CHARACTERIZATION OF $O^6$-ALKYLGUANINE-DNA ALKYLTRANSFERASE DEGRADATION AND LABELING WITH FLUOROPHORES

A Dissertation in
Genetics
by

Candice L. Lewis

© 2010 Candice L Lewis

Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

December 2010
The dissertation of Candice L. Lewis was reviewed and approved* by the following:

Kristin A. Eckert  
Professor of Pathology, and Biochemistry and Molecular Biology  
Dissertation Co-Advisor  
Co-Chair of Committee

Sreenivas Kanugula  
Assistant Professor of Cellular and Molecular Physiology  
Dissertation Co-Advisor  
Co-Chair of Committee

Hui-Ling Chiang  
Professor of Cellular and Molecular Physiology

Ira J. Ropson  
Associate Professor of Biochemistry and Molecular Biology

Robert Levenson  
Professor of Pharmacology

David Spector  
Chair of Intercollege Graduate Degree Program in Genetics

*Signatures are on file in the Graduate School
ABSTRACT

$O^{6}$-alkylguanine-DNA alkyltransferase (AGT) is a highly conserved DNA repair protein, that removes alkyl adducts at the $O^{6}$-position of guanine in a single irreversible, stoichiometric reaction. AGT is unique among all the other DNA repair pathways as it requires no co-factors or chaperones to carry out DNA repair. $O^{6}$-alkylguanine is a highly cytotoxic lesion, and when left un-repaired, leads to G:C to A:T transition mutations. Upon alkyl transfer, AGT continues to bind to DNA and could impede other AGT molecules from repairing DNA. Therefore, degradation might be the only mechanism to remove the alkylated AGT.

To study human AGT (hAGT) degradation and intracellular dynamics, we created a Green Fluorescent Protein (GFP)-hAGT fusion and expressed this tagged AGT in Chinese Hamster Ovary (CHO) cells. GFP-hAGT expressing CHO cells enabled Fluorescent Recovery After Photobleaching (FRAP) experiments showing for the first time, the rapid mobility of AGT within the nucleus. When the entire nucleus was bleached, recovery of fluorescence occurred only after 30 minutes, indicating that AGT accumulates in the nucleus over long periods of time, probably soon after synthesis. The addition of GFP at the N-terminus of hAGT changed some physiological features of hAGT. The half life of the alkylated protein changed from 9h for the native form to 26h for the GFP-tagged form. However, the formation of putative Ub-GFP-AGT species was similar to the ubiquitination pattern and time frame, seen in wt hAGT expressed in CHO cells. This finding suggests that the N-end rule pathway for ubiquitin-mediated degradation may not be the only proteolytic pathway regulating degradation of hAGT.

Fluorophores were used to characterize a novel 16kDa AGT truncation species detected in both CHO and HeLa cells. The cleavage site of this species was determined to be at codon 50, causing the elimination of three of the four ligands required to bind a zinc atom and maintain structural stability. Despite the truncation of the N-terminus, this 16kDa hAGT species showed
significant protection from $N$-Methyl-$N'$-Nitro-$N$-Nitrosoguanidine (MNNG) induced cytotoxicity in *Escherichia coli* (*E.coli*). Together with purified protein AGT activity assays, these data indicate that this truncated AGT species retains the ability to repair DNA. This intriguing finding suggests that this truncated species is either part of a previously unknown degradation cycle, or it has a unique role in the cytosol.

The biological role of AGT is to protect cells from alkylation damage generated from endogenous as well as environmental sources. However, as a consequence of its biological role, hAGT plays an important part in tumor cells by causing resistance to some chemotherapeutic drugs such as temozolomide or BCNU, by neutralizing the drug induced alkylation damage. Therefore, in order to combat this effect, a series of AGT inhibitors were created previously to inactivate hAGT and increase the efficacy of chemotherapy. $O^6$-Benzylguanine ($O^6$-BG) is one such, potent inhibitor of AGT, which is currently in clinical trials.

In this dissertation, a series of mutant hAGT proteins were created to test for increased sensitivity to $O^6$-BG. Three codons at sites 157, 159 and 160 were targeted because of their location within the $O^6$-BG binding pocket. Our studies confirm that these sites, particularly, a tryptophan at position 160, contribute to a ten-fold lowering of $O^6$-BG ED$_{50}$. These mutants are unique and rare because a majority of previously characterized AGT mutations within the $O^6$-BG binding pocket cause resistance to $O^6$-BG.
TABLE OF CONTENTS

LIST OF FIGURES ................................................................................................. viii
LIST OF TABLES ...................................................................................................... xi
LIST OF ABBREVIATIONS .................................................................................... xii
ACKNOWLEDGEMENTS ......................................................................................... xiii

Chapter 1  INTRODUCTION ................................................................................... 1

1.1 DNA Damage
   1.1.1 DNA Repair .......................................................... 2
   1.1.2 Direct DNA Repair ............................................. 5
   1.1.3 Sources of Alkylation damage ........................................... 5

1.2 Function, Occurrence and Distribution of AGT ........................................... 7
   1.2.1 Biological Role .................................................... 7
   1.2.2 AGT Species Distribution ..................................... 8
   1.2.3 Human AGT Intracellular Localization and Tissue Distribution ......................... 10

1.3 Reaction and Structure of AGT ................................................................. 11
   1.3.1 AGT Substrates and Mechanism of Action ......................... 11
   1.3.2 Human AGT Structure ........................................... 14

1.4 Clinical relevance of AGT ............................................................................. 19
   1.4.1 $O^6$-Benzylguanine ($O^6$-BG): Mechanism of Inactivation and Use as an Adjuvant Chemotherapy drug ......................... 20

1.5 AGT Regulation ......................................................................................... 23
   1.5.1 Transcriptional Level ............................................... 23
   1.5.2 Post-translational Level ........................................... 23

1.6 AGT Degradation ......................................................................................... 24

1.7 Goal and Specific Aims ................................................................................ 25

Chapter 2  MATERIALS AND METHODS ................................................................. 27

2.1 Materials ......................................................................................... 27

2.2 Cell lines Used ................................................................................ 28
   2.2.1 Cell Culture ................................................................... 28
   2.2.2 Labeling CHO cells with BGAF ................................... 28

2.3 Construction of Plasmid Vectors ............................................................... 29
   2.3.1 Generating AGT mutants to determine sensitivity to $O^6$-BG ......................... 29
   2.3.2 pCMV-G160W and pCMV-N157G/S159E/G160W for expression of mutants in CHO cells ......................................................... 29
   2.3.3 pCMV-wtAGT His for expression of C-terminal His tagged AGT in CHO cells ................................................................. 30
   2.3.4 Generating an AGT-His CHO cell line ..................................... 30
   2.3.5 Generating GFP-AGT fusion protein expression in CHO
Chapter 2: BIOCHEMICAL CHARACTERIZATION OF AGT MUTANTS

2.3.6 Construction of 16kDa truncated hAGT in vectors pQE-30 and pCMV

2.4 Purification of wt and Mutant AGT Proteins

2.5 Biochemical characterization of AGT mutants

2.5.1 Determination of AGT activity and effects of $O^6$-BG on AGT activity using filter binding assay

2.5.2 Determination of hAGT and mutant Rate Constants

2.6 Preparation of cell culture extracts for SDS-PAGE and Immunoblotting

2.7 Degradation studies of AGT in CHO cells

2.7.1 AGT Activity Assay in CHO cells

2.7.2 Measurement of AGT and GFP-AGT Half-life

2.7.3 Detection of AGT-ubiquitin complexes

2.8 GFP-AGT mobility in CHO cells

2.8.1 FRAP: Fluorescent Recovery After Photobleaching

2.8.2 Cell Fixation using Paraformaldehyde

2.9 Characterization of 16kDa AGT species

2.9.1 Isolation of 16kDa species from CHO cells expressing AGT-His

2.9.2 Cell Fractionation

2.9.3 In vitro formation of 16kDa species using empty-vector CHO cell extracts

2.9.4 Sample preparation for Mass Spectrometry

2.9.5 Sample preparation for Edman Sequencing

2.9.6 MNNG Protection Assay

Chapter 3: MUTATIONAL ANALYSIS OF THE SENSITIVITY OF AGT TO $O^6$-BENZYLGUIANINE ($O^6$-BG)

3.1 Rationale

3.2 Characterization of purified mutant AGT proteins

3.2.1 DNA repair activity

3.2.2 Inactivation of AGT proteins by $O^6$-Benzylguanine ($O^6$-BG)

3.2.3 Inactivation of AGT proteins by $O^6$-Methylguanine ($O^6$-MG)

3.3 Sensitivity to BGFL

3.3.1 In vitro Labeling

3.3.2 Labeling AGT in CHO cells

3.4 Conclusions

Chapter 4: PATHWAYS OF POST-ALKYLATION AGT DEGRADATION

4.1 Rationale

4.2 Determination of AGT Half-life

4.2.1 Determination of wt AGT Half-Life in CHO cells

4.2.2 Effect of N-terminal GFP tag on AGT Half-life
4.3 Ubiquitination of wt AGT and GFP-AGT ................................. 77
4.4 Fluorescent Recovery After Photobleaching (FRAP) in GFP-AGT CHO Cells ................................................................. 78
4.5 Conclusions ........................................................................ 79

Chapter 5: EFFECT OF ENDOGENOUS N-TERMINAL TRUNCATION ON AGT FUNCTION ................................................................. 94

5.1 Rationale ........................................................................... 94
5.2 Detection of 16kDa species of hAGT .................................. 95
  5.2.1 CHO cells .................................................................... 96
  5.2.2 Human cell lines ......................................................... 97
5.3 In vitro studies of 16kDa species ......................................... 98
5.4 Identification of 16kDa species ............................................ 100
5.5 Biological Characterization of 16kDa AGT form ............... 100
  5.5.1 In vitro DNA repair activity ........................................ 100
  5.5.2 GWR109 E.coli MNNG Protection .............................. 101
5.6 Conclusions ...................................................................... 102

Chapter 6 GENERAL DISCUSSION .................................................. 122

BIBLIOGRAPHY ...................................................................... 130
LIST OF FIGURES

Figure 1.1: DNA Lesions and Corresponding DNA Repair and Maintenance Systems……..4
Figure 1.2: Sites on DNA bases susceptible to alkylation damage.................................6
Figure 1.3: Comparison of DNA Alkyltransferases from various organisms. ....................9
Figure 1.4: Mechanism of action for alkyltransferase mediated DNA repair..................13
Figure 1.5: Model for hAGT binding to DNA ..............................................................16
Figure 1.6: Model for hAGT binding to DNA in a cooperative manner.........................17
Figure 1.7: The secondary structure of hAGT ...............................................................18
Figure 1.8: Structures of two free base inactivators of hAGT: O6-benzylguanine (O6-BG) and O6-methylguanine (O6-MG) .................................................................21
Figure 1.9: Reaction of AGT proteins with O6-BG. .....................................................22
Figure 2.1: Diagram of plasmids. .................................................................................42-44
Figure 3.1: O6-BG docked in the binding pocket of hAGT ...........................................50
Figure 3.2: Sites within the O6-BG binding pocket of hAGT, when mutated cause O6-BG resistance .............................................................................................................51
Figure 3.3: Purified AGT proteins stained by Coomassie Blue ......................................62
Figure 3.4: DNA repair activity of purified AGT proteins .............................................63
Figure 3.5: Timecourse reaction of purified AGT with [3H]-alkylated DNA .........................64
Figure 3.6: Wt and mutant AGT rate constants ............................................................65
Figure 3.7: Structures of inhibitors used to test sensitivity of mutant AGT ....................66
Figure 3.8: Inactivation of AGT using O6-BG ............................................................67
Figure 3.9: Inactivation of AGT using O6-MG ............................................................68
Figure 3.10: Mechanism of Labeling AGT with BGAF or BGFL .................................69
Figure 3.11: Reactivity of purified wt and mutant AGT with BGFL .............................70
Figure 3.12: Conversion of BGAF to BGFL ................................................................71

Figure 3.13: CHO cells expressing wtAGT, G160W or empty-vector after 1h BGAF labeling........................................................................................................72

Figure 3.14: Wt-AGT CHO cells labeled with BGAF. ...........................................73

Figure 3.15: Conserved amino acid sequences between species, around the AGT active site. .................................................................74

Figure 4.1: Half-life determination of alkylated and unalkylated wtAGT in CHO cells........84

Figure 4.2: Protein expression levels in wtAGT CHO cells and GFP-AGT CHO cells ........85

Figure 4.3: Fluorescent imaging of GFP-AGT and GFP-P140K CHO cells....................86

Figure 4.4: Half-life determination of alkylated and unalkylated GFP-AGT in CHO cells...87

Figure 4.5: Detection of ubiquitinated AGT species in CHO cell extracts....................88-89

Figure 4.6: Formation of higher molecular weight bands corresponding to ubiquitinated GFP-AGT in GFP-AGT CHO cells ......................................................90

Figure 4.7: Fluorescent Recovery after Photobleaching (FRAP) in GFP-AGT CHO cells. ..91

Figure 4.8: FRAP in GFP-AGT CHO cells with bleached nuclei. .................................92

Figure 4.9: Imaging GFP-AGT CHO cells upon MNNG treatment..............................93

Figure 5.1: Secondary structure of hAGT highlighting residues critical for hAGT structural and catalytic activity. .................................................................107

Figure 5.2: Detection of 16kDa band in AGT-His and wt AGT CHO cell extracts..........108-109

Figure 5.3: Detection of 16kDa band in human cancer cell lines.................................110

Figure 5.4: In vitro formation of 16kDa species..........................................................111

Figure 5.5: Formation of 16kDa species using purified N-terminal His AGT in empty-vector CHO cell extracts and HeLa extracts.........................................................112

Figure 5.6: Formation of 16kDa species from N-terminal His tagged AGT by HeLa extracts .................................................................................................................113

Figure 5.7: Formation of 16kDa species from C-terminal His tagged AGT in empty-vector CHO cell extracts .................................................................114-115

Figure 5.8: Edman sequencing results identifying the cleavage site of truncated AGT species ..........................................................................................116-117
Figure 5.9: Comparison of full length hAGT-C terminal His and truncated 16kDa AGT amino acid sequences................................................................................................................................. 118

Figure 5.10: Purification of 16kDa truncated hAGT from multiple bacterial strains.............. 119

Figure 5.11: $O^6$-BG inhibition of 16kDa AGT DNA repair activity ........................................... 120

Figure 5.12: MNNG protection with 16kDa AGT species .............................................................. 121

Figure 6.1: Model depicting the degradation of hAGT and the formation of the truncated 16kDa AGT species. ................................................................................................................................. 129
LIST OF TABLES

Table 2.1: Plasmid Constructions. ................................................................. 45
Table 2.2: Primers used for plasmid construction. ........................................ 46
Table 2.3: Alkyltransferase activity in the the various CHO cell lines........... 47
Table 3.1: Rate Constants of wt and mutant AGT proteins. ......................... 60
Table 3.2: Comparison of ED_{50} values between O^δ-BG and O^δ-MG. ........ 61
Table 4.1: Comparison of yeast ubiquitination enzymes and the human homologs of these enzymes. ................................................................. 83
Table 5.1: Mass Spectrometry results of 16kDa species in-gel tryptic digest matches with purified hAGT.................................................. 106
LIST OF ABBREVIATIONS

ACNU : N’-[4-amino-2-methyl-5-pyrimidinyl)methyl]-N-(2-chloroethyl)-N-nitrosourea
AGT : O6-Alkylguanine-DNA alkyltransferase
ANOVA : Analysis of Variance
BCNU : 1,3-bis(2-chloro-ethyl)-1-nitrosourea
BER : Base excision repair
BGAF : O6-Benzylguanine deriviative with attached diacetylated fluorescein
BGFL : O6-Benzylguanine derivative with attached fluorescein
CCNU : 1-(2-chloroethyl)-3-cyclohexyl-1-nitrosourea
CHO : Chinese Hamster Ovary
DMSO : Dimethyl sulfoxide
DNA : Deoxyribonucleic acid
DTT : dithiothreitol
EDTA : ethylenediaminetetraacetic acid
FRAP : Fluorescent Recovery after Photobleaching
GFP : Green Fluorescent Protein
GGR : Global genome nucleotide excision repair
hAGT : human O6-Alkylguanine-DNA alkyltransferase
HECT : Homologous to the E6-AP Carboxyl Terminus
Mer : methyl excision repair deficient
Mer+ : methyl excision repair competent
MG132 : N-(benzyloxycarbonyl)leucinylleucinylleucinal
MNNG : N-methyl-N’-nitro-N-nitrosoguanidine
MNU : N-methyl-N-nitroso urea
NER : Nucleotide excision repair
O6-BG : O6-Benzylguanine
O6-MG : O6-Methylguanine
PAGE : Polyacrylamide gel electrophoresis
RNA : Ribonucleic acid
SDS : Sodium dodecyl sulfate
SNP : Single nucleotide polymorphism
TCR : Transcription coupled repair
UFD : Ubiquitin Fusion Degradation
Ub : Ubiquitin
ACKNOWLEDGEMENTS

I would like to thank Dr. Anthony Pegg for giving me the opportunity to work in his lab. Thanks to Dr. Kristin Eckert for helping me see the ‘big picture’, and to all my committee members for your invaluable input.

To my parents, Dr. Antony Lewis and Mrs. Loretta Lewis, I am deeply grateful for all that you have done, and all that you continue to do. You are my inspiration.

To my husband, Souvik De, your support, encouragement and spirit have been the wind beneath my wings. Thank you!
Chapter 1

Introduction

1.1 DNA Damage

Deoxyribonucleic acid (DNA) is the blueprint containing the genetic instructions upon which the development and functioning of all organisms depend. DNA is the only biological molecule in the cell which does not get re-manufactured, in its entirety, upon damage. Over the lifetime of the organism, it must be repaired. Other biological molecules such as RNA or protein, if damaged, get targeted for degradation and are re-synthesized. DNA is susceptible to attack from a plethora of endogenous and exogenous damaging agents. Within the cell, reactive oxygen species such as hydroxyl radicals, superoxide anions etc, and reactive nitrogen species cause significant lesions (Kirkwood, 2005; Samson, 1992). These compounds cause damage to all biological molecules including DNA, RNA and protein. Evidence shows that a hundred oxidative modifications occur in DNA (Cadet, 1997). Alkylating agents are found endogenously, as well as in the environment. Within cells, S-Adenosylmethionine is a weak methylating agent and has been shown to cause DNA mutations (Drablos, 2004). In addition, some products of the metabolic pathway generate reactive electrophilic species which alkylate DNA (Marnett, 1993).

DNA can get hydrolyzed under normal physiological conditions leaving abasic sites. During DNA replication, errors can occur resulting in pairing mismatches or spontaneous mutations (Lahue, 1996). From the external environment, DNA is exposed to UV radiation, x-rays, and genotoxic chemicals. These agents can cause single or double strand breaks, interstrand cross-links, and numerous adducts on DNA. If left un-repaired, damage to DNA causes inhibition of transcription, replication and aberrant chromosome segregation. If there is extensive DNA damage and it cannot be repaired prior to cell division, the cell undergoes apoptosis. If
mutations resulting from DNA damage accumulate, this leads to aging or cancer. Therefore, it is vital to maintain genomic integrity in order for the cell to maintain viability (Hoeijmakers, 2001; Hoeijmakers, 2009).

1.1.1. DNA Repair

There are five broad mechanisms that the eukaryotic cell utilizes for genomic maintenance. Each is designed to repair a specific category of lesions on DNA (Figure 1.1). Base excision repair (BER) repairs non-bulky base adducts, caused by oxidized, reduced or fragmented bases and alkylation. There are two mechanisms of BER: (a) Long patch BER in which a 2-6 nucleotide gap is re-synthesized and (b) Short patch BER which replaces a single nucleotide (Hoeijmakers, 2001). Nucleotide excision repair (NER) recognizes a wide array of bulky lesions that cause DNA helix distortion (Slupphaug, 2003; Friedberg, 2006). For example, the cyclobutane pyrimidine dimers formed by UV radiation are repaired by NER (Sarasin, 1997). NER has two categories: (a) global genome NER (GGR) which repairs transcriptionally silent DNA and (b) Transcription-coupled repair (TCR) which targets lesions which obstruct transcription. Replication and recombination errors originating from an incorrect pairing of bases, insertions or deletions, are repaired by the Mismatch repair system. Microsatellites are sequences of DNA with short tandem repeats, and during replication of these regions, DNA polymerases are prone to slippage, causing microsatellite instability. The mismatch repair system plays a vital role in correcting these errors (Hoeijmakers, 2001). Nonhomologous end joining and double strand break repair target the various forms of double strand breaks which are caused by ionizing radiation. Homologous recombination occurs after replication and uses the intact copy on the sister chromatid to correctly align and seal the broken ends. If damage occurs leading to stalled DNA polymerase δ or ε, either due to strand breaks or cross-links, it can be repaired by homologous recombination. In addition, this type of damage can be repaired by translesional
synthesis which is a mechanism of bypassing specific lesions and allowing replication to continue (Hoeijmakers, 2009). Finally, direct DNA repair reverses DNA damage by breaking the chemical bond between DNA and the adduct. It is well documented that defects in DNA repair systems cause dramatic increases in mutation, and consequently, oncogenesis (Modrich, 1994). When components of the DNA repair pathway are defective, it leads to severe disorders such as Cockayne’s Syndrome or trichothiodystrophy, where the patient’s lifespan is greatly reduced (Hoeijmakers, 2009).
DNA repair pathways target specific lesions in DNA caused by various DNA damaging agents such as x-rays, oxygen radicals, or genotoxic chemicals. Double and single stranded breaks are repaired by non-homologous end-joining or homologous recombination. Small base adducts are repaired by base excision repair, while bulky, helix distorting adducts are repaired by nucleotide excision repair. Errors in replication caused by a stalled DNA polymerase are repaired by transcription coupled repair. Translesional synthesis and template switching are two mechanisms for repairing DNA damage leading to stalled DNA polymerase δ/ε (e.g., strand breaks or cross-links). Adapted from Figure 2, Hoejimakers, New England Journal of Medicine, 2009.
1.1.2 Direct DNA Repair

There are three known mechanisms of direct DNA repair. The first example is photolyase, which removes pyrimidine cyclobutane dimers caused by UV radiation (Sato, 1993; Sancar, 1994). While photolyases have not been described in mammals, it has recently been found in *Arabidopsis thaliana* and *Drosophila melanogaster* (Deisenhofer, 2000). The second is oxidative demethylase, which is responsible for reversing methylation damage in DNA and RNA (Falnes, 2007). Finally, $O^6$-Alkylguanine-DNA alkyltransferase (AGT) is a reactive protein which transfers alkyl adducts from the $O^6$-position of guanine, to a cysteine residue at its active site, and consequently inactivates itself. This DNA repair protein is the focus of this dissertation.

1.1.3 Sources of Alkylation damage

The most common alkylating carcinogens in the environment are nitrosoureas and nitrosamines. $N$-ethyl-$N$-nitrosourea, $N$-methyl-$N$-nitrosourea (MNU), $N$-methyl-$N'$-nitro-$N$-nitrosoguanidine (MNNG) are all examples of simple alkylating agents (Kusmierek, 1982). Some nitrosourea compounds are used in chemotherapy to treat a variety of cancers including lymphoma, malignant melanoma, neuroblastoma, pancreatic cancer, glioblastoma and astrocytoma. Examples of anti-cancer alkylating agents are procarbazine, dacarbazine, streptozotocin, temozolomide, and chloroethylating agents carmustine (BCNU), lomustine (CCNU) and nimustine (ACNU) (Kaina, 2007). Evidence shows that mutagens such as nitrosamines in tobacco smoke and food can also alkylate DNA (DeMarini, 2004; Jagerstad, 2005). There are at least 12 sites in DNA which are susceptible to alkylation: positions 1, 3 and 7 on adenine, positions 3, 7 and $O^6$ on guanine, positions 3 and $O^2$ on cytosine and positions 3, $O^2$ and $O^4$ on thymine and phosphate (Figure 1.2) (Essigmann, 1991). Methylation at the N7 position of guanine is the most abundant adduct formed due to alkylating agents. However, since this site
does not participate in base pairing, 7-Me-dGuo is only weakly mutagenic (Robins, 1983). A biologically significant lesion formed by alkylating agents is at the $O^6$ position of guanine. Alkyl adducts at this site have been shown to form G:C to A:T transition mutations during DNA replication *in vitro*, in bacteria and in mammalian cells (Topal, 1991; Singer, 1989; Dosanjh, 1990; Ellison, 1989; Pauly, 1991).

Figure 1.2 Sites on DNA bases susceptible to alkylation damage.
1.2 Function, Occurrence and Distribution of AGT

1.2.1 Biological Role

AGT removes alkyl adducts at the $O^6$ position of guanine, and at the $O^4$ position of thymine. While the bacterial alkyltransferase Ogt efficiently repairs the adduct at the $O^4$ position of thymine, human AGT (hAGT) is inefficient in repair at this site (Pegg, 2000). A cysteine residue in the active site of all known AGTs from different species acts as an alkyl group acceptor site. AGT requires no co-factors or chaperones and carries out repair, in a single irreversible step. Unlike an enzyme, AGT can act only once; therefore, once the alkyl transfer is completed, AGT is inactivated and degraded (Pegg, 2002; Pegg, 1993). Thus, each AGT molecule can repair only one alkyl adduct in DNA. When AGT is overexpressed, either in cell culture or transgenic mice, the protection from alkylating agents is increased. Inversely, cells and mice lacking AGT are more sensitive to alkylation and consequently to carcinogenesis (Dolan, 1997; Kaina, 1998; Kaina, 1993; Sakumi, 1997). AGT protects cells against cytotoxicity and mutation. When $O^6$-methylguanine is left un-repaired, it results in the formation of G:C to A:T transition mutations in DNA (Singer, 1989; Dosanjh, 1990). $O^6$-meG pairs with T instead of C, during replication. This mismatched pair is recognized by the mismatch repair proteins and excised. However, the placement of T opposite $O^6$-meG in the parent strand sets up a futile cycle of repair, which stalls DNA replication and leads to cell death. Only when the mismatch repair proteins are inactivated do cells develop tolerance to methylation damage (Kawate, 1998; Karran, 1996). The NER pathway does recognize and repair alkylation damage at the $O^6$ position of guanine, however, the repair is inefficient (Sancar, 1996; Samson, 1988). $O^6$-methylguanine is likely to be the physiological substrate of AGT, and is small enough to cause no helix distortion; thus it is unrepaired by the NER pathway (Pegg, 2000). The impact of AGT is illustrated by that fact that
when AGT is absent, there is a tremendous increase in mutation rate caused by MNNG or MNU (Pegg, 1995). Transgenic animals expressing high levels of AGT show a lower incidence of tumors induced by carcinogens (Gerson, 1994). The presence of $O^6$-methylguanine can lead to sister-chromatid exchange and chromosome breakage. In 1993, Kaina et. al. showed that the expression of alkyltransferases prevented this type of chromosomal breakage (Kaina, 1993). Thus AGT has a vital role in protecting cells from chromosomal aberration. Furthermore, $O^6$-methylguanine has also been implicated as an apoptosis-inducing lesion (Kaina, 1997). Therefore, by removing this cytotoxic adduct, AGT mediates protection from apoptosis caused by alkylating agents.

1.2.2 AGT Species Distribution

AGT is expressed in several hundred species from three domains of life: Archaea, Eubacteria and Eukarya (Sedgwick, 1991; Sekiguchi, 1996; Lindahl, 1988; Pegg, 2001; Xiao, 1991). However, AGT has been undocumented in plants. There is substantial homology between the alkyltransferases across species, shown in Figure 1.3. A conserved -PCHR- motif exists at the active site of all known alkyltransferases. Many species of bacteria such as Escherichia coli (E.coli) or Salmonella typhimurium (S.typhimurium) have two alkyltransferase genes. In E. coli, the alkyltransferase Ogt is constitutively expressed, while the second alkyltransferase Ada is inducible (Pegg, 1995). Ada has a second domain which repairs methylation of methylphosphotriesters, by transferring the methyl adduct onto a cysteine within a --PCKR- sequence. The C-terminal domain of Ada repairs alkylation damage at the $O^6$ position of guanine and the $O^4$ position of thymine (Lindahl, 1988). Eukaryotic AGT closely resembles the C-terminal domain of Ada, while the N-terminal domain of Ada shows little homology. The alkyltransferase of Ferroplasma acidarmanus is fused with an endonucleaseV domain, and completely lacks the N-terminal domain seen in eukaryotic AGT (Kanugula, 2005). This feature
is also seen in the alkyltransferase found in Caenorhabditis elegans (C. elegans), which lacks an N-terminal domain, but has a long C-terminal domain with 25% sequence homology to human Histone1C (Kanugula, 2001).

Figure 1.3 Comparison of DNA Alkyltransferases from Various Organisms.

The first structure shows the two domains of E.coli’s Ada alkyltransferase. The left half of the protein contains a –PCKRC- acceptor site which facilitates repair of methylphosphotriester DNA. The black box common to all alkyltransferases is the helix-turn-helix motif used to bind to DNA. The conserved sequence –PCHRVI- is shown in gray. In C. elegans, this sequence is –PCHPV-, and is shown in white. The shaded area indicates regions of the protein with conserved sequences among species. The numbers in parentheses indicate the number of amino acids present in the alkyltransferase of that species. Modified from Figure 1 in Sekiguchi et al, 1996, J. Cancer Res. Clin. Oncol. 122:199-206.
1.2.3 Human AGT Intracellular Localization and Tissue Distribution

Human AGT is found primarily in the nucleus; however, a small percentage is also present in the cytoplasm. Immunostaining of human biopsy tissues showed that AGT was a nuclear protein (Ayi, 1992). This observation is consistent with the biological role of AGT. The proportion of nuclear and cytosolic distribution is dependent on the cell-type (Sekiguchi, 1996). Studies have shown that AGT contains nuclear localization signals which allows its translocation and retention in the nucleus. Residues 124-128, -PKAAR- have been shown to be critical for AGT nuclear retention (Li, 1996).

AGT distribution in tissues varies greatly depending on organs and the tissues within organs (Citron, 1991; Citron, 1992; Citron, 1993; Citron, 1994; Citron, 1995). The highest AGT expression levels are found in the liver and spleen and the lowest are found in the brain (Pegg, 1990). In addition to the tissue-related variation, there is also inter-individual variation (Grafstrom, 1984). Furthermore, there is tremendous variation in AGT activity between normal tissue and tumors. There are human tumor cell lines deficient in AGT activity, referred to as Mer− (methyl excision repair deficiency). Cell lines capable of alkyltransferase-mediated repair are called Mer+. The Mer− cells are extremely sensitive to killing by alkylating agents (Pegg, 2000). Aberrant methylation at the promoter of the AGT gene has been demonstrated to cause gene silencing. A correlation between low AGT content and colorectal cancers has been identified, and is partially linked with the lack of methylation at specific sites within the AGT gene (Esteller, 1999).
1.3 Reaction and Structure of AGT

1.3.1 AGT Substrates and Mechanism of Action

AGT recognizes alkyl lesions at the O\textsuperscript{6}-position of guanine and transfers the alkyl group onto its active site, thus restoring DNA to its original form (Figure 1.4). Human AGT repairs methyl, ethyl, 2-chloroethyl, butyl, and more bulky adducts such as benzyl and pyridyloxobutyl at the O\textsuperscript{6}-position of guanine. The relative rates of repair are: benzyl >> methyl > ethyl > n-propyl, n-butyl (Pegg, 1995; Goodtzova, 1997; Coulter, 2007; Mijal, 2004). The cysteine responsible for mediating this repair reaction (Cys 145) is highly reactive (Guengerich, 2003). This reaction results in an S-alkylcysteine, which is extremely stable. Evidence suggests that AGT repairs DNA in a base-flipping mechanism, whereby the alkylguanine is flipped out of the DNA helix and into the active site pocket (Tubbs, 2007). Figure 1.5 shows the schematic of the repair reaction.

Tyrosine at position 114 plays a critical role in this reaction as it promotes the rotation of DNA’s 3’ phosphate. Arginine at site 128, then stabilizes the protein-DNA complex by replacing the base in the DNA helix (Daniels, 2004). AGT has modest DNA binding affinity and low sequence specificity (Chan, 1993; Fried, 1996; Rasimas, 2003). AGT has been shown to interact with an 8bp region, within the minor groove of DNA and has a higher affinity for alkylated DNA (Chan, 1993; Takahashi, 1990). Evidence shows that AGT proteins bind to DNA in a cooperative manner. Models created based on these experiments, and the crystal structure of AGT bound to DNA, indicate that AGT is densely packed together at certain regions of DNA and this is the mechanism with which it scans DNA and efficiently repairs lesions, \textit{in vitro}. Densities of cooperative AGT binding can reach 1 protein every 4 bp for DNA templates \(\leq 50\text{bp} \) long (Melikishvili, 2008). AGT forms multiprotein complexes, arranged in a three start helical array in which there is strong protein interaction between the amino terminal face of protein \(n\) and the carboxy-terminal face of protein \(n+3\) (Figure 1.6) (Melikishvili, 2008; Rasimas, 2007; Adams,
2009). A single array of proteins consists of an \( n, n+1 \) and \( n+2 \) set of AGT proteins, each offset by 138°. This model accounts for the 4bp separation between AGT proteins and minor groove binding on the DNA helix.
Figure 1.4 Mechanism of action for alkyltransferase mediated DNA repair

The diagram shows the conserved sequence –PCHR- at the active site of AGT. The alkyl group ‘R’ is transferred onto cysteine forming S-alkylcysteine in an irreversible, stoichiometric reaction.
1.3.2 Human AGT Structure

Human AGT is a small protein consisting of 207 amino acids. The crystal structure of hAGT reveals the existence of two distinct domains in human AGT. The N-terminal domain extends from residues 1-85, while the C-terminal domain extends from residues 92-176 (Daniels 2000). The most highly conserved residues, and those critical for AGT activity are found within the C-terminal domain. The crystal structure of hAGT revealed that the N-terminal domain exists as a β-sheet, with three β strands, packed against a helix, H1 (Figure 1.7). This domain contains a bound zinc atom, which plays a role in stabilizing the structure of AGT (Figure 1.7) (Daniels 2000). This bound Zn ion is absent in prokaryotic AGTs. Experiments done with the Zn ion removed have shown that AGT still repairs DNA; however, the repair rate constant is greatly reduced (Rasimas, 2003). The two domains are connected by a loop. The C-terminus domain consists of four α-helices. As shown in Figure 1.7, helices 3 (H3) and 4 (H4) form the helix-turn-helix motif responsible for DNA binding. This motif is common to many DNA binding proteins. Between helices H4 and H5 is a region containing a coil and loop where the conserved sequence PC$_{145}$HR or the active site (C145) of hAGT is found (Wibley, 2000).

hAGT is the only known example of a mammalian alkyltransferase crystal structure. Comparisons of this structure with the three other solved AGT structures from E.coli (Ada-C), archaeabacterium Pyrococcus kodakaraensis and thermophile Methanococcus jannaschii show similar domains, α/β fold and dimension to human AGT (Tubbs, 2007). This structural similarity is remarkable considering there is little amino acid sequence homology between these proteins (Moore, 1994; Hashimoto, 1999; Daniels, 2000). Of the known AGT sequences, there are 25 highly conserved residues, and 13 invariant residues, including the PCHR active site motif. The greatest diversity in AGT sequences is found in the N-terminal domain. For example, the N-terminal domain of E. coli’s Ada or S. typhimurium alkyltransferase is much longer than hAGT.
Ada even has a second methyl acceptor site, capable of repairing methyl phosphotriester DNA damage. *C. elegans* is the only eukaryote known to have two types of AGT. One of these AGTs, completely lacks the N-terminal domain seen in human AGT (Kanugula, 2001).
Figure 1.5 Model for hAGT binding to DNA (taken from Tubbs et al., 2007)

Panel A shows unreacted AGT and the DNA helix. The N-terminal domain is in green and the C-terminal domain is shown in yellow. The DNA binding motif, helix turn helix is in blue. Critical residues for repair are also represented. Panel B shows AGT bound to DNA and with a base flipped out into the active site pocket of AGT.
Figure 1.6. Model for hAGT binding to DNA in a cooperative manner. Taken from Melikishvili et al., 2008. AGT proteins are juxtaposed to form a three-start helical array, with the $n^{th}$ AGT protein interacting with the $n+3$ protein. Each AGT protein is separated by 4 bp and is rotated approximately 138°. A single array with three AGT proteins is depicted in magenta, blue and green. The DNA double helix is shown in gray and black.
Figure 1.7 The secondary structure of hAGT. Modified from Daniels et al. 2000. The two domains of hAGT are depicted, with \( \beta \)-sheets (in orange) represented B 1-3, and \( \alpha \)-helices (in blue) represented H1-H4. The zinc atom critical for structural integrity is shown in purple.
1.4 Clinical relevance of AGT

Under normal conditions, the role of AGT is to preserve the integrity of DNA, by repairing alkyl adducts and preventing the formation of mutations. However, an important consequence of it’s function is that AGT causes resistance to some chemotherapeutic strategies by removing the alkylation caused by the alkylating agents (Gerson, 2004). Therefore, AGT serves as a target for improving the efficacy of cancer treatment drugs. Tumors expressing AGT are resistant to treatment with alkylating agents. Studies of AGT in human tumors have demonstrated high expression levels in melanoma, colon cancer, pancreatic carcinoma, lung cancer and glioma (Lee, 1992; Moriwaki, 1992; Redmond, 1991; Zaidi, 1996; Silber, 1999; Chen, 1999; Bobola, 2001; Citron, 1993; Kokkinakis, 1997). Interestingly, AGT activity in these tumors was higher than in the surrounding normal tissue (Gerson, 2002). In early studies, it was shown that in astrocytoma patients receiving BCNU treatments, low AGT content in tumors correlated with positive treatment response and greater survival. Patients with tumor cells expressing high levels of AGT had a poorer prognosis (Belanich, 1996). In a study of malignant gliomas, once again, tumors with high AGT expression were less responsive to therapy with chloroethylnitrosoureas (Hotta, 1994). Furthermore, it was shown that using an $O^6$-methylguanine substrate or streptozotocin (which forms $O^6$-methyl lesions) in combination with BCNU, increased the efficacy of treatment in colon cancer (Wilson, 1995). The use of methylating agents such as MNNG or streptozotocin successfully depleted cellular alkyltransferase activity and in combination with BCNU, caused interstrand cross-linking and a 2-3 log enhancement of cytotoxicity (Futscher, 1989; Pieper, 1991). Since the AGT reaction is stoichiometric, any inhibitor of AGT would irreversibly inactivate the AGT molecule. Therefore, studies were conducted to create a direct substrate inhibitor of AGT to deplete cellular levels of AGT and allow its use in clinical settings to enhance chemotherapy.
1.4.1 \( O^6 \)-Benzylguanine (\( O^6 \)-BG): Mechanism of Inactivation and Use as an Adjuvant

Chemotherapy drug

\( O^6 \)-BG (Figure 1.8) is one of the first generation of AGT inhibitors and was first described in 1990 (Dolan, 1990). It is a potent, non-toxic inhibitor of AGT, which has an ED\(_{50}\) of 0.5 \( \mu \)M against hAGT. Figure 1.9 shows the reaction of \( O^6 \)-BG with AGT. The benzyl group of \( O^6 \)-BG is transferred onto AGT, by the formation of \( S \)-benzylcysteine (Pegg, 1993). \( O^6 \)-BG is readily taken up by cells and treatment of human cell extracts with 2.5 \( \mu \)M \( O^6 \)-BG reduced alkyltransferase activity by 90%, within 10 min (Dolan, 1990). When \( O^6 \)-BG was administered to mice with tumor xenografts, AGT was inactivated for up to 24 h, and tumors previously resistant to BCNU, became sensitive (Dolan, 1990; Mitchell, 1992). In vivo, \( O^6 \)-BG is metabolized to 8-oxo- \( O^6 \)-BG (Figure 1.8) which is very stable and maintains its inactivation properties (Dolan, 1997). These positive results in cell culture and animal model studies led to clinical trials. Early results from clinical trials with \( O^6 \)-BG showed that myeloma patients, when treated with 2-8 cycles of \( O^6 \)-BG and BCNU, had a 30% response rate to the treatment (Bahlis, 2003). In another study, T-cell lymphoma patients had a 60% positive response rate using \( O^6 \)-BG intravenously and BCNU administered every two weeks (Apisarnthanarax, 2002). \( O^6 \)-BG successfully passed through Phase I of clinical testing and is currently in Phase II of clinical trials (Quinn, 2002; Kaina, 2010).

Recent work has demonstrated that short oligodeoxyribonucleotides with a central \( O^6 \)-BG are more potent AGT inactivators than free base \( O^6 \)-BG (Luu, 2002). However, this approach is not conducive to the clinical setting. The dosage of \( O^6 \)-BG and similar free base AGT inactivators, in combination with alkylating drugs, is limited by toxicity due to myelosuppression. Therefore, local drug administration is the best alternative (Verbeek, 2008). Some studies have shown that human AGT polymorphisms cause resistance to \( O^6 \)-BG (Pegg, 2007). The currently
known AGT variants, caused by single nucleotide polymorphisms (SNPs), are W65C, L84F, I143V/K178R and G160R (Pegg, 2007). This raises concern about selection of AGT-resistant variants during chemotherapeutic regimens. The use of $O^6$-BG and other AGT inhibitors does not have side effects for patients receiving chemotherapy, however, the AGT inhibitors reduce AGT levels in normal tissue, which exacerbates the toxic effect of alkylating agents used for chemotherapy (Kaina, 2010). Therefore, the need to improve the current design of AGT inactivators is urgent, and the search for potent AGT inhibitors will continue.

Figure 1.8 Structures of two free base inactivators of hAGT: $O^6$-benzylguanine ($O^6$-BG) and $O^6$-methylguanaine ($O^6$-MG). $O^6$-BG is oxidized to $O^6$-Benzyl-8-oxoguanine within human cells.
Figure 1.9 Reaction of AGT protein with $O^6$-BG. Modified from Dolan and Pegg, 1997. $O^6$-BG acts as a pseudo-substrate to AGT and serves as an inactivator by transferring its benzyl group to the reactive cysteine of AGT.
1.5 AGT Regulation

1.5.1 Transcriptional Level

The human AGT gene is located at the telomeric end, 26q of the long arm of chromosome 10. The gene consists of five exons, spanning 300 Kb of genomic DNA. The promotor extends from the first exon into the first intron. It contains CpG regions and lacks both TATA and CAAT boxes. In addition, there are 10 Sp1 transcription factor binding sites, two glucocorticoid response elements and two AP-1 sites within the promoter region (Harris, 1991). These features indicate that AGT is constitutively expressed in all tissues (Pegg, 2000). A 59bp element present at the first exon/intron boundary has been demonstrated to be an enhancer (Harris, 1991). The chromatin structure of the AGT promoter may also play a role in AGT transcriptional regulation (Costello, 1994). Most regulation of mammalian AGT appears to be at the transcription level (Pieper, 1997). The half life of the AGT mRNA transcript from various cell lines has been shown to be approximately 10h (Fritz, 1991; Erickson, 1995). DNA hypermethylation has been demonstrated to repress transcription of the AGT gene, in some cancer cell lines (Pieper, 1997). Mammalian AGT appears to be modestly inducible in response to alkylation damage, while the *E.coli* Ada alkyltransferase shows significant induction (Fritz, 1991; Sedgwick, 2002).

1.5.2 Post-translational Level

The role of AGT phosphorylation is unclear. AGT has been shown to be the target of phosphorylation in human brain tumor cells and the phosphorylated form of AGT has lower efficiency in $O^6$-Methylguanine repair. (Srivenugopal, 2002; Mullapudi, 2000). One report shows that phosphorylation at residue 204 causes resistance to proteolytic degradation (Lim, 2000).
1.6 AGT Degradation

AGT is stable in vivo when unalkylated. Upon alkylation, it is rapidly ubiquitinated and degraded (Pegg, 1991; Xu-Welliver, 2002; Srivengopal, 1996). Little is known about the ubiquitin machinery operating on mammalian AGT upon alkylation, or the sequence of release of AGT from DNA. Evidence shows that an inactive AGT mutant C145A, and alkylated AGT actually impede DNA repair by other AGT molecules, by continuing to bind to DNA (Edara, 1999). Experiments have shown that alkylation causes a conformational change in AGT, which sensitizes the alkylated AGT to degradation (Oh, 1996; Kanugula, 1998; Federwisch, 1997). It is possible that this conformational change allows recognition by the Ub enzymes, which then allows release of the alkylated protein from DNA. The specific site of AGT ubiquitination or the sequence of events in release of the alkylated AGT from DNA remains unknown. In 2009, Hwang et al described two synergistic ubiquitin degradation pathways that target Mgt1 (Saccharomyces cerevisiae alkyltransferase). The first pathway is the Ubr1/Rad6 dependent N-end rule pathway, and the second is the Ufd4/Ubc4-dependent ubiquitin fusion degradation (UFD) pathway. The N-end rule pathway functions through E3 ubiquitin ligases recognizing degradation signals (degrons) at the N-terminus of the substrate protein (Varshavsky, 2008). Destabilizing residues at the N-terminus constitute an N-degron. The UFD pathway in S.cerevisiae operates through a HECT-domain E3 enzyme (Ufd4), together with E2 enzymes Ubc4 or Ubc5. This pathway is responsible for increased polyubiquitin processing of its target substrate. Preliminary experiments have shown that the mouse homologs of the N-end rule pathway interact with mouse AGT in vivo, and mouse fibroblasts deficient in Ubr2 (mouse homolog for yeast Ubr1) were hypersensitive to the alkylating agent mitomycin C. These indirect data suggests that the N-end rule might play a role in mammalian AGT degradation (Hwang, 2009).
Experiments were performed to find the primary ubiquitination site of human AGT in which all twelve of the lysine residues were mutated individually to arginine. However, none of these alterations caused a change in ubiquitination pattern (Xu-Welliver, 2002). This strongly suggests that multiple ubiquitination sites exist in AGT.

Intriguingly, hAGT has been found to interact with BRCA2, and undergoes degradation upon BG treatment (Philip, 2008). BRCA2 has been demonstrated to be a part of a multiprotein complex involving ubiquitin E3 ligase activity (Dong, 1998). A 29 amino-acid deletion in BRCA2, in mouse embryonic fibroblasts, abolishes binding with AGT and increases sensitivity to alkylating agents. Therefore, BRCA2 might play a role in signaling ubiquitination of AGT.

In the clinical realm hAGT is targeted for degradation as a strategy for enhancing the effectiveness of alkylating agent chemotherapy. Use of $O^6$-BG and other AGT inhibitors in conjunction with alkylating agents in clinical trials have not been able to greatly increase the therapeutic index of these drugs (Quinn, 2002; Quinn, 2009; Batt, 2007; Ryan, 2006). Therefore, it is critical that we understand the regulatory mechanisms leading to AGT degradation and the turnover of AGT in mammalian systems.

1.7 Goal and Specific Aims

AGT plays an important dual role in preventing mutation in DNA, as well as causing resistance to chemotherapy in tumors. Our overall goal was to understand the fate of the alkylated AGT within mammalian cells, and determine the timecourse within which ubiquitination occurs after alkylation. Our hypothesis was that alkylation was the critical event leading to AGT degradation. In addition, we postulated that there was a primary ubiquitination site within the AGT sequence that was recognized by the Ub machinery. We used Chinese Hamster Ovary (CHO) cells for these experiments because endogenous AGT is silenced in these cells; therefore, we could study the effect of expression and regulation with no ambiguity (Loktionova, 1996). In addition, we addressed the role of the amino terminus in the degradation of the protein. We used a
unique, sensitive and specific method to create a fluorescent AGT at the point of alkylation and visualize living cells. These studies have not been used previously to characterize AGT, and provide new information and methodology. We also used a GFP-AGT construct to visualize and study the effect of alkylation on intracellular distribution. The goal of this dissertation was to provide insight into the physiological regulation of AGT post-alkylation, in the hope of augmenting the efficacy of strategies for inactivation of the protein in chemotherapy.
Chapter 2

Materials and Methods

2.1 Materials

All primers were synthesized at the Penn State Hershey Core Research Facilities. Isopropyl β-D-thiogalactopyranoside (IPTG), ampicillin, kanamycin, hemocyanin, MNNG, Coomassie Blue, paraformaldehyde, cycloheximide, lactacystin and most other biochemical reagents were purchased from Sigma Chemical Co., St. Louis, MO. MG132 and the protease inhibitor cocktail used in cell extracts was purchased from Calbiochem, La Jolla, CA. O\textsuperscript{6}-Benzylguanine, O\textsuperscript{6}-Methylguanine, BGFL and BGAF were synthesized as described by Dolan et al and Keppler et. al. (Keppler, 2004; Dolan, 1990), and generously provided by Dr. R.C. Moschel, ABL-Basic Research Program, National Cancer Institute-Frederick Cancer Research and Development Center, Frederick, MD. Talon Metal Affinity Resin was purchased from Clontech, Palo Alto, CA. E.coli strain GWR109 was kindly provided by Dr. L. Samson, Department of Molecular and Celluar Toxicology, Harvard School of Public Health, Boston, MA. \[^{3}H\]methylated DNA, along with bacterial vector pQE-30-hAGT, mammalian vector pCMV-G160R, and wtAGT CHO cell line were previously made in the laboratory (Loktionova, 1996; Edara, 1996). The GFP-P140K CHO cell line and MFGS-GFP-P140K vector was generously provided by Dr. H. L. Malech, National Institute of Health, Bethesda, MD. Monoclonal AGT Ab, MT3.1 was purchased from NeoMarkers, Fremont, CA. AGT antibodies ATO-1 and ATO-3 were developed as described by Pegg et al (Pegg, 1991), and were previously prepared in the laboratory. Polyclonal AGT Ab was purchased from Novus Biologicals, Littleton, CO. Polyclonal GFP Ab and penta-His Ab was purchased from Santa Cruz Biotechnology Inc., Santa Cruz, CA. β-actin and GAPDH antibodies were purchased from Cell Signaling technology, Danvers, MA.
2.2 Cell Lines Used

Empty-vector CHO cells were used as a control in most experiments. CHO cells expressing wt-AGT and AGT-His were used to study the in vivo properties of AGT. Stable cell lines of CHO cells expressing G160W and triple mutant N157G/S159E/G160W were made to study the effect of AGT mutants in vivo. The G160W and N157G/S159E/G160W CHO cell lines were created using Lipofectamine 2000 (Invitrogen, Carlsbad, CA), according to the manufacturer’s instructions. For intracellular localization studies and protein dynamics, GFP-AGT CHO cells were used. Finally, a transient expression of 16kDa in CHO cells was studied to complement in vitro studies.

2.2.1 Cell Culture

CHO cells expressing wt-AGT and AGT-His, were grown in α-MEM media (Gibco BRL, Gaithersburg, MD), supplemented with 10% fetal bovine serum and 100 U/mL Penicillin and 100 U/mL streptomycin, and maintained at 37°C and 5% CO₂. Stably transfected cells were grown in media containing 1mg/mL geneticin/G418 sulfate (Gibco BRL, Gaithersburg, MD) to prevent plasmid loss.

2.2.2 Labeling CHO cells with BGAF

CHO cells were treated with 50μM BGAF for 1h, then washed extensively with PBS, before being imaged under a fluorescent microscope, in OPTI-MEM (Gibco BRL, Gaithersburg, MD).
2.3 Construction of Plasmid Vectors

2.3.1 Generating AGT mutants to determine sensitivity to O\textsuperscript{6}-BG

Six mutants were constructed to test for sensitivity to O\textsuperscript{6}-BG. The primers used for vector construction are listed in Table 2.1. *Pfu* polymerase (Stratagene, La Jolla, CA) was used to generate the PCR product. The PCR reaction was carried out according to the manufacturer’s instructions. PCR product DNA was incubated with restriction enzymes *Dra*\textsubscript{III} and *Age*I (New England Biolabs Inc, Beverly, MA) to create an 85bp fragment (Figure 2.1A). *Dra*\textsubscript{III} was embedded into the forward primers, while *Age*I was present in the AGT gene. N157G, S159E, G160W, N157G/S159E, N157G/S159E/G160A, N157G/S159E/G160W fragments were then ligated into pQE-30 (Qiagen, Chatsworth, CA), for expression in XL1-Blue *E.coli* (Stratagene, La Jolla, CA). Clones were isolated, and the vector DNA was verified by sequencing at the Penn State Hershey Core Research Facilities.

2.3.2 pCMV-G160W and pCMV-N157G/S159E/G160W for expression of mutants in CHO cells

pQE vectors expressing either G160W or triple mutant, N157G/S159E/G160W were digested using restriction enzymes *Eco*\textsubscript{NI} and *Age*I (New England Biolabs Inc, MA) to produce a 422 bp fragment. The pCMV vector (Baker, 1990) has unique *Eco*\textsubscript{NI} and *Age*I restriction sites (Figure 2.1B). The purified fragment was ligated into digested pCMV vector, and used to transform *E.coli*. The pCMV-G160W and pCMV-N157G/S159E/G160W vectors purified from *E.coli* clones were sequenced to confirm the absence of any unintended mutations.

These pCMV constructs (8\mu g) were transfected into CHO cells using Lipofectamine 2000 (Invitrogen, Carlsbad, CA), according to the manufacturer’s instructions.
2.3.3 pCMV-wt-hAGT-His for expression of C-terminal His tagged AGT in CHO cells

The bacterial expression vector pQE-30-hAGT-His (Fang, 2005) expresses AGT with 6 Histidines in place of the last six amino acids of hAGT (Figure 2.1A). A 526 bp fragment containing the AGT-His gene was isolated from pQE-30-hAGT-His using restriction enzymes EcoNI (New England BioLabs Inc, Beverly, MA) within the AGT sequence and KpnI (Stratagene, La Jolla, CA) downstream of AGT. These restriction sites are unique in both the pQE30 and pCMV, the mammalian expression vector. The digested DNA fragment was purified from a 1% agarose gel using a Qiaquick Gel Extraction kit (Qiagen, Chatsworth, CA). pCMV G160R vector backbone was also digested with EcoNI and KpnI and purified in the same manner as the PCR product. The purified fragment was then ligated into the pCMV vector (Figure 2.1B). Since the EcoNI restriction site lies upstream of AGT residue 160, the exchange of sequences removes the G160R mutation from the vector.

2.3.4 Generating an AGT-His CHO cell line

CHO cells were transfected with pCMV-wt-hAGT-His using Lipofectamine 2000 (Invitrogen, Carlsbad,CA). After a 3h incubation, cells were trypsinized and transferred to a T-75 flask. After 24h, 40µM O6-BG was added for 2h. 20µM BCNU was then added for 2h. Media was changed and cells were diluted 1:20 and split into 100mm plates. Cells were grown in media containing G418 sulfate (Gibco BRL, Gaithersburg, MD) and colonies were picked and grown over two weeks. Western blots were used to determine which colonies had the highest wt-AGT-His protein expression. RNA was isolated from three positive colonies, and RT-PCR was used to generate a PCR product spanning the AGT site. This product was sequenced to ensure no mutations were created in AGT due to BCNU treatment.
**2.3.5 Generating GFP-AGT fusion protein expression in CHO cells:**

GFP-AGT, was generated using the MFGS-GFP-P140K vector (Figure 2.1C) (Choi, 2004), a Moloney murine leukemia virus replication-incompetent vector (Cell Genesys, Foster City, CA), using restriction enzymes *Nco*I and *Bam*HI (New England Biolabs, Beverly, MA)(Figure 2.1C). In this vector, GFP is present as an N-terminal fusion to AGT. The vector construct contains a five amino acid linker between GFP and AGT. The linker sequence was TCCGGCCTCAGATCT. GFP-P140K AGT was used as a template, and mutated to the wildtype sequence using a Stratagene Quikchange site directed mutagenesis kit (Stratagene, La Jolla, CA). The primers used for this reaction are listed in Table 2.2. The vector was sequenced to confirm that no mutations were introduced. Viral titre production and transduction was carried out by Dr. U. Choi, National Institute of Health, Bethesda, MD, as described by Choi et al (Choi, 2004) Transduced cells were analyzed for fluorescence using a FACSort (BD Immunocytometry System, with CellQuest Software).

**2.3.5 Construction of 16kDa truncated hAGT in vectors pQE-30 and pCMV**

A Methionine was added to the cleavage site, at position 51 of hAGT to facilitate translation in *E.coli*. The primers used for the PCR reaction are listed in Table 2.2. All the primers had the restriction enzyme sites embedded in them, except for the reverse primer used for pQE30. This reverse primer complemented a region downstream of the AGT gene. The PCR reaction was carried out using an AccuPrime *Pfx* DNA polymerase (Invitrogen, Carlsbad, CA) according to the manufacturer’s instructions. The PCR product was digested with restriction enzymes *Bse*RI and *Kpn*I (New England Biolabs, Beverly, MA), resulting in a 476bp fragment. Within the pQE-30 vector, the AGT gene spans from position 115 to 738. The use of restriction enzymes *Bse*RI and *Kpn*I on the pQE30 vector, excises the entire AGT gene. In order to clone 16KDa AGT into the mammalian expression vector pCMV, restriction enzymes *Sal*I and *Bam*HI
(New England Biolabs, Beverly, MA), were used to excise the PCR product resulting in a 487bp fragment. Similar to the cloning of pQE-30, the use of SalI and BamHI in pCMV excises the entire AGT gene from the vector. The digested DNA fragments were run on a 1% agarose gel, and purified as described above. This vector was then transformed into a heat sensitive XL1-Blue E.coli strain (Stratagene, La Jolla, CA). Colonies were selected on LB-Ampicillin plates (50µg/mL) and tested for insertion by a further restriction enzyme digestion of the plasmid DNA, and then tested for correct alignment and sequence by DNA sequencing at the Penn State Hershey Core Facilities.

2.4 Purification of wt and Mutant AGT Proteins

XL-1 Blue E.coli cells (Stratagene, La Jolla, CA) transformed vector constructs (Table 2.1) were grown in 500mL LB medium supplemented with 50µg/mL ampicillin and the AGT proteins purified by affinity chromatography, as described by Edara et. al. (Edara, 1996) Briefly, 150µM IPTG was added to logarithmically growing cells, which were harvested 4 h later. The pellet was resuspended in purification buffer (20 mM Tris-HCl (pH 8.0) and 500 mM NaCl) and sonicated. After centrifugation at 4°C, the supernatant was applied to a Talon metal affinity resin equilibrated with purification buffer. The column was washed extensively and the AGT protein was eluted using 200 mM imidazole. A small aliquot of protein fraction was tested for protein concentration using a Bio-Rad dye (Bio-Rad, Hercules, CA), and the fractions with the highest concentration, were pooled and dialyzed with a dialysis buffer (50mM Tris-HCl, pH 7.6, 250 mM NaCl, 5mM dithiothreitol and 0.1mM EDTA) for 18 h at 4°C. The protein was then concentrated with a Centricon 30 device (Amicon, Beverly, MA).
2.5 Biochemical characterization of AGT mutants

2.5.1 Determination of AGT Activity and Effects of $O^6$-BG on AGT Activity using Filter Binding

The AGT filter binding assay was done essentially as described by Kanugula et al (Kanugula, 1995). AGT activity of purified protein was measured by adding different concentrations of the protein in 0.5 ml 50mM Tris-HCl (pH 7.6), 5 mM DTT, 0.1mM EDTA, containing 50µg haemocyanin and 10µg calf-thymus DNA and incubating it with $[^3]H$ methylated DNA substrate for 30min at 37°C. The $[^3]H$ methylated DNA substrate was prepared by allowing calf-thymus DNA to react with $[^3]H$ methyltnitrosourea, as described by Dolan et al (Dolan, 1990). The AGT activity was calculated by measuring $[^3]H$ methylated protein formed after collection on 0.45µm HA nitrocellulose filters (Millipore, Billerica, MA). The reaction was terminated by adding 5mL of denaturation buffer (8M urea, 0.5M NaCl, 10mM Tris HCl, pH 7.6). The reaction was collected onto 25mm$^2$ nitrocellulose filters in a 12-well sample holder (Millipore, Billerica, MA), and these filters were washed twice with 5mL denaturation buffer and twice with 5mL 10% ethanol in water. The filters were dried and the radioactivity counted in Econofluor scintillation fluid using a Beckman Coulter LS 6500 multipurpose scintillation counter. Each experiment contained blanks, to determine the radioactivity present without AGT, and the positive control had AGT in excess to determine the maximum amount of radioactive methyl adduct removed from the DNA. Each sample was tested in duplicate. The results were expressed as a percentage of the AGT activity remaining.

To study the inactivation effect of $O^6$-BG on AGT activity, the purified protein was pre-incubated with different drug concentrations for 30min at 37°C before the incubation with $[^3]H$ methylated DNA substrate as described above. Plots of remaining AGT activity vs $O^6$-BG
concentration were used to calculate an ED$_{50}$ value representing the amount of inhibitor needed to produce a 50% loss of activity. These experiments were repeated twice.

### 2.5.2 Determination of hAGT and mutant Rate constants

Reactions of AGT and [$^3$H] methylated DNA were carried out as described by Goodtzova et. al (Goodtzova, 1998). Briefly, [$^3$H] methylated DNA is used as a substrate for AGT, and allowed to react with AGT for varying time periods. The incubation times of the purified protein and the DNA substrate, at 37°C, varied from 0.5min to 3.5min. The reaction is quenched by the addition of denaturation buffer as described above. The quantity of [$^3$H] methyl groups transferred onto AGT is determined by measuring radioactivity trapped on AGT bound to a nitrocellulose filter. The rate constant,

\[
KT = \frac{1}{C_0^a - C_0^b} \ln \left[ \frac{C_0^b (C_0^a - C_c)}{C_0^a (C_0^b - C_c)} \right].
\]

\(C_0^a\) = initial concentration of AGT = 1pmole or 0.03x10$^{-6}$ moles/L

\(C_0^b\) = initial concentration of [$^3$H] DNA = 3.03x10$^{-9}$ moles/L

\(C_c\) = Concentration of product at time t

\(C_c\) was calculated by using the mean cpm at each timepoint from four experiments.

### 2.6 Preparation of cell culture extracts for SDS-PAGE and immunoblotting

Cells were trypsinized and pelleted at 4°C, 1500 rpm for 5 min. The cells were then resuspended in harvest buffer (50mM Tris-HCl pH 8, 5mM EDTA, 5mM NEM, 1mM PMSF) with 1X concentration of protease inhibitor cocktail. The cell suspension was sonicated and centrifuged at 4°C. Protein estimation was done using the Bio-Rad Protein Assay system (Bio-Rad, Hercules, CA). Loading dye (250mM Tris-HCl pH 6.8, 10% SDS, 50% glycerol, 0.02% bromophenol blue) was added to each sample and boiled for 5 min, and then loaded onto a 15% SDS-polyacrylamide gel for SDS PAGE. Gels were run at 25mAmp and transferred to a PVDF
transfer membrane (Pall Corporation, Port Washington, NY) with 100mAmp for 2h, or 20V overnight.

Preparations separated by SDS-PAGE and transferred to the PVDF membrane were treated for 1h with blocking buffer [either 5% milk in TBS-T (20mM Tris-HCl pH 7.6, 150mM NaCl, 0.1% Tween 20) or 3% BSA in TBS-T], then washed twice for 10 min, with TBS-T. The membrane was incubated with 1:1000 primary Ab dilution in TBS-T for 1h or overnight, then washed again with TBS-T twice for 10 min. 1:2000 Secondary Ab (horseradish-peroxidase conjugated species specific from Cell Signaling Technology, Danvers, MA) was added in blocking buffer for 1h. The membrane is finally washed with TBS-T. Bound Ab was detected using chemiluminescence as described by Cell Signaling (HRP Western Blot detection system, Danvers, MA).

2.7 Degradation studies of AGT in CHO cells

2.7.1 AGT Activity Assay in CHO cells

AGT activity in the CHO cell lines was tested, using HPLC as described by Dolan et al. (Dolan, 1991). Reactions contained 11µL [³H] methylated DNA, 1mg/mL calf thymus DNA, 0.5 M Tris, pH 7.8, 0.5 M DTT and varying cell extract concentrations in a total reaction volume of 1mL. Reactions were incubated at 37°C for 30 min, and Cold perchloric acid was added in a final concentration of 0.25N. 100mg of BSA was used in place of AGT and served as a blank. All the tubes were vortexed, then centrifuged at 15,000 g for 10min. The supernatant was discarded and the pellet was resuspended in 1mL of 0.1N HCl. The pellet was hydrolyzed at 70°C for 30min, before centrifuging for 10min at 15,000 g. The supernatant was saved and the pellet was re-hydrolyzed in 500µL 0.1N HCl at 70°C for 30min, and centrifuged once more at 15,000 g. The supernatants were then combined and filtered through a 0.45µm HA nitrocellulose membrane (Millipore, Billerica, MA). The supernatant was run through an HPLC system and activity
calculated by comparing $O^6$-methylguanine adducts in untreated samples with the AGT treated samples. The results from two experiments are shown in Table 2.3. All the cell lines had comparable alkyltransferase activity to wt AGT.

### 2.7.2 Measurement of AGT and GFP-AGT Half-life

CHO cells expressing wtAGT or GFP-AGT were treated with 0.2mM cycloheximide in the presence or absence of 80µM $O^6$-BG in order to alkylate and inactivate AGT. Cells were harvested in PBS at various timepoints up to 28h following drug application. Cells were centrifuged at 1500 rpm and resuspended in 200 µL of ice cold harvest buffer (50 mM Tris-HCl pH 7.5, 0.1 mM EDTA, 5mM dithiothreitol and 1 U/ml of protease inhibitor cocktail) and sonicated on ice for 1 min with 30s pulses on setting three, using a Sonicator Ultrasonic Processor (Mosonix, Farmington, NY). Protein concentration in the cell lysates were determined using Bio-Rad Protein Assay (Bio-Rad, Hercules, CA). 60µg of extract from each timepoint was analyzed by Western analysis for AGT content.

### 2.7.3 Detection of AGT-ubiquitin complexes

CHO cells expressing wtAGT, AGT-His or GFP-AGT were treated with 80µM $O^6$-BG or BGAF to alkylate AGT and 25µM MG132 (Calbiochem, La Jolla, CA) or 30µM lactacystin to inhibit proteasomal degradation and increase the ubiquitin-AGT conjugates within the cells. The cells were harvested in cold PBS, centrifuged at 1500 rpm for 5 min and resuspended in 250µL harvest buffer (50mM Tris-HCl pH 8, 5mM EDTA, 5mM NEM and 1mM PMSF). Cells were then sonicated as described above and centrifuged for 5min at 5000 rpm at 4°C.
2.8 GFP-AGT mobility in CHO cells

2.8.1 FRAP: Fluorescent Recovery After Photobleaching

A Leica TCS SP2 AOBS Confocal Microscope was used to visualize and study CHO cells expressing GFP-AGT. The protocol was developed for use of CHO cells expressing hAGT, based on assays described by Gomez et al (Gomez, 2006). For FRAP experiments, CHO cells were grown on glass-bottom cell culture plates up to 80% confluence, placed on a heated stage at 37°C, with a CO₂ chamber and stained with Hoechst at 1μg/ml. The Hoechst stain binds to DNA and is used to visualize the nucleus. Cells were visualized in OPTI-MEM media (Gibco BRL, Gaithersburg, MD), with the 488 He/Ar laser with 6% output. The 405 laser was used to visualize the Hoechst stain. The mode setting used was xyt, PMT 2 was 517. The objective lens used was 63x oil immersion. For FRAP experiments only the He/Ar laser was used, i.e., the 405 laser was turned off, to minimize photobleaching. The settings for bleaching the entire nucleus, or half-nucleus using the FRAP application software were: Prebleach for 5 frames at 0.5 s, Bleach for 25 frames at 0.5s, Postbleach for 20 frames at 0.5s, 10 frames at 30s and 30 frames at 60s. In order to bleach the fluorescence, the He/Ar laser output was increased to 50% and the laser speed was increased to 800Hz. These settings were optimized for GFP-AGT CHO cells. All quantification and collection of data was done using the Leica imaging software.

2.8.2 Cell Fixation using Paraformaldehyde

Fixed cells were used as a control to the FRAP experiments. 4% Paraformaldehyde solution was freshly prepared in ddH₂O. 1M NaOH is used to dissolve paraformaldehyde and the solution is heated to 65°C until clear. 0.1% 10x PBS was added and the pH adjusted to 7.2-7.5 using 1M HCl. Cells grown on coverslips were first stained with Hoechst at 1μg/ml for 5 min at 37°C in PBS. Cells were then washed in PBS. Next, 4% paraformaldehyde was added and
incubated for 5 min at room temperature. The coverslips were drained and inverted onto a single drop of AquaPolyMount (Polysciences Inc, Warrington, PA) on a microscope slide. Slides were visualized after 24h and stored at 4°C.

2.9 Characterization of 16kDa AGT species

2.9.1 Isolation of 16kDa AGT species using AGT-His CHO cells

Pellets of AGT-His cells were suspended in 5mL of resuspension buffer (50mM Tris-HCl pH 8, 500mM NaCl, 5% glycerol, 7.5mM imidazole). The cells were sonicated for 2min at setting 3 with 10 second pulses, then centrifuged for 10minutes at 10,000 g at 4°C. 400µL of Talon bead suspension was added to a 20mL protein purification column (Bio-Rad, Hercules, CA). The cell lysate (approximately 25mg) was passed over the Talon bead resin, then the beads were washed extensively with wash buffer (50mM Tris-HCl pH 8, 150mM NaCl, 5% glycerol, 7.5mM imidazole). 100mM imidazole was used to elute the protein from the beads, and collected into fractions, which were concentrated using a 10,000 MWCO concentrator (Amicon, Beverly, MA), in a centrifuge at 5000 g at 4°C. Once the volume was reduced to 100µL, 2mL of dialysis buffer was added (50mM Tris-HCl pH 7.6, 150mM NaCl, 5% glycerol, 1mM DTT). This suspension was concentrated again and analyzed with SDS-PAGE. Proteins were visualized either by staining with Coomassie Blue or by Western analyses.

2.9.2 Cell Fractionation

40mg of cells were used for cell fractionation, using the NE-PER Nuclear and Cytoplasmic Extraction reagents (Pierce Biotechnology, Rockford, IL), according to the manufacturer’s instructions.
2.9.3 *In vitro* formation of 16kDa species using empty-vector CHO cell extracts

Empty-vector CHO or HeLa cell extract was prepared as described above, and 1x protease inhibitor cocktail was added. To pre-alkylate the AGT, 50µM O<sup>6</sup>-BG or BGFL was used at 37°C for 30 min. 20µL of cell extract was added to a reaction buffer (40mM Tris pH 7.6, 5mM MgCl<sub>2</sub>, 2mM ATP, 2mM DTT), followed by the addition of 1.4µg AGT for each reaction.

The reaction was stopped at various times after AGT addition by adding SDS loading buffer and boiling for 5 min. The sample was then run on a 15% polyacrylamide gel and imaged between the glass plates using a 488nm laser in a FX-Pro Plus fluorimager (Bio-Rad, Hercules, CA), using Quantity One ID software. The gel was then subjected to Western blot analysis, as described above.

2.9.4 Sample preparation for Mass Spectrometry

The protocol was adapted from the methods described by Speicher et al. (Speicher, 2000). Isolation of the 16kDa species was done as described above. Proteins were separated on a 15% polyacrylamide gel allowed to polymerize over-night. The gel was then stained with Coomassie Blue and washed extensively with ddH<sub>2</sub>O. The bands of interest were then cut and subjected to an in-gel trypsin digestion, as follows. Eppendorf tubes were washed with 50% acetonitrile, 0.1% trifluoroacetic acid (TFA). The gel bands were added into these tubes, destained twice with 200µL of 100mM NH₄CO₃ (pH 8.0) and 50% acetonitrile for 45 min at 37°C, and dried in a SpeedVac. 100µL of solution containing 10mM DTT, 25mM NH₄CO₃ (pH 8.0) was added to the gel slices for 15 min at 37°C. Next, the gel slices were treated with 100µL of 20mM iodoacetamide in 25mM NH₄CO₃ (pH 8.0) for 15min, with agitation, and dried in a SpeedVac. The gel slices were rehydrated with 20µL of 0.02µg/µL of sequencing grade modified trypsin (Promega Corporation, Madison, WI) in 10% acetonitrile, 40mM NH₄CO₃ (pH 8.0), 0.1%w/v n-octylglucoside, and incubated for 1 h at room temperature. The residual liquid was
removed and the gel slices were covered with 50µL 10% acetonitrile, 40mM NH₄CO₃ (pH 8.0), 0.1% w/v n-octylglucoside, and incubated for 4-6 h at 48°C. The supernatant was removed and saved, while 50µL 0.1% TFA was added to the gel fragment and incubated for 1h at 37°C. The supernatant was collected and combined with the previously saved supernatant. This solution was dried in a SpeedVac and resuspended in 200µL of water. This step was repeated three times, with the final resuspension volume of 10µL.

Samples were analyzed in an ESI linear trap ion mass spectrometer (LTQ, Thermo-Fisher, San Jose, CA) connected to a Waters Acquity UPLC system using an Aquity UPLC BEH C18 column (1.7 µm, 1.0 mm x 100 mm), at the Penn State Hershey Core Facilities.

2.9.5 Sample preparation for Edman Sequencing

The protocol was adapted from the methods described by Mozdzanowski et al. (Mozdzanowski, 1992) Pulldown of 16kDa and purified hAGT were run on a 15% polyacrylamide gel allowed to polymerize overnight. The gel was pre-run at 3mM constant current for 2h prior to loading the sample. Tris-glycine buffer (25mM Tris, 190mM glycine, 0.1% SDS, 20% methanol, pH 8.5) was used for the transfer of the proteins to the PVDF membrane. After the transfer, the PVDF membrane was saturated with methanol for a few seconds, then stained with 0.1% Coomassie Blue R250 in 40% methanol, 1% acetic acid, for 30 s. The membrane was de-stained with 50% methanol and finally rinsed with water. The bands of interest were then sent for sequencing to the Protein Microanalytical Laboratory, University of Pittsburgh, Pittsburgh, PA.

2.9.6 MNNG Protection Assay

The assay was done essentially as described by Fang et al. (Fang, 2005) The GWR109 E.coli strain lacks endogenous AGT activity (Rebeck, 1991) at 37°C, in a temperature-controlled
shaker. 1% of this overnight culture was used to inoculate 10mL of LB-Amp-Kan media at and allowed to grow to an OD$_{600}$ of 0.6. To induce AGT expression, 150µM IPTG was added, and the cultures were incubated at 37°C for 30 minutes. 1 mL cultures were pipetted into Falcon 2059 tubes (BD Biosciences, Franklin Lakes, NJ) and MNNG solution and DMSO (solvent control) was added in varying concentrations for 30 minutes, on the shaker at 37°C. The culture was then centrifuged at 3000 g for 5 minutes and resuspended in 1 mL M9 media. Dilutions were plated in triplicate onto LB plates with 50µg/mL ampicillin and 50µg/mL kanamycin and incubated at 37°C for 16 h. The colonies were then counted and expressed as colonies/mL to calculate the percent survival. This experiment was repeated twice and a two-way ANOVA was used to evaluate the data.
Figure 2.1 Diagram of plasmids (A) pQE-30 bacterial vector (Qiagen, Chatsworth, CA) expressing wt hAGT with a C-terminal His tag. This vector has a pT5 promoter and an ampicillin resistance gene. The restriction sites used for plasmid construction are indicated in black, with the specific plasmid location in parentheses.
Figure 2.1B. Mammalian expression vector, pCMV-Neo Bam. This vector was a kind gift from the Vogelstein lab (Baker, 1990). The vector contains a CMV promoter and enhancer sequences. In addition it contains polyadenylation sites to ensure efficient transcription. The origin of replication (pBR322) and beta-lactamase facilitates growth of the plasmid in *E.coli*. 
Figure 2.1C. MFGS viral vector expressing GFP-P140K AGT. GFP is fused to the N-terminus of AGT. A five amino acid linker is present between GFP and AGT. This vector was a kind gift from the Malech Lab.
<table>
<thead>
<tr>
<th>Plasmids Constructed</th>
<th>Experimental Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>pQE-30 N157G</td>
<td>Purification of N157G mutant AGT</td>
</tr>
<tr>
<td>pQE-30 S159E</td>
<td>Purification of S159E</td>
</tr>
<tr>
<td>pQE-30 G160W</td>
<td>Purification of G160W</td>
</tr>
<tr>
<td>pQE-30 N157G/S159E</td>
<td>Purification of N157G/S159E</td>
</tr>
<tr>
<td>pQE-30 N157G/S159E/G160A</td>
<td>Purification of N157G/S159E/G160A</td>
</tr>
<tr>
<td>pQE-30 N157G/S159E/G160W</td>
<td>Purification of N157G/S159E/G160W</td>
</tr>
<tr>
<td>pQE-30 16kDa AGT</td>
<td>Purification of 16kDa AGT</td>
</tr>
<tr>
<td>pCMV- G160W</td>
<td>Stable Expression of G160W in CHO cells</td>
</tr>
<tr>
<td>pCMV N157G/S159E/G160W</td>
<td>Stable Expression of N157G/S159E/G160W in CHO cells</td>
</tr>
<tr>
<td>pCMV 16kDa AGT</td>
<td>Stable Expression of 16kDa AGT in CHO cells</td>
</tr>
<tr>
<td>MFGS GFP-AGT</td>
<td>Creation of viral titre expressing GFP-AGT</td>
</tr>
</tbody>
</table>

Table 2.1. Plasmid Constructions
<table>
<thead>
<tr>
<th>Primer/Construction</th>
<th>Primer sequence</th>
<th>Embedded restriction site</th>
</tr>
</thead>
<tbody>
<tr>
<td>N157G forward primer</td>
<td>5’ CCATCCAGATGGTGCAGCTGAGCAGGGATCTCTCGAGGACCTCTCATG 3’</td>
<td>DraIII</td>
</tr>
<tr>
<td>S159E forward primer</td>
<td>5’ CCATCCAGATGGTGCAGCTGAGCAGGGATCTCTCGAGGACCTCTCATG 3’</td>
<td>DraIII</td>
</tr>
<tr>
<td>G160W forward primer</td>
<td>5’ CCATCCAGATGGTGCAGCTGAGCAGGGATCTCTCGAGGACCTCTCATG 3’</td>
<td>DraIII</td>
</tr>
<tr>
<td>N157G/S159E forward primer</td>
<td>5’ CCATCCAGATGGTGCAGCTGAGCAGGGATCTCTCGAGGACCTCTCATG 3’</td>
<td>DraIII</td>
</tr>
<tr>
<td>N157G/S159E/G160A forward primer</td>
<td>5’ CCATCCAGATGGTGCAGCTGAGCAGGGATCTCTCGAGGACCTCTCATG 3’</td>
<td>DraIII</td>
</tr>
<tr>
<td>N157G/S159E/G160W forward primer</td>
<td>5’ CCATCCAGATGGTGCAGCTGAGCAGGGATCTCTCGAGGACCTCTCATG 3’</td>
<td>DraIII</td>
</tr>
<tr>
<td>Mutants reverse primer</td>
<td>5’ GGAATCTACAGGAGAGGC 3’</td>
<td></td>
</tr>
<tr>
<td>16KDα- pQE-30 forward primer</td>
<td>5’ ATTAAAGAGAAATTAGTGTCTCAGGTACG 3’</td>
<td>BseRI</td>
</tr>
<tr>
<td>16KDα- pQE-30 reverse primer</td>
<td>5’ ACCGGGCCGCAACGCGACG 3’</td>
<td></td>
</tr>
<tr>
<td>16KDα-pCMV forward primer</td>
<td>5’ CGGATCATCAAATGGCTGGCTCTCGAGG 3’</td>
<td>BamH1</td>
</tr>
<tr>
<td>16KDα-pCMV reverse primer</td>
<td>5’ GAAACCGACGTCGACGTG 3’</td>
<td>SalI</td>
</tr>
<tr>
<td>GFP-AGT forward primer</td>
<td>5’ GACGAGGGTCTCTCCTATCATAACCATTCCGTCCCATCCCTCTCCGTCCCTCC 3’</td>
<td></td>
</tr>
<tr>
<td>GFP-AGT reverse primer</td>
<td>5’ GCGACGGTATGAGGCGACGATGTCGTCTC 3’</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.2 Primers used for plasmid construction. Underlined sequences indicate the restriction site embedded into the primer.
<table>
<thead>
<tr>
<th>Cell lysate protein Concentration</th>
<th>AGT form</th>
<th>% $O^6$-Alkyl Adduct Removed</th>
</tr>
</thead>
<tbody>
<tr>
<td>25$\mu$g</td>
<td>Wt</td>
<td>82</td>
</tr>
<tr>
<td>25$\mu$g</td>
<td>AGT-His</td>
<td>84</td>
</tr>
<tr>
<td>25$\mu$g</td>
<td>GFP-AGT</td>
<td>76</td>
</tr>
<tr>
<td>25$\mu$g</td>
<td>GFP-P140K</td>
<td>72</td>
</tr>
<tr>
<td>100$\mu$g</td>
<td>Wt</td>
<td>90</td>
</tr>
<tr>
<td>100$\mu$g</td>
<td>AGT-His</td>
<td>90</td>
</tr>
<tr>
<td>100$\mu$g</td>
<td>GFP-AGT</td>
<td>84</td>
</tr>
<tr>
<td>100$\mu$g</td>
<td>GFP-P140K</td>
<td>89</td>
</tr>
<tr>
<td>100$\mu$g</td>
<td>Empty-vector</td>
<td>8</td>
</tr>
</tbody>
</table>

**Table 2.3.** The alkyltransferase activity in the various CHO cell lines. Results show the average data collected from two experiments using HPLC. Empty-vector CHO cell extract was used as a control; in addition to blanks in which BSA replaced cell extracts.
3.1 Rationale

Alkyl nitrosoureas such as carmustine (BCNU) or procarbazine are used to treat malignant gliomas, myelomas and lymphomas. However, the effectiveness of these agents is limited due to the acquisition of resistance from tumor cells over-expressing AGT (Pegg, 1993; Dolan, 1998; Pegg, 1995). $O^6$-Benzylguanine ($O^6$-BG) was one of the first generation of inhibitors created to combat resistance by selectively inactivating hAGT (Friedman, 2002; Pegg, 1997). A major concern with using $O^6$-BG in chemotherapeutic strategies is the potential for resistance of AGT to $O^6$-BG. Numerous studies have investigated the molecular basis of AGT resistance to $O^6$-BG, particularly as it relates to the context of chemotherapy (Xu-Welliver, 1998; Pegg, 2000). The crystal structure of AGT has revealed the existence of a binding pocket at the active site (Val130 through Gly173), into which any inhibitor of the protein must be able to bind, in order to inactivate AGT. Some of the critical residues known to affect $O^6$-BG sensitivity are shown in Figure 3.1. The effects of amino acids within this pocket have been studied. These investigations reveal that the majority of alterations within the binding pocket of AGT lead to resistance towards $O^6$-BG (Figure 3.2)(Encell, 1998; Christians, 1997).

There are 27 sites in this binding pocket region of AGT which when altered, by a single amino acid change, cause resistance to $O^6$-BG (Loktionova, 2002). In addition, there have been reports of rare polymorphic forms of AGT in human populations. One of these, G160R, was shown to have an ED$_{50}$ for $O^6$-BG 40 times that of the wild type hAGT (Edara, 1996). In 1999, Xu-Welliver et al. carried out a detailed study of mutations at codon 160 and its effect on $O^6$-BG resistance. Only two mutants at the 160 position, G160W and G160A were sensitive to $O^6$-BG,
while 14 alterations caused resistance to $O^6$-BG (Xu-Welliver, 1999; Edara, 1996). It has been well documented that amino acids at sites 140 and 160 are important in the conformation of the active site pocket (Crone, 1994; Pegg, 1993). An additional mutation at position 140, P140A/G160R led to a highly $O^6$-BG resistant form of AGT (Xu-Welliver, 1999). In 2003, Johnsson et al., created a library of hAGT mutants using directed evolution based on phage display, in order to find the most sensitive mutants to $O^6$-BG derivatives. Their studies showed that N157G, S159E and G160W had the highest reactivity against $O^6$-BG, as GST-fusion AGT proteins. These mutant AGT fusion proteins showed greater sensitivity to $O^6$-BG derivatives than the wtAGT (Juillerat, 2003; Keppler, 2004).

The goal of this chapter was to investigate the extent of $O^6$-BG sensitivity of mutations at three sites within AGT (157, 159 and 160) and determine what the effect of multiple alterations would be towards a reaction with $O^6$-BG. Our hypothesis was that altered steric hindrance between residues at codons 157, 159 and 160 of the binding site as well as increased hydrophobic interactions between AGT and $O^6$-BG would cause sensitivity of these mutants to the inhibitor. To directly test this hypothesis, we created 6 AGT mutants, point mutants as well as double and triple combination mutants, and measured DNA repair activity and $O^6$-BG sensitivity. Our long-term goal is to use the $O^6$-BG-sensitive AGT mutants to study AGT dynamics within cells.
Figure 3.1 $O^\delta$-BG docked in the binding pocket of hAGT (Juillerat, 2003) Critical sites are highlighted in red. The cysteine at codon 145, is the alkyl group acceptor site.
Figure 3.2. Sites within the $O^6$-BG binding pocket of human AGT, when mutated cause $O^6$-BG resistance. Highly conserved residues across species are represented in bold type. Stars indicate amino acid mutation sites resulting in $O^6$-BG resistance. 50-fold increases in ED$_{50}$ are represented by two stars and 1000-fold increases in ED$_{50}$ are represented by three stars. Taken from Pegg et. al. 2000. DNA Alterations in Cancer.
3.2 Characterization of purified mutant proteins

Wildtype hAGT and six mutant AGT proteins, N157G, S159E, G160W, double mutant N157G/S159E, triple mutants N157G/S159E/G160A and N157G/S159E/G160W were expressed from bacterial expression vector pQE30. All constructs had a six-histidine tag at the C-terminus of the protein and were purified from E.coli using metal affinity chromatography, as described in Section 2.4. To assess the purity of the proteins, 1µg of the proteins were separated on a 15% polyacrylamide gel and stained with Coomassie Blue. All the recombinant proteins had comparable homogeneity to the wt AGT protein, except for G160W, which had a few non-specific bands (Figure 3.3).

3.2.1 DNA repair activity

To assess the DNA repair activity of each protein, we measured the transfer of [3H] from alkylated DNA to AGT in a filter-binding assay. All the mutants were active (Figure 3.4). In order to further establish the DNA repair activity of these purified proteins and quantitate reaction rates, we carried out timed incubations between the [3H]-methylated DNA substrate and 1pmole of protein (Figure 3.5). A two-way ANOVA test was used to compare the differences in mutant AGT DNA repair activity to wt AGT. All the results were statistically significant (p < 0.05). The rate constants were calculated for each time point, then plotted against time to illustrate the differences between the purified proteins (Figure 3.6). The rate constants for all the AGT proteins are listed in Table 3.1. All mutations in AGT reduced its rate of alkyl transfer. Among the mutant AGT proteins, triple mutant N157G/S159E/G160W had the highest rate of repair followed by G160W. The second triple mutant N157G/S159E/G160A had half the DNA repair activity as wtAGT. S159E had the lowest repair rate, approximately 8 times lower than wtAGT.
3.2.2 Inactivation of AGT proteins by O\textsuperscript{6}-benzylguanine (O\textsuperscript{6}-BG)

After establishing the range of activity with the mutant AGT proteins, we next used these purified proteins to assess sensitivity to the inhibitor, O\textsuperscript{6}-BG (Figure 3.7). In order to determine the extent of sensitivity of the purified proteins to the O\textsuperscript{6}-BG inhibitor, we treated AGT with various concentrations of O\textsuperscript{6}-BG. The purified proteins were treated with up to 10µM O\textsuperscript{6}-BG to calculate the ED\textsubscript{50} for each protein (Figure 3.8). After O\textsuperscript{6}-BG treatment, the remaining activity was then measured using \[^3\text{H}\]-alkylated DNA substrate in a filter-binding assay. The ED\textsubscript{50} is the concentration of drug needed to inactivate 50% of the protein activity, and was the measurement used to characterize O\textsuperscript{6}-BG sensitivity of each mutant AGT. We observed that all the mutants had a lower ED\textsubscript{50} than wt AGT, indicating an increase in O\textsuperscript{6}-BG sensitivity, with the exception of N157G (Table 3.2). The most significant results were observed for the G160W and N157G/S159E/G160W mutants. The ED\textsubscript{50} of G160W is 10 fold lower than wt AGT, and the ED\textsubscript{50} of N157G/S159E/G160W is 100 fold lower than wt AGT (Table 3.2). In contrast, the N157G mutant had increased O\textsuperscript{6}-BG resistance, with an ED\textsubscript{50} that is 3 fold greater than the ED\textsubscript{50} value of wt AGT.

3.2.3 Inactivation by O\textsuperscript{6}-methylguanine (O\textsuperscript{6}-MG)

O\textsuperscript{6}-BG is known to be a potent and specific inactivator of hAGT, partly due to efficient binding of the free base inside the active site of AGT. Based on the findings that G160W and N157G/S159E/G160W had such high sensitivity to O\textsuperscript{6}-BG, we wanted to determine whether the mutants would be sensitive to a weaker inhibitor, O\textsuperscript{6}-MG. O\textsuperscript{6}-BG and O\textsuperscript{6}-MG differ structurally only at the substitution at the O\textsuperscript{6} position (Figure 3.7). A smaller methyl adduct would allow us to understand the interplay between AGT amino acid residues and the binding of the inhibitor to the active site. The results in Figure 3.9 demonstrate that the N157G/S159E and
N157G/S159E/G160A mutants display increased $\text{O}^6$-MG resistance, relative to wt AGT. However, the G160W and N157G/S159E/G160W proteins exhibited slightly more sensitivity to $\text{O}^6$-MG than wt AGT. Interestingly, wt AGT had an ED$_{50}$ value of 0.8mM, while G160W had an ED$_{50}$ of 0.5mM, and N157G/S159E/G160W, an ED$_{50}$ of 0.6mM (Table 3.2). The substantial difference in reaction to the inhibitor seen for these mutants with $\text{O}^6$-BG was not observed for $\text{O}^6$-MG.

### 3.3 Sensitivity to BGFL

A unique technique of labeling AGT fusion proteins in living cells with fluorophores was described by the Johnsson group, as a method of studying dynamic intra-cellular processes (Keppler, 2004). The labeling was done using synthetic $\text{O}^6$-BG derivatives linked with fluorophores such as fluorescein, BGAF (Figure 3.7), oregon green and tetramethylrhodamine (Keppler, 2004). The fluorescein is attached to the benzyl group, which gets transferred to the active site cysteine of AGT, as shown in Figure 3.10. Our goal was to use this fluorescent $\text{O}^6$-BG derivative (BGAF) to study AGT turnover in CHO cells expressing wtAGT as well as mutant AGT. Because BGAF allows AGT to become fluorescent at the point of alkylation, we reasoned that we could use this drug to study the post-alkylation fate of the AGT protein. However, it was unknown if the presence of the large fluorescein moiety attached to the benzyl group would affect their reactivity with AGT and alter sensitivity to the inhibitor. Therefore, we examined the reactivity of the fluorescent $\text{O}^6$-BG derivatives (BGAF/BGFL) with our mutant proteins.

#### 3.3.1 In Vitro Labeling

In order to address the sensitivity of wt AGT to the fluorescent inhibitor BGFL, and the sensitivity of our detection system, we examined fluorescence using a gradient of wt AGT protein concentration and 10µM BGFL (Figure 3.11 A). Our results show that wt AGT reacts with BGFL, and surprisingly there was no reduction in reactivity due to the attached fluorescein group.
The fluorescence labeling technique is highly sensitive as we were able to detect as low as 0.001µM content of full length alkylated AGT protein.

Next, we tested the reactivity of BGFL with our $\mathrm{O}^6$-BG-sensitive mutant AGT proteins, and also a C145S mutant to confirm that the reaction occurred at the reactive cysteine (Figure 3.11 B). We found that all the mutants reacted with BGFL, and that the C145S mutant did not react.

### 3.3.2 Labeling AGT in CHO cells

We examined whether G160W and N157G/S159E/G160W could be labeled in CHO cells with fluorescent $\mathrm{O}^6$-BG derivatives. Mutant AGT genes were cloned into a pCMV mammalian vector and stably transfected into CHO cells. In these experiments, BGAF was used instead of BGFL for labeling. The additional acetyl groups result in increased diffusion through the cell membrane, and thus facilitates *in vivo* AGT labeling. The conversion from BGAF to BGFL occurs intracellularly, as shown in Figure 3.12. From *in vitro* experiments, we determined that AGT is sensitive to BGFL. However, we used excess BGAF concentrations for AGT labeling *in vivo*, that were optimized for wt AGT CHO cells. These concentrations enabled strong detection of fluorescently labeled AGT both visually under a fluorescent microscope, as well as in SDS-PAGE gels.

BGAF is specific for AGT, as we observed no labeling in empty-vector CHO cells that do not express AGT (Figure 3.13). In addition, BGAF labeling indicated that the wtAGT molecules were primarily localized in the nucleus. However, a small percentage of BGAF labeled AGT molecules are present in the cytosol. CHO cells expressing G160W had lower expression levels than wt AGT; moreover, the pattern of fluorescent labeling was significantly different from that of wtAGT (Figure 3.13). G160W was localized to specific locations within the nucleus, suggestive of being sequestered at the nuclear membrane. CHO cells expressing
N157G/S159E/G160W had very low expression and BGAF labeling could not be detected (data not shown).

Fluorescent labeling of wt AGT was stable for up to 3h. After this time, the pattern of labeling transitioned into punctate spots evenly dispersed over the cell (Figure 3.14).

3.4 Conclusions

Previous studies have shown that even minor structural alterations within the O₆-BG binding pocket of AGT cause resistance to the inhibitor (Pegg, 1993; Crone, 1994; Edara, 1996). These studies narrowed down a few critical residues within this binding pocket, whereby a single amino acid residue mutation resulted in O₆-BG resistance. One of these residues was a glycine at position 160. This site within AGT was important because a polymorphism exists in some human populations, resulting in a G160R variant AGT form (Imai, 1995). While G160R is highly resistant to O₆-BG, G160W and G160A were two rare mutants reported to show sensitivity to the inhibitor (Xu-Welliver, 1999).

There is significant homology in AGT around the active site of AGT between species, as shown in Figure 3.15. Interestingly, the corresponding amino acid residue at position 160 of the hAGT protein is different in the bacterial alkyltransferase Ada-C. There seems to be greater accessibility for the O₆-BG substrate in the binding pocket of hAGT, than in Ada-C. The presence of a tryptophan at site 161 in *E.coli’s* Ada-C alkyltransferase, while the presence of glycine at the corresponding site in hAGT is believed to be the cause of Ada-C resistance to O₆-BG (Rafferty, 1997).

Based on the crystal structure of hAGT, the sensitivity of G160W to O₆-BG was postulated to be caused by hydrophobic interactions between the protein and the inhibitor (Rafferty, 1997). In mammalian AGTs starting at Y158, there is a conserved sequence motif YSGG in humans and YSGGG in other mammals such as mouse or rat (Figure 3.15) (Rafferty,
1997; Wibley, 2000). Our experiments showed that the G160W mutant was tenfold more sensitive to $O^6$-BG than wt AGT. One possibility, for the cause of G160W sensitivity, is that the large ring of tryptophan might interact with the Y158 and stabilize the YSGG motif. It would be interesting to see if the sensitivity to $O^6$-BG caused with a tryptophan at position 160 would be retained if Y158 was mutated, disrupting this motif. Studies with G160R causing resistance to $O^6$-BG suggest that charge of the residue at position 160 might be a factor in inhibition, in addition to the size of the residue side chain (Edara, 1996). Fourteen of nineteen possible amino acid substitutions at G160 cause $O^6$-BG resistance in hAGT (Xu-Welliver, 1999). From our observations, the other point mutants N157G and S159E, did not have increased sensitivity to $O^6$-BG to the same extent as G160W. In the case of N157 and S159, their side chains were suggested to form a hydrogen bond with the N7 of $O^6$-BG and the Cβ of S159 made contact with $O^6$-BG (Xu-Welliver, 1999). Our results support the findings of Juillerat et al. that AGT sites 157,159 and 160 are critical to $O^6$-BG binding (Juillerat, 2003). The combination of tryptophan at position 160, along with glycine at 157 and glutamic acid at 159 probably create the least steric hinderance and ideal binding surface for the benzyl group of $O^6$-BG. This conclusion is supported by our ED$_{50}$ values for triple mutants N157G/S159E/G160A and N157G/S159E/G160W. While triple mutant N157G/S159E/G160A has a similar ED$_{50}$ value to G160W, triple mutant N157G/S159E/G160W had an ED$_{50}$ value one hundred fold more than wtAGT. This result is a rare one as the vast majority of alterations to AGT’s binding pocket cause resistance to $O^6$-BG.

This information might be useful in improving the design of future AGT inhibitors, using molecular modeling within the binding site based on the known crystal structure of hAGT. In this way, these findings could improve the efficacy of current chemotherapeutic strategies.

In order to determine which part of the $O^6$-BG molecule, either the benzyl group or the guanine group, reacts with the tryptophan and causes sensitivity we tested the mutant proteins with AGT inhibitor $O^6$-MG. $O^6$-MG was developed as the first AGT inactivator as it is an analog
of a major physiological AGT substrate. It was demonstrated that wt hAGT has a low affinity for $O^6$-MG and a slow rate of reaction, and therefore, not effective enough for clinical testing (Pegg, 1995). $O^6$-BG was designed based on a biomolecular displacement reaction between AGT and the leaving group at the $O^6$ position of guanine. Benzyl groups enter biomolecular reactions more readily than methyl groups because of the stabilization offered by the electron density of the phenyl ring in the transition state (Dolan, 1990). The sensitivity of only two mutants, G160W and N157G/S159E/G160W to $O^6$-methylguanine indicates that the purine group of $O^6$-BG is indeed the moiety causing sensitivity in the mutants. A tryptophan is present at the equivalent site of the Ada alkyltransferase found in E.coli, and it is resistant to $O^6$-BG. This suggests that the tryptophan in hAGT has important interactions with other residues of the binding pocket and this is what causes the sensitivity to both inactivators observed in our experiments. The two other mutant proteins tested N157G/S159E and N157G/S159E/G160A were both resistant to $O^6$-MG. Thus, these results of the differing sensitivities of the mutants to the two AGT inhibitors provide evidence that the purine ring of $O^6$-BG allows an interaction with the tryptophan side chain, possibly better stacking of the aromatic rings, and causes the mutants G160W and N157G/S159E/G160W to have greater sensitivity to $O^6$-BG than wt AGT. The sensitivity of mutants S159E and N157G/S159E to $O^6$-BG highlights the required interaction between the benzyl group and the amino acids at codons 157 and 159. Optimal binding of $O^6$-BG occurs because of mutations at codons 157 and 159, but this sensitivity is then further increased in combination with the addition of a tryptophan at site 160.

Our results have shown that mutating AGT at positions 157, 159 and 160 does change the rate constant of AGT. The DNA repair activity of mutants at these sites is less efficient than the wt protein. However, reaction of these mutants with the pseudo-substrate, $O^6$-BG is significantly better than wt AGT. Changes to all three sites, as in the case of N157G/S159E/G160W caused the most dramatic change in sensitivity to the inhibitor. Another
important finding of these results is the effect of a benzyl group in the inactivating agent. When \( O^6\)-BG is replaced with \( O^6\)-MG, we observed that all the AGT proteins are more resistant to inactivation. The smaller methyl adduct is probably not recognized and stabilized within the active site, as a free base.

BGFL and BGAF have proven to be extremely useful tools in labeling AGT, both \textit{in vitro} and \textit{in vivo}. We were successful in detecting an AGT reaction within seconds. In addition, we have shown that BGAF and BGFL are specific for AGT, and react well with all the mutant AGT proteins. For our purposes, however, we could not use BGAF to study AGT degradation in living cells because BGAF gets photobleached easily and cannot be detected past 2-3h. It will be interesting to investigate the cause of the speckled feature seen when wtAGT expressing CHO cells were treated with BGAF. It might be a symptom of how the cells process excess \( O^6\)-BG. Co-localization studies of BGAF alkylated AGT and intra-cellular organelles would be helpful to understanding this observation.

The presence of a tryptophan at position 160 causes a change in AGT localization in CHO cells which might be due to the instability of the protein \textit{in vivo}, or the recognition by a co-factor that keeps it sequestered in pockets within the nucleus. The difference seen in labeling between the wt AGT and G160W CHO cells did not allow us to use the mutant AGT cells for further investigation into the \textit{in vivo} properties of AGT. We instead used CHO cells expressing wt AGT for these studies.
Table 3.1. Rate constants of Wt and Mutant AGT proteins determined from mean values from 4 experiments indicated in Figure 3.5 and depicted in Figure 3.6.

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Wt</th>
<th>N157G</th>
<th>S159E</th>
<th>G160W</th>
<th>N157G/ S159E</th>
<th>N157G/ S159E/ G160A</th>
<th>N157G/ S159E/ G160W</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.32</td>
<td>0.181</td>
<td>0.0</td>
<td>0.24</td>
<td>0.32</td>
<td>0.78</td>
<td>0.42</td>
</tr>
<tr>
<td>0.25</td>
<td>5.4</td>
<td>1.78</td>
<td>0.0</td>
<td>2.05</td>
<td>0.70</td>
<td>1.64</td>
<td>2.3</td>
</tr>
<tr>
<td>0.50</td>
<td>7.5</td>
<td>3.75</td>
<td>0.118</td>
<td>3.46</td>
<td>2.12</td>
<td>2.28</td>
<td>5.09</td>
</tr>
<tr>
<td>1.00</td>
<td>14.3</td>
<td>7.75</td>
<td>1.22</td>
<td>7.39</td>
<td>3.2</td>
<td>5.00</td>
<td>8.75</td>
</tr>
<tr>
<td>1.67</td>
<td>25.9</td>
<td>7.18</td>
<td>1.45</td>
<td>11.5</td>
<td>4.98</td>
<td>7.5</td>
<td>15.3</td>
</tr>
<tr>
<td>3.33</td>
<td>40.1</td>
<td>15.6</td>
<td>3.45</td>
<td>20.8</td>
<td>7.95</td>
<td>14.7</td>
<td>22.8</td>
</tr>
<tr>
<td>hAGT Protein</td>
<td>$O^6$-Benzyliguanine $\text{ED}_{50}$</td>
<td>$O^6$-Methylguanine $\text{ED}_{50}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------------</td>
<td>-------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wildtype hAGT</td>
<td>0.4 mM</td>
<td>0.8 mM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N157G</td>
<td>1.2 mM</td>
<td>*nd</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S159E</td>
<td>0.2 mM</td>
<td>*nd</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G160W</td>
<td>0.04 mM</td>
<td>0.5 mM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N157G/S159E (Double Mutant-GE)</td>
<td>0.09 mM</td>
<td>&gt; 5 mM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N157G/S159E/G160A (Triple Mutant-GEA)</td>
<td>0.05 mM</td>
<td>&gt; 5 mM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N157G/S159E/G160W (Triple Mutant-GEW)</td>
<td>0.005 mM</td>
<td>0.6 mM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*nd: not determined

**Table 3.2.** Comparison of ED$_{50}$ values between $O^6$-BG and $O^6$-MG.
Figure 3.3 Purified AGT proteins stained by Coomassie Blue. 1µg was loaded onto a 15% polyacrylamide gel. Point mutants: N157G, S159E, G160W; Double mutant (GE): N157G/S159E; Triple mutants (GEA): N157G/S159E/G160A and (GEW): N157G/S159E/G160W.
Figure 3.4 DNA repair activity of purified AGT proteins. Activity for each protein was calculated by using the activity of 1µg of protein, which is an excess of AGT protein, as the highest value of activity or 100%. The lower concentration activity percentages were then standardized against this value. (A) DNA repair activity of the point mutants: N157G, S159E and G160W. (B) DNA repair activity of double mutant (GE): N157G/S159E, triple mutants: (GEA): N157G/S159E/G160A and (GEW): N157G/S159E/G160W.
Figure 3.5 Timecourse reaction of purified AGT proteins with $[^3]$H-alkylated DNA. Data points are a mean of 4 experimental results, except for N157G, which was based on 8 experiments. The error bars represent standard deviation across results. A two-way ANOVA test shows statistical significance. *  p<0.05  
**  p<0.0005
Figure 3.6. Wt and Mutant AGT rate constants. The mean values from Figure 3.5 were used to calculate the rate constants for each protein.

The rate constant, $K_T = \frac{1}{C_{oa} - C_{ob}} \ln \left[ \frac{C_{ob} (C_{oa} - C_c)}{C_{oa} (C_{ob} - C_c)} \right]$. 

$C_{oa} =$ initial concentration of AGT = 1 pmole or $0.03 \times 10^{-6}$ moles/L 

$C_{ob} =$ initial concentration of $[^3]$H DNA = $3.03 \times 10^{-9}$ moles/L 

$C_c =$ Concentration of product at time t 

$C_c$ was calculated by using the mean cpm at each timepoint
**Figure 3.7.** Structures of inhibitors used to test sensitivity of mutant AGT. The only difference between all the inhibitors was the size of the adduct at the $O^6$ position of guanine.
Figure 3.8 (A) Inactivation of AGT using O\textsuperscript{6}-BG. Protein concentrations were adjusted to account for differences in rate constant, in order to determine an accurate ED\textsubscript{50} value. For wt AGT, 10ng of protein consistently produced an ED\textsubscript{50} value of 0.4µM. 10ng of N157G, 40ng of S159E, 10ng of G160W, 30ng of double mutant N157G/S159E (GE), 35ng of triple mutant N157G/S159E/G160A (GEA) and 20ng of N157G/S159E/G160W (GEW) was used to determine the ED\textsubscript{50} value. The proteins were treated with O\textsuperscript{6}-BG for 30min before the [\textsuperscript{3}H]-DNA substrate was added. The graph shows the mean results of two experiments, with each time point run in duplicate. (B) Magnification of O\textsuperscript{6}-BG inactivation curves. A line at 50% AGT activity is indicated to give an approximation of the ED\textsubscript{50} value for each mutant.
Figure 3.9 Inactivation of AGT using $O^6$-methylguanine ($O^6$-MG). Mutants determined to have highest sensitivity to $O^6$-BG were tested for $O^6$-MG activity. Only one point mutant, G160W was tested, along with all the combination mutants: double mutant (GE): N157G/S159E; triple mutant (GEA): N156G/S159E/G160A and (GEW): N156G/S159E/G160W.
Figure 3.10 Labeling AGT with BGAF or BGFL. Upon reaction with cysteine at the active site of AGT, the benzyl group of BGFL/BGAF forms a covalent bond. The fluorescein is transferred along with the benzyl group to the active site causing AGT to be fluorescent upon alkylation.
Figure 3.11 Reactivity of purified wt and mutant AGT with BGFL. (A) Gradient labeling of wtAGT with BGFL. A range of AGT concentrations from 0.001µM to 1µM was allowed to react with 10µM BGFL for 5min to test the reactivity of AGT to BGFL. (B) Labeling Wt AGT and mutant AGT with BGFL. 500ng of protein was treated with 1mM BGFL for 5min at 37°C. The reaction was stopped by added loading SDS buffer to the sample and then boiling it for 5 min.

Figure 3.12. Conversion of BGAF to BGFL
Figure 3.13 CHO cells expressing wtAGT, G160W or empty vector after 1h BGAF labeling.
Figure 3.14. Wt-AGT CHO cells labeled with BGAF. To label a majority of AGT molecules, 80µM BGAF was added to the media and incubated for 1h. The media was removed and cells washed extensively with PBS, before being visualized in OPTI-MEM, under a fluorescence microscope. Images were taken at 0h and 3h after incubation.
Figure 3.15 Conserved Amino acid sequences between species, around the AGT active site. The PCHR motif at the active site of AGT is highly conserved between species. Stars (*) indicate the conserved YSGG motif in mammalian AGTs.
Chapter 4
Pathways of Post-alkylation AGT Degradation

4.1 Rationale

In the last 30 years, our understanding of many aspects of AGT function and regulation have been greatly expanded. However, one area that remains unclear is the fate of hAGT upon alkylation. When AGT is unalkylated, it is stable; however upon alkylation is rapidly degraded (Xu-Welliver, 2002). The sequence of events which lead to AGT degradation are unknown. Particularly, it is unknown if once AGT is alkylated, ubiquitination causes release from DNA. Evidence suggests that DNA-bound alkylated AGT, or an inactive AGT mutant (C145A) impede repair by active AGT proteins (Edara, 1999). A further question in the field is whether ubiquitin plays a role in releasing AGT from DNA. Studies in mammalian cell lines have shown that AGT, once alkylated, is degraded via the ubiquitin-proteasome pathway (Srivengopal, 1996; Xu-Welliver, 2002; Major, 1997). Daniels et al showed that a conformational change occurs when an alkyl group is added onto AGT (Daniels, 2000). However, other investigators have revealed that this conformational change does not greatly reduce AGT binding to DNA (Rasimas, 2003). More specific studies trying to decipher the ubiquitination site of AGT have been unsuccessful thus far. When lysine to arginine point mutations and combination mutations were introduced into human AGT, ubiquitination still occurred, suggesting that secondary ubiquitination sites exist (Xu-Welliver, 2002). In 2009, Hwang et al described two pathways that target AGT degradation in S. cerevisiae (Hwang, 2009). The two pathways, which act synergistically, are the Ubr1/Rad6 dependent N-end rule pathway and the ubiquitin fusion degradation (UFD) pathway. The N-end rule implies that certain residues at the amino terminus of a protein are destabilizing and is the primary determinant of the half-life of that protein. These amino acids at the amino terminus can
act as degradation signals. The N-end rule pathway has been described in all organisms examined from yeast, bacteria, plants and mammalian cell lines (Varshavsky, 1997). In addition, as was hypothesized earlier by Xu-Welliver (Xu-Welliver, 2002), the UFD pathway implies that an internal degron exists within AGT that is recognized by the ubiquitination machinery.

Addition of GFP to the N-terminus of AGT produces an active fusion protein (Choi, 2004; Remington, 2008). We therefore used this construct to study the localization of AGT before and after alkylation, as well as its degradation. In addition, a comparison of the half-life of GFP-AGT to the wt-AGT was performed to determine the effect of GFP on AGT degradation.

### 4.2 Determination of AGT Half-Life

#### 4.2.1 Determination of wt AGT Half-Life in CHO cells

Sequences corresponding to the wt hAGT cDNA were expressed from a pCMV-Neo-Bam vector (Figure 2.1B). wtAGT expressing CHO cells were treated with cycloheximide to inhibit de novo protein synthesis. Cells were then followed for 28 hours in the presence or absence of O$_6$-BG. Upon alkylation due to O$_6$-BG, we measured the half life of wt AGT to be 9h (Figure 4.1). However, when wt AGT is unalkylated, it is stable and has a half-life >24h. This is consistent with results seen using the same CHO cell expression system by Xu-Welliver et al, in a shorter time frame of 8 hours (Xu-Welliver, 2002).

#### 4.2.2 Effect of N-terminal GFP tag on hAGT Half-Life

CHO cells stably expressing GFP-AGT were created to study intracellular dynamics of AGT. GFP has a molecular weight of 28 kDa and therefore a GFP-AGT fusion protein is double the size of AGT alone (Figure 4.2). Fluorescent imaging of GFP-AGT shows that the distribution of GFP-AGT is similar to that seen in Chapter 3 of wt AGT labeled with BGAF. The fusion protein is located primarily in the nucleus with a small percentage present in the cytosol (Figure
4.3). In contrast, GFP-P140K AGT, an $O^6$-BG resistant AGT mutant (Choi, 2004), is entirely nuclear. While the use of GFP allows for novel localization studies, the effect of such a large tag on the normal intracellular processing of AGT was not known. We conducted the same experiment as described above to study the half-life of alkylated GFP-AGT in CHO cells. We determined that GFP-AGT is much more stable upon alkylation, than wt-AGT. The half-life of alkylated GFP-AGT was found to be 26h (Figure 4.4). Unalkylated GFP-AGT had a half-life >26h.

**4.3 Ubiquitination of wt hAGT and GFP-AGT**

We wanted to expand knowledge of the known ubiquitination pattern of wt AGT in CHO cells. Therefore, we treated wt AGT CHO cells with an alkylating agent, $O^6$-BG and a proteasomal inhibitor, MG132 or lactacystin. MG132 functions by preventing protein hydrolysis at the 26S proteasome. In the absence of an alkylating agent, but when MG132 or lactacystin is present, we see the formation of higher molecular weight AGT species after 16h. These bands, based on their size (all greater than 30kDa) correspond with Ub-conjugated AGT species (Figure 4.5A). When cells are exposed to $O^6$-BG, in tandem with MG132, we see the presence of these species at earlier times (8h) indicating that the alkylation event leads to an increase in the formation of Ub-AGT. Similarly, when the same wt AGT CHO cells were treated with BGAF and MG132 we detected fluorescent, higher molecular weight AGT species at the 8 and 16 h timepoints (Figure 4.5B).

Experiments were conducted to conclusively prove that the higher molecular weight species detected upon $O^6$-BG and MG132 treatment, were indeed Ub-AGT forms. Ub-AGT isolation from CHO cells was carried out using immunoprecipitation with AGT polyclonal antibodies, and double immunoprecipitation using AGT antibodies first followed by Ub
antibodies. Both these techniques were unsuccessful in clearly detecting the Ub-AGT forms due to technical reasons.

Next, we examined whether the same high molecular weight species would be formed in CHO cells expressing GFP-AGT. Treatment of GFP-AGT CHO cells with 80µM O6-BG and MG132, did not produce any higher molecular weight bands, despite numerous repetitions. However, when cells were treated with 80µM BGAF and MG132, we were able to detect higher molecular weight bands corresponding to Ub-conjugated GFP-AGT. These species formed in the same time frame as did wt AGT (Figure 4.6).

**4.4 Fluorescent Recovery after Photobleaching (FRAP) in GFP-AGT CHO cells**

In order to repair DNA, AGT must have a mechanism by which it can scan the chromatin within the nucleus. It was unknown if AGT was bound to DNA and therefore immobile within the nucleus or if it was subject to diffusion. We used FRAP to study the mobility and dynamics of AGT within the nucleus as well as transport from the cytoplasm into the nucleus. The GFP-AGT CHO cells have a very robust fluorescent signal (Figure 4.4). In a FRAP experiment, a defined region of a cell is irreversibly bleached by a laser pulse for 12.5 seconds. Photobleaching damages the fluorophores within a specific region of the cell rendering them non-fluorescent. In our GFP-AGT CHO cell line, photobleaching small areas within the nucleus produced recovery within a few seconds and could not be accurately measured. Therefore, we expanded the bleaching area to half the nucleus and then determined how quickly equilibrium was reached. Cells had a pre-bleached fluorescence of 120 units in both the bleached and unbleached regions of the nucleus (Figure 4.7A). Upon bleaching by the 488 He/Ar laser, equilibrium is reached within 100 seconds. This result implies that the AGT molecules are moving rapidly within the nucleus. Single cells in which half the nucleus was bleached showed fluorescence recovery after 15min (Figure 4.7B).
To understand how quickly AGT molecules are transported from the cytosol into the nucleus, we bleached the entire nucleus and measured fluorescent recovery. We found that the rate of movement from the cytosol into the nucleus is very slow, as is determined by the fact that recovery occurs only after 30 minutes (Figure 4.8A and B). This is evidence that the majority of AGT proteins reside in the nucleus and are probably transported there soon after synthesis. Therefore, if the entire nucleus is bleached, the rate of recovery is equal to the amount of time it takes for newly synthesized AGT molecules to get transported into the nucleus. Several control experiments were conducted using the same parameters used for bleaching the nucleus. Bleaching in the cytoplasm did not cause a detectable loss of fluorescence in the nucleus, indicating that fluorescent molecules were saturated in the nucleus. In addition, a second control demonstrated that there was no fluorescence recovery in paraformaldehyde fixed cells.

4.5 Conclusions

The use of GFP as an epitope tag at the N-terminus of AGT has provided us with a unique method of studying AGT. It enabled us not only to visualize AGT in living cells, but also to understand how the amino terminus is involved in AGT degradation. Initially, we used the GFP-AGT CHO cell line to determine if AGT distribution and intracellular localization changed, upon alkylation. Recent evidence has demonstrated that proteasomes are homogeneously distributed between the nucleoplasm and cytosol (Salomons, 2010). Our expectation was to detect a percentage of alkylated GFP-AGT being transported to the cytosol for proteasomal degradation. Treatments with O\(^6\)-BG or MNNG revealed that there were no such changes in distribution upto 16 hours post-alkylation (Figure 4.9). These results are probably due to the fact that the GFP fluorescent signal is so robust in these cells that subtle changes in localization could not be detected. From fractionation studies done using wt-AGT expressing CHO cells, we detected Ub-
wtAGT species in both the nuclear and cytosolic fractions (Figure 5.2) This supports the formation of these species within the nucleoplasm.

One of our most significant findings was that the addition of GFP to AGT changes the half-life of the alkylated protein from 9 h to 26 h. The N-end rule pathway states that there is a hierarchy of amino acids that are de-stabilizing and act as degradation signals when present at the amino terminus of a protein. A primary destabilizing residue, such as histidine, arginine, tryptophan or lysine is recognized directly by the E3 ubiquitin ligases along with an internal degradation signal, while secondary and tertiary residues require secondary modifications such as oxidation or arginylation to become a primary destabilizing residue (Varshavsky, 2008).

Secondary residues such as aspartic acid or glutamic acid get arginylated and are then recognized as primary destabilizing residues. Tertiary amino acid residues, asparagine and glutamine are deaminated and converted to aspartic acid and glutamic acid respectively and then arginylated to become a substrate for the E3 ubiquitin ligases (Varshavsky, 2008). AGT has an aspartic acid at it’s amino terminus, and it is likely that the aspartic acid gets arginylated and then functions as a primary destabilizing residue. In our GFP-AGT construct, the N-terminal residue is a valine that functions as a stabilizing residue (Tasaki, 2007). Therefore, these results are evidence for the N-end rule functioning on hAGT. These results are consistent with the N-end rule and indicate that inhibition of AGT degradation by the N-end rule pathway, through a fusion with GFP causes a stabilization of the alkylated form of AGT. This preliminary evidence suggests that the N-end rule does have an important role in the degradation of mammalian AGT. An experiment to conclusively prove the N-end rule pathway targeting hAGT is to mutate aspartic acid of AGT to a stabilizing residue in the CHO cell construct and we should detect a stabilization of the alkylated AGT similar to the half life of alkylated GFP-AGT. Previously, this pathway had been described only in yeast (Hwang, 2009). The detection of a putative Ub-conjugated GFP-AGT species after BGAF treatment, is evidence of the presence of an internal degron in the GFP-AGT protein.
which is recognized by Ub ligases. Therefore, we conclude that hAGT is targeted by two proteolytic pathways: the N-end rule pathway and a second ubiquitination pathway, possibly the UFD pathway. There might be redundancy in AGT degradation because of the need to regulate this constitutively expressed protein. This is indicated by the observation that we did not detect cytotoxicity in CHO cells that express GFP-AGT cells. With the N-end rule pathway blocked in these CHO cells, there is no deleterious effect on cell viability. Based on the findings of the two pathways which regulate AGT degradation in yeast (Hwang, 2009), we can extrapolate the possible players in mammalian AGT degradation (Table 4.1) (Ciechanover, 1984; Hwang, 2009; Park, 2009; Koken, 1991; Finley, 1984). A simple method to test if the HECT domain of TRIP-12 is indeed an E3 ligase of the human AGT protein would be to use siRNA and knock down this protein in cells containing a high expression of AGT, and determine if there is a reduction in the development of Ub-AGT. Ideally, if the purified E1, E2 and E3 enzymes of the human ubiquitination pathway are available, an in vitro ubiquitination reaction could be performed using alkylated AGT as a substrate. The human E1 and E2 homologs, UBE1 and UBE2D1 or E2D1 are already commercially available (Millipore, Billerica, MA).

The detection of putative Ub-GFP-AGT species with BGAF and MG132 treatment indicates that GFP-AGT has greater reactivity with BGAF rather than $O^6$-BG. This is determined by the presence of higher molecular bands using Western analysis (Figure 4.6). Similar to results in Chapter 3, wt AGT and mutant AGT proteins had greater reactivity with the larger benzyl moiety attached to guanine ($O^6$-BG) rather than the methyl group of $O^6$-MG. This result is unique to the GFP-AGT fusion form of the protein, as wt-AGT showed no reactive disparity between $O^6$-BG and BGAF.

The results of our FRAP experiments illustrated for the first time the mobility of AGT within the nucleus. The bleaching of small regions of the nucleus was undetectable to the human eye, i.e. fluorescent recovery was above the range of measurement. Similar findings of rapid
diffusion have been reported for GFP-fusion nuclear proteins HGM-14 (nucleosome binding protein), SC35 (pre-mRNA splicing factor), and nucleolar proteins, nucleolin and B23 (Misteli, 2000). From FRAP experiments done in HeLa or baby hamster kidney cells, these GFP-nuclear protein fusions, displayed fluorescence recovery in 30 seconds. This is preliminary evidence for the hypothesis that AGT is constantly scanning the DNA for alkylation damage (Tubbs, 2007). FRAP experiments conducted with a GFP-fusion protein to NER endonuclease ERCC1/XPF, indicated that in the absence of DNA damage, the fluorescent fusion protein had free diffusion within the nucleus. However, upon damage the fusion protein displayed immobility, as mobility was only regained after 4 minutes (Houtsmuller, 1999). Thus, in the case of ERCC1/XPF the protein does not follow a processive scanning mechanism, but rather participates in NER in a distributive manner.

Our results of GFP-AGT transport from the cytosol into the nucleus showed that the fluorescent recovery was dramatically different compared to the recovery within the nucleus (Figure 4.7, 4.8). This suggests that AGT is transported into the nucleus over a long period of time, probably soon after its synthesis. The use of fluorophores in addition to the GFP-AGT CHO cell model might provide further insight into the dynamics of AGT within the cell.
Table 4.1 Comparison of yeast ubiquitination enzymes and the human homologs of these enzymes.

<table>
<thead>
<tr>
<th>Ub enzyme</th>
<th>Yeast (S. cerevisiae)</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>Uba1</td>
<td>UBE1</td>
</tr>
<tr>
<td>E2</td>
<td>Rad6 or Ubc4</td>
<td>UBE2D1 or E2D1</td>
</tr>
<tr>
<td>E3</td>
<td>Ufd4 Ubr1</td>
<td>TRIP12 (HECT domain)</td>
</tr>
</tbody>
</table>
Figure 4.1 Half-life determination of alkylated and unalkylated wt-AGT in CHO cells.

(A) Cells were treated with O6-BG and cycloheximide and harvested at various timepoints. AGT protein content was analyzed using Western analysis and plotted as a percent protein remaining against time. (B) Cells were treated with cycloheximide alone and harvested at various timepoints.
**Figure 4.2** Protein expression levels in wtAGT CHO cells and GFP-AGT CHO cells. 60µg of cell extract was loaded onto SDS-PAGE gels. PVDF membrane was probed with antibodies to (A) hAGT; (B) GFP; (C) GAPDH. The fusion GFP-AGT protein has a molecular weight of 50kDa.
**Figure 4.3** Fluorescent image of GFP-AGT and GFP-P140K CHO cells. Fluorescent imaging of untreated GFP-AGT and GFP-P140K showed different intracellular localization between the two cell lines. GFP-P140K was entirely nuclear, while GFP-AGT was found in the nucleus as well as the cytosol.
**Figure 4.4** Half-life determination of alkylated (A) and unalkylated (B) GFP-AGT in CHO cells.

Determined as described in Figure 4.1.
Figure 4.5 (A) Detection of ubiquitinated AGT species in CHO cell extracts. Western analysis showing the formation of higher molecular weight bands in wt-AGT CHO cells upon treatment with proteasomal inhibitor, MG132. Monoclonal AGT Ab MT3.1 was used to detect AGT.
**Figure 4.5 (B)** Formation of higher molecular weight bands corresponding to AGT-Ub, when wt hAGT CHO cells are treated with BGAF and MG132, and isolated at various timepoints. The lower panel shows the Western analysis of a membrane probed by AGT Ab, MT3.1.
Figure 4.6 Formation of higher molecular weight bands corresponding to GFP-AGT-Ub, when GFP-AGT CHO cells are treated with BGAF and MG132/lactacystin. 60µg of cell extract was loaded to each lane. 1: empty vector CHO cell extract; 2: wt AGT CHO cell extract; 3: GFP-AGT CHO cell extract; 4: GFP-AGT + BGAF + MG132, 0h; 5: GFP-AGT + BGAF + MG132, 8h; 6: GFP-AGT + BGAF + MG132, 16h; 7: GFP-AGT + BGAF + lactacystin, 8h; 8: GFP-AGT + BGAF + lactacystin + MG132, 8h.
**Figure 4.7(A)** Fluorescent recovery after photobleaching (FRAP) in GFP-AGT CHO cells. Fluorescence was measured before and after photobleaching on a confocal microscope. Region for bleaching (half the nucleus of untreated GFP-AGT CHO) was selected and photobleaching was done for 12.5 sec. (B) Images taken on the confocal microscope, pre-bleach, post-bleach and after nuclear fluorescent recovery after 15 min. The white arrow indicates the nucleus undergoing photobleaching.
Figure 4.8 (A) FRAP in GFP-AGT CHO cells with bleached nuclei. This is an average result of experiments done in 24 cells. Fluorescent recovery increased within the nucleus over time, while the fluorescence decreases in the cytosol, indicating the flow of AGT molecules is towards the nucleus. The slow recovery of fluorescence imples that AGT accumulated in the nucleus over a very long period of time. (B) Images taken on the confocal microscope of the cell pre-bleach, post-bleach and after fluorescence recovery (30 minutes).
Figure 4.9 Imaging GFP-AGT CHO cells upon MNNG treatment. GFP-AGT expressing CHO cells were treated with 10µM MNNG for 0.5h and visualized over 16h. No change in intracellular distribution was detected.
Chapter 5

Effect of Endogenous N-terminal truncation on AGT function

5.1 Rationale

The crystal structure of hAGT reveals that the overall structure of the protein consists of two domains, an N-terminal domain (residues 1-85) and a C-terminal domain (residues 86-207) (Daniels, 2000), (Figure 5.1). The alkyltransferases across species show little sequence homology at the N-terminal domain, but have similar three dimensional structures. AGT variants found in C.elegans and Ferroplasma acidarmanus (F.acidarmanus) have DNA repair capacity and stable structures without the presence of an N-terminal domain (Kanugula, 2005; Kanugula, 2001). The function of the N-terminal domain other than to maintain the structural integrity of AGT, is unknown. The N-terminal domain is the site of four ligands bound to a Zinc atom, which is critical to maintaining overall structure of AGT and enhancing AGT DNA repair activity (Daniels, 2000; Rasimas, 2003). Most of the critical residues enabling AGT functionality are present in the C-terminal domain, including the reactive cysteine, at position 145 (Daniels, 2000; Wibley, 2000). In early studies, truncations at the N-terminal domain of hAGT were made to understand the effects on structure, DNA binding and repair function. Removal of the first 56 amino acids resulted exclusively in the formation of inclusion bodies, while the repair rate constant was unaffected (Bender, 1996). Other studies showed that a 31 amino acid truncation C-terminus had no effect on AGT DNA repair activity or O\textsuperscript{6}-BG sensitivity (Crone, 1994). However, when the C-terminal truncations were combined with further truncations at the N-terminus, the protein became unstable. A NΔ8CΔ31 construct had low expression and low activity. Removal of an additional residue from either end reduced AGT activity below detectable
range (Hazra, 1997). More recent studies showed that a 1-91 amino acid N-terminal domain construct surprisingly had mild activity, while a 92-207 amino acid C-terminal domain was completely inactive. Only when the two domains were combined in freeze-thaw techniques, did the C-terminus regain DNA repair activity. It has been postulated that the source of the weak alkyltransferase activity shown with the N-terminal domain is due to an activated cysteine within this domain; Cys5 or Cys24. These cysteines are involved with zinc binding within the N-terminal domain and when either is mutated, alkyl transfer was abolished (Fang, 2005). This result indicates the critical role of zinc in maintaining the catalytic activity of the active site of AGT within the C-terminal domain.

In our studies using CHO cells, we detected a novel truncated AGT product in one of our cell lines. The results of characterizing this truncated AGT species are presented in this chapter. Our hypothesis was that a truncated form of AGT within cell extracts must be formed post-alkylation and as a part of the degradation cycle. We tested this using different reagents and conditions to determine the formation and stability of this truncated form. In addition, we carried out BGFL treatments to determine if the truncated species reacts with the pseudo-substrate. We replicated the formation of the truncated AGT species in in vitro experiments. Finally, we determined the cleavage site of the truncated species and cloned, purified and characterized the activity of the truncated species.

5.2 Detection of 16kDa species of hAGT

A CHO cell line expressing AGT with a His tag at the C-terminus of the protein was created with the intention of isolating AGT-Ub (See Figure 5.2). This AGT-His CHO cell line had comparable protein expression with the CHO cell line expressing wt-AGT. Six histidines replaced the last six amino acids of wt AGT; therefore, AGT-His has the same molecular weight as wt-AGT and migrates similar to the wt protein, on polyacrylamide gels.
5.2.1 CHO cells

A 16kDa species was first detected when CHO cell extracts expressing AGT-His were compared with the wt-AGT CHO cell extract. This species was detected by a penta-His antibody, in almost equal concentration to the full length AGT-His protein (Figure 5.2A). The 16kDa band was detected only by the penta-His antibody, and was found in untreated AGT-His cell extracts, as well as the lysate of cells that were treated with 80µM O6-BG and 25µM MG132 (Figure 5.2A). The AGT antibody ATO-1 did not detect this band. The epitope of ATO-1 is AGT residues 8-20 (Pegg, 1991). The AGT-His construct has a C-terminal His tag, therefore, the size and detection of the 16kDa band suggested that there was a cleavage occurring at the N-terminus. This 16kDa band was undetected in both the treated and untreated wt AGT CHO cell extracts. Both AGT antibodies detected the full length AGT in both cell lines, and in the presence and absence of O6-BG and MG132 treatment (Figure 5.2A and 5.2B). Western blot analysis showed the band running at around 15-16kDa.

To determine whether the 16kDa band is an AGT species, we used multiple AGT antibodies in Western analysis. We also wanted to discover if the formation of this species was specific to the His-tagged cell line, or if it could be detected in the wt-AGT CHO cell line. The polyclonal Novus AGT Ab detected a 16kDa band in untreated wt AGT cell extracts (Figure 5.2B). In addition, we loaded cell pellets lysed by boiling in SDS-PAGE loading buffer, directly to the gel to discover if an artifact was being created during experimental manipulation. However, the truncated species was still present in wt AGT CHO cells (Figure 5.2B), suggesting that this species is created within cells prior to lysis. We probed cell pellets from both wt AGT and AGT-His cell lines with a C-terminal AGT antibody, ATO-3. The ATO-3 AGT antibody was raised to AGT residues 197-207 (Pegg, 1991). The ATO-3 antibody detected a band at the 16kDa position in the wt-AGT cell pellet (Figure 5.2B). However, because of the presence of a histidine tag in
the AGT-His cell extract we could not use the C-terminal antibody to detect a corresponding band in the AGT-His sample.

The 16kDa band also was detected in fractionated samples of AGT-His, in both the nuclear and cytoplasmic fractions (Figure 5.2C). This result shows that the truncated AGT species is present in almost equal concentrations to the full-length protein in both sub-cellular fractions.

5.2.2 Human cell lines

Our analysis focused thus far on the existence of the truncated AGT species in a recombinant AGT construct, which highly overexpresses AGT within CHO cells. These cells have no endogenous alkyltransferase activity. Also, we wanted to determine if we could detect such a species in human cell lines which also express high levels of AGT. It has been well documented that human cancer cells resistant to chemotherapy express high levels of AGT (Santibanez-Koref, 2002; Dolan, 1997; Gerson, 2002). We probed KB, a variant of HeLa cells, and HeLa cells with the polyclonal Novus AGT Ab. Both these cell lines showed the presence of a 16kDa band (Figure 5.3A). In a similar experiment, we also probed various cell lines with the C-terminal AGT Ab, ATO-3 (Figure 5.3B). An additional cell line, HT-29 was used, which is a human colon carcinoma cell line. Unfortunately, there were many non-specific bands detected with the ATO-3 antibody around 15kDa, highlighted by the presence of a band seen at this position in control empty-vector CHO cells. However, a strong band at the 16kDa position was detected in wtAGT-CHO cells and in KB cells. In summary, our results demonstrate the existence of a truncated AGT species in HeLa and KB cells. The existence of the truncated form in other human cell lines was not definitively determined, due to the lack of specificity with the available AGT antibodies.
5.3 In vitro studies of 16kDa species

We created an *in vitro* system to study the stability and formation of the 16kDa AGT species, as well as its reactivity with the fluorescein tagged BG derivative, BGFL. This system consisted of empty-vector CHO cell extract or HeLa cell extract, to which 1μg of pure AGT-His was added and incubated for various times at 37°C. Using this system, we were able to test the effect of a His tag at the N-terminus versus the C-terminus of AGT, and determine the effect of the His tag on cleavage of the full-length protein and the formation of the truncated AGT species.

Our result demonstrates formation of the 16kDa polypeptide, under *in vitro* conditions, from pure AGT. A 15 min incubation at 37°C of hAGT (C-terminal His tag) with the empty vector CHO lysate produced a truncated species of the exact size as seen in the cell extracts made from CHO cells expressing AGT-His (Figure 5.4). We used this *in vitro* system to determine if alkylation enhances the formation of truncated AGT species. For these experiments, we pre-alkylated the pure protein (N-terminal His tag) by incubation with BGFL at 37°C for 30 min. This alkylated protein was then added to the empty-vector cell lysates for varying times, and the proteins were separated by SDS-PAGE. The gel was scanned first for fluorescent-labeled AGT bands and the proteins were then transferred onto a PVDF membrane. The membrane was then probed with a monoclonal AGT antibody (MT3.1). The result shows the formation of a fluorescent 16kDa band from full-length AGT at all the timepoints examined, from 15 to 60 min, in the presence of both extracts of CHO and HeLa cells (Figure 5.5). This band was not detected by MT3.1, as seen in the cell culture results (Figure 5.2A and B).

To examine whether the 16kDa species was created by an N-terminus cleavage event, we also probed the membrane with a RGS-His antibody to serve as a negative control. As expected the 16kDa species was not detected. We repeated this experiment by reacting N-terminal His AGT with BGFL in HeLa cells at shorter incubation periods, from 5 to 15 min (Figure 5.6).
Similar to the previous result, we observed a fluorescent band corresponding to the 16kDa species in all the samples. The polyclonal AGT antibody also detected the 16kDa band corresponding to the fluorescent band. Therefore, we determined that a His tag at the N-terminus does not inhibit the formation of the truncated 16kDa species.

We also studied a C-terminus tagged AGT in the *in vitro* system. Figure 5.7A shows the formation of the truncated AGT species, when pre-alkylated AGT is added to empty-vector CHO cell extract. In all the samples, we found that the pre-alkylated AGT was rapidly truncated upon addition to CHO cell extracts. To expand on this timecourse, we pre-alkylated the pure AGT-C terminal His tag, with O\(^6\)-BG and incubated it from 2 to 60 min, in the presence or absence of calf-thymus DNA. The use of calf-thymus DNA was an attempt to stabilize AGT and determine if this would affect the formation of the truncation species (Figure 5.7B). Unalkylated AGT added to the extracts, also formed the 16kDa band, within 60 minutes. The absence or presence of calf thymus DNA does not alter the formation of the 16kDa AGT species.

Finally, we wanted to assess if the truncated species can react directly with the pseudosubstrate BGFL. We designed an experiment in which unalkylated C-terminal His AGT was added to the extracts for 5 min to generate truncated AGT species and then incubated with BGFL for 1 min or 10 min. These samples were compared with the BGFL treated AGT added to cell extracts. As shown in Figure 5.7C (lanes 3,4,8 and 9) a fluorescein labeled 16kDa band appeared in the samples that contained unalkylated AGT incubated with empty-vector CHO cell extract. These results suggest that the truncated AGT species still retains alkyltransferase activity and it reacted with the fluorescent inhibitor, and the formation of the truncated 16kDa AGT species was not due to alkylation of AGT.
5.4 Identification of 16kDa species

In order to conclusively prove that the truncated species was formed from hAGT and to identify the cleavage site of the protein, we used mass spectrometry analysis. We did a pulldown of the truncated band from AGT-His CHO cells, using the histidine tag at the C-terminus of AGT and metal affinity chromatography. We subjected the protein to in-gel trypsinization, followed by mass spec analysis. The truncated species was compared with the full-length purified AGT protein. Table 5.1 shows the mass spectrometry results, which showed the peptides of the truncated species matched with wt AGT. This analysis identified the truncated species as hAGT.

Next, we analyzed the N-terminus of the truncated species using Edman sequencing, and compared the pure hAGT protein (Figure 5.8A) with the pulldown sample from AGT-His CHO cells (Figure 5.8B). The results demonstrated the first 50 amino acids of AGT were not present in the 16kDa AGT species (Figure 5.9).

5.5 Biological Characterization of 16kDa AGT form

Once we had determined the cleavage site of AGT resulting in the truncated species, we wanted to establish whether it was capable of repairing DNA. In order to test this the sequence coding for the 16kDa polypeptide was cloned into a pQE30 bacterial expression vector, and the recombinant protein was purified from E.coli. We analyzed the purified protein preparation for alkyltransferase activity and also tested the ability of the 16kDa protein to protect E.coli from MNNG cytotoxicity as means to assess the in vivo DNA repair activity.

5.5.1 In vitro DNA repair activity

A single methionine was placed directly upstream of the cleavage site determined for the 16kDa species, in order to express it in E.coli using the pQE30 vector. However, we discovered that this species was unstable, the yields were extremely low with significant background
contamination. In an attempt to increase the protein yields, we transformed the expression plasmid into two other bacterial strains, BL21-DE3 and GWR109. BL21-DE3 lacks two bacterial proteases, Lon and OmpT. GWR109 lacks the bacterial alkyltransferases Ada and Ogt and allows us to detect any alkyltransferase activity due to expressed proteins (Samson, 1991). All the bacterial strains used gave similar low protein yields (Figure 5.10A and 5.10B).

A filter binding alkyltransferase assay was performed using these partially purified 16kDa proteins, as described in Section, 2.5.1. Briefly, the proteins were incubated with $[^3]H$ methylated DNA for 30 minutes. The concentration of $[^3]H$ methyl groups transferred onto AGT is determined by measuring radioactive protein collected on nitrocellulose filters. Protein purified from BL21-DE3 showed greater activity than those purified from the XL1-Blue and GWR109 bacterial strains (Figure 5.11A). Importantly, the alkyltransferase activity of the 16kDa protein preparation from all three strains was inhibited by pre-incubation with $O^6$-BG. Since the bacterial alkyltransferases are highly resistant to $O^6$-BG (Elder, 1994), this result indicates that the activity observed in this assay is specific to the 16kDa truncated AGT. While the expression level of the 16kDa was very low, we were still able to detect DNA repair activity, in both the BL21-DE3 and the GWR109 bacterial strains. Therefore, we next tested if this activity would provide protection to *E.coli* GWR109 from MNNG induced cytotoxicity.

5.5.2 GWR109 *E.coli* MNNG Protection

The GWR109 bacterial strain lacks alkyltransferase expression, making it susceptible to damage induced by alkylating agents, such as MNNG. Our result showed that the empty-vector GWR109 strain was highly sensitive to MNNG (Figure 5.12A). The GWR109 strain expressing wt AGT had 100% survival at the lower doses of MNNG tested. The 16kDa AGT showed statistically significant protection of the GWR109 strain at the low doses of 1mg/mL and 2.5mg/mL, compared to the empty-vector strain. Two-way ANOVA of the dose response curves
resulted in a p value <0.0001. This result confirms the DNA repair activity of the 16kDa AGT species, in spite of low expression. Figure 5.12 B shows the induction of the three GWR109 bacterial lines used in the MNNG cytotoxicity test.

5.6 Conclusions

In this chapter we describe the formation of a novel 16kDa truncated human AGT species that has intrinsic DNA repair activity, both in vitro and in vivo. This truncated AGT species lacks a large portion of the N-terminal domain. Absence of this domain is similar to the alkyltransferases found in C.elegans and the archaeon F.acidarmanus (Kanugula, 2005; Kanugula, 2001). The biological significance of the 16kDa AGT truncation species is that it has the capability to repair DNA, as supported by the finding that the 16kDa AGT species protects E.coli from MNNG, as well as its sensitivity to O6-BG (Figure 5.11A, 5.12A). This 16kDa AGT truncated species was undetected in all prior work using CHO cell lines because most AGT antibodies have epitopes at the N-terminus of AGT (i.e., MT3.1 and ATO-1 AGT antibodies). It is unlikely that the truncation is an artifact of AGT overexpression in CHO cells because of two reasons. First, we were able to detect formation of this truncation product in CHO cell extracts as well as HeLa cell extracts (Figures 5.5 and 5.6). A 16kDa band was detected in KB and HeLa cells despite these cells having a much lower AGT expression level than our AGT CHO cell lines (Figure 5.3). Second, this species was undetected in CHO cells which overexpress GFP-AGT, indicating that a specific recognition site on AGT was altered with the creation of a fusion protein.

This 16kDa truncated AGT species appears to form independently of alkylation, as supported by our in vitro results (Figure 5.7). Therefore, the 16kDa AGT species is constantly being created in the cell. Our in vitro results showed that a histidine tag either at the amino or carboxyl terminus does not affect the formation of the 16kDa species (Figure 5.4, 5.7). In
addition, we successfully showed that the 16kDa species, once formed, can rapidly react with BGFL. This evidence disproved our initial assumption that the truncated AGT occurs due to alkylation of AGT.

We did not detect a truncation product in GFP-AGT CHO cell lines. The presence of GFP may inhibit the putative protease from recognizing the cleavage site. This is supported by the fact that we were never able to find an intermediate ubiquitinated 16kDa AGT species in any of our studies. Conjugation of a single Ub monomer to the 16kDa AGT species would result in a 24kDa species that would migrate very close to the full length AGT protein. We would need better separation during electrophoresis in order to detect a single Ub monomer linked truncation species. Adding MG132 did not affect the expression level of the 16kDa AGT species. If ubiquitinated forms of the truncated AGT species are not found, this raises the possibility that the 16kDa AGT species is part of a Ub-independent degradation pathway. This observation would also explain why only a small percentage of Ub-AGT is detected in cells, despite having high AGT expression (Figure 5.2A).

It is likely that the cleavage of hAGT is caused by a specific protease rather than a broad family of proteases, as the CHO cell extracts that the 16kDa species was first identified in, had a broad protease inhibitor cocktail. A search on the MEROPS database (http://merops.sanger.ac.uk) revealed that there are three possible proteases that can cleave amino acid sequence A/AVL. They are a matrix metallopeptidase-9, and 2 signal peptidase complexes. Similarly, there are 6 possible proteases that can cleave amino acid sequence PA/A. These proteases are calpain-2, elastase-2, cathepsin G, Htra2 peptidase, aarA-type peptidase, signal peptidase complex. Further investigation testing these proteases with inhibitors might reveal the mechanism of the formation of this 16kDa AGT species \textit{in vivo}.

An interesting future direction for expanding our current understanding of the 16kDa AGT species would be to explore the effect of residues surrounding the truncation site, around
codon 50. A mutational analysis of these amino acids would reveal the conditions under which AGT remains a substrate to the putative protease. In addition, such an analysis would help identify the protease causing the truncation, as well as determine if there are one or more proteases involved.

It is not surprising that this truncated AGT species retains DNA repair activity, and was also shown to be sensitive to O6-BG, as many studies indicated that the critical residues for AGT DNA repair activity reside in the C-terminus (Daniels, 2000). A surprising finding is that in this AGT truncated species 3 of the 4 ligands needed to bind the zinc atom are missing, yet the 16kDa species is still active. The Zn ion enhances the DNA repair reaction rate of hAGT and contributes to protein stability (Rasimas, 2003). While the residues involved in Zn binding are conserved in mammalian AGTs, they are absent among the amino acid sequences of microbial AGTs. Zn was absent in the crystal structures of E. coli’s Ada-C terminal domain or the alkyltransferase from Pyrococcus kodakaraensis (Moore, 1994; Hashimoto, 1999). However, this lack of the Zn atom and the N-terminus, is probably the reason that the protein was unstable and had extremely low yields upon purification in the E.coli system. In addition, in the pQE30 vector, there was a single methionine added to the 16kDa sequence for purification. This could be another factor contributing to the low purification yields. If a construct were made which resulted in the exact 16kDa sequence found in vivo, for example using an AGT fusion protein and a Tev protease, we may increase the yield of this 16kDa AGT polypeptide.

The existence of a 16kDa AGT that is capable of DNA repair sheds light onto AGT regulation and degradation. It will be interesting to study if the formation of this polypeptide exists as part of a separate degradation cycle or if it serves a specific role within the cell. A future line of experimental investigation would be to determine if this AGT truncation species can repair RNA, in which case it might be biologically relevant in the cytosol. In addition, a second
cytosolic substrate for this species is mitochondrial DNA. Both these results would help elucidate the role of this novel AGT truncation species within the cell.
Table 5.1. Mass Spectrometry results of 16kDa species in-gel tryptic digest peptide matches with purified hAGT.

<table>
<thead>
<tr>
<th>Tryptic Peptide</th>
<th>hAGT amino acid residue position</th>
</tr>
</thead>
<tbody>
<tr>
<td>(K)/FGEVISYQQLAALAGNPK/(A)</td>
<td>108-125</td>
</tr>
<tr>
<td>(R)/GNPVPILIPCHR/(V)</td>
<td>136-147</td>
</tr>
<tr>
<td>(R)/VVCSSGAVGNYSGGLAVK/(E)</td>
<td>148-165</td>
</tr>
<tr>
<td>(K)/EWLLAHEGHR/(L)</td>
<td>166-175</td>
</tr>
<tr>
<td>(R)/LGKPGGLGGSSGLAGAWLK/(G)</td>
<td>176-193</td>
</tr>
<tr>
<td>(K)/GAGATSGSHHHHHH</td>
<td>194-207</td>
</tr>
</tbody>
</table>
Figure 5.1. Secondary Structure of hAGT highlighting residues critical for hAGT structural and catalytic activity (Daniels, 2000). The zinc atom is shown in purple. The reactive cysteine at the active site of AGT is at position 145.
**Figure 5.2.** Detection of 16kDa band in AGT-His CHO and wt AGT CHO cell extracts. Samples were run on a 15% polyacrylamide gel and examined by Western analysis.

**A** CHO cells: AGT-His. Lane 1: 10ng pure AGT protein; Lane 2: 60µg untreated AGT-His CHO cell extract; Lane 3: 60µg AGT-His CHO cells +BG+ MG132 extract; Lane 4: 60µg untreated wt AGT CHO cell extract.

**B** Detection of 16kDa band in wt AGT CHO cell extract. Duplicate samples of wt-AGT CHO lysate probed by a monoclonal AGT Ab, MT3.1 and a polyclonal Novus AGT Ab. Lane 1: 80ng pure AGT; Lane 2 and 3: 60µg wt AGT CHO cell extract; Lane 4: wt AGT CHO cell pellet.

**C** Detection of 16kDa species in CHO cell extracts using ATO-3 AGT antibodies (C-terminal epitope AGT residues 197-207).
Figure 5.2 (D) Fractionation of AGT-His CHO cells. Western analysis of two fractionation experiments, run on the same gel. Empty-vector CHO cells were fractionated and run on the gel, and probed with penta-His Ab, as a control. Lane 1: wt AGT nuclear fraction; Lane 2: wt AGT cytosolic fraction; Lane 3: empty-vector CHO cell nuclear fraction; Lane 4: empty-vector CHO cell cytosolic fraction.
Figure 5.3. Detection of 16kDa band in human cancer cell lines (A) Polyclonal AGT (Novus) Antibody was used to probe 60µg (Lanes 1 and 2) and 100µg (Lanes 3 and 4) of KB and HeLa cell extract.  (B) Detection of 16kDa band in human cancer cell lines using a C-terminus AGT antibody (ATO-3). Stars (*) indicate non-specific bands.
**Figure 5.4.** *In vitro* formation of 16kDa species. Lane 1: 1.4 µg of pure AGT (C-His tag); Lane 2: 1.4µg of pure AGT (C-His tag) incubated for 15min with empty vector CHO lysate at 37°C; Lane 3: 100µg AGT-His CHO cell extract. The incubation produced a truncated species of the exact size as seen in the cell extracts. The membrane was probed with a penta-His Ab.
**Figure 5.5.** Formation of 16kDa species using purified N-His AGT in empty-vector CHO extracts and HeLa extracts. This figure shows the N-terminal His tagged AGT, treated with BGFL. The gel was scanned before the western transfer, to image the fluorescence bound to AGT Panel(A). The PVDF membrane was then probed with MT3.1 AGT Ab Panel(B) and an RGS-His Ab Panel(C).

1: N-His AGT unalkylated  
2: pre-BGFL alkylated N-His AGT  
3: empty-vector CHO+unalkylated AGT 60min  
4:empty-vector CHO+ AGT alkylated with BGFL 15min  
5:empty-vector CHO+ AGT alkylated with BGFL 30min  
6:empty-vector CHO+ AGT alkylated with BGFL 60min  
7: HeLa+ AGT alkylated with BGFL 15min  
8: HeLa+ AGT alkylated with BGFL 30min  
9: HeLa+ AGT alkylated with BGFL 60min
Figure 5.6. Formation of 16kDa species from N-terminal His tagged hAGT by HeLa extracts. 16kDa species was detected in HeLa extract alone which corresponds to the same size and location of the 16kDa species seen in wtAGT CHO and *in vitro* incubation truncation species.
Figure 5.7(A) Formation of 16kDa species using C-terminal His tagged hAGT in empty-vector CHO extracts. Lane 1: pre-BGFL alkylated AGT; Lane 2: 5 min incubation of alkylated AGT in empty-vector CHO cell extract; Lane 3: 15min incubation of alkylated AGT in empty-vector CHO cell extract; Lane 4: 5 min incubation of alkylated AGT in empty-vector CHO cell extract, - ATP, - kinase (B) Timecourse of hAGT (C-terminal His) in empty-vector CHO extracts upto 60 minutes incubation, in the presence or absence of calf-thymus DNA.
Figure 5.7(C) Reaction of BGFL added to C-terminal His-hAGT in CHO cell extracts directly, leads to formation of 16kDa fluorescent band. 1.4µg AGT was added in each sample.

Lane 1: Marker; Lane 2: untreated pure hAGT; Lane 3: Unalkylated hAGT added to cell extracts with 1 minute addition of 50µM BGFL; Lane 4: Unalkylated hAGT added to cell extracts with 1 minute addition of 1mM BGFL; Lane 5: hAGT alkylated with BGFL; Lane 6: AGT alkylated with BGFL added to cell extracts for 5 minutes; Lane 7: AGT alkylated with BGFL added to cell extracts for 15 minutes; Lane 8: Unalkylated hAGT added to cell extracts with 10 minute addition of 50µM BGFL; Lane 9: Lane 4: Unalkylated hAGT added to cell extracts with 1 minute addition of 1mM BGFL.
Figure 5.8 Edman sequencing results identifying the cleavage site of the truncated AGT species. (A) Edman sequencing results showing the first 6 residues of full length AGT. The results corresponded with the known sequence of wt hAGT i.e. MDKDCE. Blanks and standards were run between samples.
Figure 5.8 (B) Edman sequencing results showing the first six residues of the truncated AGT species. The residues at the N-terminal cleavage site are AVLGGP.
Figure 5.9. Comparison of full-length hAGT-C terminal His and truncated 16kDa AGT amino acid sequences. The underlined sequence was the region of full length AGT which was truncated. The residues shown in bold red are the residues identified by Edman sequencing of the 16kDa truncation species. The numbers indicate amino acid residue position within AGT protein sequence.
Figure 5.10(A). Purification of 16kDa truncated hAGT from bacterial strains, XL1-Blue and BL21-DE3. (B). Protein expression of N-His AGT, 16kDa AGT, and Empty-vector in GWR109. Lane 1: Pulldown from AGT-HisCHO cell extract; Lane 2: 2.5µg GWR109 empty vector; Lane 3: 2.5µg GWR109 16kDa AGT; Lane 4: 50ng GWR109 N-his wtAGT.
Figure 5.11(A) $O^6$-BG inhibition of 16kDa AGT DNA repair activity. 100µg of purified 16kDa truncated AGT protein derived from the three bacterial strains, XL-1 Blue, BL21-DE3, GWR109, was pre-incubated with 10µM $O^6$-BG for 30 min, before addition of [$^3$H]-methylated DNA substrate. Measurement of [$^3$H]-methyl transfer to AGT was conducted using a filter binding assay, as described in Section 2.5.1. (B) DNA repair activity of 16kDa AGT in GWR109 and empty-vector GWR109.
Figure 5.12(A) MNNG Protection with 16kDa AGT species. GWR109 strains expressing wt-His AGT, 16kDa AGT and empty vector were in the logarithmic growth phase, and exposed to varying doses of MNNG for 30 min. The bacteria were then pelleted, resuspended in M9 media, diluted and plated in triplicate. The colonies were then counted and the data plotted as a percentage of cell survival. Error bars represent the standard deviation between two experiments. A two-way ANOVA of this result had a p value < 0.0001. (B) Western analysis of GWR109 bacterial strains, before and after IPTG induction. The positive control was His-tagged polypeptides pulled down with His-tag binding resin, from AGT-His expressing CHO cell lysate. Wt-AGT GWR109 showed positive induction after four hours. 16kDa GWR109 was uninducible and empty-vector GWR109 was used as a control.
Chapter 6

General Discussion

AGT is a unique mechanism for DNA repair and is highly conserved among species within the bacterial and animal kingdoms. Over the last 20 years, we have expanded our understanding of the extent of AGT homology between species, its mechanism of action, its preferred substrates and its physical structure. One aspect of hAGT that remains to be delineated is its regulation and degradation. Upon alkylation, hAGT continues to bind to DNA, and evidence suggests that this binding impedes repair by other AGT proteins or the enzymes of the NER pathway (Edara, 1999; Samson, 1988). Thus, degradation might be the only mechanism to remove the alkylated AGT from DNA.

The goal of this dissertation was to determine the fate of hAGT upon alkylation, and broaden our understanding of ubiquitination as a pathway for AGT degradation. As summarized in Chapter 1, hAGT has an important role to play as a chemotherapeutic target. In this context, we set out to discover under what conditions the AGT protein becomes sensitized to the specific inhibitor, $O^6$-BG. Our mutational analysis of three codons within the $O^6$-BG binding pocket, illustrate protein conditions under which alternative inhibitors might have increased efficacy. Our first major finding, described in Chapter 3, is that tryptophan at residue 160, in place of a glycine, causes a tenfold increase in $O^6$-BG sensitivity; in combination with the N157G and S159E mutations, G160W replacement causes a one hundredfold increase in $O^6$-BG sensitivity. These results emphasize that the stacking of aromatic rings between the tryptophan side chain and the benzyl ring of $O^6$-BG facilitate enhanced interactions at the active site resulting in increased sensitivity of AGT to $O^6$-BG. As described in Chapter 3, these results can be used in modeling
studies together with the known crystal structure of AGT, to improve the design of $O^6$-BG and create a more potent inhibitor of hAGT.

Two observations in this thesis suggest that further investigation must be done to study the effect of AGT mutation towards $O^6$-BG reactivity, \textit{in vivo}. First GFP-P140K, with a single amino acid residue change, becomes entirely nuclear (Figure 4.4). Second, G160W while having promising reactivity to $O^6$-BG, showed a change in intracellular distribution to the nuclear membrane (Figure 3.13). Since, most studies on AGT mutations have been conducted almost entirely \textit{in vitro}, our results provide insight into the possible impact of these mutations on the \textit{in vivo} functioning and role of these mutations on AGT function. An interesting study would be to take the known polymorphic variants of AGT, and compare their intracellular distribution with the wt AGT. This experiment might demonstrate novel mechanisms of resistance to $O^6$-BG.

Our second major finding is that the addition of GFP to the amino terminus of hAGT causes stabilization in the degradation of AGT, in that the half-life changes from 9h with wt hAGT to 26h for the fusion protein (Chapter 4). The half life of alkylated GFP-AGT is similar to the half-life of unalkylated hAGT, which was greater than 24h. These data suggest a dramatic shift in degradation pattern of the GFP-AGT from the wt AGT. However, there were other aspects of GFP-AGT cell physiology that were unchanged from wt AGT. When CHO cells expressing GFP-AGT were studied, we observed the same pattern of intracellular distribution (Figure 4.4), as observed with the wt-AGT treated with BGAF (Figure 3.13). This specifically indicates that the transport of GFP-AGT from the cytosol into the nucleus, and retention of GFP-AGT within the nucleus, was unaffected by the increase in molecular size caused by the GFP fusion. Furthermore, alkyltransferase activity of GFP-AGT and GFP-P140K proteins expressed in CHO cells were comparable to wt AGT (Figure 2.2). Finally, the putative ubiquitination pattern of GFP-AGT seen upon alkylation and proteasomal inhibition was not altered with the addition of GFP (Figure 4.6).
Extrapolating from results found in yeast by Hwang et al., we hypothesize that GFP at the amino terminus blocks the functioning of the N-end rule degradation pathway (Hwang, 2009). Wild-type hAGT has an aspartic acid at its amino terminus, which functions as a secondary destabilizing residue. However, our GFP-AGT construct had a valine at its amino terminus, which is a stabilizing residue. Consistent with the N-end rule, we observed the half-life of wt AGT to be 9h, whereas the half-life of GFP-AGT is 26h. This is preliminary evidence that the N-end rule pathway targets AGT in mammalian cells. The physiological role of the N-end rule pathway is unclear at present, because N-end rule deficient strains of *S. cerevisiae* are viable (Tasaki, 2007). The formation of Ub-GFP-AGT is evidence that an internal degron also is present within AGT that is recognized by a ubiquitin conjugating enzyme upon alkylation. Furthermore, this result suggests that there is redundancy with the ubiquitin proteolytic systems targeting AGT.

This observation suggests that it is critical to remove AGT bound to DNA in order for DNA repair systems to continue scanning DNA for damage. We currently know all the degradation pathways and ubiquitin enzymes targeting the alkyltransferase (Mgt1) in *S. cerevisiae* (Hwang, 2009). An analysis of known homologous ubiquitin enzymes in humans reveals that UBE1 could function as the E1 enzyme, UBE2D1 could function as the E2 enzyme, and the TRIP12 HECT domain could function as the E3 enzyme (Table 4.1). We could test the biological significance of the N-end rule pathway on AGT degradation in mammalian cells by knocking down TRIP 12, using siRNA. The identify whether the UFD pathway is the second proteolytic pathway targeting mammalian AGT, we could knock down the mammalian homolog of Ufd4, E6AP, and determine if there is a reduction in Ub-AGT species (Johnson, 1995). A possible complication is that there might be multiple mammalian homologs of Ufd4 that have not yet been described.

The construction of GFP-AGT expressing CHO cells enabled us to perform FRAP studies for the first time, which highlighted the rapid mobility of AGT within the nucleus. This result supports the DNA binding studies which indicate that AGT is constantly scanning DNA for
damage (Rasimas, 2007; Adams, 2009; Melikishvili, 2008). FRAP experiments using GFP fusions with other nuclear proteins such as HGM-14 (nucleosome binding protein), SC35 (pre-mRNA splicing factor), and nucleolar proteins, nucleolin and B23 have shown similar rapid diffusion rates in mammalian cells (Misteli, 2000).

Our third major finding is the detection of a truncated species of AGT within CHO and Hela cells (Chapter 5). This 16kDa species shows DNA repair activity (Figures 5.10, 5.11) and reactivity with O\(^6\)-BG (Figures 5.5-5.7), despite the absence of 50 amino acids at the amino terminus. This result adds information to the numerous studies previously performed using truncations to understand the effect of protein structure on AGT function (Fang, 2005; Bender, 1996; Hazra, 1997; Crone, 1994).

The truncation of the first 50 amino acids raises questions about the effect of the cooperative binding of AGT to DNA. As depicted in Figure 1.6, there is a strong interaction between AGT protein \(n\) and \(n+3\) in this model. One future direction to test the effect of this truncation on DNA binding would be to do a biochemical reconstruction of the purified 16kDa AGT species and DNA. An important difference in our result and many previous truncation studies is that the truncation event occurs endogenously and post-translationally. Therefore, the protein is probably folded in a manner that retains DNA repair activity. However, the truncation would be expected to impact cooperative binding. Therefore, an interesting study would be to determine the comparison in efficiency of DNA repair on short DNA oligonucleotides with purified 16kDa AGT vs wt hAGT.

As summarized in Chapter 5, this 16kDa AGT species could be part of a previously unknown degradation cycle (Figure 6.1). Our results, using cell culture extracts, demonstrated that alkylation is not a pre-requisite for the formation of this species. The fate of this 16kDa truncated form of AGT is as yet unknown. This raises interesting questions about how the cell degrades both the truncated and full-length versions of AGT. We were unable to detect a Ub-
16kDa intermediate species, due to the overlap in size with the full-length AGT protein. Proteins that are not ubiquitinated can still be degraded by the proteasome (Jariel-Encontre, 2008). Therefore, if the intermediate ubiquitinated-16kDa species does not exist, this does not eliminate the proteasome as its proteolytic endpoint. A few examples have now been described of proteins that are degraded by a ubiquitin-independent pathway. The first and best characterized is orinithine decarboxylase (ODC), which is targeted for degradation by an antizyme protein (Bercovich, 1989; Murakami, 1992). Many examples in which a single protein gets degraded by two different degradation pathways, have been described. The guardian of the genome, p53, is degraded both by ubiquitin-dependent and independent pathways (Brooks, 2006; Asher, 2005). This has also been demonstrated in the case of cyclin-dependent kinase inhibitor p21\textsuperscript{WAF1/CIF1}, and transcription factor c-Fos (Sheaff, 2000; Chen, 2004; Chen, 2007; Blagosklonny, 1996; Basbous, 2008; Piechaczyk, 1994). Intriguingly, c-Fos contains two destabilizing elements. The element at the N-terminal region controls ubiquitin-dependent degradation while the destabilizing element at the C-terminus is responsible for ubiquitin-independent degradation (Jariel-Encontre, 2008). The possibility of such elements at the N-terminus of hAGT is likely since its yeast homolog, Mgt1 has been proven to undergo ubiquitin-dependent degradation through the N-end rule pathway (Hwang, 2009). However, the existence of C-terminal destabilizing elements in AGT remains to be investigated.

Our results have demonstrated that the truncated species readily reacts with the pseudo-substrate $O^6$-BG. The 16kDa species should also repair DNA, as we observed protection from MNNG induced cytotoxicity in \textit{E.coli}. Therefore, a second possibility for the function of the truncated AGT is that it could play a biological role in the cytosol. The 16kDa AGT species is present in both the nuclear and cytosolic extract (Figure 5.2C). While RNA is not the physiological substrate of AGT, nevertheless, experiments have demonstrated that AGT does have the capability of repairing RNA alkylation damage (Pegg, 1988). Thus, RNA remains as a
possible substrate for the truncated AGT species within the cytosol. A simple test to determine if 16kDa AGT can repair RNA, would be to treat methylated tRNA with the purified 16kDa species and test for the free base using HPLC as described by Pegg et. al. (Pegg, 1988). A more provocative theory is that the 16kDa species repairs DNA in the mitochondria. From previous fractionation studies in rat liver, 1% AGT activity was found to be in the mitochondria (Pegg, 1983). We might be able to test if this AGT species occurs in the mitochondria by comparing the known internal signal sequences for mitochondrial translocation with that of the AGT sequence (Mokranjac, 2009). If we can find common mitochondrial translocation sequences, we would mutate these and do similar fractionation studies to eliminate AGT activity in the mitochondria. This would be a critical function since mitochondrial DNA undergoes extensive damage, and almost every genomic DNA repair pathway is represented in the mitochondria (Liu, 2010).

An experimental approach to assess the biological significance of the 16kDa AGT species would be to mutate the AGT sequence around the truncation site. If the mutant AGT cannot be recognized by the proteases we might see a disruption of degradation and accumulation of AGT within the cell. Another experimental approach would be to express the truncated AGT species in CHO cells and assess the level of protection afforded, from alkylating agents. This study would expand our understanding of the DNA repair capacity of the truncated AGT species. A 16kDa truncated AGT CHO cell line could be treated with BGAF and imaged to determine if there is a unique pattern of distribution within the cell. As described in Chapter 5, there are a few possible proteases that recognize the AGT sequence around codon 50. Some of these proteases, such as calpain-2 or cathepsin G are well characterized. To determine specifically which protease targets AGT as a substrate, we could incubate both the purified protease and purified wt hAGT in vitro. This investigation would decipher under what conditions AGT truncation occurred.

In summary, the studies presented in this dissertation describe methodologies of fluorescently tagging hAGT in vitro and in vivo. These experiments provide insight into the
broader mechanisms underlying AGT degradation. Our data supports information known about AGT ubiquitination in mammalian cells. Two proteolytic degradation pathways target hAGT for degradation; the N-end rule pathway and a second unknown pathway. In addition, we confirmed that three codons within AGT’s $O^6$-BG binding pocket cause greater sensitivity to reaction with $O^6$-BG, and could lead to sensitivity enhancement of future AGT inhibitors.
Figure 6.1. Model depicting the degradation of hAGT and the formation of the truncated 16kDa AGT species. The full-length hAGT is shown in blue. The alkyl adduct added onto AGT upon repair is indicated in yellow. A monomer of ubiquitin (Ub) is represented in orange. The molecular weight of a single Ub monomer is 8kDa, so its addition onto AGT results in a 30kDa AGT-Ub species. The truncated 16kDa AGT species is shown in purple. It can be formed from the truncation of the unalkylated or alkylated full-length AGT. The unalkylated truncated AGT species retains DNA repair capability.
Bibliography


tumors to DNA alkylating agents. *Cancer Research* 57, 5360-68.


Pegg, A., Wiest, L., Mummert, C., Stine, L., Moschel, R., and Dolan, M. (1991) Use of antibodies to human $\text{O}^6$-alkylguanine-DNA alkyltransferase to study the content of this protein in
cells treated with $O^6$-benzylguanine or $N$-methyl-$N'$-nitro-$N$-nitrosoguanidine. *Carcinogenesis* 12, 1679-1683.


Srivengopal, K., Mullapudi, S., Shou, J., Hazra, T., and Ali-Osman, F. (2000) Protein Phosphorylation is a regulatory mechanism for $O^6$-Alkylguanine-DNA Alkyltransferase in


VITAE

Candice L. Lewis

Name
Candice Loretta Lewis

Birth Place and Date
Mangalore, India
October 11, 1981

Marital Status
Married

Education
9/2003-Present:
Pennsylvania State University,
College of Medicine
IGDP in Genetics
Ph.D. Candidate

9/2001-2003
University of Minnesota,
Twin Cities, MN
BS in Genetics, Cell Biology
and Development

9/1999-2001
College of Saint Catherine, St Paul, MN
Candidate for BS in Biology,
transferred to University of Minnesota.

Work Experience
• Research (Pennsylvania State University, College of Medicine, Dr. AE Pegg’s Laboratory)

  The research focused on a DNA repair protein, $O^6$-Alkylguanine-DNA-Alkytransferase (AGT) and its degradation upon alkylation with a specific inhibitor $O^6$-Benzylyguanine (BG).

  (University of Minnesota, Dr. RG Johnson’s Laboratory)

9/2002- 5/2003. The research project focused on studying the effect of mutated
connexion-43 on gap junction mediated cell communication.

• Teaching