 173 in human AGT. This region was chosen because, along with previously examined residues 138–140 [31], it makes up a large part of the binding pocket for the substrate base. A plasmid-expression library was constructed in which each of these individual residues in the 150–173 region were randomly mutated, and the AGT was expressed in *E. coli*. The bacteria containing active but BG-resistant AGT mutants were then selected by exposure to BG and *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG). This screening revealed that many sites could be mutated to yield BG-resistant AGTs. So many mutants were obtained in this screen that it was necessary to use libraries with more limited alterations in order to get a more complete collection of such BG-resistant mutants at any particular site. Such detailed studies were carried out with mutants formed at Cys¹⁵⁰, Gly¹⁵⁶, Asn¹⁵⁷, Leu¹⁶⁸ and Leu¹⁶⁹. These results, and those in the accompanying paper in which mutants at Lys¹⁶⁵ were examined [34], show that numerous amino acid substitutions affect the ability of human AGT to accept BG at the active site, without impairing the ability to repair alkylated DNA. The facile appearance of BG-resistant mutants may therefore be a significant problem in the clinical development of this drug.

EXPERIMENTAL

Materials

All oligodeoxyribonucleotides were purchased from Life Technologies, *E. coli* strain TRG8 was derived from strain GWR109 [35], which was kindly provided by Dr L. Samson, Department of Molecular and Cellular Toxicology, Harvard School of Public Health, Boston, MA, U.S.A., XL1-Blue bacterial strain and competent cells were purchased from Stratagene (La Jolla, CA, U.S.A.). The DNA isolation kits were obtained from Qiagen (Chatsworth, CA, U.S.A.). Restriction enzymes and *Pfu* polymerase were purchased from New England Biolabs (Beverly, MA, U.S.A.) and Stratagene. T4-ligase with the 'ligation express kit' was purchased from Clontech Laboratories. Ampicillin, kanamycin, isopropyl β -D-thiogalactoside, haemocyanin,

MNNG and most other biochemical reagents were purchased from Sigma. BG was generously provided by Dr R. C. Moschel (ABL-Basic Research Program, National Cancer Institute-Frederick Cancer Research and Development Center, Frederick, MD, U.S.A.). Plasmids pUC-AGTm [31] and pUC-inAGT2 [30] were made as previously described.

Construction and selection of random libraries containing individual mutations from codon Cys¹⁵⁰ to Gly¹⁷³

The random library covering codons 150–173 was designed to contain only single amino acid changes. The DNA insert was created by PCR covering the region from nt 420–540, including the *Dra*III site (at 436) and the *Age*I site (at 521). Twenty-four individual PCR products were synthesized using pUC-AGTm as a template, and each contained amino acid sequences replaced with random nucleotide sequences using the primers listed in Table 1. Eleven PCR products, including C150X to G160X, used the 5'-GGCTTCCCCAACCGGTGGCC-3' (*Age*I site in italics) as the common antisense primer, while the other 13 PCR products, including G161X to G173X, used the 5'-CCGTGC-CACAGAGTGG-3' (*Dra*III site in italics) as the common sense primer. The PCR reactions were carried out using the following conditions: initial denaturation for 2 min at 92 °C; 30 cycles of denaturation (1 min at 92 °C); annealing (1 min at 54 °C); extension (1 min at 72 °C); followed by a final extension at 72 °C for 5 min. The 24 PCR products were individually gel purified using the Gel Extraction Kit (Qiagen), quantified by UV spectrophotometry, combined in equal amounts, digested with *Dra*III and *Age*I enzymes, and ligated into pUC-inAGT2 digested with the same enzymes. The random libraries in which one or two amino acids were substituted with random nucleotide sequences (C150X, S159X, L168X, L169X and G156X/N157X) were made by inserting two (for G156X/N157X) or one (for the rest of the libraries) PCR products corresponding to the above changes into the pUC-inAGT2 plasmid digested with the same enzymes. The ligated products were transformed into XL1-Blue cells; an aliquot of the transformation mixture was plated on Luria-Bertani (LB) plates, supplemented with 50 μ g/ml ampicillin to determine the number of plasmid-containing bacteria (3.6×10^6 for the C150-G173 library, 1.0×10^6 for the C150X, S159X, L168X, L169X libraries and 1.2×10^6 for the G156X/N157X library), and the remainder were amplified by growing overnight. The plasmid DNA was isolated from the overnight cultures,

Mismatches are underlined, restriction sites are shown in italics.

Mutation	Primer
C150X	5'-CCGTGCCACAGAGTGGTCNN <u>S</u> AGCAGCGGAGCCG-3'
S151X	5'-CCGTGCCACAGAGTGGTCTGCNN <u>S</u> AGCGGAGCCGTGGGC-3'
S152X	5'-CCGTGCCACAGAGTGGTCTGCAGCNN <u>S</u> GGAGCCGTGGGC-3'
G153X	5'-CCGTGCCACAGAGTGGTCTGCAGCAGCNN <u>S</u> GCCGTGGGCAACTACTCC-3'
A154X	5'-CCGTGCCACAGAGTGGTCTGCAGCAGCGGANN <u>S</u> GTGGGCAACTACTCC-3'
V155X	5'-CCGTGCCACAGAGTGGTCTGCAGCAGCGGAGCCNN <u>S</u> GGCAACTACTCCG-3'
G156X	5'-CCGTGCCACAGAGTGGTCTGCAGCAGCGGAGCCGTGNN <u>S</u> AACTACTCCGAGG-3'
N157X	5'-CCGTGCCACAGAGTGGTCTGCAGCAGCGGAGCCGTGGGCNN <u>S</u> TACTCCGGAGGG-3'
Y158X	5'-CCGTGCCACAGAGTGGTCTGCAGCAGCGGAGCCGTGGGCAACNN <u>S</u> TCCGGAGGGCTAGC-3'
S159X	5'-CCGTGCCACAGAGTGGTCTGCAGCAGCGGAGCCGTGGGCAACTACNN <u>S</u> GGAGGGCTAGCC-3'
G160X	5'-CCGTGCCACAGAGTGGTCTGCAGCAGCGGAGCCGTGGGCAACTACTCCNN <u>S</u> GGGCTAGCCGTGAAGG-3'
G161X	5'-GGCTTCCCCA <u>ACCGGT</u> GGCCTTCATGGGCCAGAAGCCATTCTTCACGGCTAGSN <u>N</u> TCCGGAGTAGTTGC-3'
L162X	5'-GGCTTCCCCA <u>ACCGGT</u> GGCCTTCATGGGCCAGAAGCCATTCTTCACGGCSNNCCCTCCGGAGTAGTTGC-3'
A163X	5'-GGCTTCCCCA <u>ACCGGT</u> GGCCTTCATGGGCCAGAAGCCATTCTTCACSN <u>N</u> TAGCCCTCCGGAGTAGTTGC-3'
V164X	5'-GGCTTCCCCA <u>ACCGGT</u> GGCCTTCATGGGCCAGAAGCCATTCTTSNN <u>N</u> GGCTAGCCCTCCGG-3'
K165X	5'-GGCTTCCCCA <u>ACCGGT</u> GGCCTTCATGGGCCAGAAGCCATTCSNN <u>N</u> CACGGCTAGCCCTCC-3'
E166X	5'-GGCTTCCCCA <u>ACCGGT</u> GGCCTTCATGGGCCAGAAGCCASN <u>N</u> CTTCACGGCTAGCC-3'
W167X	5'-GGCTTCCCCA <u>ACCGGT</u> GGCCTTCATGGGCCAGAAGSN <u>N</u> TTCTTCACGGCTAGC-3'
L168X	5'-GGCTTCCCCA <u>ACCGGT</u> GGCCTTCATGGGCCAGSN <u>N</u> CCATTCTTCACGG-3'
L169X	5'-GGCTTCCCCA <u>ACCGGT</u> GGCCTTCATGGGCCNN <u>N</u> AAGCCATTCTTCACG-3'
A170X	5'-GGCTTCCCCA <u>ACCGGT</u> GGCCTTCATGSNNCAGAAGCCATTCC-3'
H171X	5'-GGCTTCCCCA <u>ACCGGT</u> GGCCTTCSNN <u>N</u> GGCCAGAAGCCATTCC-3'
E172X	5'-GGCTTCCCCA <u>ACCGGT</u> GGCCSN <u>N</u> TATGGGCCAGAAGCC-3'
G173X	5'-GGCTTCCCCA <u>ACCGGT</u> GSNN <u>N</u> TTATGGGCCAGAAGC-3'

© 2000 Biochemical Society

Table 2 Mutations in Cys¹⁵⁰–Gly¹⁷³ region of AGT imparting resistance to BG

All mutants found from the Cys¹⁵⁰–Gly¹⁷³ random library after the MNNG-BG selection are shown. The mutants with BG ED₅₀ values higher than 0.3 μ M are shown in bold, and the mutants with BG ED₅₀ values over 10 μ M are shown in bold and with an asterisk. †Highly conserved residues in the AGT sequence are shown in bold and underlined.

AGT residue	Mutants found
Cys-150	L E F Y
Ser-151	A C R D G
Ser-152	R D K Q V W
Gly-153	C C
Ala-154	F F L P
Val-155	L P T Y Y E
Gly-156	H* H* S* W* R* R* T* L*
Asn-157	R E W G G G*
Tyr-158	
Ser-159	F F L V I*
Gly-160	C Q K L L L L P P P R
Gly-161	S
Leu-162	T F
Ala-163	D F F L
Val-164	W W
Lys-165	A* S* S*
Glu-166	K N S S
Trp-167	A H K
Leu-168	
Leu-169	A S T W E
Ala-170	G
His-171	S L
Glu-172	
Gly-173	S C I I

and contained at least 3.6×10^6 transformants, which provided a sufficient source to include all 20 amino acids for each residue.

After selection of this library with 40 μ g/ml MNNG and 50 μ M BG, the AGT cDNA inserts in the plasmids isolated from 107 of the surviving colonies were analysed. Two of these clones did not produce AGT [an amber mutation (E166Stop) and religated vector sequence with no functional AGT gene]. Of the other 105 clones, 70 different AGT mutants (Table 2) and WT AGT were found. AGT activity was measurable using < 200 μ g of crude protein extract for every mutant. The sensitivity of the mutants to BG was determined as ED₅₀ values. Because of the large number of candidate mutants, it was not feasible to purify the AGT protein. Therefore, the assays were carried out using crude protein extracts. In order to check that the other proteins present in these extracts did not influence the results, studies were conducted using purified WT AGT (ED₅₀ 0.1 μ M) and two known mutants with moderate resistance to BG, G160R (ED₅₀ 4.7 μ M) and P140A (ED₅₀ 4 μ M). Extracts from the control TRG8 *E. coli* strain that lacks endogenous AGT were added to these proteins and assays were then carried out. Homogeneous preparations of these AGT proteins were tested in the standard assay system with the addition of no extract, 6 μ g of *E. coli* extract, or 200 μ g of *E. coli* extract. The ED₅₀ values were not altered significantly (by less than a factor of two, which is probably within experimental error) by the presence of the *E. coli* extract.

There were 17 clones expressing WT AGT sequences and 8 clones (6 mutants) that had ED₅₀ values for inactivation by BG of 0.1 μ M, which are identical with WT AGT. These are clearly false positives. There were also 12 clones (12 mutants), which had

Table 3 BG-resistant mutants found after screening the C150–G173 AGT library with BG and MNNG

Position	Mutant	μ g of protein*	ED ₅₀ (μ M)
Wild-type		1.5	0.1
Cys-150	Phe	2.6	2.6
	Tyr	1.1	2.6
	Glu	3.9	0.9
	Leu	1.4	0.7
Ser-151	Gly	112.0	3.7
	Asp	5.8	0.5
Ser-152	Trp	5.6	0.5
	Asp	1.1	0.3
	Lys	180.0	0.3
	Gln	1.3	0.3
	Val	2.6	0.3
Gly-153	Cys	2.4	0.3
Ala-154	Pro	1.2	0.5
	Phe	0.9	0.3
	Leu	1.4	0.3
Val-155	Glu	2.8	1.8
	Thr	1.0	0.3
	Tyr	6.5	0.3
Gly-156	Leu	39.0	110
	Thr	25.0	90
	Arg	16.0	86
	Trp	†	80
	Ser	8.2	62
	His	10	42
Asn-157	Gly	1.5	1.3
	Trp	2.1	0.5
	Glu	2.0	0.4
Ser-159	Ile	3.5	11
	Val	3.0	8.5
	Leu	4.8	0.4
Gly-160	Arg	2.5	4.6
	Pro	15	4.2
	Leu	1.7	4.1
	Lys	18	2.9
	Gln	6.5	2.5
	Cys	5.0	0.8
Gly-161	Ser	3.7	1.0
Leu-162	Phe	1.2	0.3
Val-164	Trp	2.4	0.5
Lys-165	Ser	100	62
	Ala	200	10
Glu-166	Asn	1.2	0.3
	Ser	1.4	0.3
Trp-167	Lys	0.9	0.3
Leu-169	Glu	47	8.6
	Trp	23	6.7
	Thr	23	3.0
	Ser	30	1.0
	Ala	83	0.7
Ala-170	Gly	200	3.0
Gly-173	Cys	1.5	0.3
	Ile	1.0	0.3

* The amount of crude protein extract needed to get 80% removal of O⁶-methylguanine in the standard assay.

† The ED₅₀ value for this mutant was already known [28].

ED₅₀ values of 0.2 μ M, which were not placed in the resistant category, as a value of $3 \times$ WT ED₅₀ was used as a cut-off for resistance. With this definition, there were 68 clones representing 52 separate BG-resistant mutants found (Table 3). Thus 52 of the

Table 4 BG-resistant AGT mutants at glycine-156 found after screening of G156X/N157X mutant library

Amino acid	μg of protein*	ED ₅₀ (μM)
Pro	24	450
Leu	39	110
Thr	25	90
Tyr	24	90
Arg	16	86
Trp	†	80†
Ser	12	62
Ala	†	60†
Cys	12	49
His	10	42
Val	12	19

* The amount of crude protein extract needed to get 80% removal of *O*⁶-methylguanine in the standard assay.

† Previously described [28].

70 mutants found (74%) were definitively BG-resistant. These mutants were distributed widely, including 19 different amino acid positions in the 24 residues between Cys¹⁵⁰ and Gly¹⁷³. Thus mutations at most of the residues in this region may lead to significantly increased ED₅₀ values for BG.

No substitutions were found for Tyr¹⁵⁸, Leu¹⁶⁸ and Glu¹⁷², and these residues are highly conserved among all alkyltransferase proteins (Figure 1). Some substitutions were found for other well-conserved residues, such as Lys¹⁶⁵ and Leu¹⁶⁹. This indicates that changes can be made at these positions without losing all AGT activity, but these mutants were less active in comparison with WT AGT, in that a higher amount of crude bacterial extract was needed to repair the same amount of DNA substrate (Table 3).

Mutations at residues Gly¹⁵⁶ and Lys¹⁶⁵ resulted in the largest increase in resistance to BG with at least 100-fold higher ED₅₀ values. These two positions, as well as Asn¹⁵⁷, Cys¹⁵⁰, Leu¹⁶⁹ and Leu¹⁶⁸ (see below), were studied in more detail. Since so many alterations render AGT resistant to BG, this analysis was conducted using libraries in which a more limited number of residues were altered.

Mutants at Gly¹⁵⁶

Five alterations at Gly¹⁵⁶ found in the general screen rendered AGT highly resistant to BG (Table 3). These results are consistent with previous studies in which two mutations (G156A and G156W) at Gly¹⁵⁶, including one found in the current screen (G156W), were reported to cause the AGT protein to become resistant to BG [28]. In order to examine this position in more detail, a library was constructed containing random mutations for each of the two residues, Gly¹⁵⁶ and Asn¹⁵⁷. Mutants from this library of a pool of 1.2×10^6 members were selected with BG and MNNG. The AGT cDNA inserts in the plasmids isolated from 22 of the surviving colonies were analysed. Sixteen clones (73%) contained mutations at Gly¹⁵⁶, and the remainder contained mutations at Asn¹⁵⁷. The mutants at position 156 are shown in Table 4.

A total of 11 mutants at Gly¹⁵⁶ were identified, and all were highly BG-resistant with at least 190-fold higher ED₅₀ values than WT AGT. The replacement with proline produced the biggest increase in ED₅₀ with a > 4000-fold increase. However,

Table 5 AGT mutants at asparagine-157 found after screening of G156X/N157X mutant library

Amino acid	μg of protein*	ED ₅₀ (μM)
Gly	1.5	1.3
Thr	7.6	1.3
Val	10	0.8
Tyr	5	0.7
Trp	2.1	0.5
Glu	2.0	0.4
Arg	2.8	0.1

* The amount of crude protein extract needed to get 80% removal of *O*⁶-methylguanine in the standard assay.

substitutions found included hydrophobic and hydrophilic, including positively charged, residues. No acidic residues were found. Glutamine and asparagine residues were also absent. It is possible that these changes are not compatible with AGT activity. In all cases, the amount of bacterial protein needed to get 80% repair of the methylated DNA substrate used was increased 10–30-fold over that needed with extracts expressing WT AGT (Table 4).

Mutants at Asn¹⁵⁷

In the general screen three mutants, N157G, N157E and N157W, were found to impart increases in resistance to BG with ED₅₀ values increased by up to 13-fold. Eight mutants at Asn¹⁵⁷ were identified in the screening of the joint G156X/N157X library. These provided another 3 alterations resulting in resistance. Most of the mutants at position Asn¹⁵⁷ were almost as active as WT AGT in repairing methylated DNA (Table 5). Six of these mutants were BG-resistant; however, none resulted in reductions in the sensitivity to BG as large as those found with alterations at position 156. The increases in ED₅₀ values from N157X ranged from 4 to 13-fold (Table 5). This may be relevant to the observations that the mouse AGT is less sensitive (by approx. 5-fold) to BG than the human or the rat AGT, even though these proteins are very similar over the entire coding sequence [24,37]. Position 157 is one of the few places in the region studied which varies in these species in a manner related to the sensitivity to BG, being asparagine in the rat and human AGTs, but histidine in the mouse AGT.

Mutants at Cys¹⁵⁰

Four BG-resistant mutants at position Cys¹⁵⁰ were found in the initial screen. A further ten mutants (plus WT AGT) were isolated by selection with BG and MNNG from a C150X library (Table 6). All of the mutants were resistant to BG with increases in ED₅₀ values from 3- to 26-fold. The largest increases were produced with bulky amino acid replacement such as tyrosine, phenylalanine and tryptophan, whereas alterations to glycine, serine and threonine produced the least increase in BG-resistance. The alterations at Cys¹⁵⁰ produced only small increases in the amount of protein needed to repair 80% of the methylated DNA substrate used in the standard assay, suggesting that there was little, if any, alteration in the stability of the protein or of the ability to repair alkylated DNA.

Table 6 BG-resistant AGT mutants at cysteine-150

Amino acid	μg of protein*	ED_{50} (μM)
Tyr	1.1	2.6
Phe	2.6	2.6
Trp	6.2	2.6
Met	6.8	1.7
Pro	3.8	1.7
Val	1.1	1.5
Glu	3.9	0.9
Leu	1.1	0.8
Asn	3.2	0.7
Gln	0.8	0.6
Lys	1.1	0.6
Thr	3.5	0.5
Gly	0.8	0.4
Ser	1.1	0.3

* The amount of crude protein extract needed to get 80% removal of O^6 -methylguanine in the standard assay.

Table 7 AGT mutants at leucine-169 found after screening of L169X mutant library

Abbreviation: ND, not determined.

Amino acid	μg of protein*	ED_{50} (μM)
Glu	47	8.6
Trp	23	6.7
Thr	23	3.0
Lys	200† (48%)	2.9
Asn	10	2.1
Arg	20	2.0
Ser	30	1.0
Tyr	5	0.8
Ala	83	0.7
Phe	5	0.7
Ile	3	0.5
Asp‡	inactive	ND
Cys‡	inactive	ND
Gly‡	200† (20%)	ND

* The amount of crude protein extract needed to get 80% removal of O^6 -methylguanine in the standard assay.

† The maximum of protein that can be used in the assay is 0.2 mg. This gave less than 80% removal of O^6 -methylguanine and the amount removed is shown in parentheses.

‡ The AGT activity level was too low for accurate assays with crude extracts prepared under the standard conditions. Therefore, the bacterial cultures were grown in the presence of 0.15 mM isopropyl β -D-thio galactoside to increase AGT expression. These extracts were then used for the assay. The expression of AGT obtained in the presence of isopropyl β -D-thio galactoside with these conditions is increased about 4-fold.

Mutants at Leu¹⁶⁹

Five BG-resistant mutants at position Leu¹⁶⁹ were found in the initial screen of the library (Table 3). A further 5 mutants were isolated from a L169X library by selection with BG and MNNG, and 4 additional mutants were obtained by taking clones from the library without selections. All of these mutants assayed were also BG-resistant (Table 7). These 11 mutants had ED_{50} values increased 5–86-fold. Most of the mutants found at this position were substantially less active than WT AGT. Three of the mutants isolated (L169D, L169C and L169G) had insufficient activity for assay of an ED_{50} value for BG inactivation. Since these mutants were all isolated after selection with MNNG and BG, it is likely that they did have activity *in vivo* and can protect

Table 8 AGT mutants at leucine-168 found after screening of L168X mutant library

Abbreviation: N.D., not determined.

Amino acid	μg of protein*	ED_{50} (μM)
Cys	7	0.5
Phe	20	0.5
Ile	8.6	0.4
Met	5.5	0.4
Ser	200† (< 10%)	ND
Leu (WT)	—	—

* The amount of crude protein extract needed to get 80% removal of O^6 -methylguanine in the standard assay.

† Even when extracts were grown in the presence of isopropyl β -D-thiogalactoside there was insufficient AGT activity for determination of an ED_{50} value. The maximum protein that can be used in the assay is 0.2 mg. This gave less than 80% removal of O^6 -methylguanine. The amount removed is shown in parentheses.

the host *E. coli*. However, this activity is too low or too unstable to be measured in crude cell extracts. The only 3 mutants with activities towards methylated DNA close to WT were replacements with isoleucine, phenylalanine and tyrosine. These alterations were among those having the lowest increases in BG-resistance, but there was still a 5–8-fold increase in the ED_{50} value.

Mutants at Leu¹⁶⁸

There are 3 highly conserved residues (Tyr¹⁵⁸, Leu¹⁶⁸ and Glu¹⁷²) in the Cys¹⁵⁰–Gly¹⁷³ region in which no mutants were found in the screen of this library (Figure 1). It is likely that only a few permissible substitutions that lead to active AGT exist for these residues, and the even more limited BG-resistant subset may not have been found due to the limited numbers of clones screened. Positions Tyr¹⁵⁸ and Glu¹⁷² have been investigated previously by site-directed mutagenesis. Glu¹⁷² appears to be important for the structure/function of AGT, and even replacement with glutamine greatly reduces AGT activity and stability [38,39]. The only replacements for Tyr¹⁵⁸ that give active AGT are phenylalanine (which is not BG-resistant) [40] and histidine (which is highly BG-resistant) [30].

Position Leu¹⁶⁸ had not been investigated previously so a library in which only the codon 168 was mutated was screened using MNNG alone to identify any active AGT mutants. Twenty-one clones were isolated. These included 5 WT sequences. The mutants found were changes to cysteine (6 clones), methionine (4 clones), isoleucine (3 clones), serine (2 clones) and phenylalanine (1 clone). The L168S mutant gave a very low activity in the *E. coli* extracts (Table 7). These results suggest that a hydrophobic amino acid at position 168 is needed to give active AGT. The 4 mutants that could be assayed were resistant to BG but had only small (4–5-fold) increases in ED_{50} values (Table 8).

DISCUSSION

The procedure used for the selection of active but BG-resistant mutants from the library of plasmids expressing AGT altered in the Cys¹⁵⁰–Gly¹⁷³ region was successful in that most of the surviving clones contained AGT that was both active and BG-resistant. However, there were a number of false positives. All but two of these had either WT or AGT mutants with ED_{50} values of 0.2 μM or less, and it is likely that they were found because the stringency of the selection was not high enough to

exclude them. Higher stringency could be achieved by increasing the concentration of BG used in selection but this may also eliminate the selection of mutants that are only moderately resistant to BG.

Although all of the studies reported in this paper were carried out with crude cell extracts containing the mutant AGTs, the results are highly consistent with the limited number of previous studies in which purified AGT protein was used or the AGT was expressed in mammalian cells. Thus mutants G165A, G165W, G160R and K165A have all been shown to have similar ED_{50} values when assayed in crude extracts or with purified protein, and to protect mammalian cells from alkylating agents even in the presence of BG [29,34,41,43]. Many sites were found to lead to BG-resistance when mutated. For this reason it would be necessary to isolate and sequence a very large number of clones in order to have a complete representation of mutants with BG-resistant AGT activity. It should be noted that two of the positions (Leu¹⁶⁸ and Tyr¹⁵⁸), not found in general screening to be sites at which BG-resistance can occur, are actually positions in which BG-resistant mutants have been found. As shown in Table 8, the conservative alterations at position Leu¹⁶⁸ that yielded active AGT protein did impart a small degree of resistance to BG. The alteration of Tyr¹⁵⁸ to histidine (which is one of only two alterations tolerated at this position) produces a huge increase in ED_{50} for BG (from 0.1 μ M to 620 μ M) [30]. The reason that these mutants were not found is probably that they represent only a small percentage of the total mutants that provide protection under the screening conditions used.

Some of the AGT mutants found in the screen appeared to have a reduced activity towards methylated DNA, as indicated by the amount of bacterial extract needed for the assay, whereas others were close to WT in activity. In general, those mutants that were very highly resistant to BG had low activity and those that had only moderate resistance were more active. This may be a consequence of the screening procedure, where the residual activity in the presence of BG forms the major parameter for survival. This residual activity would be related to both the activity of AGT mutant (which is discussed below) and its sensitivity to BG. This is well illustrated by the screening of the complete library and the mixed G156X/N157X library. Of 38 mutations at the two sites, 24 were at Gly¹⁵⁶ and 14 at Asn¹⁵⁷. Even though mutants at Gly¹⁵⁶ were 100-fold more resistant to BG than mutants at Asn¹⁵⁷, they were also much less active when crude cell extracts were tested. Previous experiments with AGT mutants altered in the DNA-binding domain have shown that mutants in which the rate constant is reduced substantially are still able to provide protection from MNNG [11]. WT AGT repairs alkylated DNA very rapidly and this rate may be in excess of that needed to promote cell survival.

It should be noted that a reduction in AGT activity in crude bacterial extracts might be due to a number of factors in addition to a reduced rate of methyl transfer. Other possibilities include: incorrect folding of part of the AGT protein leading to an inactive fraction of the population; an increased rate of degradation of the AGT protein in *E. coli* leading to a reduced level of protein; and a decreased stability of the protein leading to partial denaturation during preparation of the extracts. All of these factors may contribute to the need for greater amounts of crude bacterial protein in the assays. Previous studies have shown that mutations of residues involved in AGT structural integrity do cause an incorrect folding and/or rapid degradation of the protein [39,44].

The results of the library screening indicate that mutations of many residues (19 out of 24) in the Cys¹⁵⁰–Gly¹⁷³ region may alter the sensitivity of AGT to BG. The BG-resistance of the Y158H

mutant and the L168 mutants described above raises this number to 21 out of 24. The three-dimensional structure of human AGT [14] shows that many of these 24 residues are located near the active site of the protein and alterations would tend to exclude BG by steric effects and/or by disrupting the hydrophobic region that accommodates the BG. Virtually all alterations of Gly¹⁵⁶ or Lys¹⁶⁵ would be expected to cause minor structural distortions of the binding pocket that exclude BG.

The abundance of BG-resistant mutants occurring in this region raises concern in using BG to enhance chemotherapy by alkylating agents. Since these agents are themselves mutagenic, the possibility that BG-resistant forms of AGT may arise is a serious issue. Even though many of the BG-resistant mutants identified in our screen may have somewhat impaired AGT activity (as discussed above), the ability of the *E. coli* clones expressing these mutants to survive treatment with a methylating agent, and the ability of the G156A [41,42,45,46], P140K [43] and K165A [34] mutant AGTs to protect mammalian cells from BCNU, strongly suggest that the mutants would be effective in counteracting the actions of chemotherapeutic agents such as temozolomide and BCNU, even in the presence of BG. Furthermore, even modestly resistant mutants of AGT such as P140A and G160R, which have ED_{50} values increased by approx. 40-fold, have been shown to provide a very high level of survival in mammalian cells exposed to BCNU and BG at levels that far exceed the maximal concentrations of BG and its active metabolite 8-oxo-BG reached in clinical trials [22,47]. Therefore most, if not all, of the BG-resistant mutants from the Cys¹⁵⁰–Gly¹⁷³ region described here, and other mutants found in the Pro¹³⁸–Pro¹⁴⁰ region [31], are likely to impart resistance to BG therapy if produced in treated tumours. Thus at least 23 codons in the AGT gene are potential sites for the generation of resistance as a result of single amino acid changes. The recently published crystal structure of human AGT reveals that other regions in the AGT protein, such as part of the H4 helix (residues 129–136), are also close to the BG-binding pocket [14] and may also be targets for such mutants rendering resistance to BG. The development of alternative AGT inhibitors able to inactivate these resistant mutants should therefore be given a high priority.

This work was supported in part by grants CA-18137, CA-57725 and CA-71976 from the National Cancer Institute.

REFERENCES

- Lindahl, T., Sedgwick, B., Sekiguchi, M. and Nakabeppu, Y. (1988) *Annu. Rev. Biochem.* **57**, 133–157
- Samson, L. (1992) *Mol. Microbiol.* **6**, 825–831
- Pegg, A. E. and Byers, T. L. (1992) *FASEB J.* **6**, 2302–2310
- Pegg, A. E., Dolan, M. E. and Moschel, R. C. (1995) *Progr. Nucleic Acid Res. Mol. Biol.* **51**, 167–223
- Mitra, S. and Kaina, B. (1993) *Progr. Nucleic Acid Res. Mol. Biol.* **44**, 109–142
- 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
- Kanugula, S., Goodtzova, K. and Pegg, A. E. (1998) *Biochem. J.* **329**, 545–550
- 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
- Kanugula, S., Goodtzova, K., Edara, S. and Pegg, A. E. (1995) *Biochemistry* **34**, 7113–7119
- Goodtzova, K., Kanugula, S., Edara, S. and Pegg, A. E. (1998) *Biochemistry* **37**, 12489–12495
- Moore, M. H., Gulbus, J. M., Dodson, E. J., Demple, B. and Moody, P. C. E. (1994) *EMBO J.* **13**, 1495–1501
- Hashimoto, H., Inoue, T., Nishioka, M., Fujiwara, S., Takagi, M., Imanaka, T. and Kai, Y. (1999) *J. Mol. Biol.* **292**, 707–716

- 14 Wibley, J. E. A., Pegg, A. E. and Moody, P. C. E. (2000) *Nucleic Acids Res.* **28**, 393–401
- 15 Vora, R., Pegg, A. E. and Ealick, S. E. (1998) *Proteins* **32**, 3–6
- 16 Spratt, T. E., Wu, J. D., Levy, D. E., Kanugula, S. and Pegg, A. E. (1999) *Biochemistry* **38**, 6801–6806
- 17 Dolan, M. E., Moschel, R. C. and Pegg, A. E. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5368–5372
- 18 Pegg, A. E., Boosalis, M., Samson, L., Moschel, R. C., Byers, T. L., Swenn, K. and Dolan, M. E. (1993) *Biochemistry* **32**, 11998–12006
- 19 Gerson, S. L., Liu, L., Phillips, W. P., Zaidi, N. H., Heist, A., Markowitz, S. and Wilson, J. K. V. (1994) *Proc. Am. Assn. Cancer Res.* **35**, 699–700
- 20 Dolan, M. E. and Pegg, A. E. (1997) *Clinical Cancer Res.* **3**, 837–847
- 21 Friedman, H. S., Kokkinakis, D. M., Pluda, J., Friedman, A. H., Cokgor, I., Haglund, M. M., Ashley, D. M., Rich, J., Dolan, M. E., Pegg, A. E. et al. (1998) *J. Clin. Oncol.* **16**, 3570–3575
- 22 Dolan, M. E., Roy, S. K., Fasanmade, A., Paras, P. R., Schilsky, R. L. and Ratain, M. J. (1998) *J. Clin. Oncol.* **16**, 1803–1810
- 23 Dolan, M. E., Pegg, A. E., Dumenco, L. L., Moschel, R. C. and Gerson, S. L. (1991) *Carcinogenesis* **12**, 2305–2310
- 24 Elder, R. H., Margison, G. P. and Rafferty, J. A. (1994) *Biochem. J.* **298**, 231–235
- 25 Crone, T. M., Kanugula, S. and Pegg, A. E. (1995) *Carcinogenesis* **16**, 1687–1692
- 26 Goodtzova, K., Kanugula, S., Edara, S., Pauly, G. T., Moschel, R. C. and Pegg, A. E. (1997) *J. Biol. Chem.* **272**, 8332–8339
- 27 Crone, T. M. and Pegg, A. E. (1993) *Cancer Res.* **53**, 4750–4753
- 28 Crone, T. M., Goodtzova, K., Edara, S. and Pegg, A. E. (1994) *Cancer Res.* **54**, 6221–6227
- 29 Edara, S., Kanugula, S., Goodtzova, K. and Pegg, A. E. (1996) *Cancer Res.* **56**, 5571–5575
- 30 Xu-Welliver, M., Leitao, J., Kanugula, S. and Pegg, A. E. (1999) *Cancer Res.* **59**, 1514–1519
- 31 Xu-Welliver, M., Kanugula, S. and Pegg, A. E. (1998) *Cancer Res.* **59**, 1936–1945
- 32 Encell, L. P., Coates, M. M. and Loeb, L. A. (1998) *Cancer Res.* **58**, 1013–1020
- 33 Christians, F. C., Dawson, B. J., Coates, M. M. and Loeb, L. A. (1997) *Cancer Res.* **57**, 2007–2012
- 34 Xu-Welliver, M., Kanugula, S., Loktionova, N. A., Crone, T. M. and Pegg, A. E. (2000) *Biochem. J.* **347**, 527–534
- 35 Rebeck, G. W. and Samson, L. (1991) *J. Bacteriol.* **173**, 2068–2076
- 36 Reidhaar-Olson, J. F., Bowie, J. U., Breyer, R. M., Hu, J. C., Knight, K. L., Lim, W. A., Mossing, M. C., Parsell, D. A., Shoemaker, K. R. and Sauer, R. T. (1991) *Methods Enzymol.* **208**, 564–586
- 37 Liu, L., Lee, K., Wasan, E. and Gerson, S. L. (1996) *Cancer Res.* **56**, 1880–1885
- 38 Rafferty, J. A., Tumulty, J., Skorvaga, M., Elder, R. H., Margison, G. P. and Douglas, K. T. (1994) *Biochem. Biophys. Res. Commun.* **199**, 285–291
- 39 Crone, T. M., Goodtzova, K. and Pegg, A. E. (1996) *Mutation Res.* **363**, 15–25
- 40 Edara, S., Goodtzova, K. and Pegg, A. E. (1995) *Carcinogenesis* **16**, 1637–1642
- 41 Davis, B. M., Reese, J. S., Koc, O. N., Lee, K., Schupp, J. E. and Gerson, S. L. (1997) *Cancer Res.* **57**, 5093–5099
- 42 Loktionova, N. A. and Pegg, A. E. (1996) *Cancer Res.* **56**, 1578–1583
- 43 Loktionova, N. A., Xu-Welliver, M., Crone, T., Kanugula, S. and Pegg, A. E. (1999) *Biochem. Pharmacol.* **58**, 237–244
- 44 Pieper, R. O., Morgan, S. E. and Kelley, M. R. (1994) *Carcinogenesis* **15**, 1895–1902
- 45 Koc, O. N., Reese, J. S., Davis, B. M., Liu, L., Majczenko, K. J. and Gerson, S. L. (1999) *Hum. Gene Ther.* **10**, 1021–1030
- 46 Reese, J. S., Koc, O. N., Lee, K. M., Liu, L., Allay, J. A., Phillips, W. P. and Gerson, S. L. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14088–14093
- 47 Stefan, T. L., Ingalls, S. T., Minkler, P. E., Willson, J. K. V., Gerson, S. L., Spiro, T. P. and Hoppel, C. L. (1997) *J. Chromatogr.* **704**, 289–298

Received 9 November 1999/18 January 2000; accepted 8 February 2000